NUMERICAL AND STATISTICAL METHODS FOR THE COARSE-GRAINING OF MANY-PARTICLE STOCHASTIC SYSTEMS

MARKOS A. KATSOULAKIS*, PETR PLECHÁČ†, AND LUC REY-BELLET‡

Abstract. In this article we discuss recent work on coarse-graining methods for microscopic stochastic lattice systems. We emphasize the numerical analysis of the schemes, focusing on error quantification as well as on the construction of improved algorithms capable of operating in wider parameter regimes. We also discuss adaptive coarse-graining schemes which have the capacity of automatically adjusting during the simulation if substantial deviations are detected in a suitable error indicator. The methods employed in the development and the analysis of the algorithms rely on a combination of statistical mechanics methods (renormalization and cluster expansions), statistical tools (reconstruction and importance sampling) and PDE-inspired analysis (*a posteriori* estimates). We also discuss the connections and extensions of our work on lattice systems to the coarse-graining of polymers.

1. Introduction. In this paper we provide an overview of mathematical techniques, developed in a series of papers [27, 28, 35, 29, 34, 32, 31, 33, 3, 49], for the construction and the numerical analysis of coarse-graining schemes for stochastic, many-body microscopic systems, primarily motivated from problems in materials science and chemical engineering. We focus on mathematical formulation and give only brief summaries of our results; we refer the reader to the cited works for more technical details.

Modelling and simulation of materials properties at small spatio-temporal scales primarily relies on detailed microscopic models such as Molecular Dynamics (MD) or Monte Carlo methods (MC) that typically account for atomistic or molecular information. On the other hand, understanding large-scale, macroscopic properties of materials requires simulations with microscopic systems at prohibitively large spatial, as well as temporal scales. In the direction of such considerations, an important class of computational tools developed in recent years, in the physics, applied sciences and engineering literature, is the method of coarse-graining. The idea behind this approach is to reduce the complexity of molecular systems by lumping together degrees of freedom into coarse-grained variables, thus yielding accelerated simulation methods capable of reaching mesoscopic length scales. Such coarse-grained models have been developed for the study and simulation of crystal growth, surface processes, polymers, proteins and complex fluids, among others [38, 44, 12, 45]. At the same time the importance of coarse-graining in industrial processes has also been recently recognized, [8]. The existing approaches can give unprecedented speed-up to molecular simulations and can work well in certain parameter regimes, for instance at high temperatures or low density. On the other hand, they can also give wrong predictions on important features such as diffusion, crystallization and phase transitions. Along these lines, some relevant mathematical and statistical goals are:

(a) to derive coarse-graining schemes in a systematic manner,
(b) to understand the validity regimes of existing coarse-graining methods by developing a mathematical and statistical error analysis, and
(c) to develop adaptive algorithms, which have the capacity to automatically adjust if substantial deviations are detected during simulation.

From a mathematical point of view, a (perfect) coarse-graining can be understood, in a natural way, as a *renormalization group* map such as the ones widely used in statistical physics [26, 20]. One of our point of view is that a numerical coarse-graining scheme is an approximate computation of a renormalization group map. The strategy we use in this context is to expand around a properly chosen first coarse-grained guess; for this task the techniques of *cluster expansions* turn out to be a very useful and powerful tool. On one hand, cluster expansions can be used to derive systematically a hierarchy of schemes of increasing accuracy, and, on the other hand, they provide both a priori error estimates, usually expressed in terms of *relative entropy*, and also *a posteriori error estimates*, as used in the numerical analysis of PDE's.

*Department of Mathematics and Statistics, University of Massachusetts (markos@math.umass.edu).
†Department of Mathematics, University of Tennessee and Oak Ridge National Laboratory (plechac@math.utk.edu).
‡Department of Mathematics and Statistics, University of Massachusetts(luc@math.umass.edu).

We also develop a complementary statistical point of view, by viewing and analyzing the coarse-graining schemes as a version of *importance sampling*. The coarse-graining scheme is viewed as a proposal used in the sampling of the Monte-Carlo method. It appears that these two strategies have complementary strengths and their synergy has the potential to provide a fairly systematic framework for developing flexible coarse-grained (CG) algorithms for microscopic systems. The common theme behind both approaches is the observation that long-range interactions can be handled very efficiently by coarse-grained MC developed recently [27, 35, 32]. On the other hand short-range interactions are relatively inexpensive and can be handled by Direct Numerical Simulation (DNS), provided there is a suitable *splitting* of the algorithm into short and long-range parts.

Furthermore, even if coarse-grained simulators are available, in many applications such as the simulation of polymer melts, it is necessary to obtain genuinely microscopic information. This amounts to essentially reversing the coarse-graining by carrying out a *microscopic reconstruction*. Here we discuss how to mathematically formulate this issue, as well as we demonstrate some probabilistic reconstruction methodologies.

Based on these distinct points of view, we have developed a systematic way to assess properties of coarse-graining algorithms from a numerical analysis perspective, focusing on error quantification. In this paper we primarily discuss two classes of models: (i) Stochastic lattice systems with combinations of short and long range interactions, and (ii) Coarse-graining and reconstruction in off-lattice macromolecular systems, such as polymers. The paper is organized as follows. In Section 2 we introduce equilibrium and non-equilibrium stochastic lattice models and discuss a first set of coarse-grainings. In Section 2.2 we present error estimates and develop higher order coarse-graining schemes, in the analytically more tractable case of lattice systems with long/intermediate range interactions. In Section 2.3 we discuss coarse-grained models for lattice systems with combined short and long range interactions. In Section 3 we focus on the issue of microscopic reconstruction. In Section 4 we combine importance sampling and coarse-graining methods using the example of short and long range interactions as a paradigm. In Section 6 we discuss the connections and extensions of our work on lattice systems to the extensive engineering literature on coarse-graining of polymer systems. Finally, in Section 5 we present recent progress towards developing adaptive coarse-graining schemes based on a posteriori error estimates.

Acknowledgments. The research of M.A.K. was partially supported by the National Science Foundation under grants NSF-DMS-0413864 and NSF-DMS-0715125 and the U.S. Department of Energy under grant DE-FG02-05ER25702. The research of P.P. was partially supported by the National Science Foundation under grant NSF-DMS-0303565 and by the Office of Advanced Scientific Computing Research, U.S. Department of Energy; the work was partly done at the ORNL, which is managed by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725. The research or L.R-B. was partially supported by the National Science Foundation under grant NSF-DMS-0605058.

2. Coarse-graining schemes for extended lattice systems. We discuss the coarse-graining of stochastic lattice systems such as Ising-type models. These models are more accessible both mathematically and computationally as there is a vast related literature that includes analytical methods and explicitly solvable models that serve as benchmarks for the numerics. The direct numerical simulation in such models is usually carried out using Monte Carlo methods. This class of stochastic models is employed in simulations of adsorption, desorption, reaction and diffusion of chemical species in numerous applied science areas such as catalysis, microporous materials, biological systems, etc. [4, 39]. The fundamental principle on which this type of modeling is based can be easily formulated as follows: when the binding of species on a surface or within a pore is relatively strong, these physical processes can be described as jump processes from one site to another or to an adjacent gas-phase (see Figure 2.1(a)) with a transition probability that is expected to be calculated from even smaller scales using quantum-mechanical calculations, and transition state theory, or from detailed experiments, see for instance [4].

2.1. Mathematical formulation. We consider Ising-type systems on a periodic lattice Λ_N which is a discretization of the interval $\mathbb{I} = [0, 1)$. We divide $\mathbb{I}$ in N (micro-)cells and consider the microscopic grid

$\Lambda_N = \frac{1}{N}\mathbb{Z} \cap \mathbb{I}$ in Figure 2.1(a). Throughout this discussion we concentrate on one-dimensional models, however, our results readily extend to higher dimensions, [32]. With each $x \in \Lambda_N$ we associate an order parameter $\sigma(x)$; for instance, when taking values 0 and 1, it describes vacant and occupied sites. The energy H of a configuration $\sigma \in \Sigma = \{\sigma(x) : x \in \Lambda_N\}$ is given by the Hamiltonian,

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

where h is the external field and $J(\cdot)$ is the inter-particle potential. The equilibrium states at temperature T are described by the (canonical) Gibbs measures

$$\mu_N(d\sigma) = Z_N^{-1} \exp\left(-\beta H(\sigma)\right) P_N(d\sigma)\,, \tag{2.2}$$

where $\beta = 1/kT$ (k is the Boltzmann constant), Z_N is the partition function, and the product Bernoulli distribution $P_N(\sigma)$ is the uniform *prior distribution* on Σ. The interparticle potentials J account for interactions between occupied sites. In Section 2.2 we consider systems with intermediate or long-range interactions. Such systems are more tractable analytically, while at the same time they pose a serious challenge to conventional MC methods due to the large number of neighbors involved in each MC step. More general potentials with combinations of short and long-range interactions are discussed in Section 2.3.

2.1.1. Microscopic lattice dynamics. The dynamics of Ising-type models consists of order parameter flips and/or exchanges that correspond to different physical processes, [39]. More specifically a flip at the site $x \in \Lambda_N$ is a spontaneous change in the order parameter, 1 being converted to 0 and vice versa, while a spin exchange between the neighboring sites $x, y \in \Lambda_N$ is a spontaneous exchange of the order parameters at the two locations. For instance, a spin flip can model the desorption of a particle from a surface described by the lattice to the gas phase above and conversely the adsorption of a particle to the surface [2, 53]. Such a model has also been proposed recently for modeling unresolved features of tropical convection [42].

For a configuration σ we denote by σ^x the configuration which differs from σ by an order parameter flip at site x. The dynamics is described by a continuous time Markov Chain with state space Σ: the configuration update $\sigma \to \sigma^x$ occurs with a rate $c(x,\sigma)$, i.e., the order parameter at x changes over the time interval $[t, t+\Delta t]$ with the probability $c(x,\sigma)\Delta t + o(\Delta t)$. The resulting stochastic process $\{\sigma_t\}_{t\geq 0}$ is a continuous time jump Markov process with a generator defined in terms of the rate $c(x,\sigma)$, [37]. The condition of *detailed balance*,

$$c(x,\sigma)e^{-\beta H(\sigma)} = c(x,\sigma^x)e^{-\beta H(\sigma^x)}\,, \tag{2.3}$$

imposed on the rates implies that the jump dynamics leaves the Gibbs states invariant.

The simplest example is the Metropolis-type dynamics [18], which define a Monte-Carlo (MC) relaxation algorithm for sampling from the Gibbs measure (2.2). However, in *non-equilibrium* chemical applications one also considers *Arrhenius dynamics*. Such dynamics is justified when the binding of species on a surface or within a pore is relatively strong, desorption and diffusion can be modeled as a hopping process from one site to another or to the gas phase, with a transition probability that depends on the potential energy surface, e.g., [2]. The Arrhenius rate for the adsorption/desorption mechanism is

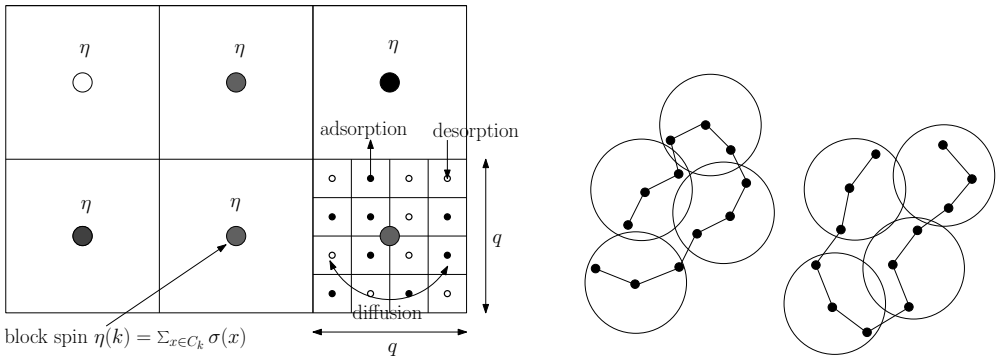
$$c(x,\sigma) = d_0\big(1-\sigma(x)\big) + d_0\sigma(x)\exp\big[-\beta U(x,\sigma)\big]\,, \tag{2.4}$$

where the potential

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

is the total energy of a particle located at the lattice site $x \in \Lambda_N$. Typically, an additional term corresponding to the energy associated with the surface binding of the particle at x, can be also included into the external

field h in U; the rate constant d_0 is related to the pre-exponential factor of the microscopic processes. Similarly, one can define Arrhenius particle diffusion [52], where the configuration updates are carried out via spin exchange dynamics, i.e., $\sigma^{x,y}$ is the new configuration that resulted from the exchange between the order parameter at the sites x and y.



(a) Coarse-grained lattice system (b) Coarse-grained chains

FIG. 2.1. *Schematics of coarse-graining for lattice and off-lattice systems: (a) an example of a lattice system coarse-grained by using block-spins, (b) coarse-grained polymer chains: starting from united atom model three particles are grouped into a meta-particle.*

2.1.2. Coarse-grained stochastic processes. In the series of papers [27, 28, 35] the authors initiated the development of mathematical strategies for the *coarse-graining* (CG) of stochastic lattice dynamics. One constructs the coarse grid $\bar{\Lambda}_M$ by dividing $\mathbb{I} = [0, 1)$ in M coarse cells, each of which contains q (micro-)cells, see Figure 2.1(a). Each coarse cell is denoted by $C_k\,, k = 1, ..., M$. A typical choice for the coarse variable in the context of Ising-type models is the block-spin over each coarse cell C_k. We consider the process

$$(\mathbf{T}\sigma_t)(k) := \sum_{y \in C_k} \sigma_t(y)\,, \quad k = 1, ..., M\,. \tag{2.5}$$

If $\mathbf{T}\sigma_t$ is a Markov process then $\{\sigma_t\}_{t \geq 0}$ is called *lumpable*. However, in general, $\mathbf{T}\sigma_t$ is *not* a Markov process, which creates serious mathematical and computational difficulties in the study and the usefulness of the coarse-grained process.

The perspective in [27, 28, 35] is to derive an *approximating* Markov process $\{\eta_t\}_{t \geq 0}$, for the true microscopic average $\mathbf{T}\sigma_t$ and control the ensuing numerical errors. For example, for the microscopic Arrhenius dynamics (2.4), we obtain a *birth-death* Markov process with states $\eta = \{\eta(k) \equiv \sum_{x \in C_k} \sigma(x) \,:\, k \in \bar{\Lambda}_M\}$ as a Markovian coarse-grained approximations of (2.5). The process is defined by adsorption and desorption rates of a single particle in the coarse cell C_k

$$c_a(k, \eta) = d_0(q - \eta(k))\,, \quad c_d(k, \eta) = d_0\eta(k) \exp\big[-\beta \bar{U}(k)\big]\,, \tag{2.6}$$

where the CG interaction potential is given by

$$\bar{U}(l) = \sum_{k \neq l\,,\, k \in \bar{\Lambda}_M} \bar{J}(l, k)\eta(k) + \bar{J}(0, 0)(\eta(l) - 1) - \bar{h}(l)\,,$$

and the two-body CG potential $\bar{J}$ is defined by the local average

$$\bar{J}(k, l) = \frac{1}{q^2} \sum_{x \in C_k} \sum_{y \in C_l} J(|x - y|) \;\text{ and } \bar{J}(k, k) - \frac{1}{q(q-1)} \sum_{x \in C_k} \sum_{y \in C_k} J(|x - y|)\,. \tag{2.7}$$

The potential $\bar{J}$ can be interpreted as a local mean-field approximation as it describes interactions between the cells C_k and C_l under the assumption that fluctuations of individual spins in the cells are small. The effective external field $\bar{h}$ is defined by a similar averaging. While such choice may appear natural from the physical point of view we show later that it also can be used as a suitable starting point for an error expansion. The CG diffusion dynamics can be derived in a similar way, see [35].

Computationally, the CG Markov process $\{\eta_t\}_{t\geq 0}$ is advantageous over the underlying microscopic $\{\sigma_t\}_{t\geq 0}$, since it has a smaller state space and can be simulated much more efficiently. An added computational advantage of the CG process is that the long-range interaction potential (2.10) is *compressed* through the CG procedure into the potential (2.7). We refer to all such CG algorithms as *Coarse-Grained Monte Carlo* (CGMC) methods. It is important to note that, by construction, the invariant measure for the CG process $\{\eta_t\}_{t\geq 0}$ is again a Gibbs measure given by

$$\bar{\mu}_M^{(0)}(d\eta) = \frac{1}{\bar{Z}_M^{(0)}} \exp\left(-\beta \bar{H}^{(0)}(\eta)\right) \bar{P}_M(d\eta)\,, \tag{2.8}$$

where the product binomial distribution $\bar{P}_M(d\eta)$ is the exact coarse-graining of the prior measure $P_N(d\sigma)$, and the associated CG Hamiltonian $\bar{H}^{(0)}$ is derived from the microscopic Hamiltonian H,

$$\bar{H}^{(0)}(\eta) = -\frac{1}{2} \sum_{l\in\bar{\Lambda}_M} \sum_{k\neq l\,,\, k\in\bar{\Lambda}_M} \bar{J}(k,l)\eta(k)\eta(l) - \frac{1}{2}\bar{J}(0,0) \sum_{l\in\bar{\Lambda}_M} \eta(l)(\eta(l)-1) + \sum_{k\in\bar{\Lambda}_M} \bar{h}\eta(k)\,. \tag{2.9}$$

Note that (2.9) defines a hierarchy of MC models spanning from Ising ($q = 1$) to the mean-field ($q = N$) models.

A closely related coarse-grained Hamiltonian was suggested independently in [24, 25], where it was constructed in an equilibrium context using a wavelet expansion. A different CGMC version of (2.9) was proposed in [11] for the CG of the nearest-neighbor Ising model. There the authors replaced the local mean field type approximations in the coarse cells with a quasi-chemical approximation, equivalent to a local Bethe approximation of the Ising model.

2.2. Lattice systems with intermediate/long-range interactions. Coarse-graining methods can provide a powerful computational tool in molecular simulations, however, it has been observed that in some regimes important macroscopic properties may not be captured properly. For instance, (over-)coarse graining in polymer systems may yield wrong predictions in the melt structure [1]; similarly wrong predictions on crystallization were also observed in the CG of complex fluids, [45]. In CGMC for lattice systems, hysteresis and critical behavior may not be captured properly for short and intermediate range potentials, [28], see also Figure 2.2. On the other hand, it has also been demonstrated computationally that CGMC performs well in the case of long-range interactions, where traditional MC methods experience a serious slow-down.

Motivated by such observations, in our recent work we analyzed under what conditions CG methods perform satisfactorily, and how to quantify the CG approximations from a numerical analysis perspective. We focus here on systems with *intermediate or long range* interactions, i.e., we consider symmetric potentials such that a particle at a given site on Λ_N interacts with neighboring sites at the distance less or equal to L and J has the form

$$J(x-y) = \frac{1}{L} V\left(\frac{N(x-y)}{L}\right)\,, \quad x\,, y \in \Lambda_N\,, \tag{2.10}$$

where $V(r) = V(-r)$, and $V(r) = 0$, $|r| \geq 1$. The interaction length L along with the inverse temperature β and properties of V will impose restriction on the coarse-graining size q. A similar analysis can be carried out also for long-range potentials with specific decay/blow-up conditions, see [3]. However, the assumption of a fixed potential length L simplifies the presentation of the estimates, see the definition of the parameter in (2.12).

Lattice systems characterized by (2.10) are good test cases for analyzing coarse-graining strategies since on one hand they exhibit complex behaviors such as phase transitions, nucleation and hysteresis, while on the other they present a serious challenge to conventional MC methods due to the large number of interacting neighbors involved in each MC step. A key point though is that systems with long-range interactions are usually more tractable analytically hence we are able to carry out detailed error analysis for coarse-graining approximations from a number of different perspectives. Next we review some of these results.

2.2.1. Error quantification in CGMC approximations. In [34] we studied CGMC algorithms from a numerical analysis perspective in the *non-equilibrium* context, by estimating the error between microscopic, (2.4) and coarse-grained, (2.6), adsorption/desorption lattice dynamics. The key step in this direction was to use the concept of *relative entropy*. We obtained an estimate between the time-dependent probability measures on the path space between CGMC, $\{\eta_t\}_{t\geq 0}$, and the exact projection on the coarse variables of the microscopic probability measure, $\mathbf{T}\sigma_t$

$$\mathcal{R}\left(\mathcal{D}^{\mathbf{T}\sigma,\rho}_{[0,T]} \,|\, \mathcal{D}^{\eta,\rho}_{[0,T]}\right) = O_T(\epsilon)\,, \quad t\in[0,T]\,, \tag{2.11}$$

where $\mathcal{D}^{\eta,\rho}_{[0,T]}$ (resp. $\mathcal{D}^{\mathbf{T}\sigma,\rho}_{[0,T]}$) is the distribution of $\{\eta_t\}_{t\in[0,T]}$ (resp. $\{\mathbf{T}\sigma_t\}_{t\in[0,T]}$) with fixed initial condition ρ. The specific relative entropy is defined as

$$\mathcal{R}\left(\mathcal{D}^{\mathbf{T}\sigma,\rho}_{[0,T]} \,|\, \mathcal{D}^{\eta,\rho}_{[0,T]}\right) = \frac{1}{N}\int \log\left(\frac{d\mathcal{D}^{\mathbf{T}\sigma,\rho}_{[0,T]}}{d\mathcal{D}^{\eta,\rho}_{[0,T]}}\right) d\mathcal{D}^{\mathbf{T}\sigma,\rho}_{[0,T]}$$

and ϵ is a "small" parameter given by

$$\epsilon \,\equiv\, \beta\|\nabla V\|_1\Big(\frac{q}{L}\Big)\,. \tag{2.12}$$

Such a relative entropy estimates give a mathematical criterion for analyzing the parameter regime, i.e., the degree of coarse-graining versus the interaction range, temperature and oscillations in the potential, for which CGMC is expected to give errors within a given tolerance. Note that the scaling factor N^{-1} is related to the fact that the system is extended, i.e., it has a large number of N particles. Consequently, the proper error quantity that needs to be tracked is the loss of information *per particle* (2.11).

Although (2.11) provides a bound on the entire time-dependent probabilities, estimates on specific observables are also desirable as they could be accurately simulated with less stringent CG strategies. This point of view necessitates the use of a weak convergence framework for the study of the error between CGMC and the direct numerical simulation. For instance, in [29] we proved the second order accuracy of CGMC in terms of the parameter ϵ in (2.11):

$$|\mathbb{E}[\psi(\mathbf{T}\sigma_T)] - \mathbb{E}[\psi(\eta_T)]| \leq C_T\epsilon^2\,, \quad t\in[0,T]\,. \tag{2.13}$$

In order to obtain that the constant C_T is independent of the system size N it is essential to obtain an upper bound on the total number of jumps up to time T. This is a key point related to the extensivity of the system and it is shown rigorously using a Bernstein-type argument on the discrete derivatives of the solutions to the backward equation. The technique is similar to obtaining Bernstein estimates for parabolic PDEs. In [29] we used these analytical results to guide CGMC algorithms and we demonstrated a CPU speed-up in demanding computational regimes that involved nucleation, phase transitions and metastability; see Figures 2.2, 2.3 and Table 2.1 for related simulations and comparisons with the higher order methods developed in Scheme 2.1.

2.2.2. Higher-order CG schemes and cluster expansions. In [32] our goal was to develop more accurate coarse-graining schemes than those proposed in [27, 35], and quantify their effectiveness in terms of a priori and a posteriori error analysis. The work [32] can also be viewed as a blueprint for tackling more

complex problems described in the subsequent sections. We briefly outline the method and the results. We first recall the renormalization group map, [26, 20]

$$e^{-\beta \bar{H}_M(\eta)} = \int e^{-\beta H_N(\sigma)} P_N(d\sigma|\eta)\,, \tag{2.14}$$

where $\bar{H}_M(\eta)$ is, by definition, the *exactly coarse-grained* Hamiltonian and $P_N(d\sigma|\eta)$ is the conditional probability (with respect to the prior distribution P_N) of having a microscopic configuration σ given a CG configuration η. Note that, due to the high-dimensional integration, $\bar{H}_M(\eta)$ cannot be calculated explicitly and used in numerical simulations. Our perspective is to approximate it by viewing it as a perturbation of $\bar{H}^{(0)}$ in (2.9). Using this first approximation we have

$$\bar{H}_M(\eta) \;=\; \bar{H}_M^{(0)}(\eta) - \frac{1}{\beta} \log \int e^{-\beta (H_N(\sigma) - \bar{H}_M^{(0)}(\eta))} P_N(d\sigma|\eta)\,. \tag{2.15}$$

The fact that the conditional probability $P_N(d\sigma|\eta)$ factorizes at the level of the coarse cells allowed us to use *cluster expansion* techniques to write a series expansion for $\bar{H}_M(\eta)$ around $\bar{H}_M^{(0)}$

$$\bar{H}_M(\eta) = \bar{H}_M^{(0)}(\eta) + \bar{H}_M^{(1)}(\eta) + \cdots + \bar{H}_M^{(p)}(\eta) + N \times \mathrm{O}(\epsilon^{p+1})\,,\; p = 1, \ldots \tag{2.16}$$

uniformly in η; also recall that Hamiltonians scale linearly with N, hence the $N \times O(\epsilon^{p+1})$ term. The small parameter ϵ is given again by (2.12).

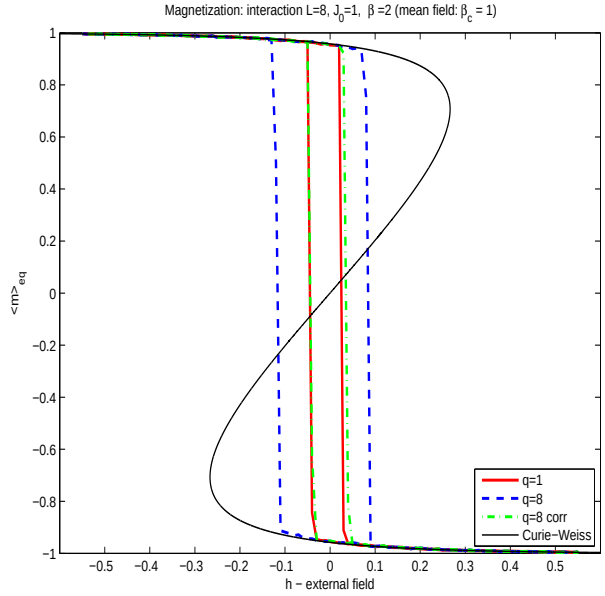


FIG. 2.2. *Comparison of CG schemes for $L = 8$: the direct numerical simulation (DNS) $q = 1$, $q = 8$ (CGMC without corrections), $q = 8$ (CGMC with corrections) and the mean-field Curie-Weiss model $L = N$ (solid line). Note the agreement of the (higher-order) CGMC with corrections and the DNS.*

The error estimates for the improved CG schemes are calculated in terms of the specific relative entropy of the corresponding equilibrium Gibbs measures, which here is used to assess the *information loss* for a given level of coarse-graining.

An immediate outcome of our analysis is that the CG scheme using $H^{(0)}$ (2.8) is actually second-order accurate, i.e.,

$$\mathcal{R}(\bar{\mu}_M^{(0)} | \mu_N \circ \mathbf{T}^{-1}) = O(\epsilon^2)\,.$$

This is due to cancellations induced by our choice of $H^{(0)}$ which is selected in (2.9) so that

$$\bar{H}^{(0)}(\eta) \;=\; \int H_N(\sigma) P_N(d\sigma|\eta)\,. \tag{2.17}$$

The correction terms $\bar{H}_M^{(1)}(\eta)$, $\bar{H}_M^{(2)}(\eta)$ etc. include multi-body interactions and are calculated explicitly in [32]. The term $\bar{H}_M^{(0)}(\eta)$ consists of only two-body interactions as does $\bar{H}_M^{(1)}(\eta)$ whereas $\bar{H}_M^{(2)}(\eta)$ consists of three-body interactions. As an example we define the improved CG scheme with $p = 3$.

SCHEME 2.1 (3^{rd} order accurate).

1. *Hamiltonian:* $\bar{H}_M^{(0)} + \bar{H}_M^{(1)} + \bar{H}_M^{(2)}$ *,where the corrections are*

$$\bar{H}_M^{(1)}(\eta) = \sum_{k<l} \Lambda_2^{(1)}(k,l;\,\eta(k),\eta(l)) + \sum_k \Lambda_1^{(1)}(k;\,\eta(k)) \tag{2.18}$$

where,

$$\begin{aligned}
\Lambda_2^{(1)}(k,l;\,\eta(k),\eta(l)) = \frac{\beta}{2}\big(& j_{kl}^1\,[E_2(\eta(k))E_2(\eta(l))E_1(\eta(k))E_2(\eta(l)) \\
& - E_2(\eta(k))E_2(\eta(l)) + E_1(\eta(k))E_1(\eta(l))] \\
& + j_{kl}^2\,[-2E_2(\eta(k))E_2(\eta(l)) + E_2(\eta(k))E_1(\eta(l)) \\
& + E_1(\eta(k))E_2(\eta(l))]\,\big) \\
\Lambda_1^{(1)}(k;\,\eta(k)) = \frac{\beta}{8}\big(& 4j_{kk}^2\,[-E_4(\eta(k)) + E_3(\eta(k)] \\
& + 2j_{kk}^1\,[E_4(\eta(k)) + E_2(\eta(k)) - 2E_3(\eta(k))]\big)\,,
\end{aligned}$$

and

$$\begin{aligned}
\bar{H}_M^{(2)}(\eta) = & \sum_{k\neq l} \Lambda_2^{(2)}(k,l,k;\,\eta(k),\eta(l),\eta(k)) \\
& + \sum_{k<l<m} \Lambda_3^{(2)}(k,l,m;\,\eta(k),\eta(l),\eta(m))\,,
\end{aligned} \tag{2.19}$$

where

$$\begin{aligned}
\Lambda_2^{(2)}(k,l,k;\,\eta(k),\eta(l),\eta(k)) = -\frac{\beta}{2}\big(& j_{kkl}^2\,[-E_3(\eta(k))E_1(\eta(l)) + E_2(\eta(k))E_1(\eta(l)) \\
& + E_1(\eta(k))E_2(\eta(l)) - E_3(\eta(l))E_1(\eta(k))\,])
\end{aligned}$$

$$\begin{aligned}
\Lambda_3^{(2)}(k,l,m;\,\eta(k),\eta(l),\eta(m)) = \beta\big(& j_{klm}^2\,[E_1(\eta(k))E_1(\eta(m))(1 - E_2(\eta(l)))] \\
& + j_{lmk}^2\,[E_1(\eta(k))E_1(\eta(l))(1 - E_2(\eta(m)))] \\
& + j_{mkl}^2\,[E_1(\eta(m))E_1(\eta(l))(1 - E_2(\eta(k)))]\big)
\end{aligned}$$

The terms E_i are defined in (2.20-2.23) and the quantities $j_{kl}^1, j_{kl}^2, j_{k_1k_2k_3}^2$ are defined in (2.24-2.26).

2. *CG Gibbs measure* $\bar{\mu}_{M,\beta}^{(2)}(d\eta) = \frac{1}{\bar{Z}_M^{(2)}} e^{-(\bar{H}_M^{(0)} + \bar{H}_M^{(1)} + \bar{H}_M^{(2)})} \bar{P}_M(d\eta)$.

The quantities E_i defined in the scheme are computed explicitly using conditional expectation of the

prior P_N, conditioned on $\eta(k) = \eta$

$$E_1(\eta) := \mathbb{E}[\sigma(x)|\eta] = \frac{\eta}{q}\,, \tag{2.20}$$

$$E_2(\eta) := \mathbb{E}[\sigma(x)\sigma(y)|\eta] = \frac{\eta(\eta-1)}{q(q-1)}\,, \tag{2.21}$$

$$E_3(\eta) := \mathbb{E}[\sigma(x)\sigma(y)\sigma(z)|\eta] = \frac{\eta(\eta-1)(\eta-2)}{q(q-1)(q-2)}\,, \tag{2.22}$$

$$E_4(\eta) := \mathbb{E}[\sigma(w)\sigma(x)\sigma(y)\sigma(z)|\eta] = \frac{\eta(\eta-1)(\eta-2)(\eta-3)}{q(q-1)(q-2)(q-3)}\,, \tag{2.23}$$

assuming in each case that all spin sites w, x, y, z are different. Furthermore, the interparticle potential covariances are defined as

$$j^1_{kl} := \sum_{\substack{x\in C_k\\ y\in C_l}} (J(x-y) - \bar{J}(k,l))^2 \tag{2.24}$$

$$j^2_{kl} := \sum_{\substack{x\in C_k\\ y,y'\in C_l}} (J(x-y) - \bar{J}(k,l))(J(x-y') - \bar{J}(k,l)) \tag{2.25}$$

$$j^2_{k_1k_2k_3} := \sum_{\substack{x\in C_{k_1}\\ y\in C_{k_2}, z\in C_{k_3}}} (J(x-y) - \bar{J}(k_1,k_2))(J(y-z) - \bar{J}(k_2,k_3)) \tag{2.26}$$

For the scheme (2.1) we have

$$\mathcal{R}(\bar{\mu}^{(2)}_{M,\beta}|\mu_{N,\beta}\circ\mathbf{T}^{-1}) = \mathrm{O}(\epsilon^3)\,, \tag{2.27}$$

where ϵ is given by (2.12).

Remark 2.1.

1. Clearly the choice of $\bar{H}^{(0)}$ in (2.16) is crucial to our method and it should be such that, (i) it is explicitly computable, and (ii) it provides a good estimate to the microscopic model in order to initiate the expansion (2.16). In Section 2.3 we extend this approach to more realistic systems with both short and long-range interactions and demonstrate how the small parameter ϵ (2.12) is determined.
2. Notice that we do not perform the usual high-temperature expansion (at the microscopic level) using the Bernoulli product measure $P_N(d\sigma)$ as it is standard in numerous applications of cluster expansions in statistical mechanics, e.g., [17, 21, 47]. We rather expand around (2.8) using the product properties of the conditional distribution $P_N(d\sigma|\eta)$ at the coarse level. This leads to a very substantial gain in the domain of applicability of the expansion. Cluster expansions around mean-field models have been carried out in the statistical physics literature (e.g. [48, 9, 6, 40]) for the purpose of studying critical behavior using a known (mean field) model as a starting point for the cluster expansions. These approaches are closely related to ours, although our expansion is around a better initial approximation than the mean-field model, as can be seen in a comparison of the mean-field and the CG (without corrections) simulations in Figure 2.2. This is true in part because (2.8) form a hierarchy in the CG level q, spanning from DNS ($q = 1$) to mean field, thus giving added flexibility with which the compression of the interaction range L can be handled. Furthermore, our focus is on the computational schemes and related numerical analysis questions such as a priori (2.27) and a posteriori estimates and adaptivity, see Section 5.
3. Although we have described the procedure leading to higher-order schemes for the case of Ising spins the derivation is applicable to systems with multiple chemical species or components. Obviously as

the number of components increases the number of inter-component interactions in multi-body terms of the expansions will grow. However, having explicit quantification of the error helps to balance computational complexity, the level of coarse-graining q and accuracy for estimated observables.

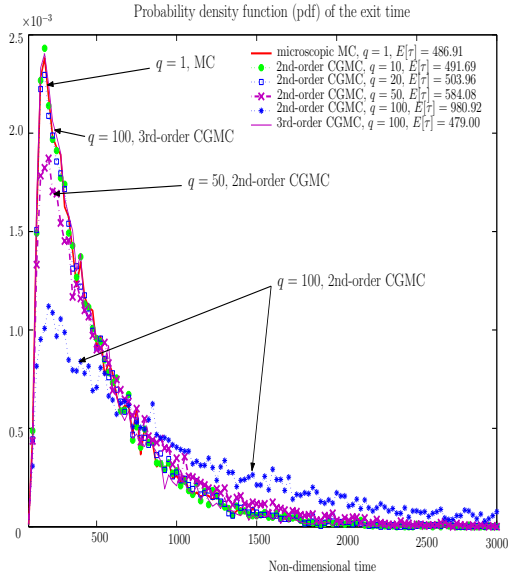


FIG. 2.3. *Probability density function (PDFs) for the switching times between two uniform states; comparisons between different coarse-grainings q. The estimated mean times for each PDF are shown in the figures. All PDFs comprised of 10000 samples and the histogram is approximated by 200 bins. The potential is a piecewise constant potential with interaction range is* $L = 100$ *and* $\beta J_0 = 5.0$.

TABLE 2.1

Approximation of $\bar{\tau}_T$, $\|\rho_\tau^q - \rho_\tau\|_{L^1}$ *and relative error. Measurements based on averaging over 10000 independent realizations for each q.*

$N = 1000$, $\beta J_0 = 6.0$, $h = 0.4406$

CGMC without corrections

L	q	$\bar{\tau}_T$	$\|\rho_\tau^q - \rho_\tau\|_{L^1}$	Rel. Err.
100	1	486.91	0	0
100	10	491.69	0.0022	1.16%
100	20	503.96	0.0025	3.68%
100	25	511.67	0.0032	5.27%
100	50	584.08	0.0074	20.17%
100	100	980.92	0.0246	101.82%

CGMC with corrections

L	q	$\bar{\tau}_T$	$\|\rho_\tau^q - \rho_\tau\|_{L^1}$	Rel. Err.
100	50	480.78	0.0025	1.08%
100	100	479.00	0.0028	1.45%

2.2.3. Computational implications of (2.27) and the effect of multi-body corrections. The multi-body interactions given by $\bar{H}_M^{(2)}(\eta)$ in Scheme 2.1 are temperature-dependent and hence they can be

crucial in low temperature regimes; similarly the dependence of the small parameter (2.12) in (2.27) on the interaction potential suggests that multi-body corrections are also important when $\|\nabla V\|_1$ is not small. The primary practical advantage of the higher-order corrections – more than higher accuracy – is in allowing us to extend the regime of validity of the expansion and to obtain an accurate CG of the Gibbs measure even if the parameter ϵ in (2.11) is not necessarily much smaller than one.

For instance in systems with intermediate (but not necessarily long) range interactions L. It has been already observed that hysteresis and critical behavior are not captured properly for short and intermediate range potentials, [28]. In Figure 2.2 we present an example with relatively short range interactions and high coarse-graining up to the interaction range, i.e., $q = L = 8$; we compare schemes with one term (CGMC without corrections) and two terms from the expansion (2.16). Note that here we are far from the $\epsilon \sim \frac{q}{L} \ll 1$ limit suggested by (2.27); the results are dramatically improved when the next-order corrections $\bar{H}_M^{(1,2)}$ are added. In contrast, and in agreement with the error estimate (2.27), for smoother interaction potentials such as long-range power laws corrections do not offer much improvement in the simulations, see the hysteresis diagram in Figure 2.4. Similar issues arise in the simulation of switching times and rare events, see Figures 2.3 and Table 2.1, where, again, the addition of higher-order corrections provide a dramatic improvement.

The multi-body interactions such as the three-body term $\bar{H}_M^{(2)}$ in Scheme 2.1 can be computationally expensive when implemented directly, due to the potentially large number of three-body terms. Hence it is important to understand (a) when the multi-body interactions are necessary in a CG scheme in order to achieve an error for a given tolerance in (2.27), and (b) if multi-body corrections are needed how they can be compressed via suitable truncations. For example, (2.27) and (2.12) suggest that when coarse-graining smooth long-range potentials the scheme based on (2.9) is an accurate approximation and does not require the computationally expensive multi-body interactions in Scheme 2.1. On the other hand we can make use of decay properties of the interaction potentials to show that the multi-body interactions can be compressed by truncating them to a given tolerance. Further strategies to decrease the computational cost of implementing the multi-body interactions are discussed in [3].

In the coarse graining literature of polymer chains essentially all existing CG schemes do not include the multi-body interactions. In view of our analysis it is not surprising that they perform well at high-temperature regimes while they tend to break down at lower temperatures. We will revisit this issue in Section 6, where we discuss the McCoy-Curro scheme for CG of polymer chains, which typically accounts only for two-body terms in the CG Hamiltonian.

2.3. Lattice systems with short and long-range interactions. Lattice and off-lattice systems characterized by competing interactions arise in numerous applications, for instance in micromagnetics, models of epitaxial growth, macromolecules, etc. We focus on lattice systems, and consider a Hamiltonian (2.1) where in addition to the long-range potential we incorporate a short-range potential

$$K(x-y) = \frac{1}{S} V_s \left(\frac{N|x-y|}{S} \right) , \quad x, y \in \Lambda_N , \tag{2.28}$$

where $S \ll L$ and V_s has similar properties as V in (2.10). When $S = 1$ we have the usual nearest-neighbor interaction. The new Hamiltonian includes long and short-range interactions, (2.10) and (2.28) respectively

$$H_N^{(L,S)} = H_N^{(L)} + H_N^{(S)} . \tag{2.29}$$

The corresponding Gibbs state (2.2) is defined accordingly. Our goal is to develop CGMC algorithms for this class of problems in equilibrium and non-equilibrium. To achieve this task, in the remainder of Section 2 and in Section 4 we explore two separate strategies, using our earlier work as a starting point:

1. An applied math/statistical mechanics perspective of expanding (using cluster expansions) around a "carefully" chosen first CG guess, in analogy to (2.16).
2. A statistics perspective that performs importance sampling with the transformation of measure derived from a "carefully" chosen CG guess, given, for instance, by (2.4) or (2.8).

The common theme behind both ideas is the observation that long-range interactions can be handled very efficiently by CGMC with or without corrections, as demonstrated in the previous section. On the other hand short-range interactions are relatively inexpensive and one would be tempted to simulate them with Direct Numerical Simulation (DNS) provided there is a suitable *splitting* of the algorithm into short and long-range parts, that can reproduce equilibrium Gibbs states and transient dynamics such as domain switching, nucleation, etc. Both approaches are in part analytical, relying on explicitly derived rates and, in part computational and statistical, when dealing with terms involving short-range interactions.

2.3.1. Coarse-graining and cluster expansions. The proposed approach follows the strategy outlined in Section 2.2 and described in more detail in [33]: *Step 1: Identify a suitable new prior distribution* $P_N^{(S)}$. We augment the product Bernoulli measure P_N in (2.2), adding short-range interactions in the form,

$$P_N^{(S)}(\sigma) = \otimes_k p^{(S)}(\sigma|_k)\,, \quad p^{(S)}(\sigma|_k) = \frac{1}{Z_{C_k}} e^{-H_{C_k}^{(S)}(\sigma|_k)} P_N(\sigma|_k)\,, \tag{2.30}$$

where on each coarse cell C_k,

$$H_{C_k}^{(S)}(\sigma|_k) = \sum_{\substack{x,y\in C_k,\\ y\neq x}} K(x-y)\sigma(x)\sigma(y)$$

is defined as the restriction of $H_N^{(S)}$, *without* including boundary interactions from the neighboring coarse cells; here $\sigma|_k$ denotes the restriction of σ on C_k.

Step 2: CG long-range Hamiltonian. In analogy to (2.17) we may now define the new CG long-range Hamiltonian

$$\bar{H}_M^{(L)}(\eta) \;=\; \int H_N^{(L)}(\sigma) P_N^{(S)}(d\sigma|\eta)\,, \tag{2.31}$$

where $P_N^{(S)}(d\sigma|\eta)$ is the conditional probability obtained from (2.30) by conditioning $P_N^{(S)}(d\sigma)$ on the CG state η. Note that under the approximation of periodic boundary conditions on each coarse cell or by a suitable approximation of the potential J by (2.7), the expression (2.31) is identical to (2.9). On the other hand, the short-range interactions enter in the CG prior (2.33) below.

Step 3: CG Gibbs state. We formulate the analogue of (2.8) by first considering the exact coarse-graining of (2.30). We obtain

$$\bar{P}_M^{(S)}(d\eta) \;=\; \prod_k \frac{1}{\bar{Z}_{C_k}} e^{-\bar{H}_{C_k}^{(S)}(\eta(k))} \bar{P}_M(d\eta) \tag{2.32}$$

where the CG Hamiltonian $\bar{H}_{C_k}^{(S)}$ is given by the exact renormalization group map on a *single* coarse cell C_k

$$e^{-\bar{H}_{C_k}^{(S)}(\eta(k))} = \int e^{-H_{C_k}^{(S)}(\sigma)} P_N(d\sigma|_k|\eta(k))\,.$$

Finally, the first approximation of the CG Gibbs state is defined as

$$\bar{\mu}_M^{(L,S)}(d\eta) = \frac{1}{Z_M^{(L,S)}} e^{-\beta \bar{H}_M^{(L)}(\eta)}\, \bar{P}_M^{(S)}(d\eta)\,. \tag{2.33}$$

Step 4: Error analysis The error between (2.33) and the projection of the microscopic Gibbs state on the coarse variables is shown in [33] to be

$$\mathcal{R}\left(\mu_M^{(L,S)} \,|\, \mu_N \circ \mathbf{T}^{-1}\right) = \mathrm{O}\left(\beta\left[\frac{S}{q}\|V_s\|_\infty + \frac{q}{L}\|\nabla V\|_1\right]\right), \quad \text{where } S \ll L. \tag{2.34}$$

This expression provides, at least qualitatively, an estimate on the regimes of validity of the method, and on the "optimal" CG, $q = q_{\text{opt}}$. The corresponding error according to (2.34) is then

$$q_{\text{opt}} \sim \sqrt{SL\frac{\|V_s\|_\infty}{\|\nabla V\|_1}}, \quad \mathcal{R}\left(\mu_M^{(L,S)} \,|\, \mu_N \circ \mathbf{T}^{-1}\right) = \mathrm{O}\left(\beta\sqrt{\frac{S}{L}\|V_s\|_\infty\|\nabla V\|_1}\right). \tag{2.35}$$

REMARK 2.2.

1. The choice of (2.30) can be ultimately justified by the error analysis in (2.34) and extensive numerical exploration. However, the original motivation rests on two related points. First, (2.30) is chosen so that $P_N^{(S)}(\sigma|\eta)$ is a *product measure* over coarse cells thus, applying (2.17) gives rise to a computable formula similar to (2.9). Secondly, we need to split short and long-range interactions in our simulations. We can carry out the former with DNS in a *single* coarse cell C_k during the evaluation of $\bar{H}_{C_k}^{(0)}$, and the latter with CGMC using (2.33).
2. We also note that the optimal CG $q = q_{\text{opt}}$ suggested by (2.34), allows us to coarse-grain beyond the range S, in contrast with the estimate (2.27), where the size of CG was restricted by the potential radius.
3. We have presented a choice for the initial approximation of the coarse-grained Hamiltonian such that the prior measure (2.30) preserves the product structure. The choice exploits explicit splitting of short and long-range interactions in the microscopic Hamiltonian. Consequently, the approximate CG Gibbs state (2.33) has only limited region of validity which is controlled by the error estimate (2.34). As a part of the general methodology explained in this section the expansion of coarse-graining error allows for identifying parameter regimes which lead to satisfactory approximation. More refined higher order estimates such as (2.27) are possible; they require the use of a cluster expansion and are discussed in [33].

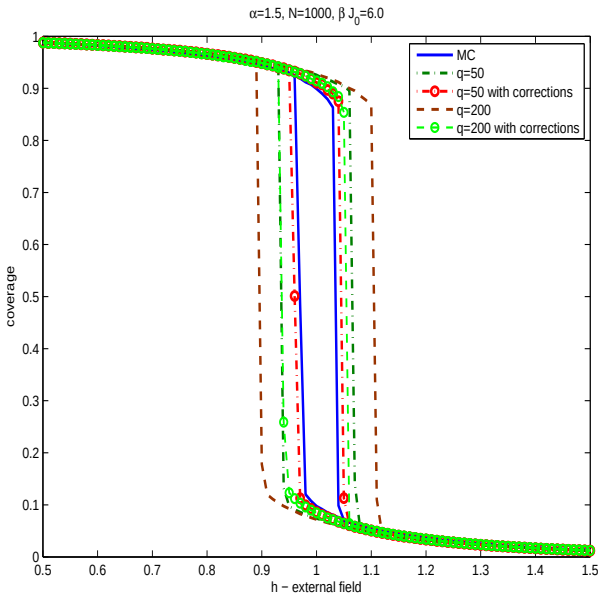


FIG. 2.4. *Comparison of different fully resolved ($q = 1$), coarse-grained $q = 50$, coarse-grained $q = 200$, coarse-grained $q = 50$ with corrections and coarse-grained $q = 100$ with corrections. The potential here is the decaying power-law potential with exponent $\alpha = 0.8$. Clearly, corrections terms have minimal impact as (2.6) is performing well for $q = 50$.*

2.3.2. Improved CG schemes. The source of error in (2.34) originates in two approximations: the $O(q/L)$ term is due to the approximation of the long-range interactions, see (2.27), while the error $O(S/q)$ is due to neglecting the coarse cell/cell interaction terms when considering the prior (2.30). It is precisely the latter omission that makes (2.30) a product measure. A closer inspection of the calculations reveals that this latter error depends on the strength of the nearest-neighbor interactions as described by $K = K(x-y)$, thus it is *non-negligible* for strong interactions or low temperatures. This is the regime where our product measure approximation (2.30) is expected to be inadequate. In this case we need to include correction terms that improve the CG approximation. In [33] we introduce a different prior by making a more educated guess than (2.30) which already includes *multi-body* coarse cell correlations. An expansion in this spirit was already suggested for the nearest-neighbor Ising model in [5]. The ensuing cluster expansion is again not the usual high-temperature series around the uniform prior P_N in (2.2) and as in [5] it allows us to carry through the expansion to much lower temperatures.

2.3.3. Computational strategies. For systems with long-range interactions, CGMC allows for fast and accurate simulation in spite of their presence: computational savings are obtained through the compression of the interaction potential (2.7) similarly to wavelet or fast multipole methods. Significant *additional* computational savings result due to the fact that CGMC involves only CG observables defined on a coarser lattice than the microscopic lattice Λ_N, [28]. The resulting reduction in the effective radius of interactions in (2.7) leads also to better KMC algorithmic properties when considering lists of neighbors and possibly easier development of parallel implementations.

In the case of short and long-range interactions the CGMC algorithm is based on the CG Gibbs measure (2.33), which suggests an efficient splitting of the algorithm into a CGMC long-range piece and a short-range simulation over a single coarse cell. Here we are faced with two options: (a) to pre-compute the CGMC rates by calculating $\bar{H}^{(0)}_{C_k}(\eta(k))$ in (2.33) with microscopic DNS and subsequently use them in CGMC; or (b) calculate $\bar{H}^{(0)}_{C_k}(\eta(k))$ on-the-fly during the CGMC simulation only for the $\eta(k)$'s involved in the *particular* simulated path. In the second case we move back and forth between a micro-solver (for the short-range) and a coarse step (using CGMC), as in general micro-solver methodologies proposed in [15, 36]. However, our derived formula (2.33) suggests precisely how the simulation should be split into a macro- and a micro-solver step. In conclusion, we remark that whatever approach is selected pre-computation or on-the-fly, the micro-simulation is performed in micro-canonical ensemble – for a fixed $\eta(k)$ – on a *single* coarse cell C_k, using only inexpensive short-range interactions. A related multigrid approach combined with CGMC was introduced earlier in [11] for the simulation of nearest-neighbor Ising models. There a splitting between microscopic and CG steps is also performed. The short-range microscopic simulations run on a single coarse cell with periodic boundary conditions are employed to calculate computationally CG rates such as (2.6).

3. Microscopic reconstruction. Reversing the coarse-graining, i.e., reproducing "atomistic" properties, directly from CG simulations is an issue that arises extensively in the polymer science literature, [51, 44]. The principal idea is that computationally inexpensive CG simulations will reproduce the large-scale structure and subsequently microscopic information will be added through *microscopic reconstruction*. One such an example is the calculation of diffusion of penetrants through polymer melts, reconstructed from CG simulation, [44]. Current approaches address the equilibrium case and rely on a semi-empirical approach by conditioning on most CG variables and carrying out a local relaxation of the microscopic system.

We first provide a general mathematical framework for the problem of reconstruction of the microscopic equilibrium properties of the system. We recall that $\mu_N(d\sigma)$ is the Gibbs measure (2.2) and we define

$$\bar{\mu}_M(d\eta) = \frac{1}{\bar{Z}_M} e^{-\beta \bar{H}(\eta)} \bar{P}_M(d\eta)$$

as the exact coarse-grained measure with $\bar{H}(\eta)$ given by (2.14). Then we have the relation

$$\mu_N(d\sigma) \;=\; e^{-\beta(H(\sigma)-\bar{H}(\eta))} P_N(d\sigma|\eta)\bar{\mu}_M(d\eta) \equiv \; \mu_N(d\sigma|\eta)\bar{\mu}_M(d\eta)\,. \tag{3.1}$$

We can think of the conditional probability $\mu_N(d\sigma|\eta)$ as (perfectly) reconstructing $\mu_N(d\sigma)$ from the exactly CG measure $\bar{\mu}_M(d\eta)$. Although many fine-scale configurations σ correspond to a single CG configuration η, the "reconstructed" conditional probability measure $\mu_N(d\sigma|\eta)$ is *uniquely* defined, given the microscopic and the coarse-grained measures $\mu_N(d\sigma)$ and $\bar{\mu}_M(d\eta)$ respectively.

A coarse-graining scheme provides us with an approximation $\bar{\mu}_M^{\rm app}(d\eta)$ for $\bar{\mu}_M(d\eta)$, at the coarse level. The approximation $\bar{\mu}_M^{\rm app}(d\eta)$ could be, for instance, any of the schemes discussed in Section 2.2 or 2.3. To provide a reconstruction we need to lift the measure $\bar{\mu}_M^{\rm app}(d\eta)$ to a measure $\mu_N^{\rm app}(d\sigma)$ on the microscopic configurations. That is, we need to specify a conditional probability $\nu_N(d\sigma|\eta)$ and set

$$\mu_N^{\rm app}(d\sigma) \,:=\, \nu_N(d\sigma|\eta)\bar{\mu}_M^{\rm app}(d\eta)\,. \tag{3.2}$$

In the spirit of our earlier discussion, it is natural to measure the efficiency of the reconstruction by the specific relative entropy $\mathcal{R}\,(\mu_N^{\rm app}|\mu_N)$. A simple computation shows that

$$\mathcal{R}\,(\mu_N^{\rm app}\,|\,\mu_N) \,=\, \mathcal{R}\,(\bar{\mu}_M^{\rm app}\,|\,\bar{\mu}_M) + \int \mathcal{R}\,(\nu_N(\cdot|\eta)\,|\,\mu_N(\cdot\,|\,\eta))\,\bar{\mu}_M^{\rm app}(d\eta)\,, \tag{3.3}$$

i.e., relative entropy splits the total error at the microscopic level into the sum of the error at the coarse level and the error made during the reconstruction.

The first term in (3.3) can be controlled, for example, by our cluster expansion results, see (2.27), (2.34). In order to obtain a successful reconstruction we then need to construct $\nu_N(d\sigma\,|\,\eta)$ such that (a) it is easily computable and implementable, and (b) the error $\mathcal{R}\,(\nu_N(d\sigma\,|\,\eta)\,|\,\mu_N(d\sigma\,|\,\eta))$ should be of the same order as the first term in (3.3).

Example: The simplest example of reconstruction is obtained by considering a microscopic system with intermediate/long range interactions (2.10), simulated in terms of the measure (2.8). We set

$$\bar{\mu}_M^{\rm app}(d\eta) \,=\, \bar{\mu}_M^{(0)}(d\eta)\,, \quad \nu_N(d\sigma\,|\,\eta) = P_N(d\sigma\,|\,\eta)\,, \tag{3.4}$$

i.e., we first sample the CG variables η involved in (2.8), using the CGMC algorithm (2.6); then we reconstruct the microscopic configuration σ by distributing the particles uniformly on the coarse cell, conditioned on the value of η. Since $P_N(d\sigma|\eta)$ is a product measure this can be done numerically in a very easy way, without communication between coarse cells and only at the coarse cells where an update has occurred in the CGMC algorithm. In this case the analysis in [31] yields the estimates

$$\mathcal{R}\left(\bar{\mu}_M^{(0)}\,|\,\bar{\mu}_M\right) = O(\epsilon^2)\,,\;\; \mathcal{R}\,(\mu_N(\cdot\,|\,\eta)\,|\,P_N(\cdot\,|\,\eta)) \,=\, \frac{\beta}{N}\left(\bar{H}^{(0)}(\eta) - \bar{H}(\eta)\right) \,=\, O(\epsilon^2)\,,$$

hence the reconstruction is second order accurate and of the same order as the coarse-graining given by (2.9).

Note that (3.4) implies that if we use Scheme 2.1, which is third-order accurate, and carry out the reconstruction using $P_N(d\sigma|\eta)$, we will obtain a reconstruction (3.2) which is only second order accurate. This problem is treated in the recent paper [49] where the authors decompose the coarse lattice into odd and even alternating coarse cells (denoted $\mathcal{O}$ and $\mathcal{E}$). The conditional probabilities are then decomposed in the form

$$\mu_N(\sigma|\eta) = \mu_{N,\mathcal{E}}(\sigma^{\mathcal{E}}|\eta)\;\mu_{N,\mathcal{O}}(\sigma^{\mathcal{O}}|\sigma^{\mathcal{E}},\eta). \tag{3.5}$$

and similarly for $\nu_N(\cdot|\eta)$. The second factor $\mu_{N,\mathcal{O}}(\sigma^{\mathcal{O}}|\sigma^{\mathcal{E}},\eta)$ factorizes nicely and can be computed through direct simulation and used for reconstruction. The first term $\mu_{N,\mathcal{E}}(\sigma^{\mathcal{E}}|\eta)$ is approximated in [49] using cluster expansions.

The reconstruction in the case of short and long-range interactions can be done in a similar way to (3.4), at least as a first approximation, using the conditional prior distribution $P_N^{(S)}(d\sigma|\eta)$ arising from (2.30). It is discussed in more detail in [33].

Dynamics and reconstruction. The problem of reconstruction can also be formulated for the case of dynamics. Interestingly the issue of microscopic reconstruction arises naturally in the mathematical error analysis in [34, 29]. One of the mathematical difficulties in carrying out error estimates lies in the fact that $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ defined in (2.5) is (in general) *non-Markovian*, while it needs to be compared to the approximating Markov process $\{\eta_t\}_{t\geq 0}$. To circumvent this mathematical difficulty in [34] the authors construct an auxiliary microscopic Markov process $\{\gamma_t\}_{t\geq 0}$ from $\{\eta_t\}_{t\geq 0}$, which is an approximation of $\{\sigma_t\}_{t\geq 0}$. The reconstruction is carried out by distributing particles uniformly on each coarse cell, using precisely (3.4). This enforces a local equilibrium, which is crucial in obtaining the 2nd order accuracy in (2.13). It is shown in [29] that for any *microscopic* observable ϕ, one has an analogous error estimate to (2.13) for $\{\sigma_t\}_{t\geq 0}$ and $\{\gamma_t\}_{t\geq 0}$

$$|\mathbb{E}[\phi(\sigma_T)] - \mathbb{E}[\phi(\gamma_T)]| \leq C_T\epsilon^2\,. \tag{3.6}$$

It is also conceivable that $\{\gamma_t\}_{t\geq 0}$ can be used as a systematic procedure for model refinement or adaptivity. We discuss such tools in Section 5.

4. Importance sampling and coarse-graining. In the study of lattice systems with short and long-range interactions in Section 2.3 our approach was essentially based on asymptotics carried around a reasonable first guess (e.g. (2.27) or (2.34)). This reference state was further improved by including more terms in the cluster expansion. It is possible to pursue a different but complementary perspective on CG in both equilibrium and non-equilibrium setting, which is partly based on a statistics approach. The proposed method *replaces* the explicit asymptotics with statistics, by carrying out importance sampling (IS) around a first CG guess that serves as a "trial process" or "trial distribution". As in all importance sampling methods, the goal is to sample from the trial distribution which is chosen in such a way that it is easy to simulate. On the other hand, the careful choice of trial distributions is important as it can lead to more efficient sampling, [41]. Finally, we note that a wealth of sophisticated IS techniques have been long employed in the simulation of equilibrium problems in statistical mechanics. Examples include protein folding, polymer chains, etc., see for instance [41, Sections 3.1, 4.1] and the extensive references there in. Similarly, importance sampling is used extensively in finance, e.g. [19], and in the study of stochastic systems exhibiting rare events, e.g., [14].

In our presentation we focus on the case of dynamics, since the equilibrium sampling is conceptually same but significantly simpler to implement. For the sake of definiteness we present the method for the case of adsorption/desorption dynamics (2.4) used as a demonstrating example in Section 2. We assume the potential

$$U(x,\sigma) = U^{(S)}(x,\sigma) + U^{(L)}(x,\sigma)\,, \tag{4.1}$$

corresponding the the short and long-range interactions in the Hamiltonian (2.29).

Importance sampling and change of measure. The principal tool behind the importance sampling algorithm is a change of measure. We employ a change of measure in order to replace the evaluation of observables for the lattice dynamics σ_t with the evaluation of *modified* observables for a numerically more tractable dynamics γ_t. Since both processes are continuous-time Markov jump processes we can calculate explicitly, in terms of the rates, see, e.g., Appendix 1 in [37], the change of measure (i.e., the Radon-Nikodym derivative) $(d\mathcal{D}^{\sigma,\rho}_{[0,T]}/d\mathcal{D}^{\gamma,\rho}_{[0,T]})$ on any path $((\rho_t)_{t\in[0,T]})$, where $\mathcal{D}^{\sigma,\rho}_{[0,T]}$ (resp. $\mathcal{D}^{\gamma,\rho}_{[0,T]}$) is the distribution of $\{\sigma_t\}_{t\in[0,T]}$ (resp. $\{\gamma_t\}_{t\in[0,T]}$) with the fixed initial condition ρ. More precisely,

$$\frac{d\mathcal{D}^{\sigma,\rho}_{[0,T]}}{d\mathcal{D}^{\gamma,\rho}_{[0,T]}}((\rho_t)_{t\in[0,T]}) = \exp\left\{\int_0^T [\lambda_\sigma(\rho_s) - \lambda_\gamma(\rho_s)]\,ds - \sum_{s\leq T}\log\frac{\lambda_\sigma(\rho_{s-})p_\sigma(\rho_{s-},\rho_s)}{\lambda_\gamma(\rho_{s-})p_\gamma(\rho_{s-},\rho_s)}\right\} \tag{4.2}$$

in terms of the rates of σ_t and γ_t, where $(\lambda_\sigma, p_\sigma)$ and $(\lambda_\gamma, p_\gamma)$ are the skeleton processes, see Appendix 1 in [37]. The sum in the last term runs over the jump times of $\{\gamma_t\}_{t\in[0,T]}$. Then for any observable ϕ we

have the transformation formula that relates expected values computed with respect to the measure of the process $\{\sigma_t\}_{t\in[0,T]}$ in terms of the averaging on the process $\{\gamma_t\}_{t\in[0,T]}$

$$\mathbb{E}^{\sigma}_{\rho}[\phi(\sigma_T)] = \mathbb{E}^{\gamma}_{\rho}\left[\phi(\gamma_T)\left(\frac{d\mathcal{D}^{\sigma,\rho}_{[0,T]}}{d\mathcal{D}^{\gamma,\rho}_{[0,T]}}((\gamma_t)_{t\in[0,T]})\right)\right] \tag{4.3}$$

Relation (4.3) implies that the evaluation of observables for σ_t at time T reduces to the evaluation of modified observables that depend on the entire path $\{\gamma_t\}_{t\in[0,T]}$. Essentially, the Radon-Nikodym derivative represents a memory term that provides an exact correction, *possibly of order one*, in the approximation of σ_t by γ_t.

EXAMPLE: A simple choice for a trial process, although by no means the definite one, as we discuss further below, is to choose γ_t associated only with the long-range interactions J. For this choice of the trial process we have

$$\begin{aligned}\frac{d\mathcal{D}^{\sigma,\rho}_{[0,T]}}{d\mathcal{D}^{\gamma,\rho}_{[0,T]}}((\gamma_t)_{t\in[0,T]}) = \exp\Bigg\{\int_0^T \sum_{x\in\Lambda_N} d_0 e^{-\beta\bar{U}^{(L)}(k(x),\mathbf{T}\gamma_s)}\gamma_s(x)[e^{-\beta U^{(S)}(x,\gamma_s)}-1]ds \\ + \sum_{\substack{s\le T \\ \exists x\in\Lambda_N \gamma_{s^-}(x)=\gamma_s(x)+1}} \beta U^{(S)}(x,\gamma_{s^-})\Bigg\},\end{aligned} \tag{4.4}$$

Coarse-graining & Reconstruction. At first glance the formula (4.4) does not seem to provide a useful algorithmic tool since it involves the microscopic process γ_t. We construct the trial microscopic process γ_t by reconstruction from a computationally inexpensive CG process η_t. The reconstruction could be carried out with a controlled error as derived in [34, 29] and not on the entire lattice but only on a particular path. At each jump time we locally reconstruct the configuration only on the coarse cells where the CG dynamics η_t exhibit a jump. Comparing to the direct numerical simulation performed with the conventional MC the proposed algorithm uses only the compressed long-range potential and the microscopic process does not have to be constructed on each microscopic cell. The expression (4.4) provides an abstract *splitting* of the computational effort into the CGMC piece for the long-range interactions and the DNS piece for the relatively inexpensive short-range component on a smaller configuration space. We include the exact short-range effects through the calculation of the Radon-Nikodym derivative (4.4), using the reconstructed process γ_t.

The outlined approach in its generality gives only basic guidelines and it has to be tailored to a specific case of interactions to be simulated. The mathematical and algorithmic aspects of the method, as well as extensive simulations will appear in [30]. Finally, we mention several issues that require special attention.

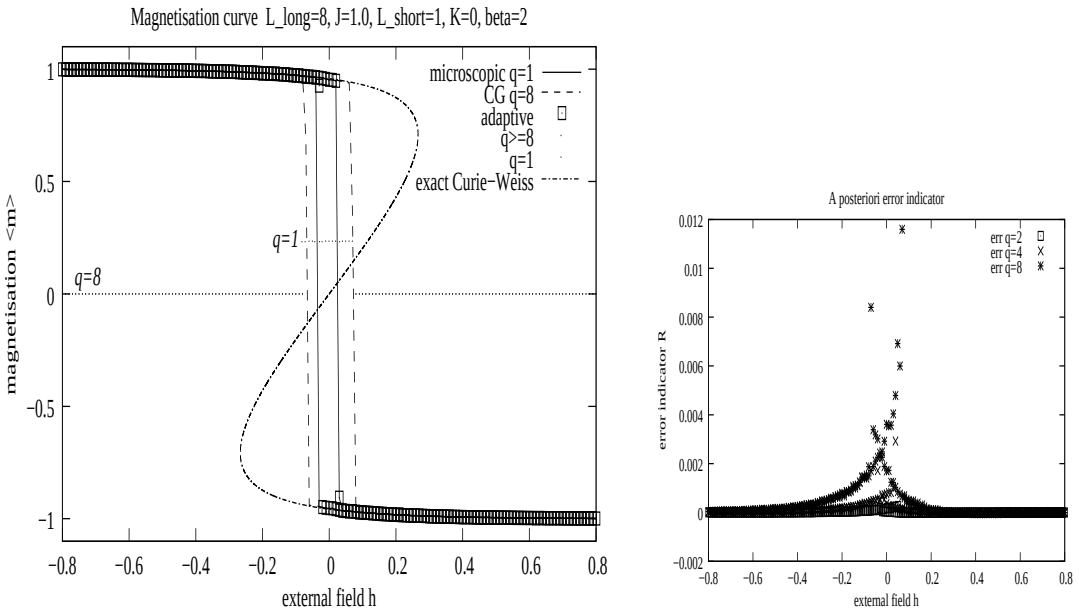
REMARK 4.1.

1. Although the formula (4.3) holds in general, the choice of the trial dynamics γ_t is a crucial step in all importance sampling methods. In this choice we have to balance two demands: (a) γ_t needs to be easy to simulate, or at least easier than DNS, and (b) the statistical sampling in the calculation of (4.3) needs to be efficient. For instance, in the previous example, γ_t is expected to work well at least when its time dependent distribution is a reasonable approximation of the original process $\{\sigma_t\}_{t\ge0}$, i.e., when the short-range interactions are not dominant. Otherwise we can improve the first guess by including a self-interaction approximation for the short range piece such as the one in (2.9). In general, an advantage of IS methods is that even a suboptimal guess for the trial distribution can give good results.
2. Sampling from a given distribution using importance sampling is a powerful statistical method, [46, 41], which directly inspired all the aforementioned material. It is worth emphasizing that our focus is however on sampling *non-equilibrium* processes. Closer to the statistical practice, we can

sample, in a similar way as the one proposed above, from the distribution (2.2) with the short/long Hamiltonian (2.29). In the case of equilibrium simulations, we replace the Radon-Nikodym derivative in (4.3) with $d\mu_N^{(L,S)}/d\hat{\mu}_N$, where $\mu_N^{(L,S)}$ is the Gibbs measure corresponding to (2.29), while $\hat{\mu}_N$ is the trial equilibrium measure from which we would prefer to sample. One such a choice, in analogy to (4.4), could be the reconstructed CG measure that includes only long range interactions, see (2.8). A better trial distribution, as suggested by (2.34), could be reconstructed from (2.33). Note that for any such choice, estimating the corresponding partition functions does not pose a problem since we can either employ the dynamics in the sampling of the Radon-Nikodym derivative, as in (4.4), or use a more sophisticated estimator for $d\mu_N^{(L,S)}/d\hat{\mu}_N$ that does not depend on prefactors, [41].

3. A possible advantage of the importance sampling approach over the asymptotics methodology discussed in Section 2.3.1 is that the trial distribution does not have to approximate the target distribution for the method to perform well. On the other hand, it seems reasonable to explore the possibility that the methods can complement each other, as the different trial processes γ_t proposed above, suggest. Similarly, for equilibrium sampling we could use the reconstruction of (2.8), while a better trial distribution, as suggested by (2.34) can be reconstructed from (2.33) or even taking further into account the coarse cell corrections, [30].

5. A posteriori estimation and adaptive coarse-graining. In this section we explore the possibility of having adaptive coarse-graining schemes, where the actions of refining and coarsening are carried out according to an error indicator that can be estimated in the course of coarse-grained simulation. Such estimates in the finite element/PDE literature are known as *a posteriori* error estimates. Here we study the error control in the context of lattice systems considered earlier in Section 2.2.



(a) Adaptive CG (b) A posteriori error indicator

FIG. 5.1. *Demonstration of adaptivity dictated by the a posteriori error indicator: no CG is performed close to the hysteresis region; the simulation uses the "crude" CGMC without corrections with the coarse-cell size* q variable. *The level of coarse-graining (the parameter* q*) used when computing the magnetization curve for different values of* h *is indicated in the phase diagram (a). The Hamiltonian consists of the long-range interactions with* $L = 8$ *only. The curve for the Curie-Weiss model represents the long-range interactions with* $L = N$*,* ($N = 1000$)*. The a posteriori indicator estimated in the CG simulation is depicted in (b).*

The error estimate (2.27) combined with the cluster expansion in Scheme 2.1 yield an explicit representation of the error in the coarse-grained numerical approximation. In fact, in [32] we showed the following a posteriori error estimate for the CG scheme based on (2.8)

$$\mathcal{R}(\bar{\mu}^{(0)}_{M,\beta}|\mu_{N,\beta}\circ\mathbf{T}^{-1}) = \frac{1}{N}\mathbb{E}_{\bar{\mu}^{(0)}_{M,\beta}}[R(\eta)] + \frac{1}{N}\log\left(\mathbb{E}_{\bar{\mu}^{(0)}_{M,\beta}}[e^{R(\eta)}]\right) + \mathrm{O}(\epsilon^3)\,, \tag{5.1}$$

where the residuum operator R is given by

$$R(\eta) = \bar{H}^{(1)}_M(\eta) + \bar{H}^{(2)}_M(\eta)\,.$$

This error representation indicates that the error in coarse-graining can be computed on-the-fly, during a coarse-grained simulation, exclusively in terms of the coarse observables η. The error involves only $\bar{H}^{(1,2)}_M(\eta)$ plus a controlled error of order $\mathrm{O}(\epsilon^3)$.

In [31] we implemented the sharp a posteriori estimates (5.1), tracking them throughout our simulations, where the on-the-fly estimated error served as a diagnostic tool for the quality of the coarse-grained simulations. It indicated when a particular level of coarse-graining can produce excessive error and needs to be refined, or when it can be safely coarsened further in order to speed up the simulation. This approach leads to adaptive coarse-graining of the configuration space and clearly relies on the fact that the coarse-grained models introduced in [27, 32] form a hierarchy. The hierarchy includes the microscopic description at the finest level, and allows for a seamless transition between different resolutions.

In [31] we demonstrated the use of such diagnostics and the ensuing adaptive coarse-grainings in the numerical calculation of phase diagrams in systems with combined short and long-range interactions. In this case it turns out that most of the phase diagram is constructed using coarse levels and hence inexpensive CG simulations are used, while the relatively fewer regimes where critical phenomena occur, require finer, or even fully resolved simulations. The transitions from finer to coarser scales and back are done on-the-fly, based on the a posteriori error computation.

Such application of our a posteriori estimates is demonstrated in Figure 5.1. The goal of those simulations is to computationally construct a phase diagram, plotting the average magnetization (or coverage) and the external field in a parameter regime where hysteresis is exhibited. We carry out a "crude" CGMC simulation based on (2.8), for each external field. The a posteriori estimate demonstrates that for most fields we only need very coarse cells. The error increases close to the hysteresis region (recall this is the region where fixed-size CG without corrections did not perform well, see Figure 2.2), and we are forced to refine our simulation almost down to the microscopic level. The computational gain against full DNS stems from the very limited use of such microscopic simulations as compared to the majority of external fields. Note also the excellent agreement of adaptive and DNS in Figure 5.1.

The refinement or coarsening are governed by the error indicator in (5.1), although this indicator does not easily relate to the absolute error of a given observable (e.g., magnetization). In the presented simulations a simple strategy has been adopted: the change of the level is controlled by the relative value of the indicator with respect to its maximal value along the simulation path. More elaborate strategies for the error control are also possible and will be discussed elsewhere. Earlier work that uses only an upper bound and not the sharp error expansion in (5.1) can be found in [10, 13]. These papers are more related to the spirit of adaptive finite element methods for PDEs, where a posteriori errors are typically used to construct spatially adaptive coarse-grainings.

6. Coarse-graining and reconstruction in macromolecular systems. The coarse-graining of polymeric and other macromolecular systems has attracted considerable attention in polymer science and engineering, [38, 44]. The primary goal is to group together, in a systematic manner, several atoms on a macromolecule, see Figure 2.1(b), creating an effective new chain, as means of reducing the degrees of freedom of the original system. Key challenges include the presence of complex short and long-range interactions, as well as the off-lattice nature of the models. In many applications, for instance in the simulation of

microscopic penetrants in a polymeric melt, it is also important to be able to reconstruct microscopic details of the chain by "reversing" the coarse-graining, [51, 44].

6.1. The microscopic model. Here we consider as our microscopic system the United Atom (UA) model. This point of view is typically taken in the coarse-graining literature in polymer science, e.g., [7], [16], [22], as it is expected that the coarse-graining of UA models captures many of the related technical and computational difficulties. We note that the UA model is not a fully resolved atomistic description of the polymeric chain but involves a coarse-graining procedure that defines the united atoms in the chain. This class of models consists of n macromolecules (e.g., polymer chains) in a simulation box with volume V at a fixed temperature T. Each molecule contains m atoms. Thus the system consists of $N = nm$ microscopic particles at the positions $X = (x_1, \dots, x_N) \in \mathbf{X}$, where $x_i \in \mathbb{R}^3$ is the position vector of the ith atom. The interactions in the system are described by the Hamiltonian

$$H_N(X) = H_{\rm b}(X) + H_{\rm nb}(X) + H_{\rm Coul}(X) + H_{\rm wall}(X) + H_{\rm kin}(X)\,. \tag{6.1}$$

The first term $H_{\rm b}$ defines *short-range* (bonded) interactions between neighboring atoms in each individual polymer chain. This term is defined in terms of a pair-wise potential $U_{\rm b}$, i.e., $H_{\rm b} = \sum U_{\rm b}$. The second term $H_{\rm nb}$ describes *long-range* (non-bonded) interactions between atoms in different chains and is typically associated with the Lennard-Jones two-body potential $U_{\rm nb}$. The Coulomb term $H_{\rm Coul}$ describes interactions associated with charged macromolecules, while $H_{\rm wall}$ interactions with walls. The term $H_{\rm kin}$ is the total kinetic energy of the system. In our analysis we focus on the canonical ensemble (also called NVT ensemble), at the inverse temperature β, given by the microscopic Gibbs measure

$$\mu(dX) = \frac{1}{Z} e^{-\beta H_N(X)} \prod_i dx_i\,, \quad Z = \int_{\mathbf{X}} e^{-\beta H_N(X)} \prod_i dx_i\,. \tag{6.2}$$

The classical microscopic dynamics is derived from the Newtonian dynamics imbedded into a heat bath. The resulting stochastic Langevin dynamics is then compatible with the canonical equilibrium measure (6.2) which is the invariant measure for the process. Therefore Langevin dynamics is widely employed for simulations and sampling of the canonical Gibbs measure; the dynamics is formulated in terms of Newtonian equations of motion

$$m_i \frac{d^2 x_i}{dt^2} = -\nabla_{x_i} H_N - \gamma \frac{dx_i}{dt} + \xi_i(t)\,, \quad i = 1, ..., N\,. \tag{6.3}$$

The damping is given by the friction constant γ and m_i denotes the mass of the i-th particle. The random forcing is the Wiener process $\xi_i(t)$ with the covariance matrix that couples β and γ through the fluctuation-dissipation relation and ensures that the Gibbs measure is the invariant measure of the resulting process $\{X_t\}_{t\geq 0}$.

6.2. Coarse-graining of polymer chains. We begin our discussion by first setting the notation and defining a set of coarse-grained variables. We assume a polymeric system with n molecules (polymer chains) in a simulation box with the volume V at a fixed temperature T. In the CG description of the system, k microscopic atoms in a single chain are grouped into one CG state (analogous to η in (2.5)), which is usually referred to as a "super-atom". The result of this procedure is a system of $M = N/k$ CG particles. The positions of CG particles are denoted by the coarse variables $Q = (q_1, \dots, q_M)$ where $q_i \in \mathbb{R}^3$ represents the position of each super-atom. The first goal of the coarse-graining procedures is to find the CG interactions. In fact, as in the lattice case (2.14), the exactly coarse-grained Hamiltonian $\bar{H}(Q)$ is defined by the relation

$$e^{-\beta \bar{H}_M(Q)} = \int_{\{X \in \mathbf{X} | \mathbf{T}X = Q\}} e^{-\beta H_N(X)}\, dX\,. \tag{6.4}$$

The evaluation of the integral involves high dimensional integration and one way to make some progress towards calculating $\bar{H}_M(Q)$ is to adopt a semi-empirical strategy that will reduce the dimension and hence

the computational cost in (6.4). Below we briefly describe two different perspectives developed recently in the literature [44, 50, 7, 16].

6.2.1. Statistical methods. In view of the CG methods for lattice systems using importance sampling presented in Section 4, we note that a *parametric statistics*-based approach for CG of macromolecules was developed earlier in the polymer science literature, [44]. In this class of methods a parametrization of a CG potential is introduced, based on an assumed functional form (e.g., a Lennard-Jones type potential), and is sequentially optimized, against the pair distribution function $g_{\rm CG}(r)$. The pair distribution function can be obtained either from the microscopic simulations in (6.3), or from experimental data. A more advanced approach involves optimization with respect to a fixed set of observables beyond just the pair distribution function. The main drawback of this approach is that it is purely empirical and does not provide any physical information about the resulting potential. In addition the method does not give a unique CG potential since it is essentially a statistical fitting that may have more than one solution which will produce the desired pair distribution function or more generally a fixed number of observables.

6.2.2. Computational statistical mechanics methods. In this direction the goal is to calculate directly the CG Hamiltonian $\bar{H}_M(Q)$ in (6.4) using MC sampling. However, as the computational cost to sample (6.4) is prohibitive, the perspective in the polymer science literature is to construct an approximation, checked eventually against direct numerical simulations (DNS). Such approximations are obtained by a series of simplifying assumptions on the structure of the CG Hamiltonian $\bar{H}_M(Q)$ that allow to reduce the integration dimension and hence the computational cost in (6.4), [50]. We briefly review this approach below. For the sake of simplicity we focus on the case where only bonded and non-bonded interactions are present in the Hamiltonian, i.e., $H_N(X) = H_{\rm b}(X) + H_{\rm nb}(X)$. Then the CG approximation involves the following steps:

Step 1 From (6.4) we have,

$$\bar{H}_M(Q) = -\frac{1}{\beta}\log\int_{\{X\in\mathbf{X}|\mathbf{T}X=Q\}} e^{-\beta H_N(X)}\,dX\,. \tag{6.5}$$

Although (6.5) is not additive, the following approximation is made

$$\bar{H}_M = \bar{H}_{\rm b} + \bar{H}_{\rm nb}\,, \tag{6.6}$$

where $\bar{H}_{\rm b}$ and $\bar{H}_{\rm nb}$ are CG bonded and non-bonded Hamiltonians (to be calculated separately below). Since β is proportional to the inverse temperature, clearly the approximation (6.6) is expected to be a reasonable one at relatively high temperatures. Furthermore, it is assumed $\bar{H}_{\rm b} = \sum \bar{U}_{\rm b}$ and $\bar{H}_{\rm nb} = \sum \bar{U}_{\rm nb}$ for corresponding CG potentials that will need to be calculated numerically.

Step 2 It is assumed that the term $\bar{U}_{\rm b}$ decouples as $\bar{U}^{\theta}_{\rm b} + \bar{U}^{\phi}_{\rm b} + \bar{U}^{r}_{\rm b}$ where each term depends *only* on torsion angle ϕ, rotation angle θ and distance r respectively between CG particles.

Step 3 The long-range term $\bar{U}_{\rm nb}$ is approximated as a binary interaction between CG particles. This interaction is further assumed to be rotationally invariant, although the latter is clearly not true in general.

These simplifying assumptions allow for the break-up of the computational evaluations. For instance, $U_{\rm b}$ is calculated through atomistic simulations, performed on *one* isolated oligomer by sampling the phase space. Thus we obtain the probability distribution function between the CG particles at a given angle or distance and the effective bonded CG potentials are calculated by taking a logarithm similarly to (6.5). Consequently, we define the CG bonded Hamiltonian as

$$\bar{H}_{\rm b}(Q) = \sum_{\langle i,j\rangle} \bar{U}_{\rm b}(q_i - q_j)\,, \tag{6.7}$$

where $\langle \cdot, \cdot \rangle$ indicates the summation is over all neighboring coarse sites i and j on the same macromolecule (chain).

The non-bonded CG Hamiltonian is assumed to be the sum of only two-body, radial CG non-bonded interactions

$$\bar{H}_{\rm nb}(Q) = \sum_{i,j} \bar{U}_{\rm nb}(|q_i - q_j|)\,. \tag{6.8}$$

A typical approach to calculate $\bar{U}_{\rm nb}$ is to use the *McCoy-Curro* scheme. This scheme was first proposed in [43] for calculating the effective non-bonded interaction between the centers of mass q_1 and q_2 of two isolated small molecules (and later extended to polymer chains in [16]). The scheme yields a formula for the effective potential

$$\bar{U}_{nb}(|q_i - q_j|) = -\frac{1}{\beta} \log \int_{\{X:\mathbf{T}X=(q_i,q_j)\}} e^{-\beta H_{2m}(X)} dX\,, \tag{6.9}$$

where $H_{2m}(X)$ is the Hamiltonian involving the detailed description of two isolated molecules, each of them containing m atoms. In other words, $\bar{U}_{\rm nb}(|q_i - q_j|)$ is a two-body CG interaction, where all other molecules are disregarded. All these calculations need to be repeated whenever the parameters N, V, T are changed. When compared to DNS, this approach yields good results in some regimes. However, deviations are also observed both in the structure and dynamics [1, 22]. Differences are attributed to assumptions such as (i)-(iii) and (6.8) that are necessary for the computational implementation of the CG potentials. From a numerical analysis perspective such simplifying assumptions are nothing else but approximations generating numerical errors. As discussed earlier this error is expected to be small in relatively high temperatures, although a detailed understanding of the parameter regimes when the approximations are valid is still lacking for the macromolecular case. However, our work in lattice systems addresses directly this issue, as we see in (2.27) and (2.12).

In the off-lattice case we can follow the strategy outlined in Section 2 for the lattice case, starting with identifying a suitable first approximation $\bar{H}_M^{(0)}(Q)$. Such an approximation can be given by the McCoy-Curro scheme (6.8), (6.9) or the off-lattice analogue of (2.17), or even a combination of the two. Subsequently, we can rewrite (6.4) as

$$\bar{H}_M(Q) = \bar{H}_M^{(0)}(Q) - \frac{1}{\beta} \log \int_{\{X|\mathbf{T}X=Q\}} e^{-\beta(H_N(X) - \bar{H}_M^{(0)}(Q))}\, dX\,. \tag{6.10}$$

Cluster expansions can be used to further improve the initial approximation $\bar{H}_M^{(0)}(Q)$, similarly to (2.16). A detailed presentation, analysis and extensive simulations for the polymers case will appear in [23].

6.3. Connections to CG methods for lattice systems. Next we make some precise connections between the two types of models and their CG strategies. We pursue the connection between the McCoy-Curro CG methodology to Scheme 2.1. We compare the two approaches by using the cluster expansion methodology we presented for the lattice systems. According to (6.9) and (6.8) the McCoy-Curro scheme for the coarse grained Hamiltonian is given by

$$\bar{H}^{mc}(\eta) = -\sum_{k,\, l \in \Lambda_M^c} \bar{U}^{mc}(\eta_k, \eta_l, |k-l|)\,, \tag{6.11}$$

where

$$\bar{U}^{mc}(\eta_k, \eta_l, |k-l|) = -\frac{1}{\beta} \log \left(\mathbb{E}\left[e^{-\beta H_{C_k, C_l}(\sigma)} | \eta_k,\, \eta_l \right] \right), \tag{6.12}$$

and

$$H_{C_k,C_l}(\sigma) = -\frac{1}{2} \sum_{x\in C_k,\, y\in C_l} J(x-y)\sigma(x)\sigma(y)\,,$$

corresponds to the totality of microscopic interactions between two isolated coarse cells C_k and C_l. A simple calculation yields

$$\bar{U}^{mc}(\eta_k, \eta_l, |k-l|) = \frac{1}{2}\bar{J}(k,l)\,\eta_k\,\eta_l - \frac{1}{\beta}\log(\mathbb{E}[e^{-\beta\Delta J_{kl}(\sigma)}\,|\,\eta_k,\,\eta_l]) \tag{6.13}$$

where

$$\Delta_{kl}J(\sigma) := -\frac{1}{2} \sum_{\substack{x\in C_k \\ y\in C_l\,, y\neq x}} (J(x-y) - \bar{J}(k,l))\sigma(x)\sigma(y)\,. \tag{6.14}$$

By using the Taylor expansion of the exponential and of the logarithm we eventually obtain that the McCoy-Curro CG Hamiltonian behaves asymptotically as

$$\bar{H}^{mc}(\eta) = \bar{H}^{(0)}_M(\eta) + \bar{H}^{(1)}_M(\eta) + (\dots) \tag{6.15}$$

where $(\dots)$ include higher order *two-body* terms only.

Remark 6.1.

1. It is clear from these calculations that the McCoy-Curro scheme is asymptotically identical to using $\bar{H}^{(0)}_M(\eta)$ and $\bar{H}^{(1)}_M(\eta)$ of Scheme 2.1, while the three-body term $\bar{H}^{(2)}_M(\eta)$ included in Scheme 2.1 is absent. As we have seen in the estimates in Section 2.2 corrections including three-body CG terms are crucial in obtaining higher accuracy, especially in low temperature regimes. A detailed study of the impact of multi-body terms in coarse-graining, using the lattice systems as a paradigm, can be found in [3].
2. Implementing the McCoy-Curro scheme on a lattice is computationally expensive as it requires calculating $\bar{U}^{CG}(\eta_k, \eta_l, |k-l|)$ using (6.12) over all possible ranges of $\eta_k, \eta_l, |k-l|$ and repeating all the calculations for each temperature β. On the other hand the Scheme 2.1 provides an analytical way of coarse graining which in this particular context is computationally more efficient.

REFERENCES

[1] C. F. Abrams and K. Kremer. The effect of bond length on the structure of dense bead-spring polymer melts. *J. Chem. Phys.*, 115:2776, 2001.

[2] J. E. Adams and J. D. Doll. Desorption from solid surfaces via generalized slater theory. *J. Chem. Phys.*, 74:1467, 1981.

[3] S. Are, M. A. Katsoulakis, L. Rey-Bellet, and P. Plecháč. Multi-body interactions in coarse-graining schemes of extended systems. *preprint*, 2007. submitted to SIAM J. Sci. Comp.

[4] S. M. Auerbach. Theory and simulation of jump dynamics, diffusion and phase equilibrium in nanopores. *Int. Rev. Phys. Chem.*, 19(155), 2000.

[5] L. Bertini, E. N. M. Cirillo., and E. Olivieri. Renormalization-group transformations under strong mixing conditions: Gibbsianness and convergence of renormalized interactions. *J. Statist. Phys.*, 97(5-6):831–915, 1999.

[6] A. Bovier and M. Zahradník. The low-temperature phase of Kac-Ising models. *J. Statist. Phys.*, 87(1-2):311–332, 1997.

[7] W. J. Briels and R. L. C. Akkermans. Coarse-grained interactions in polymer melts: a variational approach. *J. Chem. Phys.*, 115:6210, 2001.

[8] Report by The Chemical Industry Vision 2020 Technology Partnership. *Technology Roadmap for Computational Chemistry*. 1999.

[9] M. Cassandro and E. Presutti. Phase transitions in Ising systems with long but finite range interactions. *Markov Process. Related Fields*, 2(2):241–262, 1996.

[10] A. Chatterjee, M. A. Katsoulakis, and D. G. Vlachos. Spatially adaptive grand canonical ensemble Monte Carlo simulations. *Phys. Rev. E*, 71, 2005.

[11] A. Chatterjee and D. G. Vlachos. Multiscale spatial Monte Carlo simulations: Multigriding, computational singular perturbation, and hierarchical stochastic closures. *J. Chem. Phys.*, 124(6):064110, 2006.
[12] A. Chatterjee and D. G. Vlachos. An overview of spatial microscopic and accelerated kinetic Monte Carlo methods for materials' simulation. *J. Computer-Aided Materials Design*, 14(2):253, 2007.
[13] A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis. Spatially adaptive lattice coarse-grained Monte Carlo simulations for diffusion of interacting molecules. *J. Chem. Phys.*, 121(22):11420, 2004.
[14] P. Dupuis and H. Wang. Dynamic importance sampling for uniformly recurrent Markov chains. *Ann. Appl. Probab.*, 15(1A):1–38, 2005.
[15] W. E and B. Engquist. Multiscale modeling and computation. *Notices Amer. Math. Soc.*, 50(9):1062–1070, 2003.
[16] H. Fukunaga, J. J. Takimoto, and M. Doi. A coarse-grained procedure for flexible polymer chains with bonded and nonbonded interactions. *J. Chem. Phys.*, 116:8183, 2002.
[17] G. A. Gallavotti and S. Miracle-Sole. Correlation functions of a lattice system. *Comm. Math. Phys.*, 7(4):274–288, 1968.
[18] B. Gidas. Metropolis-type Monte Carlo simulation algorithms and simulated annealing. In J. Laurie Snell, editor, *Topics in Contemporary Probability and its Applications*. CRC Press, 1995.
[19] P. Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2004. , Stochastic Modelling and Applied Probability.
[20] N. Goldenfeld. *Lectures on Phase Transitions and the Renormalization Group*, volume 85. Addison-Wesley, New York, 1992.
[21] C. Gruber and H. Kunz. General properties of polymer systems. *Comm. Math. Phys.*, 22:133–161, 1971.
[22] V. A. Harmandaris, N. P. Adhikari, N. F. A. van der Vegt, and K. Kremer. Hierarchical modeling of polysterene: From atomistic to coarse-grained simulations. *Macromolecules*, 39:6708, 2006.
[23] V. A. Harmandaris, M. A. Katsoulakis, and P. Plecháč. Coarse-graining schemes for off-lattice interacting particles with internal degrees of freedom. in preparation.
[24] A. E. Ismail, G.C. Rutledge, and G. Stephanopoulos. Multiresolution analysis in statistical mechanics. I. using wavelets to calculate thermodynamics properties. *J. Chem. Phys.*, 118:4414, 2003.
[25] 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.*, 118:4424, 2003.
[26] L. Kadanoff. Scaling laws for Ising models near t_c. *Physics*, 2:263, 1966.
[27] M. Katsoulakis, A. Majda, and D. Vlachos. Coarse-grained stochastic processes for microscopic lattice systems. *Proc. Natl. Acad. Sci*, 100(3):782–782, 2003.
[28] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos. Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems. *J. Comp. Phys.*, 112:250–278, 2003.
[29] M. A. Katsoulakis, P. Plecháč, and A. Sopasakis. Error analysis of coarse-graining for stochastic lattice dynamics. *SIAM J. Numer. Anal.*, 44(6):2270–2296, 2006.
[30] M. A. Katsoulakis, L. Rey-Bellet, and P. Plecháč. Acceleration of molecular simulation methods: coarse-graining, reconstruction and importance sampling. in preparation.
[31] M. A. Katsoulakis, L. Rey-Bellet, P. Plecháč, and D. K.Tsagkarogiannis. Mathematical strategies in the coarse-graining of extensive systems: error quantification and adaptivity. *J. Non Newt. Fluid Mech.*, accepted.
[32] M. A. Katsoulakis, L. Rey-Bellet, P. Plecháč, and D. K. Tsagkarogiannis. Coarse-graining schemes and a posteriori error estimates for stochastic lattice systems. *ESAIM-Math. Model. Num. Analysis*, 41(3):627–660, 2007.
[33] M. A. Katsoulakis, L. Rey-Bellet, P. Plecháč, and D. K. Tsagkarogiannis. Coarse-graining schemes for stochastic lattice systems with short and long range interactions. preprint.
[34] M. A. Katsoulakis and J. Trashorras. Information loss in coarse-graining of stochastic particle dynamics. *J. Stat. Phys.*, 122(1):115–135, 2006.
[35] M. A. Katsoulakis and D. G. Vlachos. Hierarchical kinetic Monte Carlo simulations for diffusion of interacting molecules,. *J. Chem. Phys.*, 186:9412, 2003.
[36] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.*, 1(4):715–762, 2003.
[37] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*. Springer-Verlag, 1999.
[38] K. Kremer and F Müller-Plathe. Multiscale problems in polymer science: simulation approaches. *MRS Bull.*, page 205, March 2001.
[39] D. P. Landau and K. Binder. *A guide to Monte Carlo simulations in statistical physics*. Cambridge University Press, Cambridge, 2000.
[40] J. L. Lebowitz, A. Mazel, and E. Presutti. Liquid-vapor phase transitions for systems with finite-range interactions. *J. Statist. Phys.*, 94(5-6):955–1025, 1999.
[41] J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer Series in Statistics. Springer-Verlag, New York, 2001.
[42] A. J. Majda and B. Khouider. A numerical strategy for efficient modeling of the equatorial wave guide. *Proc. Natl. Acad. Sci. USA*, 98(4):1341–1346 (electronic), 2001.
[43] J. D. McCoy and J. G. Curro. The mapping of explicit atom onto united atom potentials. *Macromolecules*, 31:9362, 1998.
[44] F. Müller-Plathe. Coarse-graining in polymer simulation: from the atomistic to the mesoscale and back. *Chem. Phys. Chem.*, 3:754, 2002.

[45] I. Pivkin and G. Karniadakis. Coarse-graining limits in open and wall-bounded dissipative particle dynamics systems. *J. Chem. Phys.*, 124:184101, 2006.

[46] C. P. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Texts in Statistics. Springer-Verlag, New York, second edition, 2004.

[47] B. Simon. *The Statistical Mechanics of Lattice Gases, vol. I*. Princeton series in Physics, 1993.

[48] M. Suzuki, X. Hu, M. Katori, A. Lipowski, N. Hatano, K. Minami, and Y. Nonomura. *Coherent-Anomaly Method: Mean Field, Fluctuations and Systematics*. World Scientific, 1995.

[49] J. Trashorras and D. K. Tsagkarogiannis. Reconstruction schemes for coarse-grained stochastic lattice systems. *preprint*, 2008. submitted to SIAM J. Numer. Anal.

[50] W. Tschöp, K. Kremer, O. Hahn, J. Batoulis, and T. Bürger. Simulation of polymer melts. I. coarse-graining procedure for polycarbonates. *Acta Polym.*, 49:61, 1998.

[51] W. Tschöp, K. Kremer, O. Hahn, J. Batoulis, and T. Bürger. Simulation of polymer melts. II. from coarse-grained models back to atomistic description. *Acta Polym.*, 49:75, 1998.

[52] D. G. Vlachos and M. A. Katsoulakis. Mesoscopic theories for the diffusion of interacting molecules. *Phys. Rev. Lett.*, 85:3898, 2000.

[53] R. M. Ziff, E. Gulari, and Y. Barshad. Kinetic phase transitions in an irreversible surface-reaction model. *Phys. Rev. Lett.*, 56:2553, 1986.