Journal of Statistical Physics, Vol. 122, No. 1, January 2006 (© 2006)
DOI: 10.1007/s10955-005-8063-1

Information Loss in Coarse-Graining of Stochastic Particle Dynamics

Markos A. Katsoulakis[1] and José Trashorras[2,3]

Received November 11, 2004; accepted June 29, 2005

Recently a new class of approximating coarse-grained stochastic processes and associated Monte Carlo algorithms were derived directly from microscopic stochastic lattice models for the adsorption/desorption and diffusion of interacting particles[12,13,15]. The resulting hierarchy of stochastic processes is ordered by the level of coarsening in the space/time dimensions and describes mesoscopic scales while retaining a significant amount of microscopic detail on intermolecular forces and particle fluctuations. Here we rigorously compute in terms of specific relative entropy the information loss between non-equilibrium exact and approximating coarse-grained adsorption/desorption lattice dynamics. Our result is an error estimate analogous to rigorous error estimates for finite element/finite difference approximations of Partial Differential Equations. We prove this error to be small as long as the level of coarsening is small compared to the range of interaction of the microscopic model. This result gives a first mathematical reasoning for the parameter regimes for which approximating coarse-grained Monte Carlo algorithms are expected to give errors within a given tolerance.

KEY WORDS: Coarse-grained Monte Carlo methods, Markov processes, Interacting particle systems, Information loss.
MSC (2000) subject classifications: 82C80, 60J22, 94A17.

1. INTRODUCTION

Problems in scientific disciplines ranging from materials science to the dynamics of macromolecules, to the spread of epidemics and climate modeling involve

[1] Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA-01003-9305, USA.
[2] CEREMADE, Université Paris-Dauphine, 75775 Paris Cedex 16, France.
[3] Address for Correspondence: José Trashorras, CEREMADE, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France; email: trashorras@ceremade.dauphine.fr

non-linear interactions across a vast disparity of scales ranging from the microscopic to the macroscopic. While microscopic simulation methods such as Molecular Dynamics and Monte Carlo (MC) algorithms can describe aspects of such complex systems, they are limited to short scales when compared to morphological features such as vortices, traveling waves or domain walls, that typically involve much larger *mesoscopic* scales. One natural way to by-pass this obstacle is to map the microscopic model into a model defined at a coarse intermediate scale, thus defining an *exact* coarse-grained model. Unfortunately, most of the "good" properties of microscopic models, in particular the Markov property of MC dynamics, are lost in this transition. This additional difficulty urges to define *approximating* mesoscale Markovian models. The main issue is that, since stochastic fluctuations inherited from the microscopics can be important (e.g. in self-organization problems characterized by coherent structures such as pattern formation[25]) they need to be properly included in any approximating model.

In addition to the aforementioned computational and mesoscale modeling challenges posed by the disparity in scales within the same model, in numerous applications a further "disparity in models" arises: in phenomena with detailed fluid/surface or boundary layer interactions arising in catalysis, epitaxial growth and chemical vapor deposition as well as in atmosphere/ocean science[11,17,21,28], it is necessary to couple microscopic, possibly stochastic models describing small scale dynamics on a surface (e.g. atoms, molecules or an active boundary layer), along with continuum Partial Differential Equations (PDE) for species, fluid and thermodynamic variables on the fluid phase overlying to the surface/boundary layer. It is therefore inevitable that features of the microscopic model will essentially enter as a *subgrid* effect in the coupling with the coarse computational grid of the macroscopic PDE model. In this case, the proper incorporation of stochastic and deterministic effects from the subgrid microscale is once again a critical element in the modeling and simulations.

To address certain aspects of these challenges, the authors in Ref.[11––13,15] started developing systematic mathematical strategies for the approximation of exact coarse-grained dynamics and the corresponding simulators. One of the principal goals in their papers was to address a large class of applied problems which are currently intractable with conventional microscopic simulation methods such as MC algorithms. They proposed a microscopically derived hierarchy of new approximating coarse-grained stochastic models–referred to as *Coarse-Grained* MC (CGMC)–ordered by the magnitude of space/time scales. The CGMC models are reminiscent of Multi-Resolution Analysis approaches to the discretization of operators[1]. They span length/time scales from the microscopic to the mesoscopic, and involve Markovian birth-death and generalized exclusion processes. Furthermore, this new set of Markovian models involves a reduced set of observables over the original microscopic models, still incorporating microscopic details and noise, as well as the interaction of the unresolved degrees of freedom. A key feature of

the procedure is that the full hierarchy of the derived stochastic dynamics satisfies detailed balance relations and as a result yields self-consistent random fluctuation mechanisms[12,13,15].

The primary goal of this paper is to rigorously justify CGMC models in non-equilibrium by estimating the information loss between *exact* and *approximating* coarse-grained adsorption/desorption lattice dynamics for any level of coarse-graining. By exact coarse-grained process we mean the block dynamic $(F(\sigma_t))_{t\geq 0}$ resulting from the true microscopic process $(\sigma_t)_{t\geq 0}$ while the approximating dynamics $(\eta_t)_{t\geq 0}$ is the one introduced in the aforementioned works[11,12,13,15]. More specifically, for a system of size N and every finite time horizon $T > 0$ we derive an error estimate between the distribution of $(F(\sigma_t))_{t\in[0,T]}$ (denoted $\mathcal{D}^{F(\sigma)}_{[0,T]}$) and the distribution of $(\eta_t)_{t\in[0,T]}$ (denoted $\mathcal{D}^{\eta}_{[0,T]}$). In contrast to the numerical analysis of finite element/finite difference approximations for PDE where the error is calculated in a suitable norm (e.g. Sobolev, L^p, etc.), here the error is measured in terms of the specific relative entropy and represents the loss of information per particle in the transition from the exact to the approximating models. We prove that

$$\frac{1}{N} H\left(\mathcal{D}^{F(\sigma)}_{[0,T]} \middle| \mathcal{D}^{\eta}_{[0,T]}\right) = T \times O\left(\frac{q}{2L+1}\right) \tag{1}$$

i.e. that if the range of interaction L is long enough compared to the coarse-graining level q the error resulting from the approximation is small. A crucial technical step in the analysis is the introduction of an auxiliary microscopic Markovian process which is directly *reconstructed* from the CGMC process and approximates within a $T \times O(\frac{q}{2L+1})$ error in relative entropy the original microscopic process; the reconstructed process may be of independent computational and modeling interest on its own.

Some aspects of coarse-grained modeling similar to CGMC simulations have to an extent been partly introduced in a related context in the existing literature[3,6,20,22]. Indeed, in recent years there has been a growing interest in developing coarse-grained models and simulators for microscopic systems arising in a broad spectrum of applications[4,7––9,10,23,24,27]. For example in the polymer science literature the coarse-graining of atomistic models of polymer chains is expected to yield models with fewer observables than the original microscopic system making them computationally more efficient than direct numerical simulations[5,23].

Concluding the Introduction we outline the structure of this paper. In Section 2 we discuss the recent work[12,13] on CGMC which addressed the coarse-graining of adsorption/desorption dynamics. In Section 3 we discuss the error analysis between exact and approximating CG processes and in Section 4 we briefly present some indicative numerical simulation results. In Section 5 we present the details of the proofs in the paper.

2. A HIERARCHY OF MARKOVIAN MODELS

To demonstrate the basic ideas, we introduce in Section 2.1 a dynamic Ising-type system as our microscopic model. In Section 2.2 we derive a complete hierarchy of exact and approximating coarse-grained models from the microscopic one[12,13].

2.1. Microscopic model

We consider a microscopic (or "fine-grained") Ising-type system evolving on a one-dimensional torus $\mathbb{T} = [0, 1)$. More precisely, we fix a positive integer N, that represents the inverse of the distance between particles. For every $x \in \{0, \dots, N-1\}$ we denote by $\rho(x/N)$ the number of particles in the interval $[x/N, (x+1)/N)$, that is $\rho(x/N) = 1$ if $[x/N, (x+1)/N)$ is occupied and 0 otherwise. Our observable $\rho = \{\rho(0), \dots, \rho(N-1/N)\}$ is therefore an element of $\mathcal{H}_N = \{0, 1\}^{\mathcal{L}_N}$, provided $\mathcal{L}_N$ stands for the discrete one-dimensional torus with N points: $\mathcal{L}_N = \frac{1}{N}\mathbb{Z} \cap \mathbb{T}$. The equilibrium states of the system are weighted according to the Gibbs probability measure

$$P_N^{fg}(\rho) = \frac{1}{Z_N^{fg}} \exp\left\{-\beta H_N^{fg}(\rho)\right\} \prod_{x \in \mathcal{L}_N} \nu(\rho(x))$$

where ν is the prior distribution $\nu(0) = \nu(1) = 1/2$, and

$$H_N^{fg}(\rho) = -\frac{1}{2} \sum_{x \in \mathcal{L}_N} \sum_{\substack{y \in \mathcal{L}_N \\ y \neq x}} J(x, y)\rho(x)\rho(y) + \sum_{x \in \mathcal{L}_N} h\rho(x).$$

The so-called partition function Z_N^{fg} is a constant making P_N^{fg} a probability measure. The interparticle potential J accounts for interactions between occupied sites and h models an external field. We consider symmetric, translation invariant potentials where by the integer $2L$ we denote the total number of interacting neighboring sites of a given point on $\mathcal{L}_N$. Since we consider periodic boundary conditions on $\mathcal{L}_N$, then for $2L + 1 = N$ we recover the case of long-range interactions. The interaction potential can be written as

$$J(x, y) = J(x - y) = \frac{1}{2L+1} V\Big(\frac{N(x-y)}{2L+1}\Big), \quad x, y \in \mathcal{L}_N, \tag{2}$$

where $V : \mathbb{R} \mapsto \mathbb{R}$ is twice continuously differentiable with $V(r) = V(-r)$, and $V(r) = 0$ if $|r| > 1/2$, accounting for possible finite range interactions.

The dynamics of Ising-type models considered in the literature consist of order parameter flips and/or exchanges that correspond to different physical processes[19]. In the present paper we will focus on a reversible flip dynamic, the so-called *Arrhenius dynamics*. It is set-up as follows: if ρ is the configuration

prior to a flip at x, then we denote the configuration after the flip by ρ^x,

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

Such flips occur at x with a rate $c_\sigma(x, \rho)$ given by

$$c_\sigma(x, \rho) = (1 - \rho(x)) + \rho(x) \exp\big(- \beta U_N^{fg}(x, \rho)\big)$$

where

$$U_N^{fg}(x, \rho) = \sum_{\substack{y \in \mathcal{L}_N \\ y \neq x}} J(x - y)\rho(y) - h$$

is the total energy contribution from the particle interactions with the particle located at the site $x \in \mathcal{L}_N$, as well as the external field h. The resulting dynamic is a continuous time Markov chain[18] which generator L_N^σ acts on the set $\mathcal{B}(\mathcal{H}_N)$ of bounded measurable functions defined on $\mathcal{H}_N$:

$$\big(L_N^\sigma g\big)(\rho) = \sum_{x \in \mathcal{L}_N} c(x, \rho)[g(\rho^x) - g(\rho)]. \tag{3}$$

2.2. Exact and approximating coarse-grained models

Let m and q be two integers such that $N = mq$. We consider a partition of $\mathbb{T}$ into m intervals $[k/m, (k + 1)/m)$, $k = 0, \ldots, m - 1$, thus defining a *coarse lattice*. Each of these intervals contains q adjacent intervals of the form $[x/N, (x + 1)/N)$, $x = 0, \ldots, N - 1$. This is equivalent to defining a partition $D_0, \ldots, D_{m-1/m}$ of $\mathcal{L}_N$ with $D_{k/m} = \{\frac{kq+l}{N}, 0 \leq l \leq q - 1\}$.

From the microscopic model, we define a *coarse-grained observable* on the coarse lattice. One such natural choice is the coverage over each coarse cell D_k:

$$\alpha(k) = \sum_{y \in D_k} \rho(y)\,, \quad k = 0, ..., m - 1/m\,.$$

The observable $\alpha = \{\alpha(0), \ldots, \alpha(m - 1/m)\}$ at this coarse scale takes values in $\mathcal{H}_{m,q} = \{0, \ldots, q\}^{\mathcal{L}_m}$ where $\mathcal{L}_m$ stands for the discrete one-dimensional torus with m points: $\mathcal{L}_m = \frac{1}{m}\mathbb{Z} \cap \mathbb{T}$. The map F naturally establishes a relation between the fine and the coarse scales:

$$\begin{aligned} F : \mathcal{H}_N &\to \mathcal{H}_{m,q} \\ \rho &\mapsto \big(\textstyle\sum_{x \in D_0} \rho(x), \ldots, \sum_{x \in D_{m-1/m}} \rho(x)\big). \end{aligned} \tag{4}$$

Conversely, for every $x \in \mathcal{L}_N$ we will denote by $[x] \in \mathcal{L}_m$ the index of the set D_k to which x belongs. Before we continue, we remark that in general neither

the *exact* equilibrium measure at the coarse scale $F(P_N^{fg})(\alpha) = \sum_{\substack{\rho\in\mathcal{H}_N \\ F(\rho)=\alpha}} P_N^{fg}(\rho)$ is a Gibbs measure nor the *exact CG process* $(F(\sigma_t))_{t\geq 0}$ *is a Markov Process*. This means that direct numerical simulation of this block model is not straightforward. Therefore our aim is to *define an approximating Markovian dynamic* $(\eta_t)_{t\geq 0}$ at the scale $\mathcal{H}_{m,q}$ that is suitable for fast computations and at the same time keeps track of the fluctuations of the exact CG model, i.e. a Markov process which distribution is close to $(F(\sigma_t))_{t\geq 0}$'s one. As a result, even if $(\eta_t)_{t\geq 0}$ has not exactly the same distribution as $(F(\sigma_t))_{t\geq 0}$, it will be feasible to simulate it easily via a MC algorithm since *it is a Markov Process*, and this simulation will be meaningful provided we have some control on the error resulting from this approximation. Computationally $(\eta_t)_{t\geq 0}$ is advantageous over $(\sigma_t)_{t\geq 0}$, since it has a substantially smaller state space and thus can be simulated more efficiently.

The approximating dynamic is obtained as follows[12,13,15]: First we derive with a direct calculation from the microscopic process the *exact* evolution equation for the coverage $(F(\sigma_t))_{t\geq 0}$ in the coarse cell D_k. Let $g \in \mathcal{B}(\mathcal{H}_{m,q})$, $\rho \in \mathcal{H}_N$ an initial state for the microscopic dynamic $(\sigma_t)_{t\geq 0}$ and $\alpha = F(\rho) \in \mathcal{H}_{m,q}$. We have

$$\frac{d}{dt}E_\rho^\sigma g(F(\sigma_t)) = \sum_{k\in\mathcal{L}_m}\left\{\sum_{x\in D_k} c_\sigma(x,\rho)\big(1-\rho(x)\big)\right\}\times[g(\alpha+\delta_k)-g(\alpha)]$$
$$+\sum_{k\in\mathcal{L}_m}\left\{\sum_{x\in D_k} c_\sigma(x,\rho)\rho(x)\right\}\times[g(\alpha-\delta_k)-g(\alpha)],$$

where E_ρ^σ stands for the mean-value w.r.t. the distribution of $(\sigma_t)_{t\geq 0}$ with fixed initial condition ρ, and $\delta_k \in \mathcal{H}_{m,q}$ is the configuration with a single particle at the site $k \in \mathcal{L}_m$. The terms

$$\bar{c}_{\mathrm{a}}(k,\rho) := \sum_{x\in D_k} c_\sigma(x,\rho)\big(1-\rho(x)\big), \quad \bar{c}_{\mathrm{d}}(k,\rho) := \sum_{x\in D_k} c_\sigma(x,\rho)\rho(x), \tag{5}$$

are the *exact* coarse-grained rates for adsorption and desorption in a coarse cell D_k.

Next we follow the main idea in the CGMC procedure[12,13,15] and express these exact coarse-grained rates, up to a controlled error, as functions of the "coarse-grained" random variable $\alpha = F(\rho)$, rather than the microscopic ρ. It yields from (5) the generator for a Markov process $(\eta_t)_{t\geq 0}$. We refer to[12] where this derivation is formally carried-out. The obtained update rate with which the value $\alpha(k)$ is increased by 1 (adsorption rate of a single particle in the coarse cell

D_k) and decreased by 1 (desorption in D_k) respectively are:

$$c_{\eta,a}(k,\alpha) = [q - \alpha(k)]\,, \quad c_{\eta,d}(k,\alpha) = \alpha(k)\exp[-\beta(U^{cg}_{m,q}(k,\alpha))]\,, \tag{6}$$

where

$$U^{cg}_{m,q}(k,\alpha) = \sum_{\substack{l\in\mathcal{L}_m \\ l\neq k}} \bar{J}(l,k)\alpha(l) + \bar{J}(0,0)(\alpha(k)-1) - h\,.$$

The coarse-grained potential $\bar{J}$ is defined as the average of all contributions of pairwise microscopic interactions between coarse cells and within the same coarse cell:

$$\bar{J}(k,l) = \frac{1}{q^2}\sum_{x\in D_k}\sum_{y\in D_l} J(x-y) \tag{7}$$

for $k, l \in \mathcal{L}_m$, $k \neq l$, and

$$\bar{J}(l,l) = \bar{J}(0,0) = \frac{1}{q(q-1)}\sum_{x\in D_l}\sum_{\substack{y\in D_l \\ y\neq x}} J(x-y). \tag{8}$$

The first rate in (6) is obtained exactly while the second is obtained from (5) with an error of the order $O(q/2L+1)$, when replacing ρ by $F(\rho)=\alpha$. The resulting generator $L^{\eta}_{m,q}$ of the approximating coarse-grained Markov process $(\eta_t)_{t\geq 0}$ is

$$\begin{aligned}(L^{\eta}_{m,q}g)(\alpha) = \sum_{k\in\mathcal{L}_m} \big\{ & c_{\eta,a}(k,\alpha)[g(\alpha+\delta_k) - g(\alpha)] \\ & + c_{\eta,d}(k,\alpha)[g(\alpha-\delta_k) - g(\alpha)]\big\}\end{aligned}$$

for every $g \in \mathcal{B}(\mathcal{H}_{m,q})$. The stationary measure for $(\eta_t)_{t\geq 0}$ is a canonical Gibbs measure defined by

$$P^{cg}_{m,q}(\alpha) = \frac{1}{Z^{cg}_{m,q}}\exp\big(-\beta H^{cg}_{m,q}(\alpha)\big)P_{m,q}(\alpha)\,,$$

where the product binomial distribution

$$P_{m,q}(\alpha) = \prod_{k=1}^{m}\frac{q!}{\alpha(k)!(q-\alpha(k))!}\left(\frac{1}{2}\right)^q$$

is the prior distribution arising from the microscopic prior by including q independent sites, and $H^{cg}_{m,q}$ is the coarse-grained Hamiltonian defined by

$$H^{cg}_{m,q}(\alpha) = -\frac{1}{2}\sum_{l\in\mathcal{L}_m}\sum_{\substack{k\in\mathcal{L}_m\\ k\neq l}}\bar{J}(k,l)\alpha(k)\alpha(l) - \frac{\bar{J}(0,0)}{2}\sum_{l\in\mathcal{L}_m}\alpha(l)\big(\alpha(l)-1\big) + \sum_{l\in\mathcal{L}_m}h\alpha(l). \tag{9}$$

The same-cell interaction term $\alpha(l)\big(\alpha(l)-1\big)$ yields the global mean-field theory when the coarse-graining is performed beyond the interaction parameter L while at the other extreme it is consistent with the Ising case $q = 1$. As a result we obtain a complete hierarchy of models spanning from Ising ($q = 1$) to mean-field. In fact, the specific form of the self-interaction $\alpha(l)\big(\alpha(l)-1\big)$ in (9) ensures the detailed balance condition for adsorption/desorption mechanisms which guarantees the proper inclusion of fluctuations in the approximating coarse-grained model as they arise from the microscopics[15].

3. INFORMATION LOSS IN THE APPROXIMATION SCHEME

Exact and approximating coarse-grained processes share the same asymptotic mean behavior[13,15], i.e. their averages solve the same mesoscopic deterministic PDE

$$c_t = d_0\big[1 - c - \exp(\beta h)c\exp\big(-\beta V * c\big)\big]$$

in the long-range interactions case $N = 2L + 1$. In addition to comparing asymptotic mean behaviors, we would like to understand how well and in what regimes CGMC captures the fluctuations of the exact CG system. As a first step in this direction, in numerical simulations[13] it is observed almost *pathwise agreement* between approximating and exact CG processes for the dynamic introduced in Section 2 when the level of coarse-graining q is smaller than L, for instance $q/L \approx .25$, and $L = 40$ (see also the simulations in Figs. 1 and 2). These simulations suggested that in order to understand questions beyond the agreement in average behavior, we would like to have a *comparison of the entire distribution of the exact and approximating CG processes*: We compute in terms of specific relative entropy the information loss in the transition from the exact to the approximating CG process. The main difficulty we have to face is due to the fact that $(F(\sigma_t))_{t\geq 0}$ fails to be a Markov process. To overcome this obstacle we define in Section 3.1 a *microscopic* process $(\gamma_t)_{t\geq 0}$ similar to $(\sigma_t)_{t\geq 0}$ and better adapted to scale transformations. It allows to reduce the computation of our error estimate to the analysis of two Markov processes.

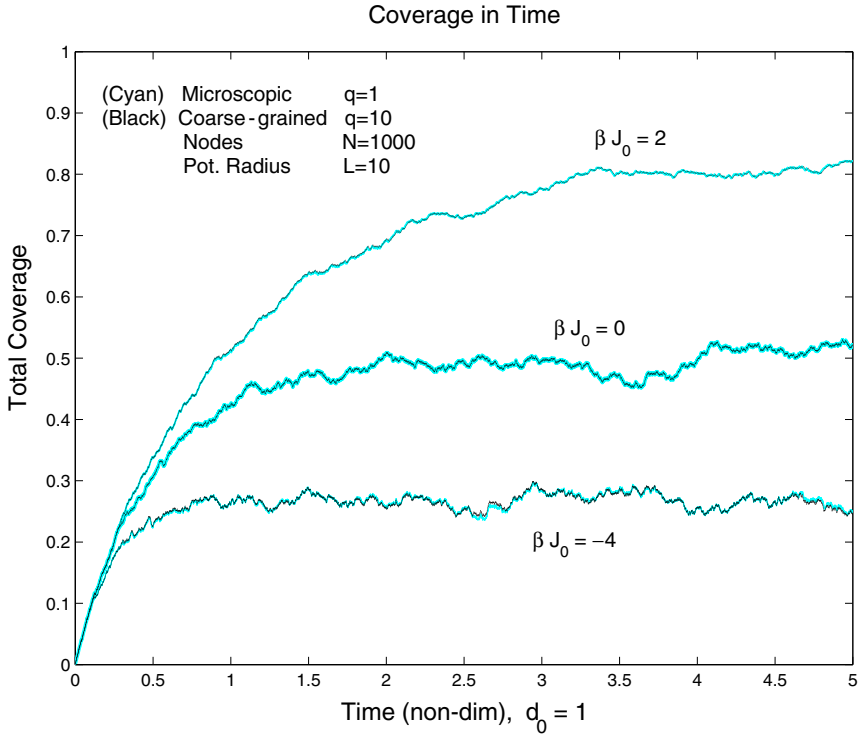

Fig. 1. Coarse-grained vs. microscopic MC adsorption/desorption simulations for the total coverage in various regimes when phase transitions do not occur (Color online).

3.1. A reconstructed microscopic Markov process

We define a new microscopic dynamics γ by the following Markov generator:

$$\left(L_N^{\gamma} f\right)(\rho) = \sum_{x \in \mathcal{L}_N} c_{\gamma}(x, \rho)\,[f(\rho^x) - f(\rho)]$$

for every $f \in \mathcal{B}(\mathcal{H}_N)$ where

$$c_{\gamma}(x, \rho) = (1 - \rho(x)) + \rho(x) e^{-\beta(U^{cg}_{m,q}([x], F(\rho)))}.$$

This dynamic is such that $(F(\gamma_t))_{t \geq 0}$ is still a Markov process. This is due to the fact that its generator is a closed function of $F(\rho)$. Moreover, as it is demonstrated below, $(F(\gamma_t))_{t \geq 0}$ and $(\eta_t)_{t \geq 0}$ have the same distribution. As a consequence a sample path of $(\gamma_t)_{t \geq 0}$ with initial configuration γ_0 can be *reconstructed* from a sample path of $(\eta_t)_{t \geq 0}$ with initial configuration $F(\gamma_0)$ by a sub-sampling procedure in the following way: Let us assume for example that the first upgrade in $(\eta_t)_{t \geq 0}$

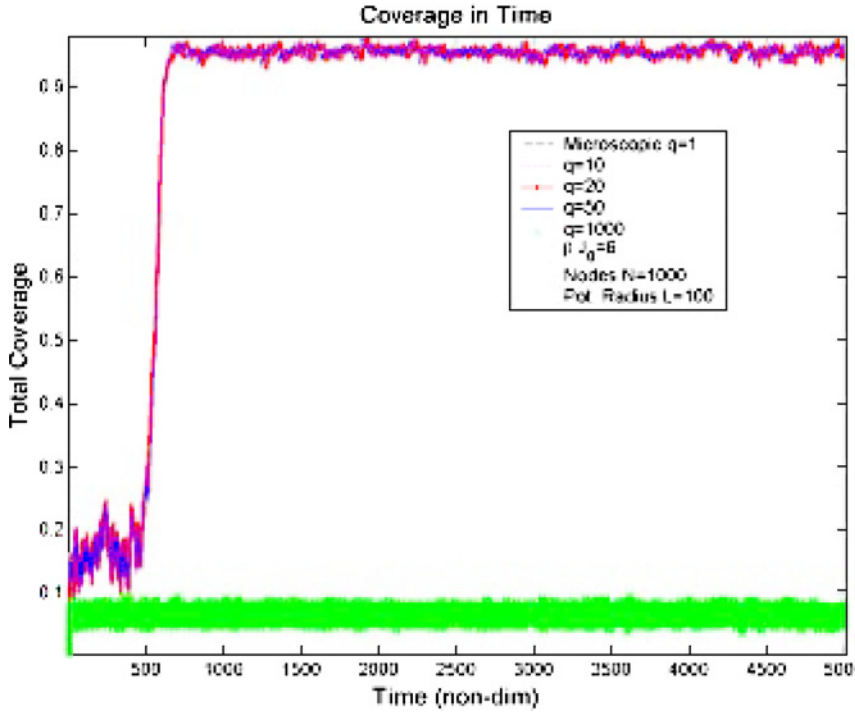

Fig. 2. Coarse-grained vs. microscopic MC adsorption/desorption simulations for the total coverage in the presence of phase transitions (Color online).

is a birth on a given coarse cell D_k at date t_0. Then we pick uniformly at random a micro-cell $x \in D_k$ among those such that $\gamma_0(x) = 0$ and change it to $\gamma_{t_0}(x) = 1$ leaving the other micro cells unchanged. It is clear how we proceed when the first change in η is a death and how the rule is extended to any time $t \geq 0$.

Now we prove that for any distribution μ on $\mathcal{H}_N$, the Markov process $(\gamma_t)_{t\geq 0}$ initially distributed according to μ and evolving according to L_N^{γ} is such that $(F(\gamma_t))_{t\geq 0}$ is still a Markov process. To this end we shall use the notation $c_\gamma(\rho, \rho')$ with $\rho, \rho' \in \mathcal{H}_N$ as a natural extension of $c_\gamma(x, \rho)$ defined above in the sense that

$$c_\gamma(\rho, \rho') = \begin{cases} c_\gamma(x, \rho) & \text{if } \exists x \in \mathcal{L}_N \ \textit{such that} \ \ \rho' = \rho^x \\ 0 & \textit{otherwise.} \end{cases}$$

Proving that $(F(\gamma_t))_{t\geq 0}$ is still a Markov process amounts to prove that the partition of $\mathcal{H}_N$ induced by F through the equivalence relation

$$\rho_1 \sim \rho_2 \quad \text{if and only if} \quad F(\rho_1) = F(\rho_2) \in \mathcal{H}_{m,q}$$

is lumpable for $(\gamma_t)_{t\geq 0}$ i.e. that for every $\varphi, \varphi' \in \mathcal{H}_{m,q}$ and every $\rho_1, \rho_2 \in \mathcal{H}_N$ such that $F(\rho_1) = F(\rho_2) = \varphi$ we have

$$\sum_{\rho' \in \varphi'} c_\gamma(\rho_1, \rho') = \sum_{\rho' \in \varphi'} c_\gamma(\rho_2, \rho') \tag{10}$$

where in the summands above, φ' is considered as an equivalence class over $\mathcal{H}_N$ (see Section 6.3 in Ref. [16] for an account on lumpability of Markov chains). Condition 10 is a straightforward extension to continuous time Markov chains of the row sum criterion, established to characterize lumpability in discrete time Markov chains (see Theorem 6.3.2 in Ref. [16]). Each of the two sums in (10) is non-zero if and only if φ and φ' differ from exactly one unit at one site $k_0 \in \mathcal{L}_m$, and are identical everywhere else. Thus

$$\begin{aligned}
\sum_{\rho' \in \varphi'} c_\gamma(\rho_1, \rho') &= \sum_{x \in D_{k_0}} c_\gamma(x, \rho_1) \\
&= \sum_{\substack{x \in D_{k_0} \\ \rho_1(x)=0}} 1 + \sum_{\substack{x \in D_{k_0} \\ \rho_1(x)=1}} e^{-\beta U^{cg}_{m,q}(k_0, F(\rho_1))} \\
&= \sum_{\substack{x \in D_{k_0} \\ \rho_2(x)=0}} 1 + \sum_{\substack{x \in D_{k_0} \\ \rho_2(x)=1}} e^{-\beta U^{cg}_{m,q}(k_0, F(\rho_2))} \\
&= \sum_{x \in D_{k_0}} c_\gamma(x, \rho_2) = \sum_{\rho' \in \varphi'} c_\gamma(\rho_2, \rho'),
\end{aligned}$$

which proves (10). In addition, it also follows easily from the raw sum criterion that the transition rates of $(F(\gamma_t))_{t\geq 0}$ and $(\eta_t)_{t\geq 0}$ are the same. Hence, whenever these processes have the same initial distribution, they induce the same probability measure on the space $\mathrm{D}(\mathcal{H}_{m,q})$ of right-continuous with left-limits $\mathcal{H}_{m,q}$-valued paths defined on $\mathbb{R}_+$.

Remark. *We refer to $(\gamma_t)_{t\geq 0}$ as a reconstructed microscopic process since it can be directly obtained from the coarse-grained process $(\eta_t)_{t\geq 0}$. We next show that for any fixed time horizon $T > 0$ $(\gamma_t)_{t\in[0,T]}$ approximates within a $T \times O(\frac{q}{2L+1})$ error in specific relative entropy the microscopic process $(\sigma_t)_{t\in[0,T]}$. The reconstructed process may be of independent computational and modeling interest, for instance in simulations that require spatially adaptive model refinement* [(2)].

3.2. Error Estimates

Now we look for a control on the relative entropy of the processes $(F(\sigma_t))_{t\in[0,T]}$ and $(\eta_t)_{t\in[0,T]}$ for any fixed time horizon $T > 0$ and any fixed initial

condition $\alpha \in \mathcal{H}_{m,q}$. We denote by $\mathcal{D}^{\eta,\alpha}_{[0,T]}$ (resp. $\mathcal{D}^{F(\sigma),\alpha}_{[0,T]}$) the distribution on the space $\mathrm{D}_{[0,T]}(\mathcal{H}_{m,q})$ of right-continuous with left-limits $\mathcal{H}_{m,q}$-valued paths defined on $[0, T]$ of $(\eta_t)_{t\in[0,T]}$ (resp. $(F(\sigma_t))_{t\in[0,T]}$) with fixed initial condition α. A direct computation of this relative entropy is difficult (if possible) due to the fact that $(F(\sigma_t))_{t\geq 0}$ fails to be a Markov process. However it follows from the variational formulation of the relative entropy[18] that

$$\begin{aligned} H\big(\mathcal{D}^{F(\sigma),\alpha}_{[0,T]}\big|\mathcal{D}^{\eta,\alpha}_{[0,T]}\big) &= H\big(\mathcal{D}^{F(\sigma),\alpha}_{[0,T]}\big|\mathcal{D}^{F(\gamma),\alpha}_{[0,T]}\big) \\ &\leq H\big(\mathcal{D}^{\sigma,\rho}_{[0,T]}\big|\mathcal{D}^{\gamma,\rho}_{[0,T]}\big) \end{aligned} \tag{11}$$

provided $\rho \in \mathcal{H}_N$ satisfies $F(\rho) = \alpha$. The problem is then reduced to the computation of the relative entropy of two Markov processes. Let us define the difference $\Delta^N_{m,q}(x, \rho)$ by

$$\Delta^N_{m,q}(x, \rho) := U^{fg}_N(x, \rho) - U^{cg}_{m,q}([x], F(\rho)).$$

We obtain

Lemma 31. *The relative entropy defined by (11) satisfies*

$$\begin{aligned} H\big(\mathcal{D}^{\sigma,\rho}_{[0,T]}\big|\mathcal{D}^{\gamma,\rho}_{[0,T]}\big) = E^\sigma_\rho \int_0^T \sum_{x\in\mathcal{L}_N} e^{-\beta U^{fg}_N(x,\rho)}\rho_s(x)\big[1 - e^{-\Delta^N_{m,q}(x,\rho_s)}\big]ds \\ + E^\sigma_\rho \sum_{\exists x\in\mathcal{L}_N\, \rho_{s-}(x)=\rho_s(x)+1}^{s\leq T} \Delta^N_{m,q}(x, \rho_{s-}). \end{aligned} \tag{12}$$

Therefore one needs a control on the $\Delta^N_{m,q}(x, \rho)$'s for $\rho \in \mathcal{H}_N$ and $x \in \mathcal{L}_N$ such that $\rho(x) = 1$.

Lemma 32. *For every $N, m, q > 0$, $\rho \in \mathcal{H}_N$, $x \in \mathcal{L}_N$ we have*

$$|\Delta^N_{m,q}(x, \rho)| \leq O\left(\frac{q}{2L+1}\right)$$

where the O-estimate is uniform in ρ and x and holds when $q/2L + 1 \to 0$.

Combining these results we obtain the announced error estimate.

Theorem 31. *The information loss in the transition from the exact to the approximating coarse-grained process satisfies*

$$\frac{1}{N} H\big(\mathcal{D}^{F(\sigma),\alpha}_{[0,T]}|\mathcal{D}^{\eta,\alpha}_{[0,T]}\big) = T \times O\left(\frac{q}{2L+1}\right)$$

for any $T > 0$, $N, m, q < L$ where the O-estimate holds uniformly when $q/2L + 1 \to 0$.

For the sake of clarity the proof of these results is postponed to the end of the paper.

Remarks.
1. The relative entropy rescaled by the size of the microscopic system in Theorem 3.1 can be regarded as the loss of information per particle. This latter interpretation is evident from an information theory point of view if one considers N independent particles.
2. The relative entropy estimate demonstrates the limitations of the coarse-graining method since it gives an order one error for nearest neighbor interactions ($L = 1$). This is not surprising in view of well-known renormalization calculations for the Ising model, as well as explicit numerical comparisons[12]. On the other hand, Theorem 3.1 rigorously identifies a small parameter in the coarse-graining process, namely the ratio $q/(2L + 1)$; as it is the case with most asymptotics, from a practical/computational point of view a small parameter can be even fairly large, see for instance the values of q and L in the simulations in Fig. 1 and even in the phase transition regime in Fig. 2. We also refer to simulations in[12,14].
3. Although the estimate in Theorem 3.1 is for finite times $[0, T]$ only, and grows as T, it is still useful; in the case of phase transitions (e.g. Fig 2 in Section 4) the estimate ensures numerical accuracy during nucleation, which is typically an initial stage of the evolution. We refer to simulations partly motivated by our rigorous results in[14] that make precisely this point. Furthermore, the simulations there demonstrate the accurate prediction of transition paths and the eventual domain switching. In addition, we have that, the coarse-grained Gibbs measure lies within a controlled error from the exact CG measure[15]. More specifically, we easily obtain the equilibrium version of Theorem 3.1, namely

$$\frac{1}{N} H\big(F\big(P_N^{fg}\big) | P_{m,q}^{cg}\big) = O\Big(\frac{q}{2L+1}\Big). \tag{13}$$

As a result of Theorem 3.1 and the aforementioned Gibbs measure estimate, the transient, as well as the long time dynamics are expected to be captured accurately by the CGMC model.
4. A simple but long computation shows that what is lost in (11) is of order $T\,O(q/2L + 1)$.
5. In the particular case $N = 2L + 1$ our analysis easily yields

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathcal{D}_{[0,T]}^{F(\sigma),\alpha}(A) = \limsup_{N\to\infty} \frac{1}{N} \log \mathcal{D}_{[0,T]}^{\eta,\alpha}(A)$$

and

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathcal{D}^{F(\sigma),\alpha}_{[0,T]}(A) = \liminf_{N\to\infty} \frac{1}{N} \log \mathcal{D}^{\eta,\alpha}_{[0,T]}(A)$$

for any measurable $A \subset \mathrm{D}_{[0,T]}(\mathcal{H}_{m,q})$ and any level of coarsening q_N such that $q_N/N \to 0$.

6. Careful inspection of the proof reveals that the smoothness assumptions on V can be suitably relaxed. However in this paper we do not pursue maximum generality. Similarly the result extends also in higher dimensions.

4. COARSE-GRAINED MONTE CARLO ALGORITHMS

We first present in Figs. 1 and 2 two comparative simulations between microscpic ($q = 1$) and coarse-grained ($q > 1$) MC simulations. We use the exact same seed for our random number generator when comparing simulations for different coarse-graining values of q. This allows us to focus on the differences attributed only to the coarse graining and not on those resulting from different realizations due to different seeds. This numerical coupling argument yields in the end the pathwise agreement in the figures. Clearly the numerics indicate there is a stronger result than the estimates in this paper.

In all figures the total coverage over the entire lattice is plotted as a time series; in particular Fig. 2 is a simulation within the phase transition regime ($\beta J_0 = 6$), while other parameters of the systems are specified in both Figures. The interparticle potential is assumed to be piecewise constant, although simulations with smooth Morse potentials give rise to similar results[12]. In Fig. 2 there is a marked disagreement with microscopics when the system is over coarse-grained ($q = 1000$). Detailed comparisons of the spatial structures, as well as of statistical quantities (e.g. escape times in Fig. 2) are given in[14].

The implementation of CGMC is essentially identical to the microscopic MC with a few modifications[19]. First, the interparticle potential J is coarse-grained at the beginning of a simulation to represent interactions between particles within each cell as well as interactions with neighboring cells. Second, the order parameter is still an integer that varies between 0 and q, instead of 0 and 1 which is typical for microscopic MC. For CGMC the CPU time in kinetic MC simulation with global update, i.e., searching the entire lattice to identify the chosen site, scales approximately as $O(m^3)$ vs. $O(N^3)$ for a conventional MC algorithm. In addition, coarse-grained potentials $\bar{J}$ are compressed and thus additional savings are made in the calculation of the activation energies. Overall in the case of adsorption/desorption processes the CPU time decreases[12] with increasing q approximately as $O(1/q^3)$. For example, a modest tenfold reduction in the number of sites ($q = 10$) results in reduced CPU by a factor of 1000, yielding a significant enhancement in performance. 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 capable of capturing large scale features, while retaining microscopic information on intermolecular forces and particle fluctuations.

5. PROOFS

5.1. Proof of Lemma 3.1

We first need the components $(\lambda_\sigma, p_\sigma)$ and $(\lambda_\gamma, p_\gamma)$ of the skeleton processes associated to the dynamics that enter into play (see Appendix 1.2 in Ref. [18] for an account on skeleton processes associated to continuous time Markov chains). First we have

$$\lambda_\sigma(\rho) = N - \sum_{x\in\mathcal{L}_N} \rho(x) + \sum_{x\in\mathcal{L}_N} \rho(x) e^{-\beta U_N^{fg}(x,\rho)}.$$

If there exists $x \in \mathcal{L}_N$ such that $\rho_{i+1} = \rho_i^x$ then

$$\lambda_\sigma(\rho_i) p_\sigma(\rho_i, \rho_{i+1}) = 1 - \rho_i(x) + \rho_i(x) e^{-\beta U_N^{fg}(x,\rho)}$$

and $p_\sigma(\rho_i, \rho_{i+1}) = 0$ otherwise. The analogous quantities for the process $(\gamma_t)_{t\geq 0}$ are

$$\lambda_\gamma(\rho) = N - \sum_{x\in\mathcal{L}_N} \rho(x) + \sum_{x\in\mathcal{L}_N} \rho(x) e^{-\beta U_{m,q}^{cg}([x],F(\rho))},$$

and if there exists $x \in \mathcal{L}_N$ such that $\rho_{i+1} = \rho_i^x$ then

$$\lambda_\gamma(\rho_i) p_\gamma(\rho_i, \rho_{i+1}) = 1 - \rho_i(x) + \rho_i(x) e^{-\beta U_{m,q}^{cg}(F(\rho),[x])}$$

while $p_\gamma(\rho_i, \rho_{i+1}) = 0$ otherwise. The Radon-Nikodym derivative of $\mathcal{D}_{[0,T]}^{\sigma,\rho}$ w.r.t. to $\mathcal{D}_{[0,T]}^{\gamma,\rho}$ is given by

$$\frac{d\mathcal{D}_{[0,T]}^{\sigma,\rho}}{d\mathcal{D}_{[0,T]}^{\gamma,\rho}}\big((\rho_t)_{t\in[0,T]}\big) = \exp\left\{\int_0^T [\lambda_\sigma(\rho_s) - \lambda_\gamma(\rho_s)]ds - \sum_{s\leq T} \log \frac{\lambda_\sigma(\rho_{s-}) p_\sigma(\rho_{s-}, \rho_s)}{\lambda_\gamma(\rho_{s-}) p_\gamma(\rho_{s-}, \rho_s)}\right\}$$

where the sum in the last term runs over the discontinuity times of $(\rho_t)_{t\in[0,T]}$, see Proposition A1.2.6 in Ref. [18]. Thus

$$\frac{d\mathcal{D}^{\sigma,\rho}_{[0,T]}}{d\mathcal{D}^{\gamma,\rho}_{[0,T]}}\big((\rho_t)_{t\in[0,T]}\big) = \exp\left\{\int_0^T \sum_{x\in\mathcal{L}_N} \rho_s(x)\big[e^{-\beta U_N^{fg}(x,\rho_s)} - e^{-\beta U^{cg}_{m,q}([x],F(\rho_s))}\big]ds \right.$$

$$\left. - \sum_{\substack{s\le T \\ \exists x\in\mathcal{L}_N \rho_{s-}(x)=\rho_s(x)+1}} \log \frac{e^{-\beta U_N^{fg}(x,\rho_{s-})}}{e^{-\beta U^{cg}_{m,q}([x],F(\rho_{s-}))}}\right\}$$

$$= \exp\left\{\int_0^T \sum_{x\in\mathcal{L}_N} e^{-\beta U_N^{fg}(x,\rho_s)}\rho_s(x)[1 - e^{-\Delta^N_{m,q}(x,\rho_s)}]ds \right.$$

$$\left. + \sum_{\substack{s\le T \\ \exists x\in\mathcal{L}_N \rho_{s-}(x)=\rho_s(x)+1}} \beta\Delta^N_{m,q}(x,\rho_{s-})\right\}.$$

Hence

$$H(\mathcal{D}^{\sigma,\rho}_{[0,T]}|\mathcal{D}^{\gamma,\rho}_{[0,T]}) = E^\sigma_\rho \log \frac{d\mathcal{D}^{\sigma,\rho}_{[0,T]}}{d\mathcal{D}^{\gamma,\rho}_{[0,T]}}$$

$$= E^\sigma_\rho \int_0^T \sum_{x\in\mathcal{L}_N} e^{-\beta U_N^{fg}(x,\rho)}\rho_s(x)\big[1 - e^{-\Delta^N_{m,q}(x,\rho_s)}\big]ds$$

$$+ E^\sigma_\rho \sum_{\substack{s\le T \\ \exists x\in\mathcal{L}_N \rho_{s-}(x)=\rho_s(x)+1}} \Delta^N_{m,q}(x,\rho_{s-}).$$

5.2. Proof of Lemma 3.2 and Theorem 3.1

We prove Lemma 3.2 and Theorem 3.1 by considering particular cases of increasing difficulty up to the general setting.

5.2.1. Curie-Weiss Model

This model corresponds to $2L + 1 = N$ with a constant potential $V(r) = J_0$ for every $r \in [-1/2, 1/2]$. Simple algebra shows that $\Delta^N_{m,q}(x,\rho) = 0$ for every $\rho \in \mathcal{H}_N$ and every $x \in \mathcal{L}_N$ such that $\rho(x) = 1$. Thus $H(\mathcal{D}^{F(\sigma),\alpha}_{[0,T]}|\mathcal{D}^{\eta,\alpha}_{[0,T]}) = 0$ for every $T > 0$, N, m, q, as it must naturally be expected in this case.

5.2.2. Long-Range Mean-Field Models

This class of models still corresponds to $2L + 1 = N$, but now the V's are not constant. In order to get a control on the $\Delta^N_{m,q}(x, \rho)$'s we first look for a control on the difference between the exact and coarse-grained interaction potentials as defined in[2,7,8]. Let $k, l \in \mathcal{L}_m, k \neq l, x, x' \in D_k, y, y' \in D_l$. Due to the smoothness of V we have

$$V(x-y) = V(x'-y') + ((x-y)-(x'-y'))V'(x'-y') + O(((x-y)-(x'-y'))^2).$$

The O-estimate above holds when $(x-y)-(x'-y') \to 0$. Actually $|(x-y)-(x'-y')| \leq |x-x'| + |y-y'| \leq 2/m$, thus

$$J(x-y) = \frac{1}{q^2} \sum_{x' \in D_k} \sum_{y' \in D_l} J(x'-y') + \frac{1}{Nq^2} \sum_{x' \in D_k} \sum_{y' \in D_l} ((x-y) - (x'-y'))V'(x'-y') + \frac{1}{Nq^2} \sum_{x' \in D_k} \sum_{y' \in D_l} O\left(\frac{1}{m^2}\right)$$

as $m \to \infty$. Hence

$$|J(x-y) - \bar{J}(k,l)| \leq \frac{1}{Nq^2} \sum_{x' \in D_k} \sum_{y' \in D_l} (|x-x'| + |y-y'|)|V'(x'-y')| + \frac{1}{N} O\left(\frac{1}{m^2}\right) \leq O\left(\frac{q}{N^2}\right)$$

as $q/N \to 0$. Accordingly, for every $k \in \mathcal{L}_m, x, y \in D_k$

$$|J(x-y) - \bar{J}(0,0)| \leq O\left(\frac{q}{N^2}\right)$$

when $q/N \to 0$. Thus for every $\rho \in \mathcal{H}_N$ and $x \in \mathcal{L}_N$ such that $\rho(x) = 1$

$$\begin{aligned}
|\Delta^N_{m,q}(x,\rho)| &= |U^{fg}_N(x,\rho) - U^{cg}_{m,q}([x], F(\rho))| \\
&= \left| \sum_{\substack{y \in \mathcal{L}_N \\ y \neq x}} J(x-y)\rho(y) - \sum_{\substack{k \in \mathcal{L}_m \\ k \neq [x]}} \bar{J}([x],k)F(\rho)(k) - \bar{J}(0,0)\Big(F(\rho)([x]) - 1\Big) \right| \\
&\leq \sum_{\substack{k \in \mathcal{L}_m \\ k \neq [x]}} \sum_{y \in D_k} |\bar{J}([x],k) - J(x-y)|\rho(y) + \sum_{\substack{y \in D_{[x]} \\ y \neq x}} |J(x-y) - \bar{J}(0,0)|\rho(y) \\
&\leq O\left(\frac{q}{N}\right).
\end{aligned}$$

On the other hand, since the potential J is summable, there exists a $K_1 < \infty$ such that

$$\sup_{N\in\mathbb{N}} \sup_{\rho\in\mathcal{H}_N, x\in\mathcal{L}_N} \left|U_N^{fg}(x,\rho)\right| \le K_1.$$

Finally, there exist $0 < K_2, K_3 < \infty$ such that

$$K_2 N \le \sup_{\rho\in\mathcal{H}_N, x\in\mathcal{L}_N} |\lambda_\sigma(\rho)| \le K_3 N$$

so the mean quantity of jumps for $(\sigma_t)_{t\ge 0}$ during a time interval $[0, T]$ is of the order NT. Combining all these controls with (11) and (12) leads to

$$\lim_{N\to\infty} \frac{1}{N} H\left(\mathcal{D}_{[0,T]}^{F(\sigma),\alpha} | \mathcal{D}_{[0,T]}^{\eta,\alpha}\right) = O\left(\frac{q_N}{N}\right)$$

for any level of coarsening q_N and m_N such that $q_N/N \to 0$.

5.2.3. Finite-Range Interaction Models

This class of models corresponds to a fixed value of L i.e. a value independent of N. We shall assume that $q < L$ since our goal is to prove that the approximation is good when q is substantially smaller than L. Due to the fact that the interaction is of finite-range type, a micro-cell $x \in \mathcal{L}_N$ "feels" L microscopic neighbors on each side, in the sense that $J(x - y) \ne 0$ for $2L$ different sites $y \in \mathcal{L}_N$. This neighborhood includes cells belonging to $D_{[x]}$. In the same time the coarse cell $[x] \in \mathcal{L}_m$ "feels" at most $\mathrm{E}((L-1)/q) + 1$ coarse cells on each side, where $\mathrm{E}(u)$ stands for the integer part of $u \in \mathbb{R}$. We denote by $Q([x])$ the union of these coarse cells. Some $x \in \mathcal{L}_N$ fail to interact with a certain number of micro-cells that belong to $Q([x])$. These micro-cells are included in at most 4 different coarse cells (depending on the location of x within $[x]$) which union we denote by $R(x)$. We shall denote by $\tau_{R(x)}$ the set of sites $y \in R(x)$ such that $|x - y| \le L/N$. Due to the smoothness of V we obtain as in the previous section that for every $k, l \in \mathcal{L}_m$, $k \ne l$, $x \in D_k$, $y \in D_l$

$$|J(x-y) - \bar{J}(k,l)| \le O\left(\frac{q}{(2L+1)^2}\right)$$

when $q/2L + 1 \to 0$ and for every $k \in \mathcal{L}_m$, $x, y \in D_k$

$$|J(x-y) - \bar{J}(0,0)| \le O\left(\frac{q}{(2L+1)^2}\right).$$

Hence for every $\rho \in \mathcal{H}_N$ and $x \in \mathcal{L}_N$ such that $\rho(x) = 1$

$$\begin{aligned}\left|\Delta^N_{m,q}(x,\rho)\right| = \left|U^{fg}_N(x,\rho) - U^{cg}_{m,q}([x],F(\rho))\right| \leq \sum_{\substack{k\in\mathcal{L}_m, k\neq[x]\\ k\in Q([x])/R(x)}} \sum_{y\in D_k} |\bar{J}([x],k)\\ -J(x-y)|\rho(y) + \sum_{\substack{y\in D_{[x]}\\ y\neq x}} |J(x-y) - \bar{J}(0,0)|\rho(y)\\ + \sum_{\substack{k\in\mathcal{L}_m, k\neq[x]\\ k\in R(x)}} |\bar{J}([x],k)|F(\rho)(k) + \sum_{y\in\tau_{R(x)}} |J(x-y)|\rho(y)\\ \leq O\left(\frac{q}{2L+1}\right)\end{aligned}$$

where the O-estimate above is uniform and holds when $q/2L + 1 \to 0$. By using the arguments employed in the previous section we obtain the announced result.

6. CONCLUSIONS

In the present paper we gave a rigorous error estimate for the approximation of exact coarse-grained processes by CGMC dynamics. We computed the information loss between exact and approximating CG adsorption/desorption lattice dynamics. In analogy to rigorous error estimates for finite element/finite difference approximations of Partial Differential Equations, our result can be viewed as an error analysis between the exact and approximating CG processes. The error is measured in terms of the specific relative entropy and represents the loss of information in the transition from one scale to the other. More specifically, we obtained

$$\frac{1}{N} H\left(\mathcal{D}^{F(\sigma)}_{[0,T]} | \mathcal{D}^{\eta}_{[0,T]}\right) = T \times O\left(\frac{q}{2L+1}\right)$$

This result gives a first mathematical reasoning for the parameter regimes for which nonequilibrium coarse grained MC algorithms are expected to give errors within a given tolerance. A crucial technical step in the analysis was the introduction of an auxiliary microscopic Markovian process which was directly reconstructed from the coarse-grained process $(\eta_t)_{t\geq 0}$ and approximated within a $O(\frac{q}{2L+1})$ error in relative entropy the microscopic process $(\sigma_t)_{t\geq 0}$.

ACKNOWLEDGMENTS

The authors would like to thank A. Sopasakis for providing the Figs. 1 and 2. They thank the Institute for Pure and Applied Mathematics of the University of

California at Los Angeles where part of this work was partly carried out during a visit in July-August 2003. The work of MAK is partially supported by the NSF under grants DMS-0219211 and DMS-0413864. Part of the work of JT was done while he was a Marie Curie Fellow at the University of Warwick (UK) under contract MEIF-CT-2003-501275. He thanks Neil O'Connell for valuable discussions on Markov processes.

REFERENCES

1. G. Beylkin, R. Coifman and V. Rokhlin, Fast wavelet transforms and numerical algorithms. I, *Comm. Pure Appl. Math.* **44:**141–183 (1991).
2. A. Chaterjee, M. A. Katsoulakis and D. G. Vlachos, Spatially adaptive grand canonical Monte Carlo simulations, *Phys. Rev. E.* **71:**026702–026708 (2005).
3. A. J. Chorin, A. P. Kast, and R. Kupferman, Optimal prediction of underresolved dynamics, *Proc. Natl. Acad. Sci. USA* **95:**4094–4098 (1998).
4. W. E and B. Engquist, Multiscale modeling and computation, Notices AMS **50:**1062-1070 (2003).
5. P. Ferrari and J. L. Lebowitz, Information loss in coarse graining of polymer configurations via contact matrices, *J. Phys. A.* **36:**5719–5730 (2003).
6. M. Hildebrand and A. S. Mikhailov, Mesoscopic modeling in the kinetic theory of adsorbates, *J. Phys. Chem.* **100:**19089–19101 (1996).
7. Q. Hou, N. Goldenfeld and A. McKane, Renormalization group and perfect operators for Stochastic Differential Equations, *Phys. Rev. E* **63:**036125–036146 (2001).
8. A. Ishikawa and T. Ogawa, Dynamics of spinodal decomposition in finite-lifetime systems: Nonlinear statistical theory based on a coarse-grained lattice-gas model, *Phys. Rev. E.* **65:**026131–026140 (2002).
9. 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:**4414–4423 (2003).
10. A. E. Ismail, G. C. Rutledge, and G. Stephanopoulos, Multiresolution analysis in statistical mechanics II: The wavelet transform as a basis for Monte Carlo simulation on lattices, *J. Chem. Phys.* **118:**4424–4431 (2003).
11. 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:**255–294 (2004).
12. M. A. Katsoulakis, A. J. Majda and D. G. Vlachos, coarse-grained stochastic processes and Monte Carlo simulations in lattice systems, *J. Comp. Phys.* **186:**250–278 (2003).
13. M. A. Katsoulakis, A. J. Majda and D. G. Vlachos, coarse-grained stochastic processes for lattice systems, *Proc. Natl. Acad. Sci. USA* **100:**782–787 (2003).
14. M. A. Katsoulakis, P. Plechac, and A. Sopasakis, "Error control and analysis in coarse-graining of stochastic lattice dynamics", Institute for Mathematics and its Applications, IMA Preprint No. 2019, February 2005, (http://www.ima.umn.edu/preprints/feb2005/feb2005.html).
15. M. A. Katsoulakis and D. G. Vlachos, Hierarchical kinetic Monte Carlo simulations for diffusion of interacting molecules, *J. Chem. Phys.* **112:**9412–9427 (2003).
16. J. Kemeny and L. Snell, *Finite Markov chains*, (Springer-Verlag, New York–Heidelberg, 1976).
17. B. Khouider, A. J. Majda and M. A. Katsoulakis, coarse-grained stochastic models for tropical convection and climate, *Proc. Natl. Acad. Sci. USA* **100:**11941–11946 (2003).
18. C. Kipnis and C. Landim, *Scaling limits of interacting particle systems*, Springer (1999).

19. D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (2000), Cambridge University Press.
20. J. S. Langer, Theory of spinodal decomposition in alloys, *Ann. Phys.* **65:**53–86 (1971).
21. A. J. Majda and B. Khouider, Stochastic and mesoscopic models for tropical convection, *Proc. Nat. Acad. Sci. USA*, **99:**1123–1128 (2002).
22. A. Milchev, D. W. Heermann and K. Binder, Monte Carlo simulation of the Cahn-Hilliard model of spinodal decomposition, *Acta Metall.* **36:**377–383 (1988).
23. F. Muller-Plathe, Coarse-graining in polymer simulation: from the atomistic to the mesoscale and back, *Chem. Phys. Chem.* **3:**754–769 (2002).
24. C. Schütte, A. Fischer, W. Huisinga and P. Deuflhard, A direct approach to conformational dynamics based on hybrid Monte Carlo, *J. Comp. Phys.* **151:**146–168 (1999).
25. T. Shinbrot and F. J. Muzzio, Noise to order, *Nature* **410:**251–258 (2001).
26. A. N. Shiryaev, *Probability*, (Springer 1996).
27. C. Theodoropoulos, Y.-H. Qian and I. G. Kevrekidis, Coarse stability and bifurcation analysis using timesteppers: A reaction diffusion example, *Proc. Natl. Acad. Sci. USA* **97:**9840–9843 (2000).
28. 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:**8306–8313 (1990).
29. D. G. Vlachos and M. A. Katsoulakis, Mesoscopic Theories for the Diffusion of Interacting Molecules, *Phys. Rev. Lett.* **85:**3898–3901 (2000).