2014

Spatial Multi-Level Interacting Particle Simulations and Information Theory-Based Error Quantification

Evangelia Kalligiannaki
Markos Katsoulakis, *University of Massachusetts - Amherst*
Petr Plechac

SPATIAL MULTI-LEVEL INTERACTING PARTICLE SIMULATIONS AND INFORMATION THEORY-BASED ERROR QUANTIFICATION.*

EVANGELIA KALLIGIANNAKI‡, MARKOS A. KATSOULAKIS ‡, AND PETR PLECHÁČ†

Abstract. We propose a hierarchy of multi-level kinetic Monte Carlo methods for sampling high-dimensional, stochastic lattice particle dynamics with complex interactions. The method is based on the efficient coupling of different spatial resolution levels, taking advantage of the low sampling cost in a coarse space and by developing local reconstruction strategies from coarse-grained dynamics. Microscopic reconstruction corrects possibly significant errors introduced through coarse-graining, leading to the controlled-error approximation of the sampled stochastic process. In this manner, the proposed multi-level algorithm overcomes known shortcomings of coarse-graining of particle systems with complex interactions such as combined long and short-range particle interactions and/or complex lattice geometries. Specifically, we provide error analysis for the approximation of long-time stationary dynamics in terms of relative entropy and prove that information loss in the multi-level methods is growing linearly in time, which in turn implies that an appropriate observable in the stationary regime is the information loss of the path measures per unit time. We show that this observable can be either estimated *a priori*, or it can be tracked computationally *a posteriori* in the course of a simulation. The stationary regime is of critical importance to molecular simulations as it is relevant to long-time sampling, obtaining phase diagrams and in studying metastability properties of high-dimensional complex systems. Finally, the multi-level nature of the method provides flexibility in combining rejection-free and null-event implementations, generating a hierarchy of algorithms with an adjustable number of rejections that includes well-known rejection-free and null-event algorithms.

Key words. kinetic Monte Carlo, coarse graining, multiple scales, phase transition, information theory, multi-level methods, relative entropy, error analysis.

AMS subject classifications. 65C05, 65C20, 82C22, 82C20

1. Introduction.

One of the widely used computational methods at atomistic scales simulating stochastic particle systems is the continuous time or kinetic Monte Carlo (kMC) method. The first implementations of rejection-free kinetic Monte Carlo methods in molecular simulations go back to the stochastic simulation algorithm (SSA) of Gillespie for well-mixed systems, [11], and the n-fold method or the BKL method of Bortz, Kalos and Lebowitz, [3], for spatially distributed Ising-type systems. Traditionally kMC algorithms are serial, explicit time-stepping methods, limiting their applicability, due to the high computational cost per event (time-step), to moderate size systems. The principal part of the cost consists of searching for an event and updating the reaction rates. In the last decade research works have been focusing on developing sophisticated search and update techniques to reduce this computational cost, [5, 30, 33]. For example, search algorithms have been proposed using binary tree pointers [10, 31] in order to obtain $\mathcal{O}(\log N)$ complexity in the size N of the simulated system. The BKL method is designed to reduce the cost of the searching step by lumping the transition states into classes of the same probability. All these techniques even though they reduce the computational cost per event are highly demanding in

*The research of E.K. was supported by the National Science Foundation under the grant NSF-CMMI-0835582 and the Department of Energy under the grant DE-SC000233; M.A.K. was partially supported by the grant NSF-CMMI-0835673; P.P. was partially supported by the grant NSF-DMS-0813893.

†Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA (plechac@math.udel.edu).

‡Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA (ekalligi@math.umass.edu, markos@math.umass.edu).

computer storage and implementation overhead. An alternative to the rejection-free methods is based on the *uniformization* of the simulated continuous Markov chain. The implementation leads to null-event algorithms, which at each step require calculation of the rate for only one, arbitrarily chosen, event that is accepted or rejected with a probability provided by the uniformization of the process, [5]. Although null-event methods reduce significantly the computational cost per Monte Carlo step they can be highly inefficient when the acceptance probability of an event becomes small, resulting to exceedingly small time-steps.

In this work we propose a *partially* rejection-free kinetic Monte Carlo method, the multilevel kinetic coarse-grained Monte Carlo (ML-KMC), for sampling high-dimensional lattice systems with complex interactions and/or lattice geometries. Our primary interest in simulating extended systems in which the system size N is large and the behavior is governed by the infinite volume or thermodynamic limit (i.e., regimes where $N \to \infty$). We distinguish between long-range interaction potentials, whose range $L \sim N^{1/d}$ is comparable to the system dimensions $N^{1/d}$, where d is the dimension of the lattice, and short-range potentials which have a fixed interaction range S independent of the system size N. Typically in the lattice-gas simulations the long-range potentials are Coulomb or Lenard-Jones and short-range ones are represented by nearest-neighbor or next nearest-neighbor potentials. Interactions that combine long and short-range potentials are difficult to be handled efficiently by existing BKL-type algorithms since the number of classes grow exponentially with the length of the interaction range and thus making an implementation intractable, [5]. A possible remedy is to simulate a system with compressed (coarse-grained) potentials that have a shorter interaction range. We have demonstrated that smooth long-range potentials can be coarse-grained with controlled errors leading to highly efficient coarse-grained Monte Carlo (CGMC) methods, [15, 19, 18, 20]. On the other hand a simple coarse-graining of singular, long-range as well as short-range potentials does not lead to an accurate approximation of the renormalized Hamiltonian and results to loss of important microscopic information, due to the presence of strong short-ranged spatial correlations. As demonstrated in [21, 2] approximations of coarse-grained Hamiltonians accounting for short-range or singular potentials necessarily involve multi-body interaction terms whose implementation can become quickly computationally expensive. The pitfalls of coarse-graining are well-known also from molecular simulations of polymeric systems where they produce spurious phase changes or incorrectly predict existing phase transitions, [8, 17]. Similarly, coarse-graining is expected to be challenging in systems with complex lattice geometries, e.g. [32], that can also induce complicated spatial correlations.

In this paper we demonstrate the use of the proposed ML-KMC method in the efficient and accurate simulation of such complex systems where coarse-graining either fails or it is computationally very expensive. We focus on the important example of systems where competing short- and long- ranged interaction forces are present and lead to complicated phase diagrams and pattern formation in various physical and chemical systems, [29, 6, 27].

2. Overview of the proposed method. The key ingredient of the proposed method is the multi-level sampling of the evolution process based on the knowledge of an even less accurate coarse-grained, meso/macroscopic dynamics. The present study is an extension of [12] from the equilibrium to dynamical sampling sharing the same principle of efficient coupling of different resolution levels. In [12] we proposed a multi-level Coarse-Grainined Metropolis-Hastings algorithm appropriate for sampling

equilibrium properties of systems. Although the embedded Markov chain generated by the Metropolis algorithm converges to the correct equilibrium distribution it does not preserve physical dynamics, a fact that motivated the present work. We present the method in a general framework to demonstrate its applicability to on- and off-lattice systems, and present in detail the application to stochastic lattice systems with short- and long-range interactions.

The dynamics of stochastic systems on a countable configuration space Σ are determined by a continuous time Markov process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$, with the infinitesimal generator $\mathcal{L} : L^\infty(\Sigma) \to L^\infty(\Sigma)$ defined by the rates $c(\sigma, \sigma')$, $\sigma, \sigma' \in \Sigma$:

$$\mathcal{L}\phi(\sigma) = \sum_{\sigma' \in \Sigma} c(\sigma, \sigma') \left(\phi(\sigma') - \phi(\sigma)\right) , \tag{2.1}$$

for every observable defined as any $\phi \in L^\infty(\Sigma)$. More specifically, in kMC we typically compute expected values of such observables, that is quantities such as

$$u(\zeta, t) := \mathbb{E}^\zeta[f(\sigma_t)] = \sum_\sigma f(\sigma) P(\sigma, t; \zeta) , \tag{2.2}$$

conditioned on the initial data $\sigma_0 = \zeta$. On the other hand, the evolution of the entire system at any time t is described by the transition probabilities $P(\sigma, t; \zeta) := \mathbb{P}\left(\sigma_t = \sigma \,|\, \sigma_0 = \zeta\right)$ where $\zeta \in \Sigma$ is any initial configuration. The transition probabilities satisfy the Forward Kolmogorov Equation (Master Equation), [9],

$$\partial_t P(\sigma, t; \zeta) = \sum_{\sigma', \sigma' \neq \sigma} c(\sigma', \sigma) P(\sigma', t; \zeta) - c(\sigma, \sigma') P(\sigma, t; \zeta) , \tag{2.3}$$

where $P(\sigma, 0; \zeta) = \delta(\sigma - \zeta)$ and $\delta(\sigma - \zeta) = 1$ if $\sigma = \zeta$ and zero otherwise. By a straightforward calculation using (2.3) we obtain that the observable (2.2) satisfies the initial value problem

$$\partial_t u(\zeta, t) = \mathcal{L} u(\zeta, t) , \qquad u(\zeta, 0) = f(\zeta) , \tag{2.4}$$

The numerical implementation of the evolution of the process is realized with the embedded Markov chain $\{X_n\}_{n\geq 0}$, $X_n = \sigma_{n\delta t}$ with transition probabilities

$$p(\sigma, \sigma') = \frac{c(\sigma, \sigma')}{\lambda(\sigma)} , \quad \lambda(\sigma) = \sum_{\sigma' \in \Sigma} c(\sigma, \sigma') , \tag{2.5}$$

where $p(\sigma, \sigma')$ is the probability of a jump from the state σ to σ'. The residence times δt_σ for which the system stays in the state σ before a jump is distributed according to an exponential law with the parameter $\lambda(\sigma)$.

Method. The proposed method generates an approximate process $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$ of the stochastic process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$. It is based on projecting the microscopic space Σ into a coarse space $\bar{\Sigma}$ with less degrees of freedom and on the knowledge of a coarse rate function $\bar{c}(\eta, \eta')$ which captures macroscopic information from $c(\sigma, \sigma')$. We denote the coarse space variables $\eta = \mathbf{T}\sigma$ defined by a projection operator $\mathbf{T} : \Sigma \to \bar{\Sigma}$. For example, for stochastic lattice systems that we elaborate on in this work, approximate coarse rate functions are explicitly known from coarse graining (CG) techniques of [15, 16]. In this work we analyze a two-level approach, i.e., coupling two configuration spaces Σ and $\bar{\Sigma}$ with different resolutions, while a multi-level extension can be considered analogously. The ML-KMC method consists of the following steps, a schematic description is demonstrated in Figure 2:

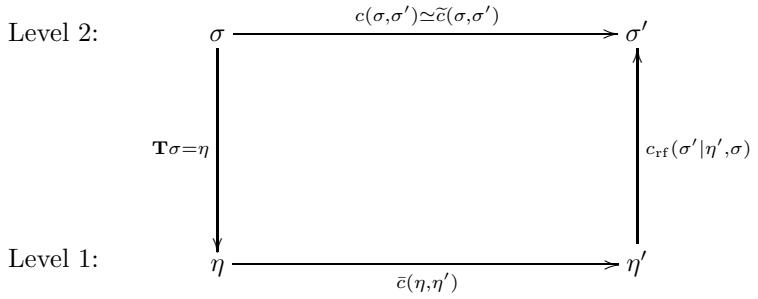

FIGURE 2.1. *Two-level decomposition, compressing with* $\mathbf{T}$ *and reconstructing with* $c_{\rm rf}$*, of the evolution process per event.*

i) Construct a computationally inexpensive approximating CG process on the coarse space $\bar{\Sigma}$ described by a coarse generator $\bar{\mathcal{L}}$ with rates $\bar{c}(\eta,\eta')$.
ii) Define the "reconstruction" rates $c_{\rm rf}(\sigma'|\eta',\sigma)$ constrained on the updated coarse state η' that are simple to simulate and such that

$$\bar{c}(\eta,\eta')c_{\rm rf}(\sigma'|\eta',\sigma) \quad \textit{approximates} \quad c(\sigma,\sigma')\,.$$

The approximation and its error is quantified in Section 5. The overall procedure can be thought as a reconstruction in dynamics of stochastic processes from a CG process. Furthermore, this procedure generates stochastic processes that are *controlled error* approximations of the process $(\{\sigma_t\}_{t\geq 0},\mathcal{L})$, determined by the reconstruction rates $c_{\rm rf}(\sigma'|\eta',\sigma)$ and the level of coarsening. The function $c_{\rm rf}(\sigma'|\eta',\sigma)$ enriches the CG procedure by re-inserting details that were smoothed out by the coarsening procedure. *Implementation.* The multilevel nature of the method provides the flexibility of combining rejection-free and null-event implementation algorithms at each resolution and consists of:

i) A rejection-free algorithm for sampling in the coarse space with $\bar{c}(\eta,\eta')$ where the reduction of the computational cost compared to microscopic sampling is significant due to the compression of spatial scales and interactions range. The rejection-free algorithm selects the most probable coarse state η' that the system will evolve.
ii) A null-event algorithm for sampling at the fine space with $c_{\rm rf}(\sigma'|\eta',\sigma)$, however with a low rejection rate due to the fact that η' was chosen as the most probable event.

This partially rejection-free implementation approach suggests a non-constant time step update in contrast to null-event methods where the time update is uniform for all system states. Non-constant time step updating algorithms have been proposed in [1] designing a class of kinetic Monte Carlo algorithms with an adaptive time step interpolating between BKL and null-event algorithm. The variable time step in these algorithms is based on the adaptively improved upper bounds of the exponential parameter controlling the time step distribution, while in our proposed method it is the result of the combination of BKL and rejection-free implementations on different resolution state spaces. However, one can further enhance the ML-KMC method with an adaptive time step kinetic Monte Carlo in the spirit of [1], in view of the freedom on the choice of implementation techniques in each level.

Analysis. We provide numerical analysis for (a) finite-time weak error estimates and

(b) long-time stationary dynamics for the proposed ML-KMC algorithm. The primary challenge in the finite-time weak error estimates is to obtain bounds that are independent of the high-dimension of the interacting particle system, and this is accomplished by focusing on suitable macroscopic observables. However, this technique leads to estimates that involve constants that grow exponentially in time, as is the case also in many of the classical numerical analysis estimates for stochastic differential equations or partial differential equations that rely on Gronwall inequality-type arguments. In order to overcome this difficulty we develop a different approach that estimates the error in long-time behavior, i.e. in stationary regimes for reversible or irreversible processes, using the relative entropy of the path measure. In kMC as well as in molecular simulations in general, we are often primarily interested in long-time, stationary regimes, including (i) the behavior of the stationary measure and its sampling, as well as (ii) stationary dynamics, i.e. dynamics where the initial measure is the stationary distribution reached after long-time integration. The latter is an especially important regime describing dynamic transitions between metastable states in complex energy landscapes, while at this regime we construct the system's phase diagrams.

The error analysis study we present reveals that the relevant quantity for assessing long-time simulations in the stationary regime is the *information loss per unit time* in the path space measure. In fact, the issues related to the error analysis in the stationary dynamics regime is the primary novelty in our paper from a numerical analysis perspective and they are not restricted only to the multilevel kMC algorithms; they are ubiquitous for numerical approximations of reversible and irreversible stochastic dynamics such as Langevin stochastic differential equations (SDEs), thermostatted dynamics, dissipative particle dynamics (DPD) methods, etc. For such systems there is a wealth of approximating schemes, the simplest being in the time stepping, e.g., explicit, implicit, predictor-corrector, operator splitting, etc. We expect that our proposed entropy-based perspective could be used to assess such numerical schemes at the stationary dynamics regime in a quantifiable manner, in a variety of stochastic, extended as well as finite-dimensional systems.

We begin this work in Section 3 presenting the method and continue with Section 3.1 proposing an efficient implementation strategy leading to a partially rejection free method. Application of the method in stochastic lattice systems is presented in Section 4 for adsorption-desorption Arrhenius dynamics. In Section 5 we prove error estimates that provide a quantitative control of the approximating process. Section 7 establishes computational efficiency of the method over conventional sampling techniques. The performance of the method is tested in Section 8 using a benchmark model of Arrhenius dynamics with a competing short- and long-range interaction potential.

3. Multilevel kinetic Monte Carlo. The ML-KMC method is a kinetic Monte Carlo method generating *controlled-error approximate* dynamics of the stochastic process $(\{\sigma_t\},\mathcal{L})_{t\geq 0}$, based on the decomposition of the rate function $c(\sigma,\sigma')$ into coarse and corresponding reconstructing terms, such that

$$c(\sigma,\sigma')\approx\widetilde{c}(\sigma,\sigma')=\prod_{i=0}^{I}\bar{c}^{(i)}(\eta_i,\eta_i')c_{\mathrm{rf}}^{(i)}(\eta_{i-1}'|\eta_i',\eta_{i-1})\,,$$

where $\eta_0\equiv\sigma$ and $\eta_i=\mathbf{T}_{i-1}\eta_{i-1}, i=1\ldots,I$ are variables in a hierarchy of coarse spaces with the decreasing numbers of degrees of freedom. Coarsening consists of

projecting the microscopic space into a coarse space $\bar{\Sigma}$ with less degrees of freedom for which a coarse rate function $\bar{c}(\eta, \eta')$ is appropriately defined, where $\eta \in \bar{\Sigma}$ denotes the coarse space variables defined by a projection operator $\mathbf{T} : \Sigma \to \bar{\Sigma}$, $\mathbf{T}\sigma = \eta$.

For the sake of simplicity we present the two-level method (ML-KMC) while every step in the study that follows can be easily adopted to the multilevel case. The construction of the ML-KMC, sketched in Figure 2, consists of two steps. In the *first step* we construct an approximating process on the coarse space $\bar{\Sigma}$ described by a generator $\bar{\mathcal{L}}$ with rates $\bar{c}(\eta, \eta')$, extracting macroscopic information from the rates $c(\sigma, \sigma')$. In the *second step* we construct rates $c_{\rm rf}(\sigma'|\eta', \sigma)$, simple to simulate, and such that $\bar{c}(\eta, \eta')c_{\rm rf}(\sigma'|\eta', \sigma)$ approximates $c(\sigma, \sigma')$ with an error quantified in Section 5. Following the above description we define a stochastic process with rates

$$\widetilde{c}(\sigma, \sigma') = \bar{c}(\eta, \eta')c_{\rm rf}(\sigma'|\eta', \sigma)\,, \tag{3.1}$$

that generates a corresponding continuous time Markov chain with transition probabilities

$$\widetilde{p}(\sigma, \sigma') = \frac{\widetilde{c}(\sigma, \sigma')}{\widetilde{\lambda}(\sigma)}\,, \quad \widetilde{\lambda}(\sigma) = \sum_{\sigma' \in \Sigma} \widetilde{c}(\sigma, \sigma')\,. \tag{3.2}$$

The corresponding continuous time process $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$ is defined by the generator $\widetilde{\mathcal{L}}$ on Σ,

$$\widetilde{\mathcal{L}}\phi(\sigma) = \sum_{\sigma' \in \Sigma} \widetilde{c}(\sigma, \sigma')\left(\phi(\sigma') - \phi(\sigma)\right)\,, \tag{3.3}$$

for every $\phi \in L^\infty(\Sigma)$. In the application we study in Section 4.4 such coarse and reconstructing rates are explicitly defined, both for exact and for controlled error sampling approaches. In Section 5 we provide estimates that quantify the approximating errors, for finite and long-time regimes, with respect to the level of resolution and the interpretation of the rate function decomposition. Furthermore, controlled approximations such as (3.1) can have significant computational advantages, see Section 7.

3.1. Partially rejection-free implementation of the ML-KMC method. Rejection-free methods are based on calculating (updating) rates $c(\sigma, \sigma')$ for all $\sigma' \in \Sigma$ at each Monte Carlo step and choosing (searching) an event that evolves the system's state based on the probabilities $p(\sigma, \sigma')$, see (2.5). In these methods every step provides a successful event but the cost of implementation becomes formidable for large complex systems where the cost increases with the system size $|\Sigma|$ and with the complexity of the transition rates. Uniformization provides a solution to this problem, suggesting a method that needs the calculation only of one transition probability at each step with the disadvantage of introducing rejections; a number of proposed events will not happen. Such methods are known as null-event methods, that generate a Markov jump process with the embedded Markov chain described by

$$p^{\rm null}(\sigma, \sigma') = \begin{cases} 1 - \frac{\lambda(\sigma)}{\lambda^*}\,, & \text{if } \sigma' = \sigma \\ \frac{\lambda(\sigma)}{\lambda^*} p(\sigma, \sigma')\,, & \text{if } \sigma' \neq \sigma \end{cases} \tag{3.4}$$

where λ^* is a uniform upper bound of $\lambda(\sigma)$ in (2.5). The time that the system stays at state σ, $\tau^{\rm null}$, is governed by an exponential distribution with the parameter $\lambda^* \geq \lambda(\sigma)$ for all $\sigma \in \Sigma$. As a result since $1/\lambda^* = \mathbb{E}[\tau^{\rm null}] \leq \mathbb{E}[\tau_\sigma] = 1/\lambda(\sigma)$ for all $\sigma \in \Sigma$, the null

event algorithm evolves the system with a smaller time step and needs more MC steps than a rejection free method. Despite this inefficiency, the significant reduction of the computational cost per MC step can be advantageous for high dimensional complex systems when compared to rejection free methods. The probability of rejecting a proposed state when the system is in the state σ is

$$p_{\mathrm{rej}}^{\mathrm{null}}(\sigma) = 1 - \frac{\lambda(\sigma)}{\lambda^*}\,, \tag{3.5}$$

and it is controlled by λ^*, indicating that the tighter the upper bound is the less rejections are introduced. The combination of rejection-free and null-event techniques at the two levels of the ML-KMC method introduces a variety of sampling techniques. Here we propose, as the more efficient approach, the rejection-free method at the coarse level and the null-event algorithm at the microscopic level, balancing between an improved rejection rate of the null-event and the computational complexity of the rejection-free method. Given $\sigma \in \Sigma$, $\eta = \mathbf{T}\sigma$ the evolution of the system to a state $\sigma' \in \Sigma$ is achieved by the following:

METHOD 1. *ML-KMC*

Coarse level. Evolve to a new coarse state $\eta' \in \bar{\Sigma}$ with the probability

$$\bar{p}(\eta, \eta') = \frac{\bar{c}(\eta, \eta')}{\bar{\lambda}(\eta)}\,, \quad \bar{\lambda}(\eta) = \sum_{\eta' \in \bar{\Sigma}} \bar{c}(\eta, \eta')\,.$$

Microscopic level. Select randomly (uniformly) $\sigma' \in \Sigma$ under the constraint $\mathbf{T}\sigma' = \eta'$, and accept it with the probability

$$p_{\mathrm{rf}}(\sigma'|\eta', \sigma) = \frac{c_{\mathrm{rf}}(\sigma'|\eta', \sigma)}{\lambda_{\mathrm{rf}}(\sigma)}\,, \quad \lambda_{\mathrm{rf}}(\sigma) = \max_{\eta'} \sum_{\{\sigma' : \mathbf{T}\sigma' = \eta'\}} c_{\mathrm{rf}}(\sigma'|\eta', \sigma)\,,$$

or reject it with the probability

$$1 - \sum_{\{\sigma' : \mathbf{T}\sigma' = \eta'\}} p_{\mathrm{rf}}(\sigma'|\eta', \sigma)\,.$$

For implementation specifics of this step we refer to Section 7.

Time update. Update time by a random time step with an exponential law with parameter

$$\widetilde{\lambda}^*(\sigma) = \bar{\lambda}(\eta)\lambda_{\mathrm{rf}}(\sigma)\,. \tag{3.6}$$

With the following lemma we prove that the ML-KMC method provides correctly a *partial uniformization* of the rejection-free method.

LEMMA 3.1. *For any $\sigma \in \Sigma$ we have*

$$\widetilde{\lambda}(\sigma) \le \widetilde{\lambda}^*(\sigma)\,.$$

Proof. The statement follows directly from the simple calculation

$$\widetilde{\lambda}(\sigma) = \sum_{\sigma' \in \Sigma} \widetilde{c}(\sigma, \sigma') = \sum_{\eta' \in \bar{\Sigma}} \bar{c}(\eta, \eta') \sum_{\{\sigma' : \mathbf{T}\sigma' = \eta'\}} c_{\mathrm{rf}}(\sigma'|\eta', \sigma) \le \bar{\lambda}(\eta)\lambda_{\mathrm{rf}}(\sigma)\,.$$

□

The real time updates, controlled by $\widetilde{\lambda}^*(\sigma)$, depend on the state of the system, while in (3.4) the time step is uniform for all states, controlled by λ^*. The multilevel method has the rejection probability at the state σ,

$$p_{\rm rej}^{\rm multi}(\sigma) = 1 - \frac{\widetilde{\lambda}(\sigma)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} = 1 - \frac{\widetilde{\lambda}(\sigma)}{\widetilde{\lambda}^*(\sigma)}\,, \tag{3.7}$$

since

$$\begin{aligned} p_{\rm rej}^{\rm multi}(\sigma) &= 1 - \sum_{\sigma'\in\Sigma} {\rm Prob}(\sigma\to\sigma') = 1 - \sum_{\eta'\in\bar{\Sigma}} \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} \frac{\bar{c}(\eta,\eta')c_{\rm rf}(\sigma'|\eta',\sigma)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} \\ &= 1 - \sum_{\sigma'\in\Sigma} \frac{\widetilde{c}(\sigma,\sigma')}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} = 1 - \frac{\widetilde{\lambda}(\sigma)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)}\,. \end{aligned}$$

Next, we show that the rejection rate of the ML-KMC method can be controlled and depends on the approximation of the coarse-grained rates. Before stating the proposition we note that a process $\{\widetilde{\sigma}_t\}_{t\geq 0}$ is defined as lumpable, ([4]), with respect to the coarsening procedure $\eta = \mathbf{T}\sigma$ when its rates satisfy

$$\sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} \widetilde{c}(\sigma,\sigma') = \bar{c}(\eta,\eta')\,.$$

This relation implies uniform reconstruction rates $c_{\rm rf}(\sigma'|\eta',\sigma) = 1/|\{\sigma':\mathbf{T}\sigma'=\eta'\}|$ for all $\sigma'\in\{\sigma':\mathbf{T}\sigma'=\eta'\}$, hence $\sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} c_{\rm rf}(\sigma'|\eta',\sigma) = 1$, $\lambda_{\rm rf}(\sigma) = 1$,

$$\widetilde{\lambda}(\sigma) = \sum_{\sigma'\in\Sigma} \widetilde{c}(\sigma,\sigma') = \sum_{\eta'\in\bar{\Sigma}} \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} \widetilde{c}(\sigma,\sigma') = \sum_{\eta'\in\bar{\Sigma}} \bar{c}(\eta,\eta') = \bar{\lambda}(\eta)\,,$$

and

$$p_{\rm rej}^{\rm multi}(\sigma) = 1 - \frac{\widetilde{\lambda}(\sigma)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} = 0\,.$$

On the other hand, the approximating process $\{\widetilde{\sigma}_t\}_{t\geq 0}$ in ML-KMC is not necessarily lumpable, nevertheless its approximation error determines the rejection probability:

PROPOSITION 3.2. *Let the coarse rates define an approximately lumpable process, that is*

$$\sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} \widetilde{c}(\sigma,\sigma') = \bar{c}(\eta,\eta') + \mathcal{O}(\epsilon)\,, \tag{3.8}$$

uniformly in $\sigma,\eta=\mathbf{T}\sigma,\eta'$ for some $\epsilon>0$. Then

$$p_{\rm rej}^{\rm multi}(\sigma) = \mathcal{O}(\epsilon)\,.$$

Proof. First, we denote by $\bar{\Sigma}_\eta = \{\eta' : c(\eta,\eta') > 0\}$, that is all coarse configurations accessible in a single step from η. Assumption (3.8) implies that

$$\bar{c}(\eta,\eta') \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} c_{\rm rf}(\sigma'|\eta',\sigma) = \bar{c}(\eta,\eta') + \mathcal{O}(\epsilon)\,, \quad \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} c_{\rm rf}(\sigma'|\eta',\sigma) = 1 + \mathcal{O}(\epsilon)\,, \tag{3.9}$$

hence $\lambda_{\rm rf}(\sigma) = \max_{\eta'} \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} c_{\rm rf}(\sigma'|\eta',\sigma) = 1 + \mathcal{O}(\epsilon)$. Then

$$\begin{aligned}\widetilde{\lambda}(\sigma) = \sum_{\sigma'} \widetilde{c}(\sigma,\sigma') &= \sum_{\eta'} \bar{c}(\eta,\eta') \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} c_{\rm rf}(\sigma'|\eta',\sigma) = \sum_{\eta'} (\bar{c}(\eta,\eta') + \mathcal{O}(\epsilon)) \\ &= \bar{\lambda}(\eta) + |\bar{\Sigma}_\eta|\mathcal{O}(\epsilon)\,,\end{aligned}$$

and

$$\frac{\widetilde{\lambda}(\sigma)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} = \frac{\bar{\lambda}(\eta) + |\bar{\Sigma}_\eta|\mathcal{O}(\epsilon)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} = 1 + \mathcal{O}(\epsilon)\,,$$

where in the last equality we have used that $\lambda_{\rm rf}(\sigma) = 1 + \mathcal{O}(\epsilon)$ and $\bar{\lambda}(\eta) \sim |\bar{\Sigma}_\eta|$ in view of the definition $\bar{\lambda}(\eta) = \sum_{\eta' \in \bar{\Sigma}} \bar{c}(\eta,\eta')$. Therefore the rejection probability (3.7) satisfies

$$p_{\rm rej}^{\rm multi}(\sigma) = 1 - \frac{\widetilde{\lambda}(\sigma)}{\bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)} = \mathcal{O}(\epsilon)\,.$$

□

REMARK 3.1. Other implementation strategies can be designed, for instance employing a rejection-free method at the second level and/or a null-event method at the first level. However, when $\lambda_{\rm rf}(\sigma,\eta') := \sum_{\{\sigma':\mathbf{T}\sigma'=\eta'\}} c_{\rm rf}(\sigma'|\eta',\sigma)$ the method can be implemented by a rejection-free algorithm at the microscopic level but the process generated will violate the Markovian property unless $\lambda_{\rm rf}(\sigma) = \lambda_{\rm rf}(\sigma,\eta')$ for all $\eta' \in \bar{\Sigma}$.Another possible modification in the algorithm is performing a null-event at the coarse level combined with a null-event algorithm at the microscopic level. In this approach the need of a uniform normalization constant due to the null event nature in both steps is a disadvantage since then the time update is small. In fact the average time update will be inversely proportional to

$$\widetilde{\lambda}^* = N \max_{\eta,\eta'}\{\bar{c}(\eta,\eta')\} \max_{\sigma',\sigma}\{c_{\rm rf}(\sigma'|\eta',\sigma)\} \leq \lambda^*\,,$$

which indicates that the method will have higher rejection rate than a conventional null event algorithm. Even though the last two implementation techniques seem to have disadvantages, we expect that application in case studies could be effective, for example when $\widetilde{\lambda}^* \simeq \lambda^*$ and/or the computational acceleration due to coarsening is significant.

4. Applications to complex interacting particle systems. We demonstrate the use of the ML-KMC method in the efficient and accurate simulation of complex systems where coarse-graining either fails or it is computationally very expensive. In this paper we focus on the important example of systems with competing short- and long- ranged interaction; such systems typically have complicated phase diagrams, may exhibit pattern formation and arise in numerous physical and chemical systems, [29, 6, 27]. We discuss such a system next and use it as a demonstration example for the ML-KMC method in the following Sections.

4.1. Microscopic dynamics. We consider an Ising-type system on a periodic d-dimensional lattice Λ with $N = n^d$ lattice points. At each $x \in \Lambda$ we can define an order parameter $\sigma(x)$. For example, when taking values 0 and 1, it can describe

vacant and occupied sites. The microscopic dynamics are described by a continuous time Markov chain with state space $\Sigma_N = \{0,1\}^\Lambda$. For a configuration σ we denote by σ^x the configuration which differs from σ by an order parameter flip at the site x. 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 $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$ is a continuous time Markov jump process with the generator defined in terms of the rate $c(x,\sigma)$ by (2.1). In this work we present as an example the dynamics with the Arrhenius rate for the spin-flip (adsorption-desorption) mechanism

$$c(x,\sigma) = d_0\big(1-\sigma(x)\big) + d_0\sigma(x)\exp\big[-\beta U(x,\sigma)\big]\,, \tag{4.1}$$

where $U(x,\sigma) = H_N(\sigma) - H_N(\sigma^x)$ is the interaction energy of a particle located at the lattice site $x \in \Lambda$. Arrhenius laws are typically used in micro-kinetic modeling of physiochemical applications, see for instance [32]. The Hamiltonian function $H_N : \Sigma_N \to \mathbb{R}$ defines the total energy of the configuration $\sigma \in \Sigma_N$ and we consider pairwise long- and short-range interactions

$$H_N(\sigma) = H^{(s)}(\sigma) + H^{(l)}(\sigma) + \sum_{x\in\Lambda} h(x)\sigma(x)\,, \tag{4.2}$$

where $h = h(x)$ is an external field and

$$H^{(l)} = \sum_{x\neq y} J(x-y)\sigma(x)\sigma(y)\,, \quad H^{(s)} = \sum_{x\neq y} K(x-y)\sigma(x)\sigma(y)\,.$$

We consider interaction potentials $J,\ K : \mathbb{R}^d \to \mathbb{R}$ depending on the distance of x and y, where the distance $|x-y|$ does not have to be necessarily measured in the Euclidean norm. Furthermore, we assume:

$$\text{(A1)}\quad J(x-y) \equiv \frac{1}{L^d} V^{(l)}\left(\frac{n}{L}|x-y|\right)\,,\ \ x,y\in\Lambda \text{ and } L>0 \tag{4.3}$$
$$V^{(l)} \in C^1(\mathbb{R})\,,\ V^{(l)}(-r) = V^{(l)}(r)\,,\ V^{(l)}(r) = 0\,,\ \text{for } r>1$$
$$\text{(A2)}\quad K \in L^1_{\mathrm{loc}}(\mathbb{R}) \text{ and } K(x-y)\neq 0 \text{ for } |x-y|\leq S \tag{4.4}$$
and S is independent of the lattice size N.

The parameters L and S then define a range of interactions, i.e., the number of particles interacting with a given particle at the site x. The parameter L can be equal to n, i.e., interactions with all particles. The scaling in (A1) ensures that in the infinite volume limit $N\to\infty$, $J \in L^1(\mathbb{R})$. Note that the regularity condition imposed on $V^{(l)}$ rules out singular potentials such as Coulomb, Lenard-Jones etc. However, such cases can be treated in our analysis by splitting the potential into a smooth part (with $L=n$) and the short-range singular part. The range parameter $S \ll n$ is fixed and independent of n. With this form of the Hamiltonian the energy difference can be expressed as

$$U(x,\sigma) = \sum_{y\in\mathcal{I}^{(s)}(x)} K(x-y)\sigma(y) + \sum_{y\in\in\mathcal{I}^{(l)}} J(x-y)\sigma(y) - h(x) = U^{(s)}(x,\sigma) + U^{(l)}(x,\sigma)\,, \tag{4.5}$$

where the size of the support $|\mathcal{I}^{(s)}| = S^d = \mathcal{O}_N(1)$ and with $L\sim n$ we have $|\mathcal{I}^{(l)}| = L^d = \mathcal{O}_N(N)$.

4.2. Coarse grained dynamics. Systems with smooth long-range interactions are well approximated by coarse-graining techniques, [16, 19, 22], and CGMC are reliable and highly efficient simulation methods with controlled error approximations. Furthermore, models where only short-range interactions appear are inexpensive to simulate with conventional methods as there exist algorithms with the complexity $\mathcal{O}_N(1)$ per time step, [25]. However, when both short and long-range interactions are present, the conventional methods become prohibitively expensive, while coarse-graining methods are either not easily applicable or very expensive to implement due to the necessity to incorporate in them multi-body interactions, [21, 2].

In the series of papers [16, 15, 19] the authors initiated the development of mathematical strategies for the *coarse-graining* (CG) in stochastic lattice dynamics. One constructs the coarse lattice $\bar{\Lambda}_M$ by dividing Λ in M coarse cells, each of which contains $Q = q^d$ (micro-)cells. Each coarse cell is denoted by C_k, $k \in \bar{\Lambda}_M$. A typical choice for the coarse variable in the context of Ising-type models is the block-spin over each coarse cell C_k,

$$\eta := \left\{ \eta(k) = \sum_{x \in C_k} \sigma(x) \, : \, k \in \bar{\Lambda}_M \right\},$$

defining the coarse graining map $\mathbf{T} : \Sigma_N \to \bar{\Sigma}_M$, $\mathbf{T}\sigma = \eta$ and the coarse state space $\bar{\Sigma}_M = \{0, 1, \dots, Q\}^{\bar{\Lambda}_M}$. The coarse grained approximating process $\{\eta_t\}_{t \geq 0}$, [16, 15, 23], is defined by adsorption and desorption rates of a single particle in the coarse cell C_k

$$\bar{c}_a(k, \eta) = d_0(Q - \eta(k)), \quad \bar{c}_d(k, \eta) = d_0 \eta(k) \exp\big[- \beta \bar{U}(k, \eta) \big], \tag{4.6}$$

where the CG interaction potential is given by

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

where for the coarse cells k, $l \in \bar{\Lambda}_M$ we define

$$\bar{J}(k, l) = \frac{1}{Q^2} \sum_{x \in C_k} \sum_{y \in C_l} J(x - y), \;\; \bar{J}(k, k) = \frac{1}{Q(Q-1)} \sum_{x \in C_k} \sum_{y \in C_k, y \neq x} J(x - y), \tag{4.7}$$

as the interaction potential on the coarse space.

4.3. Long-time behavior and the stationary measure. In many applications of interacting particle systems the long-time behavior of (ergodic) evolution is characterized by the stationary (equilibrium) measure. If the rates $c(x, \sigma)$ satisfy the reversibility condition with respect to the measure $\mu_{N.\beta} = Z_N^{-1} e^{-\beta H_N(\sigma)}$, also known as *detailed balance*,

$$c(x, \sigma) e^{-\beta H_N(\sigma)} = c(x, \sigma^x) e^{-\beta H_N(\sigma^x)}, \tag{4.8}$$

then the jump dynamics leaves the Gibbs measure

$$\mu_{N,\beta}(d\sigma) = \frac{1}{Z_N} e^{-\beta H_N(\sigma)} \, P_N(d\sigma) \tag{4.9}$$

invariant, as is the case for (4.1). The factor Z_N is the normalizing constant (partition function). Furthermore, the product Bernoulli distribution $P_N(d\sigma)$, is the prior

distribution on Λ representing distribution of states in a non-interacting system. The total energy $H_N(\sigma)$ of the system, at the configuration $\sigma = \{\sigma(x) : x \in \Lambda\}$, is given by the Hamiltonian H_N (4.2).

The coarse-grained process (4.6) satisfies detailed balance ensuring that the process, at least for long-range potentials J, has as its invariant measure an approximation of the coarse-grained Gibbs measure, [18],

$$\bar{\mu}^{(0)}_{M,\beta}(d\eta) = \frac{1}{\bar{Z}^{(0)}_M} e^{-\beta \bar{H}^{(0)}(\eta)} \bar{P}_M(d\eta)\,, \tag{4.10}$$

where

$$\begin{aligned}\bar{H}^{(0)}(\eta) = &-\frac{1}{2}\sum_{l\in\bar{\Lambda}_M}\sum_{\substack{k\in\bar{\Lambda}_M\\ k\neq l}} \bar{J}(k,l)\eta(k)\eta(l) - \frac{1}{2}\bar{J}(0,0)\sum_{l\in\bar{\Lambda}_M}\eta(l)\left(\eta(l)-1\right)\\ &+\sum_{k\in\bar{\Lambda}_M}\bar{h}(k)\eta(k)\,.\end{aligned} \tag{4.11}$$

Applying the same coarse-graining formula (4.7) to the short range potential K will introduce errors that are not well-controlled as $N \to \infty$, [21]. However, the multi-level technique provides an approach for constructing approximations that do not require higher-order (cluster) expansions of the short-range potentials developed in [21]. For use of the multi-level approach in the context of equilibrium sampling and corresponding error analysis we refer the reader to [12, 13]. We also revisit this error analysis in Section 5, noting that we do not require the reversibility condition (4.8) for the simulated process in our results.

4.4. ML-KMC method for Arrhenius dynamics. As a specific example we explain the ML-KMC method for sampling the microscopic process $\{\sigma_t\}_{t\geq 0}$ generated by the Arrhenius rate for the adsorption-desorption mechanism. For the model considered the coarse space rate functions are explicitly given by the coarse graining technique in Section 4.2. The reconstruction rates that we present rely on approximate sampling with potential splitting where the long and short-range interactions are split to the first and second level respectively and no corrections to coarse-grained rates are applied.

Approximate dynamics with potential splitting. In this variant only the sampling corresponding to the costly long-range interactions is performed at the first (coarse) level and the compression of the short-range interaction is avoided. The sampling of the short-range contributions is performed at the second (fine) level. The rates on the coarse space $\bar{\Sigma}_M$

$$\bar{c}_a(k,\eta) = d_0\left(Q-\eta(k)\right)\,,\quad \bar{c}_d(k,\eta) = d_0\eta(k)e^{-\beta\bar{U}^{(l)}(k,\eta)}\,,$$

where $\bar{U}^{(l)}(k,\eta) = \sum_{\substack{l\in\bar{\Lambda}_M\\ l\neq k}}\bar{J}(k,l)\eta(l) + \bar{J}(k,k)(\eta(k)-1) - \frac{1}{2}\bar{h}(k)$. Since $\sigma' = \sigma^x$ is a spin flip updating, the reconstruction rates in (3.1) are explicitly defined and denoted by

$$c^a_{\mathrm{rf}}(x|k,\eta) = \frac{1-\sigma(x)}{Q-\eta(k)}\,,\quad c^d_{\mathrm{rf}}(x|k,\eta) = \frac{\sigma(x)}{\eta(k)}e^{-\beta U^{(s)}(x,\sigma)}\,, \tag{4.12}$$

where

$$U^{(s)}(x,\sigma) = \sum_{y\neq x, y\in\Lambda} K(x-y)\sigma(y) - \frac{1}{2}h(x)\,,\; U^{(l)}(x,\sigma) = \sum_{y\neq x, y\in\Lambda} J(x-y)\sigma(y) - \frac{1}{2}h(x)\,.$$

Note that $c^a_{\rm rf}(x|k,\eta)$ (and $c^d_{\rm rf}(x|k,\eta)$) are well-defined since $\eta(k) \neq Q$ ($\eta(k) \neq 0$) for all $k \in \bar{\Lambda}_M, x \in C_k$ when adsorption (desorption) process is selected in the cell k. With this choice of rates the ML-KMC method generates a Markov process $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$ with the rate function defined by

$$\begin{aligned}\widetilde{c}(x,\sigma) &= \bar{c}_a(k,\eta)c^a_{\rm rf}(x|k,\eta) + \bar{c}_d(k,\eta)c^d_{\rm rf}(x|k,\eta)\\ &= d_0(1-\sigma(x)) + d_0\sigma(x)e^{-\beta\widetilde{U}(x,\sigma)}\,,\end{aligned} \tag{4.13}$$

where we define

$$\widetilde{U}(x,\sigma) = U^{(s)}(x,\sigma) + \bar{U}^{(l)}(k,\eta)\,,\quad x \in C_k\,, \eta = \mathbf{T}\sigma\,. \tag{4.14}$$

In Appendix A we prove that $\widetilde{c}(x,\sigma)$ satisfy the detailed balance condition with

$$\widetilde{\mu}_{N,\beta}(d\sigma) = \frac{1}{\widetilde{Z}_N} e^{-\beta\widetilde{H}_N(\sigma)}\, P_N(d\sigma)\,, \tag{4.15}$$

and $\widetilde{Z}_N$ is the normalization constant corresponding to the Hamiltonian

$$\begin{aligned}\widetilde{H}_N(\sigma) = &-\frac{1}{2}\sum_{x\in\Lambda}\sum_{y\neq x} K(x-y)\sigma(x)\sigma(y) - \frac{1}{2}\sum_{x\in\Lambda}\sum_{y\neq x}\bar{J}(k(x),l(y))\sigma(x)\sigma(y)\\ &+\sum_{x\in\Lambda} h(x)\sigma(x)\,.\end{aligned} \tag{4.16}$$

We define $k(x)$ to be the coarse cell $k \in \bar{\Lambda}_M$ such that $x \in C_k$.

5. Controlled-error approximations for complex systems. In this section we provide estimates that quantify the numerical error when approximating the continuous time Markov process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$ by $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$ defined by (4.13), with invariant stationary measures $\mu_{N,\beta}(d\sigma)$, (4.9), and $\widetilde{\mu}_{N,\beta}(d\sigma)$, (4.15), respectively. We prove information loss estimates in long-time stationary regimes and weak error estimates for suitably defined *macroscopic* observables in finite time intervals. Finally, as it is evident from the proofs, our results for stationary dynamics (processes) and weak error estimates, hold true for the ML-KMC approximation of general particle systems which are not necessarily reversible.

5.1. Controlled approximations at long times. We analyze approximation properties in long-time, stationary regimes, including (a) the behavior of the stationary measure and its sampling, as well as (b) the stationary process, i.e., dynamics where the initial measure is the stationary distribution reached after long-time integration. The stationary dynamics present an especially important regime describing dynamic transitions between metastable states in complex energy landscapes, while at this regime we construct the system's phase diagrams, see the simulation of hysteresis in Figure 8.1.

The error analysis in the stationary regime is the primary novelty in our paper from a numerical analysis perspective and it is not restricted only to the ML-KMC algorithms; these questions are ubiquitous for numerical approximations of reversible and irreversible stochastic dynamics such as Langevin stochastic differential equations, thermostated dynamics, dissipative particle dynamics methods, etc. We expect that our proposed entropy-based perspective could be used to assess such numerical schemes at the stationary dynamics regime in a quantifiable manner, in a variety of stochastic, extended as well as finite-dimensional systems.

(a) *Estimates for the stationary measure.* For reversible systems, the explicit knowledge of the invariant measures of the ML-KMC process allows us to compare them directly to the Gibbs states associated with reversible kMC. Error estimates are given in terms of the specific relative entropy

$$\mathcal{R}(\mu|\nu) \equiv N^{-1} \int_{\Sigma} \log \left\{ d\mu(\sigma)/d\nu(\sigma) \right\} \mu(d\sigma)$$

between the corresponding equilibrium Gibbs measures. The scaling factor N^{-1} is related to the extensivity of the system, hence the proper error quantity that needs to be tracked is the loss of information *per particle* $\mathcal{R}(\mu_{N,\beta}|\widetilde{\mu}_{N,\beta})$. Relative entropy was used as measure of loss of information in coarse-graining in [18, 20], and as means for sensitivity analysis in the context of climate modeling problems, [28]. One of the results in [18, 22] concerns derivation of the loss of information per particle estimates on the coarse space for smooth long-range interactions J, (4.3). In the following theorem we prove an analogous error estimate on the microscopic space taking into account both short and long-range interactions.

THEOREM 5.1. *Let $\widetilde{\mu}_{N,\beta}$ be the approximating measure of the microscopic equilibrium measure $\mu_{N,\beta}$ defined by (4.15) and (4.9), with Hamiltonian functions (4.16) and (4.2) respectively, then the loss of information per particle is estimated by*

$$\mathcal{R}(\mu_{N,\beta}|\widetilde{\mu}_{N,\beta}) = \mathcal{O}_N \left(\beta \frac{q}{L} \|\nabla V^{(l)}\|_{\infty} \right) .$$

Before continuing with the proof of the theorem we state a necessary estimate in Lemma 5.2 that is proved in [18].

LEMMA 5.2. *Assume the interaction potential J satisfies* (A1) *in (4.3), then the coarse-grained interaction potential $\bar{J}$, given by (4.7) at the coarsening level q approximates for any x, $y \in \Lambda$ and k, $l \in \bar{\Lambda}_M$ the potential J with the error*

$$|J(x-y) - \bar{J}(k,l)| \leq 2\frac{q}{L} \sup_{\substack{x' \in C_k, y' \in C_l \\ y' \neq x'}} \|\nabla V^{(l)}(x'-y')\| \leq C_V \frac{q}{L^2} ,$$

where the constant C_V is independent of q, L.

Proof. [Theorem 5.1]

$$\begin{aligned} \mathcal{R}(\widetilde{\mu}_{N,\beta}|\mu_{N,\beta}) &= \frac{1}{N} \int_{\Sigma_N} \log \left(\frac{d\widetilde{\mu}_{N,\beta}}{d\mu_{N,\beta}} \right) \widetilde{\mu}_{N,\beta}(d\sigma) \\ &= \frac{1}{N} \log \frac{Z_N}{\widetilde{Z}_N} + \frac{1}{N} \mathbb{E}_{\widetilde{\mu}_{N,\beta}} \left[\beta (H_N(\sigma) - \widetilde{H}_N(\sigma)) \right] . \end{aligned}$$

By the definition of the Hamiltonian and Lemma 5.2 we have the estimate

$$\frac{1}{N} |H_N(\sigma) - \widetilde{H}_N(\sigma)| = \frac{1}{N} |H^{(l)}(\sigma) - \widetilde{H}^{(l)}(\eta)| \leq C \frac{q}{L} \|\nabla V^{(l)}\|_{\infty} ,$$

where $\eta = \mathbf{T}\sigma$ and C is a positive constant independent of the system size N, the coarsening parameter q, and the short range potential $K = K(x-y)$. Therefore

$$\frac{1}{N} \log \frac{Z_N}{\widetilde{Z}_N} = \frac{1}{N} \log \mathbb{E}_{\widetilde{\mu}_{N,\beta}} \left[\exp\{-\beta (H_N(\sigma) - \widetilde{H}_N(\sigma))\} \right] = \mathcal{O}_N \left(\beta \frac{q}{L} \|\nabla V^{(l)}\|_{\infty} \right) ,$$

$$\text{and} \quad \frac{1}{N} \mathbb{E}_{\widetilde{\mu}_{N,\beta}} \left[\beta (H_N(\sigma) - \widetilde{H}_N(\sigma)) \right] = \mathcal{O}_N \left(\beta \frac{q}{L} \|\nabla V^{(l)}\|_{\infty} \right) ,$$

that concludes the proof. □

Theorem 5.1 proves that the potential splitting does not affect equilibrium properties of the system, a fact that we indeed observe in the numerical experiment, see for example Figure 8.2. The error estimate is independent of the short-range interaction potential as was expected, since the approximation of the invariant measure results only from compressing the long-range interactions.

(b) *Approximating the stationary process.* The analysis stems from attempting to understand the striking accuracy of ML-KMC in calculating phase diagrams, hysteresis simulations, Figure 8.1 and Figure 8.3, as well as in dynamic transitions between metastable states, see Figure 8.4 and Figure 8.5. We assess the approximation of the kMC process $\{\sigma_t\}_{t\in[0,T]}$ by the ML-KMC $\{\tilde{\sigma}_t\}_{t\in[0,T]}$, when the initial data are sampled from a stationary measure. We consider the relative entropy per particle formula in the time interval $[0,T]$

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

where $\mathcal{D}_{[0,T]}$ (resp. $\tilde{\mathcal{D}}_{[0,T]}$) is the distribution of a process $\{\sigma_t\}_{t\in[0,T]}$ (resp. $\{\tilde{\sigma}_t\}_{t\in[0,T]}$) on the path space $\mathcal{Q}([0,T],\Sigma_N)$, the space of right continuous with left limits Σ_N-valued functions defined on $[0,T]$. We note that the relative entropy measures the loss of information in the approximation of the kMC process $\{\sigma_t\}_{t\in[0,T]}$ by the ML-KMC $\{\tilde{\sigma}_t\}_{t\in[0,T]}$.

If the initial distribution of the process $\{\sigma_t\}_{t\in[0,T]}$ (resp. $\{\tilde{\sigma}_t\}_{t\in[0,T]}$) is the stationary measure μ (resp. $\tilde{\mu}$), then the specific relative entropy simplifies to the following relation, [7],

$$\mathcal{R}\left(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}\right)=T\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]})+\mathcal{R}\left(\mu|\tilde{\mu}\right)\,,\tag{5.1}$$

where $\mathcal{R}\left(\mu|\tilde{\mu}\right)$ is the specific relative entropy between the stationary measures, and

$$\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]})=\frac{1}{N}\mathbb{E}_{\mu}\left[\lambda(\sigma)-\tilde{\lambda}(\sigma)-\sum_{\sigma'}\lambda(\sigma)p(\sigma,\sigma')\log\frac{\lambda(\sigma)p(\sigma,\sigma')}{\tilde{\lambda}(\sigma)\tilde{p}(\sigma,\sigma')}\right]\,,\tag{5.2}$$

is given in terms of the jump rates λ, $\tilde{\lambda}$ and jump probabilities p, $\tilde{p}$ of $\{\sigma_t\}_{t\in[0,T]}$ and $\{\tilde{\sigma}_t\}_{t\in[0,T]}$ respectively. Indeed, by Girsanov's formula, [24], we obtain the corresponding Radon-Nikodym derivative

$$\frac{d\mathcal{D}_{[0,T]}}{d\tilde{\mathcal{D}}_{[0,T]}}(\rho_t)=\frac{\mu(\rho_0)}{\tilde{\mu}(\rho_0)}\exp\left\{\int_0^T[\lambda(\rho_s)-\widetilde{\lambda}(\rho_s)]ds-\int_0^T\sum_{\sigma\in\Sigma}p(\sigma,\rho_s)\log\frac{\lambda(\sigma)p(\sigma,\rho_s)}{\widetilde{\lambda}(\sigma)\widetilde{p}(\sigma,\rho_s)}dN_s(\rho)\right\}\tag{5.3}$$

on any path $\{\rho_t\}_{t\in[0,T]}$ in $\mathcal{Q}([0,T],\Sigma_N)$, where $N_s(\rho)$ is the number of jumps of the path ρ up to time s. Then for any continuous, bounded function $\phi:\Sigma_N\to\mathbb{R}$

$$\mathbb{E}_{\mathcal{D}}\left[\int_0^T\phi(\rho_s)\,dN_s(\rho)\right]=\mathbb{E}_{\mathcal{D}}\left[\int_0^T\phi(\rho_s)\lambda(\rho_s)\,ds\right]=T\mathbb{E}_{\mu}[\phi\lambda]\,,$$

where $\mathbb{E}_{\mathcal{D}}[\phi(\rho_s)]=\int\phi(\rho_s)d\mathcal{D}_{[0,T]}$, the first equality results from the fact that $N_t-\int_0^t\lambda(\rho_s)\,ds$ is a (zero mean) martingale and the second equality follows because

$\{\rho_t\}_{t\in[0,T]}$ is a stationary process, i.e., $\mathbb{E}_{\mathcal{D}}[\phi(\rho_t)] = \mathbb{E}_{\mu}[\phi(\rho_t)]$. Thus we have

$$\begin{aligned}\mathcal{R}\left(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}\right) &= N^{-1}\mathbb{E}_{\mathcal{D}}\left[\log\frac{d\mathcal{D}_{[0,T]}}{d\tilde{\mathcal{D}}_{[0,T]}}\right] = N^{-1}\mathbb{E}_{\mathcal{D}}\left[\log\frac{\mu}{\tilde{\mu}}\right]\\ &+ N^{-1}\mathbb{E}_{\mathcal{D}}\left[\int_0^T[\lambda(\rho_s)-\widetilde{\lambda}(\rho_s)]\,ds - \int_0^T\sum_{\sigma\in\Sigma}p(\sigma,\rho_s)\log\frac{\lambda(\sigma)p(\sigma,\rho_s)}{\widetilde{\lambda}(\sigma)\widetilde{p}(\sigma,\rho_s)}\,dN_s(\rho)\right]\\ &= TN^{-1}\mathbb{E}_{\mu}\left[\lambda(\sigma)-\tilde{\lambda}(\sigma)-\sum_{\sigma'}\lambda(\sigma)p(\sigma,\sigma')\log\frac{\lambda(\sigma)p(\sigma,\sigma')}{\tilde{\lambda}(\sigma)\tilde{p}(\sigma,\sigma')}\right] + \mathcal{R}\left(\mu|\tilde{\mu}\right)\\ &= T\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}) + \mathcal{R}\left(\mu|\tilde{\mu}\right),\end{aligned}$$

which is the formula (5.1).

REMARK 5.1. Formula (5.1) shows that in the stationary dynamics regime the information loss consists of two terms, one which scales as $\mathcal{O}_T(1)$ in T and is related to the stationary measures μ and $\tilde{\mu}$ and another one that captures the stationary dynamics and scales as $\mathcal{O}_T(T)$. Furthermore, we note that for the stationary process approximation the relevant quantity is the *relative entropy per unit time* $\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]})$, (5.2). On one hand, (5.1) implies that the loss of information increases linearly in time, while the stationary measure loss of information becomes irrelevant as $T\to\infty$. The fact that as T grows in (5.1) the term $\mathcal{R}\left(\mu|\tilde{\mu}\right)$ becomes unimportant is especially useful since $\tilde{\mu}$ is typically not known explicitly (contrary to the case in (4.15)), while in non-reversible systems (e.g., reaction-diffusion kMC) μ is not known either, however, due to (5.1) it is not necessary to calculate or estimate $\mathcal{R}\left(\mu|\tilde{\mu}\right)$.

We can use this information theory-based perspective to evaluate broad classes of numerical schemes for extended stochastic processes such as the ones arising in kMC, in the long-time, stationary process regime. However, here the following theorem provides the information loss estimates for approximating the microscopic process $(\{\sigma_t\}_{t\in[0,T]},\mathcal{L})$ with $(\{\tilde{\sigma}_t\}_{t\in[0,T]},\widetilde{\mathcal{L}})$ defined by (4.1) and (4.13) respectively.

THEOREM 5.3. *[A priori estimates.] Let $(\{\sigma_t\}_{t\in[0,T]},\mathcal{L})$ and $(\{\tilde{\sigma}_t\}_{t\in[0,T]},\widetilde{\mathcal{L}})$ be stationary processes with the initial distributions $\mu_{N,\beta}$ and $\widetilde{\mu}_{N,\beta}$ respectively, then*
(a) For any fixed time $T>0$

$$\mathcal{R}\left(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}\right) = T\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}) + \mathcal{R}\left(\mu_{N,\beta}|\widetilde{\mu}_{N,\beta}\right) \tag{5.4}$$

where

$$\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}) = \frac{1}{N}\mathbb{E}_{\mu}\left[\lambda(\sigma)-\tilde{\lambda}(\sigma)-\sum_{\sigma'}\lambda(\sigma)p(\sigma,\sigma')\log\frac{\lambda(\sigma)p(\sigma,\sigma')}{\tilde{\lambda}(\sigma)\tilde{p}(\sigma,\sigma')}\right]. \tag{5.5}$$

(b) For any N, coarsening parameter $q<L$ and interaction potentials $J(x-y)$, $K(x-y)$ satisfying (A1)*, (4.3), and* (A2)*, (4.4), respectively,*

$$\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}) \le \beta\frac{q}{L}C(K,V,\beta)\|\nabla V^{(l)}\|_1\,, \tag{5.6}$$

where $\|\cdot\|_1\equiv\|\cdot\|_{L^1}$ is the L^1 norm on $\mathbb{R}$ and $C(K,V,\beta)=C(\|K\|_\infty,\|V^{(l)}\|_\infty,\beta)$ is a constant independent of N.

Proof. (a) Relation (5.4) is a direct consequence of the earlier discussion.

(b) Recalling the definition of the Markov jump process for the microscopic and the approximating process we have for all $x \in \Lambda$, $\sigma \in \Sigma_N$ that $\lambda(\sigma)p(\sigma,\sigma^x) = c(x,\sigma)$, $\lambda(\sigma) = \sum_{x\in\Lambda} c(x,\sigma)$ and $\widetilde{\lambda}(\sigma)\tilde{p}(x,\sigma) = \widetilde{c}(x,\sigma')$, $\widetilde{\lambda}(\sigma) = \sum_{x\in\Lambda}\widetilde{c}(x,\sigma)$, with the rate functions $c(x,\sigma) = d_0(1-\sigma(x)) + d_0\sigma(x)e^{-\beta U(x,\sigma)}$ and $\widetilde{c}(x,\sigma) = d_0(1-\sigma(x)) + d_0\sigma(x)e^{-\beta\widetilde{U}(x,\sigma)}$ as defined in (4.1) and (4.13) respectively. Then according to formula (5.2) we have

$$\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}_{[0,T]}) = N^{-1}\mathbb{E}_{\mu_{N,\beta}}\Big[(\lambda(\sigma)-\widetilde{\lambda}(\sigma)) - \sum_{x\in\Lambda} c(x,\sigma)\log\frac{c(x,\sigma)}{\widetilde{c}(x,\sigma)}\Big].$$

We define $\Delta_{q,N}(x,\sigma) \equiv U(x,\sigma) - \widetilde{U}(x,\sigma)$. From the definition of $\widetilde{U}(x,\sigma)$, (4.14), it follows that $\Delta_{q,N}(x,\sigma) = U^{(l)}(x,\sigma) - \bar{U}^{(l)}(k,\mathbf{T}\sigma)$, for all $x \in C_k$. In view of this equality a straightforward application of Lemma 5.2 states that there exists a constant $c > 0$ such that the microscopic potential $U(x,\sigma)$ is approximated by $\widetilde{U}(x,\sigma)$ with

$$|\Delta_{q,N}(x,\sigma)| \le c\frac{q}{L}\|\nabla V^{(l)}\|_\infty, \text{ for all } \sigma \in \Sigma_N, x \in \Lambda_N\,. \tag{5.7}$$

Then

$$\begin{aligned}
\mathcal{H}(\mathcal{D}_{[0,T]}|\widetilde{\mathcal{D}}_{[0,T]}) &= N^{-1}\mathbb{E}_{\mu_{N,\beta}}\left[\sum_{x\in\Lambda} e^{-\beta U(x,\sigma)}\sigma(x)\left(1-e^{-\beta\Delta_{q,N}(x,\sigma)}\right)\right]\\
&\quad - N^{-1}\mathbb{E}_{\mu_{N,\beta}}\left[\sum_{\substack{x\in\Lambda\\ \sigma(x)=1}} d_0\sigma(x)e^{-\beta U(x,\sigma)}\log\frac{e^{-\beta U(x,\sigma)}}{e^{-\beta\widetilde{U}(x,\sigma)}}\right]\\
&\le N^{-1}C(K,V,\beta)\mathbb{E}_{\mu_{N,\beta}}\left[\sum_{x\in\Lambda}\beta\Delta_{q,N}(x,\sigma)\right]\\
&\quad + N^{-1}\mathbb{E}_{\mu_{N,\beta}}\left[\sum_{\substack{x\in\Lambda\\ \sigma(x)=1}} d_0\sigma(x)e^{-\beta U(x,\sigma)}\Delta_{q,N}(x,\sigma)\right]\\
&\le 2c\beta\frac{q}{L}C(K,V,\beta)\|\nabla V^{(l)}\|_1\,,
\end{aligned}$$

where $C(K,V,\beta) = \sup_{\sigma,x}\exp\{-\beta U(x,\sigma)\}$.
□

Remark 5.2. *[A posteriori error analysis]* Reversing the roles of μ and $\tilde{\mu}$ in the formula (5.5) we obtain an *a posteriori* calculation on the loss of information in (5.4):

$$\mathcal{H}(\tilde{\mathcal{D}}_{[0,T]}|\mathcal{D}_{[0,T]}) = \frac{1}{N}\mathbb{E}_{\tilde{\mu}}\Big[(\tilde{\lambda}(\sigma)-\lambda(\sigma)) - \sum_{\sigma'}\tilde{\lambda}(\sigma)\tilde{p}(\sigma,\sigma')\log\frac{\tilde{\lambda}(\sigma)\tilde{p}(\sigma,\sigma')}{\lambda(\sigma)p(\sigma,\sigma')}\Big]. \tag{5.8}$$

Indeed, viewing this function as an observable estimated on the approximating stationary process, we note that it can be computed *a posteriori* in the course of an ML-KMC simulation by sampling from the stationary measure $\tilde{\mu}$. We note that in [18] we derived and tested computationally a posteriori estimates for adaptive coarse-graining of extended systems, based on a similar relative entropy approach for sampling the stationary distributions. The *a posteriori* representation in (5.8) is general and applies to both reversible and irreversible processes and does not require the *a priori* estimates in Theorem 5.3 (b)(c). The complexity of numerical calculation of

(5.8) depends on the complexity of the studied model. For example, in the lattice systems we study here, the loss of information per unit time is sampled in the course of a ML-KMC simulation for $T \gg 1$ as

$$\frac{1}{T}\mathcal{R}\left(\tilde{\mathcal{D}}_{[0,T]}|\mathcal{D}_{[0,T]}\right) \approx \mathcal{H}(\tilde{\mathcal{D}}_{[0,T]}|\mathcal{D}_{[0,T]}) = \frac{1}{N}\mathbb{E}_{\tilde{\mu}}\Big[(\tilde{\lambda}(\sigma)-\lambda(\sigma))-\sum_{x\in\Lambda}\tilde{c}(x,\sigma)\log\frac{\tilde{c}(x,\sigma)}{c(x,\sigma)}\Big]$$

From the practical point of view we have to deal with the somewhat computationally costly summation over the lattice which, in principle, has the complexity of $\mathcal{O}(N)$.

5.2. Weak error estimates in finite time. In this section we prove weak error estimates of the approximation for a class of *macroscopic* observable quantities defined below. The weak error is defined by $e_w = |\mathbb{E}_{\sigma_0}[\phi(\sigma_t)] - \mathbb{E}_{\sigma_0}[\phi(\widetilde{\sigma}_t)]|$ for an observable ϕ on the microscopic configuration space Σ_N, where the expectation is defined for the path conditioned on the initial configuration σ_0. We provide a quantitative measure of the controllable approximation that depends on two features: (a) the coarsening level q and (b) the potential splitting. The explicit dependence on the strength of the short range interactions provides us a measure to control splitting of the interactions into short and long-range parts.

THEOREM 5.4. *Let $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$ be the Markov process generated by the conventional kinetic Monte Carlo method, and $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$ the process generated by the ML-KMC method, both with initial condition σ_0. For any macroscopic observable, i.e. a function $\phi \in L^\infty(\Sigma_N)$, such that when $\partial_x\phi(\sigma) := \phi(\sigma^x) - \phi(\sigma)$,*

$$\sum_x \|\partial_x\phi\|_\infty \leq C < \infty\,, \qquad \text{where } C \text{ is independent of } N\,, \tag{5.9}$$

the weak error satisfies, for $0 < T < \infty$,

$$|\mathbb{E}_{\sigma_0}[\phi(\sigma_T)] - \mathbb{E}_{\sigma_0}[\phi(\widetilde{\sigma}_T)]| \leq C(K,V,\beta)\, C_T \frac{q}{L} \tag{5.10}$$

where C_T is a constant independent of the system size and $C(K,V,\beta) = K_S J_L$, $K_S = |\sup_{x,\sigma} e^{-\beta U^{(s)}(x,\sigma)}|$ and $J_L = |\sup_{x,\sigma} e^{-\beta \bar{U}^{(l)}(k(x),\mathbf{T}\sigma)}|$.

Some typical macroscopic observables satisfying (5.9) are the coverage, $c(\sigma_t) = \frac{1}{N}\sum_{x\in\Lambda}\sigma_t(x)$, the spatial correlations $f(\sigma;k) = \frac{1}{N}\sum_{x\in\Lambda}\sigma(x)\sigma(x+k)$, and the Hamiltonian defined in (4.2).

For the proof of the theorem we will need the following Lemma 5.5 that we prove in Appendix B, see also [19]. We define $u(t,\sigma_0) = \mathbb{E}[\phi(\sigma_T)|\sigma_t = \sigma_0]$ and the function $u(t,\sigma_0)$ solves the backward Kolmogorov equation, i.e., the final value problem

$$\partial_t u(t,\sigma) + \mathcal{L}u(t,\sigma) = 0\,, \quad u(T,\cdot) = \phi\,, \;\; t < T \tag{5.11}$$

For all observables ϕ satisfying (5.9) we can estimate $\partial_x u(t,\sigma) = u(t,\sigma^x) - u(t,\sigma)$ independently of N since we have the following estimate.

LEMMA 5.5. *Let $u(t,\sigma)$ a solution of (5.11) where $\mathcal{L}$ is the infinitesimal generator $\mathcal{L}f(\sigma) = \sum_x c(x,\sigma)(f(\sigma^x) - f(\sigma))$ defined by the rate function $c(x,\sigma)$ given in (4.1). Then for any $t \leq T$*

$$\sum_x \|\partial_x u(t,\cdot)\|_\infty \leq C_T \sum_x \|\partial_x\phi\|_\infty \tag{5.12}$$

We continue with the proof of Theorem 5.4.

Proof. Using the martingale property we have for any smooth function $v(t,\sigma_0)$ and the process $\{\widetilde{\sigma}_t\}_{t\geq 0}$ with the generator $\widetilde{\mathcal{L}}$

$$\mathbb{E}_{\sigma_0}[v(T,\widetilde{\sigma}_T)] = \mathbb{E}_{\sigma_0}[v(0,\widetilde{\sigma}_0)] + \int_0^T \mathbb{E}_{\sigma_0}[(\partial_s + \widetilde{\mathcal{L}})v(s,\widetilde{\sigma}_s))]ds\,.$$

Therefore

$$\begin{aligned}
\mathbb{E}_{\sigma_0}[\phi(\sigma_T)] - \mathbb{E}_{\sigma_0}[\phi(\widetilde{\sigma})] &= \mathbb{E}_{\sigma_0}[u(0,\sigma_0)] - \mathbb{E}_{\sigma_0}[u(T,\widetilde{\sigma}_T)] \\
&= \int_0^T \mathbb{E}_{\sigma_0}[(\partial_s + \widetilde{\mathcal{L}})u(s,\widetilde{\sigma}_s)]\,ds \\
&= \int_0^T \mathbb{E}_{\sigma_0}[\widetilde{\mathcal{L}}u(s,\widetilde{\sigma}_s) - \mathcal{L}u(s,\widetilde{\sigma}_s)]\,ds \\
&= \int_0^T \mathbb{E}_{\sigma_0}\left[\sum_{x\in\Lambda}\left(\widetilde{c}(x,\widetilde{\sigma}_s) - c(x,\widetilde{\sigma}_s)\right)\partial_x u(s,\widetilde{\sigma}_s)\right]\,ds\,.
\end{aligned}$$

However, we can bound the last term by

$$\|c-\widetilde{c}\|_\infty \int_0^T \mathbb{E}_{\sigma_0}\left[\sum_{x\in\Lambda_N}|\partial_x u(s,\widetilde{\sigma}_s)|ds\right]$$

We conclude by using Lemma 5.5 and noting that

$$|\widetilde{c}(x,\widetilde{\sigma}_s) - c(x,\widetilde{\sigma}_s)| \leq d_0 e^{-\beta\bar{U}^{(l)}(k(x),\mathbf{T}\widetilde{\sigma}))}\left|e^{-\beta\bar{U}^{(l)}(x,\mathbf{T}\widetilde{\sigma}_s)} - e^{-\beta U^{(l)}(x,\widetilde{\sigma}_s)}\right|\,,$$

where using (5.7),

$$\left|e^{-\beta\bar{U}^{(l)}(x,\mathbf{T}\widetilde{\sigma}_s)} - e^{-\beta U^{(l)}(x,\widetilde{\sigma}_s)}\right| \leq Ce^{-\beta\bar{U}^{(l)}(k(x),\mathbf{T}\widetilde{\sigma}))}|\Delta_{q,N}(x,\sigma)| \leq CJ_L\frac{q}{L}\|\nabla V^{(l)}\|_\infty\,,$$

for some $C>0$. ∎

6. Exact sampling of kMC dynamics. In addition to the approximate dynamics discussed so far, the ML-KMC method can also generate the *exact* dynamics associated with the rates $c(\sigma,\sigma')$ by appropriate choice of the coarse and reconstruction rates, albeit at higher cost than the controlled-approximation dynamics $\widetilde{c}(\sigma,\sigma')$. More specifically, given $c(\sigma,\sigma')$ and $\bar{c}(\eta,\eta')$, $c_{\rm rf}(\sigma'|\eta',\sigma)$ can be selected such that

$$\bar{c}(\eta,\eta')c_{\rm rf}(\sigma'|\eta',\sigma) = \widetilde{c}(\sigma,\sigma') \equiv c(\sigma,\sigma')\,. \tag{6.1}$$

In this case the ML-KMC method generates exactly the same stochastic process with the direct kMC method achieving a perfect reconstruction. Relation (6.1) ensures that processes $\{\widetilde{\sigma}_t\}_{t\geq 0}$ and $\{\sigma_t\}_{t\geq 0}$ have the same generator $\widetilde{\mathcal{L}} = \mathcal{L}$, which is sufficient to prove that the two processes are identical, [26]. As a specific example we demonstrate exact sampling via the ML-KMC method for the microscopic process $\{\sigma_t\}_{t\geq 0}$ generated by the Arrhenius rate for the adsorption-desorption mechanism in Section 4.1. For this model the coarse space rate functions are explicitly given by the coarse graining technique in Section 4.2. The reconstruction rates that we present rely on *correcting* the error introduced by coarsening at the null-event step.

Indeed, the rates on the coarse space $\bar{\Sigma}_M$ corresponding to the compressed interactions for $U(x,\sigma)$, (4.5), are given by

$$\bar{c}_a(k,\eta) = d_0\left(Q-\eta(k)\right)\ , \quad \bar{c}_d(k,\eta) = d_0\eta(k)e^{-\beta\bar{U}(k,\eta)}\ ,$$

where

$$\bar{U}(k,\eta) = \sum_{\substack{l\in\bar{\Lambda}_M\\ l\neq k}} [\bar{K}(k,l)+\bar{J}(k,l)]\eta(l) + [\bar{K}(k,k)+\bar{J}(k,k)](\eta(k)-1) - \bar{h}(k)\,.$$

The reconstruction rates are explicitly defined by

$$c^a_{\rm rf}(x|k,\eta) = \frac{1-\sigma(x)}{Q-\eta(k)}\,, \quad c^d_{\rm rf}(x|k,\eta) = \frac{\sigma(x)}{\eta(k)}e^{-\beta(U(x,\sigma)-\bar{U}(k,\eta))}\,, \tag{6.2}$$

where we can also compare them to the reconstruction of the approximate dynamics in (4.12). This choice of rates ensures that the ML-KMC method generates the same process $(\{\sigma_t\}_{t\geq 0},\mathcal{L})$ since the two-level process has the rates $\widetilde{c}(x,\sigma) = c(x,\sigma)$. Furthermore, the quantities $U(x,\sigma)-\bar{U}(k,\eta)$ are better localized in the sense that they decay faster than $U(x,\sigma)$, hence it is easier to compress through truncation, [2]. Finally, the detailed balance condition is satisfied with invariant measure $\mu_{N,\beta}(d\sigma)$; for completeness we give the proof in Appendix A.

Remark 6.1. In analogy to the path-wise decomposition (3.1) of the stochastic process, exact or controlled error equilibrium sampling has been achieved with multilevel CGMC methods, [13], based on an analogous decomposition of the sampling probability measure $\mu(d\sigma)$, i.e., $\mu(d\sigma) = \bar{\mu}(d\eta)\nu(d\sigma|\eta)$, where ν defines the reconstruction and $\bar{\mu}$ is the measure on the coarse space.

7. Acceleration and computational complexity. The purpose of this section is to compare the efficiency of a ML-KMC method with conventional methods. In the presence of long-range interactions sampling with a rejection-free algorithm is next to impossible due to the very high number of classes in BKL-type methods, see for instance Table 7.1. Hence we may inevitably be forced to use a highly inefficient null event algorithm. With the proposed ML-KMC approach, a rather crude CG of the long-range potential gives rise to much fewer classes, thus we can sample at the first level rejection-free using the BKL algorithm while the next level can be null-event. The ML-KMC algorithm is applied to the stochastic lattice model for Arrhenius dynamics in Section 4, where we consider the global search implementation as in the conventional stochastic simulation algorithm (SSA).

Algorithm 1. *Two-level ML-KMC*

Given $\sigma,\eta=\mathbf{T}\sigma$

Coarse level. *(Level I)*

Update. *(a) Calculate transition rates* $\bar{c}_a(k,\eta)$, $\bar{c}_d(k,\eta)$, *for all* $k\in\bar{\Lambda}_M$ *and* $\bar{\lambda}^a_k(\eta)=\sum_{l<k}\bar{c}_a(l,\eta)$, $\bar{\lambda}^d_k(\eta)=\sum_{l<k}\bar{c}_d(l,\eta)$, $\bar{\lambda}(\eta)=\bar{\lambda}^a_M(\eta)+\bar{\lambda}^d_M(\eta)$.

Search. *Obtain uniform random numbers* $u_1,u_2\in[0,1)$.

If $\bar{\lambda}^a_M(\eta)<u_1$ *adsorb else desorb. Assume that adsorption is chosen, then find* $k\in\bar{\Lambda}_M$ *such that* $\bar{\lambda}^a_{k-1}(\eta)\leq\bar{\lambda}^a_M(\eta)u_2\leq\bar{\lambda}^a_k(\eta)$.

Microscopic level. *(Level II)*

Reconstruct. *Pick uniformly a site* x *in the cell* C_k.

Accept/Reject. *Select a uniform random number* $u \in [0,1)$*, and define* $c_{\rm rf}(x|k,\eta) = c_{\rm rf}^{a,d}(x|k,\eta)$ *according to the selected process in the coarse move, e.g. (4.12) or (6.2). If* $\lambda_{\rm rf}(\sigma)u \le c_{\rm rf}(x|k,\eta)$ *accept and update the state at the site* x*.*

Time update. *Update time from an exponential law with the parameter* $\widetilde{\lambda}^*(\sigma)$*, with* $\widetilde{\lambda}^*(\sigma) = \bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)$*.*

Here $\widetilde{\lambda}^*(\sigma) = \bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)$ is defined according to the sampling strategy. Specifically for the exact sampling of Section 6,

$$\bar{\lambda}(\eta) = \sum_k d_0(q-\eta(k)) + d_0\eta(k)e^{-\beta\bar{U}(k,\eta)}\,,$$
$$\lambda_{\rm rf}(\sigma) = q\max\left\{\frac{1}{q-\eta(k)}, \frac{1}{\eta(k)}e^{-\beta U^*}\right\},$$

and for the approximating sampling of Section 4.4,

$$\bar{\lambda}(\eta) = \sum_k d_0(q-\eta(k)) + d_0\eta(k)e^{-\beta\bar{U}^{(l)}(k,\eta)}\,,$$
$$\lambda_{\rm rf}(\sigma) = q\max\left\{\frac{1}{q-\eta(k)}, \frac{1}{\eta(k)}e^{-\beta U^{s*}}\right\},$$

where $U^* = \min_{x,\sigma}(U(x,\sigma) - \bar{U}(k,\eta))$ and $U^{s*} = \min_{x,\sigma} U^{(s)}(x,\sigma)$.

The ML-KMC method provides an efficient balance between benefits and limitations of the conventional null-event and rejection-free methods, that we summarize in Table 7.1. This is achieved by (a) improving the computational cost of a conventional rejection-free method, see Table 8.2 and (b) increasing the successful events of a null-event method. An event is considered successful when it is accepted and the system evolves to a new state. The cost per event of a kMC algorithm can be divided in two categories. The *search* cost, the computational cost to choose an event, and the cost of *updating* the rates when an event is performed. In Table 7.1 the updating cost is realized as the number of operations necessary to calculate energy differences appearing and the search cost as the length of the array from which the next event (site) is selected. We consider global search and update algorithms for the comparison here, but we mention that both the cost in the traditional rejection-free and ML-KMC algorithms can be improved with the use of a sophisticated search/update algorithm, for example, with binary tree methods, [5]. For the sake of comparison and completeness we describe the conventional sampling algorithms, SSA, BKL and null event, in Appendix C. The last column of Table 7.1 reveals the acceleration of the method in generating successful events, for example, when the system is at the state σ the rejection probability of a proposed event in the ML-KMC algorithm is given by (3.7).

Next we present another argument that reveals the fact that the proposed method improves the computational cost of kMC algorithms. The number of classes in a BKL algorithm is determined by the level sets of $U(x,\sigma)$, in fact we can write $U(x,\sigma) := \bar{U}(k,T\sigma) + E(x,\sigma)$. Thus the level sets of $\bar{U}(k,\eta)$ are defined on a coarser lattice, hence $U(x,\sigma)$ has many more (q^d more) level sets. For a potential decaying at the length L on the microscopic lattice, $U(x,\sigma) = \sum J(x-y)\sigma(y) = \mathcal{O}(2^{dL})$ different values (and classes), while the coarse interaction potential decays at a distance L/q and $\bar{U}(k,T\sigma) = \sum_l \bar{J}(k,l)\eta(l) = \mathcal{O}(q^{dL/q}) = \mathcal{O}(2^{dL\log(q)/q})$ different values (and classes). Hence a BKL algorithm on the coarse space has, by the factor $2^{dL(1-\log(q)/q)}$,

less classes in the implementation of the BKL algorithm. Clearly when the range of interactions L is large, the number of classes grows exponentially with L and implementation of a microscopic BKL algorithm is not feasible. Therefore sampling with a null-event algorithm is unavoidable.

TABLE 7.1
Computational complexity and event rejection rate comparison for a single kMC step in one space dimension.

	Search	Update	Rejection rate
Rejection free (SSA)	$\mathcal{O}(N)$	$\mathcal{O}(L \times L)$	0
Two-level ML-KMC (SSA)	$\mathcal{O}(M)$	$\mathcal{O}(L/q \times L/q)$	1 - $\widetilde{\lambda}(\sigma)/\widetilde{\lambda}^*(\sigma)$
Rejection free (BKL)	$\mathcal{O}(2^L)$	$\mathcal{O}(L \times L)$	0
Two-level ML-KMC(BKL)	$\mathcal{O}(2^{L\log(q)/q})$	$\mathcal{O}(L/q \times L/q)$	1 -$\widetilde{\lambda}(\sigma)/\widetilde{\lambda}^*(\sigma)$
Null - event	$\mathcal{O}(1)$	$\mathcal{O}(L)$	1- $\lambda(\sigma)/\lambda^*$

The ML-KMC method achieves acceleration of the rejection free simulations up of order q when sampling for the same finite time interval. We also note that since a transition to a new event is based on a single spin-flip, the reconstruction step is performed locally, confined to a single coarse cell, a fact that improves further the computational cost of the method. The computational times (CPU) compared next are those needed for reaching the same real time T with the conventional SSA and ML-KMC method, where we consider that CPU time is proportional to the computational complexity of the algorithms given in Table 7.1. Let n be the number of MC steps necessary in a rejection-free method to reach real time T. The corresponding necessary MC steps in a ML-KMC method are $m = n/\mathbb{E}_{\mu_{N,\beta}}\left[p_{\text{succ}}^{\text{multi}}(\sigma)\right]$, where $p_{\text{succ}}^{\text{multi}}(\sigma) = 1 - p_{\text{rej}}^{\text{multi}}(\sigma)$ is the acceptance probability of an event when the system is in the state σ. For the search algorithm the cost ratio of the microscopic SSA and the ML-KMC method is

$$r_s = \frac{\text{CPU}_{s,\text{rej}-\text{free}}}{\text{CPU}_{s,\text{multi}}} \sim \frac{Nn}{Mm} = q\mathbb{E}_{\mu_{N,\beta}}[p_{\text{succ}}^{\text{multi}}]$$

and for the update

$$r_u = \frac{\text{CPU}_{u,\text{rej}-\text{free}}}{CPU_{u,\text{multi}}} \sim \frac{L^2 n}{(L/q)^2 m} = q^2\mathbb{E}_{\mu_{N,\beta}}[p_{\text{succ}}^{\text{multi}}]\,.$$

8. Numerical experiments: an Ising-Curie-Weiss model. We consider a benchmark problem with competing short and long-range interactions that exhibits complex multi-phase behavior as captured in the phase diagrams in Figure 8.1 and Figure 8.3. Exact solutions for the free energy in the thermodynamic limit, $N \to \infty$, are known for the one-dimensional and two dimensional models, [14]. The energy of the system at the configuration $\sigma = \{\sigma(x), x \in \Lambda_N\}$ is defined by the Hamiltonian

$$\begin{aligned} H_N(\sigma) &= -\frac{K}{2}\sum_x \sum_{|x-y|=1} \sigma(x)\sigma(y) - \frac{J}{2N}\sum_x \sum_{y\neq x} \sigma(x)\sigma(y) - h\sum \sigma(x) \quad (8.1)\\ &\equiv H^{(s)}(\sigma) + H^{(l)}(\sigma) + E(\sigma)\,. \end{aligned}$$

The interactions involved in $H^{(s)}(\sigma)$ are nearest-neighbor with the constant strength K, while $H^{(l)}(\sigma)$ represents the long-range interactions given by the potential J with

the range $L = N$ and h is an external field. A closed form solution in the thermodynamic limit ($N \to \infty$) for the total coverage $c_\beta(K, J, h)$ in the one-dimensional model was derived in [14],

$$c_\beta(K, J, h) = \frac{1}{2} M_\beta(\frac{1}{4}K, \frac{1}{4}J, \frac{1}{2}h - \frac{1}{4}J - \frac{1}{4}K) + \frac{1}{2}\,, \tag{8.2}$$

where $M_\beta(K, J, h)$ is a solution (minimizer) of the problem

$$\min_m \left(\frac{J}{2} m^2 - \log(e^K \cosh(h + Jm) + (e^{2K} \sinh^2(h + Jm) + e^{-2K})^{1/2}) \right).$$

Depending on the system parameters $c_\beta(K, J, h)$ can be a multivalued function and phase transitions may occur. We are interested in sampling the dynamical behavior of the system in the bi-stable regimes as well as in constructing the phase diagram with respect to the external field h.

The computational examples demonstrate both acceleration of simulations with ML-KMC and the improved accuracy of ML-KMC contrasted with the CGMC simulations. The reference solution is obtained by the fully resolved microscopic simulation performed by the traditional null-event kMC. In numerical implementations we tested the three methods discussed previously:

(i) the direct null-event kMC, (Algorithm 4),
(ii) the developed ML-KMC (Algorithm 1) sampling on the microscopic space,
(iii) the null-event CGMC sampling on the coarse space only, i.e., with both short and long-range potentials coarse-grained and *without* corrections due to the reconstruction step in ML-KMC.

The energy difference $U(x, \sigma)$ appearing in the transition rates (4.1) is given by

$$U(x, \sigma) = K \sum_{|x-y|=1} \sigma(y) + J \sum_{y=1}^{N} \sigma(y) - h\,.$$

For the ML-KMC method we apply the potential splitting approach where the rates on the coarse space $\bar{\Sigma}_M$, at the first level of the method, are defined in (4.6) with the potential energy

$$\bar{U}^{(l)}(k, \eta) = J \sum_{k=1}^{M} \eta(k) - \frac{h}{2} = J \sum_{y=1}^{N} \sigma(y) - \frac{h}{2}\,,$$

and the reconstruction rates at the second level of the method are defined by (4.12) with

$$U^{(s)}(x, \sigma) = K \sum_{|x-y|=1} \sigma(y) - \frac{h}{2}\,.$$

To implement the null-event method we need a uniform upper bound of the rates $c_{\rm rf}(x|k, \eta)$, (4.12),

$$\lambda_{\rm rf}(\sigma) = q \max \left\{ \frac{1}{q - \eta(k)}, \frac{1}{\eta(k)} e^{\beta(\frac{h}{2} + K_*)} \right\} \tag{8.3}$$

where $\mathbf{T}\sigma = \eta$ and $K_* = |\min\{0, K\}|$. Therefore the time step of the method, that is proportional to $\widetilde{\lambda}^*(\sigma) = \bar{\lambda}(\eta)\lambda_{\rm rf}(\sigma)$, clearly varies with the system state σ since $\bar{\lambda}(\eta) = \sum_k [d_0(q - \eta(k)) + d_0\eta(k)e^{-\beta\bar{U}(k,\eta)}]$.

Note that in this example the coarse-grained Hamiltonian $\bar{H}^{(l)}$ is exact, i.e., $\bar{H}^{(l)}(\eta) \equiv H^{(l)}(\sigma)$, thus there is no approximation error due to coarse-graining the long-range potential. Therefore, while CGMC sampling is approximate, due to coarse-graining of both $\bar{H}^{(l)}(\eta)$ and $\bar{H}^{(s)}(\eta)$, the ML-KMC method samples the exact microscopic process, i.e., $\widetilde{c}(x,\sigma) = c(x,\sigma)$ for all $x \in \Lambda, \sigma \in \Sigma_N$. This allows us to quantify the effect of splitting the potential function into short and long-range parts. For example, Figure 8.7 shows that the potential splitting is not introducing errors, which verifies the theoretical estimate for the information loss of the equilibrium distribution, Theorem 5.1. The effect of the splitting is apparent only in the average acceptance rate of the method where the strength of the short-range interactions K controls the rejection rate according to (8.3).

In order to test the effect of coarse-graining in the ML-KMC method we modify the long-range potential in the Hamiltonian (8.1) and consider finite-range interactions, with the range $L < N$, and with the long-range part of the Hamiltonian

$$H^{(l)}(\sigma) = \frac{J}{2L} \sum_{x} \sum_{|y-x| \le L} \sigma(x)\sigma(y)\,.$$

For this case the proposed ML-KMC method is approximate, however, it still reduces significantly the coarse-graining error of the direct CGMC sampling, since compressing of the short-range part is avoided, see Figure 8.6 and description below.

In all simulations we consider the one-dimensional model with the coarsening parameter $q = N$ in the CGMC and ML-KMC methods, that is the coarse space consists of one cell $k = 1$ and the coarse variable η is the total coverage $\eta = \mathbf{T}\sigma = \sum_{x\in\Lambda} \sigma(x)$.

Stationary dynamics and equilibrium sampling. We demonstrate properties of the ML-KMC algorithm in the stationary regime by constructing (equilibrium) phase diagrams of the average coverage with respect to the external field h. We explore different regimes of the phase plane K-J. With the choice of the potential parameters K and J corresponding to bi-stable regimes we observe that the ML-KMC algorithm approximates properly the hysteresis behavior while the CGMC algorithm samples incorrect energy landscape and thus does not estimate the hysteresis behavior correctly. The coarse-graining parameter is set to $q = N$ both in ML-KMC and CGMC simulations. The ML-KMC method avoids compressing the short-range interactions that introduce large error in the CGMC simulations. Figure 8.1 depicts the hysteresis behavior in the case when the long-range potential is of Curie-Weiss type, i.e., the interaction range is $L = N$, and hence it is coarse-grained exactly by block-spin coarse variables. However, coarse-graining the short-range, nearest-neighbor Ising, potential introduces an error which leads to a wrong prediction of hysteresis in the CGMC simulation. In Figure 8.2 and 8.3 we chose the interaction range $L < N$ which also introduces coarse-graining error in the coarse-grained long-range potential. However, the presented error analysis for the invariant measure suggests that this error is small and thus the ML-KMC sampling, unlike the CGMC simulations, are in a good agreement with estimates from the microscopic simulations.

Transients and dynamical sampling. In the dynamical sampling we explore two quantities of interest:

(a) The *path-wise* behavior of the coverage, defined by $c(\sigma_t) = \frac{1}{N}\sum_{x\in\Lambda} \sigma_t(x)$ for the direct microscopic sampling, $\widetilde{c}(\widetilde{\sigma}_t) = \frac{1}{N}\sum_{x\in\Lambda} \widetilde{\sigma}_t(x)$ for the two-level sampling and $c(\eta_t) = \frac{1}{N}\sum_{k\in\bar{\Lambda}_M} \eta_t(k)$ for the CGMC sampling.

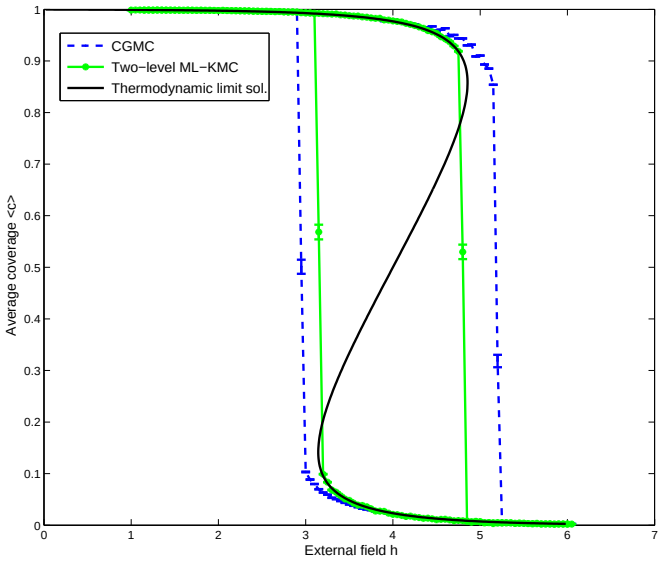

FIGURE 8.1. *Hysteresis simulation in a bi-stable regime. Potential parameters* $K = 3$, $J = 5$, $L = N$, *and the lattice-size* $N = 1024$ *and the coarsening parameter* $q = N$.

(b) The mean time to reach a transition from one equilibrium to another in the bi-stable regime, the *exit time*, $\tau = \mathbb{E}[T]$, $T = \inf\{t > 0 : c_t \geq C\}$. The probability density functions (PDFs) ρ_m, ρ_{tl}, ρ_{cg} for the exit time estimated in the microscopic, the two-level ML-KMC and the CGMC methods respectively are monitored. Starting from an initial state with the coverage $c_0 = 0$ in all methods we record the time τ when the coverage exceeds the value $C = 0.99$.

Estimating the observable τ tests both a proper approximation of the energy landscape as well as the correct time-scale in approximating dynamics. The simulation is set for the parameters K, J such that the system exhibits transition to an equilibrium which is, depending on the value of the external field h, stable or metastable (see Figure 8.1). We compare not only the expected (mean) values but also the probability density functions (PDFs) in order to demonstrate importance of error estimates in terms of the relative entropy. Probability density function was estimated from 10^4 independent samples using the MATLAB estimator `ksdensity` with a normal kernel function.

Figure 8.4 shows the comparison in the case of a single equilibrium state $c \approx 1.0$ (for the given value of h) and the long-range potential which is coarse-grained exactly, i.e., $L = N$. We observe a good agreement in all three methods, although the CGMC introduces a visible error due to coarse-graining of the short-range potential.

An additional error is introduced by coarse-graining the long-range potential with $L = 100 < N$. Comparison of the exit time PDF in the case of a single equilibrium state $c \approx 1.0$ (for the given value of h) is depicted Figure 8.5. While ML-KMC simulations are in a good agreement with the microscopic simulation the CGMC algorithm introduces significant error for the estimated PDF.

By adjusting the external field h the sampling is performed in the bi-stable regime

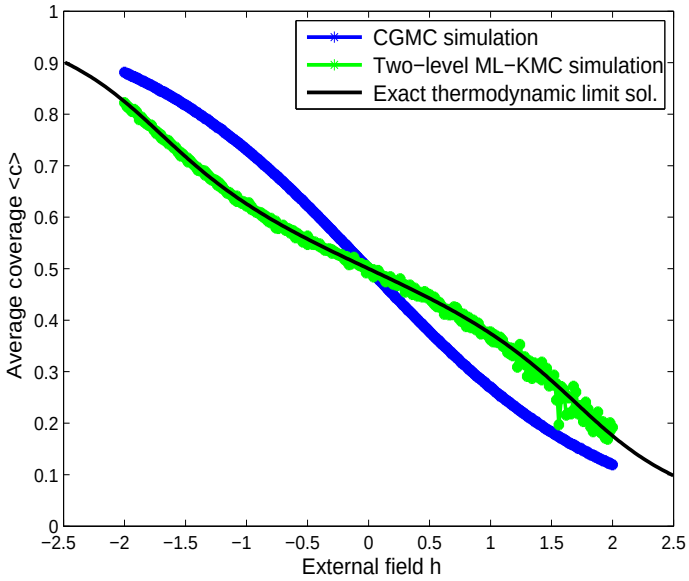

FIGURE 8.2. *Hysteresis simulation in a single phase regime. Potential parameters* $K = -5$, $J = 5$, $L = 20$, *the lattice size* $N = 256$, *and the coarsening parameter* $q = N$.

with two meta-stable equilibria. Coarse-graining both short and long-range potentials changes significantly the energy landscape and the CGMC algorithm cannot capture the transition within the simulation time-window. The exit-time probability distribution function is depicted in Figure 8.6 showing that the ML-KMC algorithm is capable of capturing the transition and approximate the exit-time PDF. The inset demonstrates that the CGMC simulation was unable to estimate the mean exit time as no transition occurred and the exit-time PDF is concentrated at the final time of the simulation window. This fact is further visualized in Figure 8.7 where evolution of the mean coverage is depicted. While the ML-KMC simulation results in a trajectory that approximates well the reference trajectory obtained from microscopic null-event kMC with a transition from $c = 0$ to $c \approx 1$ equilibrium, the trajectory averaged in the CGMC simulation does not exhibit any transition in the simulation window.

In Table 8.1 we compare numerical results for the exit time and the corresponding computational times of the three algorithms for different values of the potential parameters. For the finite range $L < N$ interactions, where coarse-graining error is present, we see that the ML-KMC method estimates are closer to the microscopic (conventional) method even when the CGMC method fails. Furthermore, we see a significant acceleration of the computational time both with the CGMC and the ML-KMC method.

Appendix A. Detailed balance.

A.1. Exact dynamics. The rate $\widetilde{c}(x,\sigma)$, (6.2), satisfies the detailed balance condition with $\mu_{N,\beta}(d\sigma)$, (4.9), i.e.,

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

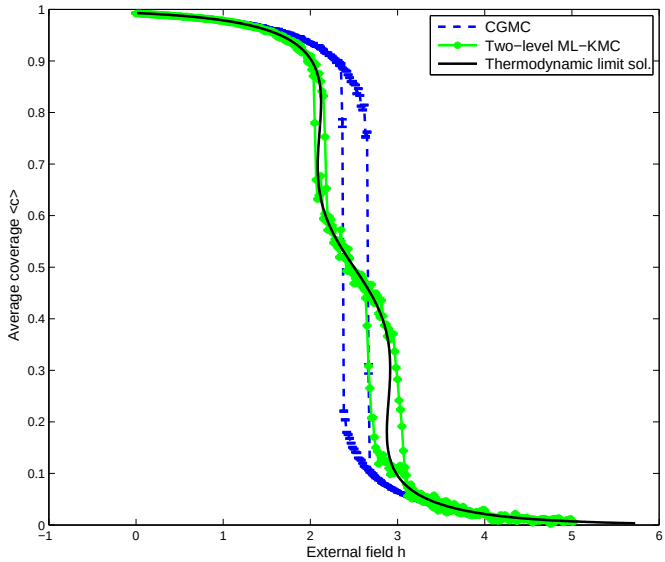

FIGURE 8.3. *Hysteresis simulation in a bi-stable regime. Potential parameters* $K = -5$, $J = 10$, $L = 100$, *the lattice size* $N = 1024$, *and the coarsening parameter* $q = N$.

TABLE 8.1

Approximation of the exit time τ. *For the statistics we use* 10^4 *samples and present the* 95% *confidence interval. The potential parameter* $J = 5$, *the coarse-graining parameter* $q = N$ *and the lattice size* $N = 1024$ *are fixed.*

Parameters	τ_m microscopic	$\tau_{\rm tl}$ ML-KMC	$\tau_{\rm cg}$ CGMC	CPU_m [sec]	$\mathrm{CPU}_{\rm tl}$ [sec]	$\mathrm{CPU}_{\rm cg}$ [sec]
$L = N$						
$K = 0, h = 1$	28.5 ±0.8	28.3±0.8	28.7±0.8	1534	9	8
$K = 2, h = 2$	6.40±0.03	6.40±0.03	6.20±0.02	884	6	5
$L = 100$						
$K = 3, h = 2.5$	6.20±0.02	6.1±0.03	5.93±0.02	158	9	7
$K = 3, h = 3.1$	11.50±0.06	12.4±0.1	44.0±0.1	526	45	100

since $\widetilde{c}(x,\sigma) = c(x,\sigma)$ for all $\sigma \in \Sigma_N$ and $x \in \Lambda$, and $c(x,\sigma)$ satisfies (4.8). We have $\widetilde{c}(x,\sigma) = c(x,\sigma)$ since

$$\begin{aligned}\widetilde{c}_a(x,\sigma) = \bar{c}_a(k,\eta)c^a_{\rm rf}(x|k,\eta) &= d_0(q-\eta(k))\frac{1-\sigma(x)}{q-\eta(k)} \\ &= d_0(1-\sigma(x)) = c_a(x,\sigma)\end{aligned}$$

and

$$\begin{aligned}\widetilde{c}_d(x,\sigma) = \bar{c}_d(k,\eta)c^d_{\rm rf}(x|k,\eta) &= d_0\eta(k)e^{-\beta\bar{U}(k,\eta)}\frac{\sigma(x)}{\eta(k)}e^{-\beta[U(x,\sigma)-\bar{U}(k,\eta)]} \\ &= d_0\sigma(x)e^{-\beta U(x,\sigma)} = c_d(x,\sigma)\,.\end{aligned}$$

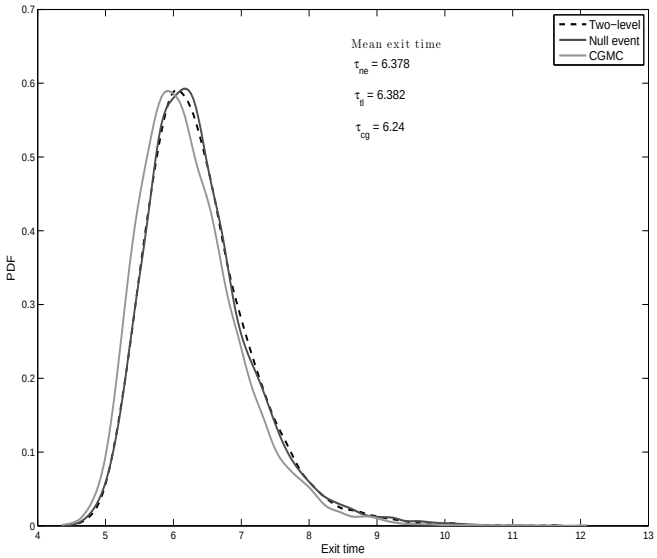

FIGURE 8.4. *Comparing the probability density function of the exit time from the initial coverage* $c_0 = 0$ *to* $c \geq 0.99$. *Potential parameters* $K = 2$, $J = 5$, $h = 2$, $L = N$, *the lattice size* $N = 1024$ *and the coarse-graining parameter* $q = N$.

TABLE 8.2
CPU time (seconds): The evolution final time $T = 20$, *the potential parameters* $K = 1$, $J = 5$, $h = 2.5$, $L = N$, *and the coarse-graining parameter* $q = N$

Lattice size N	Null event	ML-KMC
512	9	0.5
1024	33	0.9
2048	131	1.7
4096	514	4
8192	2143	13

A.2. Approximate dynamics. The approximate reaction rates $\widetilde{c}(x,\sigma)$ defined in (4.13) satisfy the DB condition with invariant measure $\widetilde{\mu}_{N,\beta}(d\sigma)$ (4.15). Indeed, if we denote $\widetilde{c}_a(x,\sigma) = \bar{c}_a(k,\eta)c^a_{\rm rf}(x|k,\eta)$ and $\widetilde{c}_d(x,\sigma) = \bar{c}_d(k,\eta)c^d_{\rm rf}(x|k,\eta)$ we have

$$\begin{aligned}
\widetilde{c}_a(x,\sigma)e^{-\beta\widetilde{H}_N(\sigma)} &= \left[\bar{c}_a(k,\eta)c^a_{\rm rf}(x|\eta)\right] e^{-\beta\widetilde{H}_N(\sigma)} \\
&= \left[d_0(q-\eta(k))\frac{1-\sigma(x)}{q-\eta(k)}\right] e^{-\beta(\widetilde{H}_N(\sigma^x)-(2\sigma(x)-1)\widetilde{U}(x,\sigma))} \\
&= \left[d_0(1-\sigma(x))e^{-\beta\widetilde{U}(x,\sigma)}\right] e^{-\beta\widetilde{H}_N(\sigma^x)} \\
&= \left[d_0\sigma^x(x)e^{-\beta\widetilde{U}(x,\sigma)}\right] e^{-\beta\widetilde{H}_N(\sigma^x)} \\
&= \widetilde{c}_d(x,\sigma^x)e^{-\beta\widetilde{H}_N(\sigma^x)}\,,
\end{aligned}$$

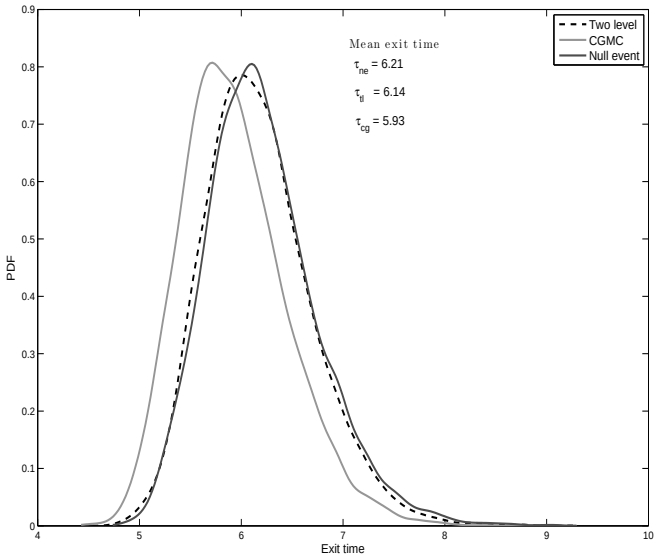

FIGURE 8.5. *Comparing the probability density function of the exit time from the initial coverage* $c_0 = 0$ *to* $c \geq 0.99$. *Potential parameters* $K = 3$, $J = 5$, $h = 2.5$, $L = 100$, *the lattice size* $N = 1024$, *and the coarse-graining parameter* $q = N$.

and similarly

$$\begin{aligned}
\widetilde{c}_d(x,\sigma)e^{-\beta\widetilde{H}_N(\sigma)} &= \left[\bar{c}_d(k,\eta)c_{\mathrm{rf}}^d(x|\eta)\right] e^{-\beta\widetilde{H}_N(\sigma)} \\
&= \left[d_0\eta(k)e^{-\beta\bar{U}_l(k,\eta)}\frac{\sigma(x)}{\eta(k)}e^{-\beta U^{(s)}(x,\sigma)}\right] e^{-\beta(\widetilde{H}_N(\sigma^x)-(2\sigma(x)-1)\widetilde{U}(x,\sigma))} \\
&= \left[d_0\sigma(x)e^{-\beta(\bar{U}_l(k,\eta)+U^{(s)}(x,\sigma))}e^{\beta\widetilde{U}(x,\sigma)}\right] e^{-\beta\widetilde{H}_N(\sigma^x)} \\
&= \left[d_0(1-\sigma^x(x))\right] e^{-\beta\widetilde{H}_N(\sigma^x)} \\
&= \widetilde{c}_a(x,\sigma^x)e^{-\beta\widetilde{H}_N(\sigma^x)}\,.
\end{aligned}$$

Appendix B. Proof of Lemma 5.5. For the sake of completeness we also give the proof of Lemma 5.5 which was proved in [19].

Proof. We denote by $\nabla_\sigma\phi(\sigma) = (\partial_x\phi(\sigma))_{x\in\Lambda}$ and $c(\sigma) = (c(x,\sigma))_{x\in\Lambda}$. The equation (5.11) can be rewritten as

$$\partial_t u(t,\sigma) + c(\sigma)\cdot\nabla_\sigma u(t,\sigma) = 0\,.$$

For the discrete difference $\partial_x u$ we obtain the equation

$$\partial_t\left(\partial_x u(t,\sigma)\right) + c(\sigma)\cdot\nabla_\sigma\partial_x u(t,\sigma) + \partial_x c(\sigma)\cdot\nabla_\sigma\partial_x u(t,\sigma^x) = 0\,.$$

From the definition of the rates $c(x,\sigma)$, (4.1), we can estimate upper bounds for $\partial_x c(\sigma)$. Using that $\partial_x U(z,\sigma) = U(z,\sigma^x) - U(z,\sigma) = K(x-z)(1-2\sigma(x)) + J(x-z)(1-2\sigma(x))$

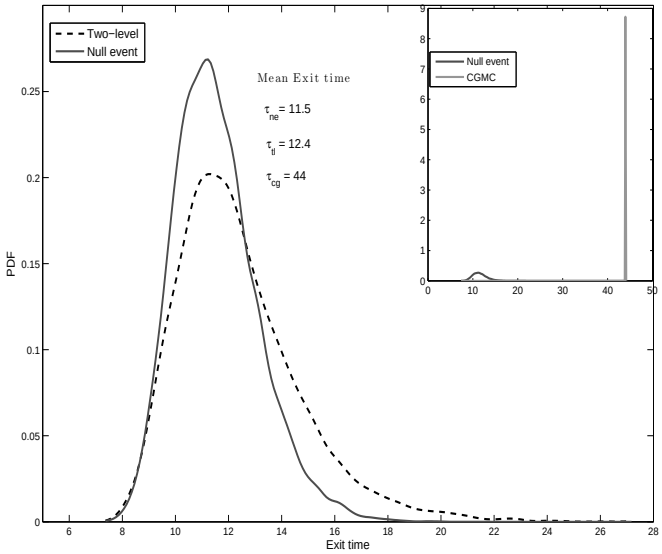

FIGURE 8.6. *Comparing the probability density function of the exit time from the initial coverage* $c_0 = 0$ *to* $c \geq 0.99$. *Potential parameters* $K = 3$, $J = 5$, $h = 3.1$, $L = 100$, *the lattice size* $N = 1024$, *and the coarse-graining parameter* $q = N$. *Coarse-graining error of CGMC that appears due to the finite-range interactions is substantially reduced with the ML-KMC method (see the text for explanation of the inset).*

when $z \neq x$, and $\partial_x U(z,\sigma) = 0$ when $z = x$ we have

$$\partial_x c(z,\sigma) = \begin{cases} \mathcal{O}(1), & \text{for } z = x\,, \\ \mathcal{O}(1), & \text{for } 0 < |z-x| \leq S\,, \\ \mathcal{O}(1/L), & \text{for } S < |z-x| \leq L\,. \end{cases}$$

Then, since $\mathcal{L}v(\sigma) = c(\sigma)\cdot\nabla_\sigma v(\sigma)$, we can write

$$\begin{aligned} &\partial_t\left(\partial_x u(t,\sigma)\right) + \mathcal{L}\partial_x u(t,\sigma) + \sum_{z\in\Lambda}\partial_x c(z,\sigma)\partial_z u(t,\sigma^x) = 0\,, \\ &\partial_t\left(\partial_x u(t,\sigma)\right) + \mathcal{L}\partial_x u(t,\sigma) + \mathcal{O}(1)\partial_x u(t,\sigma^x) \\ &\quad + \mathcal{O}(1)\sum_{|z-x|\leq S}\partial_z u(t,\sigma^x) + \mathcal{O}(\frac{1}{L})\sum_{S<|z-x|\leq L}\partial_z u(t,\sigma^x) = 0\,. \end{aligned}$$

Furthermore, we have

$$\begin{aligned} \|\partial_x u(t,\cdot)\|_\infty \leq \|\partial_x u(0,\cdot)\|_\infty + \int_t^T \mathcal{O}(1)\|\partial_x u(s,\cdot)\|_\infty\, ds + \\ + \int_t^T \mathcal{O}(1)\sum_{|z-x|\leq S}\|\partial_x u(s,\cdot)\|_\infty\, ds + \int_t^T \mathcal{O}(\frac{1}{L})\sum_{S<|z-x|\leq L}\|\partial_x u(s,\cdot)\|_\infty\, ds\,. \quad \text{(B.1)} \end{aligned}$$

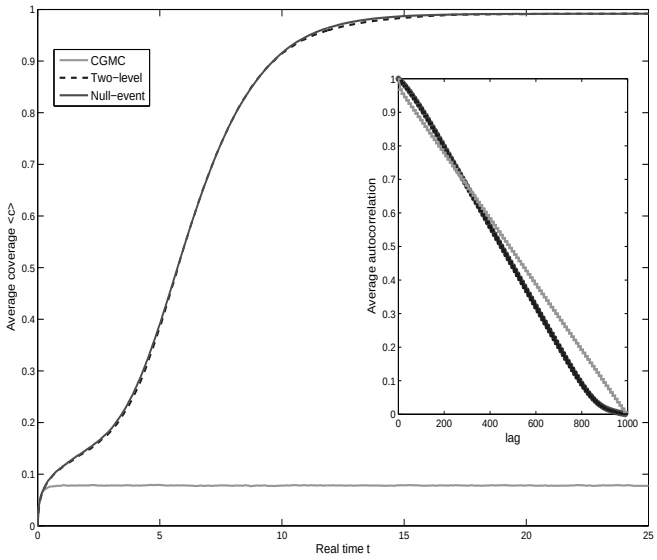

FIGURE 8.7. *Average coverage trajectory. As external field h approaches critical value $h_c = 4$ the coarse-graining error in CGMC becomes important. However, the two-level ML-KMC simulations capture correctly transitions even when CGMC fails. Potential parameters $K = 3$, $J = 5$, $h = 3.1$, $L = N$, the lattice size $N = 1024$, and the coarse-graining parameter $q = N$. The inset depicts the estimated autocorrelation function.*

Based on this relation, application of Gronwall's inequality for $\theta(t) = \sum_x \|\partial_x u(t,\cdot)\|_\infty$ and the fact that S is finite and small we conclude

$$\sum_x \|\partial_x u(t,\cdot)\|_\infty \leq e^{c(T-t)} \sum_x \|\partial_x u(0,\cdot)\|_\infty .$$

□

Appendix C. Kinetic Monte Carlo algorithms. Stochastic simulation algorithm (SSA) as proposed in [11] is described next, where the evolution from a state σ to σ^x is sampled by:

ALGORITHM 2. *Stochastic Simulation Algorithm.*

Step 1: Update. *(a) Calculate all rates $c(y,\sigma), \forall y \in \Lambda_N$ form (4.1), that are affected from the previews event.*
(b) Calculate $\lambda_x(\sigma) = \sum_{y<x} c(y,\sigma)$ and $\lambda(\sigma) = \sum_{y\in\Lambda_N} c(y,\sigma)$

Step 2: Search. *Obtain a uniform random number $u \in [0,1)$ and search for $x \in \Lambda_N$ such that*

$$\lambda_{x-1}(\sigma) < \lambda(\sigma)u \leq \lambda_x(\sigma)$$

Step 3 *Update time from an exponential law with the parameter $\lambda(\sigma)$ or equivalently with the mean $\Delta t = \frac{1}{\lambda(\sigma)}$. That is select a uniform random number $u_1 \in [0,1)$ and update the time as $t' = t + \delta t$ with $\delta t = -\log(u_1)\Delta t$.*

The n-fold way (or BKL) algorithm, [3], is equivalent to the SSA in the sense that it always leads to a successful event and requires the updating of all transition rates. The BKL algorithm was already designed to reduce the cost of the searching process by dividing the transition states into classes (the number of classes $n \ll N$) with the same probability, thus the search algorithm cost depends on the number of processes at each site, i.e., scales linearly with the reaction range (the number of interacting neighbours) in the model under consideration.

ALGORITHM 3. *n-fold algorithm (BKL).*
Given σ

Step 1: Update. *(a) Calculate all rates* $c(y,\sigma), \forall y \in \Lambda_N$ *that are affected by the previous event.*

Step 2: Search. *Group sites* $x \in \Lambda$ *in classes* D_i, $i = 1, \ldots, n$ *by classifying them with their rate values and define*

$$Q_j(\sigma) = \sum_{i=1}^{j} \sum_{y \in D_i} c(y,\sigma) = \sum_{i=1}^{j} |D_i| c(y,\sigma)$$

for some $y \in D_i$, $j = 1, \ldots, n$. *Generate a uniform random number* $u \in [0,1)$ *and search for* $i = 1, \ldots, n$*such that*

$$Q_{i-1}(\sigma) < Q_n(\sigma) u \leq Q_i(\sigma)\,,$$

then choose $x \in D_i$ *uniformly.*

Step 3 *Update time from the exponential law with the parameter* $\lambda(\sigma) = Q_n(\sigma)$, *or equivalently with the mean* $\Delta t = \frac{1}{\lambda(\sigma)}$.

The previous two algorithms are in the class of rejection-free methods as the embeded Markov chain always jumps into a new state. However, by applying the uniformization we obtaine a null-event algorithm in which the embeded chain has nonzero probability to stay at the same state in each step.

ALGORITHM 4. *Null-event algorithm.*
Find the bounds $\lambda^{*,\mathrm{loc}} = d_0 \max\{1, e^{-\beta U^*}\}$, *and* $U^* = \min_{x,\sigma} U(x,\sigma)$.
Given σ

Step 1: Search/Update. *Select a site* $x \in \Lambda$ *with the uniform probability* $\frac{1}{N}$ *and calculate* $c(x,\sigma)$.

Step 2:Accept/Reject. *Obtain a uniform random number* $u \in [0,1)$,
if $c(x,\sigma) \geq \lambda^{*,\mathrm{loc}} u$ *accept and update the state* $\sigma \to \sigma^x$ *at the site* x,
if $c(x,\sigma) < \lambda^{*,loc} u$ *assign the new state to be* σ.

Step 3 *Update time from the exponential law with the parameter* $\lambda^{*,\mathrm{loc}}$, *or equivalently with the mean* $\Delta t = \frac{1}{\lambda^{*,\mathrm{loc}}}$.

REFERENCES

[1] E. ADAM, L. BILLARD, AND F. LANCON, *Class of Monte Carlo algorithms for dynamic problems leads to an adaptive method*, Phys. Rev. E, 59 (1999), pp. 1212–1216.

[2] S. ARE, M. A. KATSOULAKIS, P. PLECHÁČ, AND L. REY-BELLET, *Multibody interactions in coarse-graining schemes for extended systems*, SIAM J. Sci. Comput., 31 (2008), pp. 987–1015.

[3] A. BORTZ, M. KALOS, AND J. LEBOWITZ, *New algorithm for Monte-Carlo simulation of Ising spin systems*, J. Comput. Phys., 17 (1975), pp. 10–18.

[4] P. BUCHHOLZ, *Exact and ordinary lumpability in finite Markov chains*, J. Appl. Probab., 31 (1994), pp. 59–75.

[5] A. Chatterjee and D. G. Vlachos, *An overview of spatial microscopic and accelerated kinetic Monte Carlo methods*, J. Comput.-Aided Mater., 14 (2007), pp. 253–308.
[6] A. Chatterjee and D. G. Vlachos, *Systems tasks in nanotechnology via hierarchical multiscale modeling: Nanopattern formation in heteroepitaxy*, Chemical Engineering Science, 62 (2007), pp. 4852–4863.
[7] J. Chazottes, C. Giardina, and F. Redig, *Relative entropy and waiting times for continuous-time Markov processes*, Electronic Journal of Probability, 11 (2006), pp. 1049–1068.
[8] 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 (2002), p. 8183.
[9] C. Gardiner, *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences*, Springer, 4th ed., 2009.
[10] M. A. Gibson and J. Bruck, *Efficient exact stochastic simulation of chemical systems with many species and many channels*, J. Phys. Chemistry A, 104 (2000), pp. 1876–1889.
[11] D. T. Gillespie, *Exact stochastic simulation of coupled chemical-reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.
[12] E. Kalligiannaki, M. A. Katsoulakis, and P. Plecháč, *Coupled coarse craining and Markov chain Monte Carlo for lattice systems*, Numerical Analysis and Multiscale Computations, Lect. Notes Comput. Sci. Eng., 82 (2011).
[13] E. Kalligiannaki, M. A. Katsoulakis, P. Plecháč, and D. G. Vlachos, *Multilevel coarse graining and nano-pattern discovery in many particle stochastic systems*, Journal of Computational Physics, 231 (2012), pp. 2599–2620.
[14] M. Kardar, *Crossover to equivalent-neighbor multicritical behavior in arbitrary dimensions*, Phys. Rev. B, 28 (1983), pp. 244–246.
[15] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos, *Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems*, J. Comp. Phys., 186 (2003), pp. 250–278.
[16] ———, *Coarse-grained stochastic processes for microscopic lattice systems*, Proc. Natl. Acad. Sci., 100 (2003), pp. 782–787.
[17] M. A. Katsoulakis, P. Plecháč, and L. Rey-Bellet, *Numerical and statistical methods for the coarse-graining of many-particle stochastic systems*, J. Sci. Comput., 37 (2008), pp. 43–71.
[18] M. A. Katsoulakis, P. Plecháč, L. Rey-Bellet, and D. K. Tsagkarogiannis, *Coarse-graining schemes and a posteriori error estimates for stochastic lattice systems*, ESAIM-Math. Model. Numer. Anal., 41 (2007), pp. 627–660.
[19] M. A. Katsoulakis, P. Plecháč, and A. Sopasakis, *Error analysis of coarse-graining for stochastic lattice dynamics*, SIAM J. Numer. Anal., 44 (2006), pp. 2270–2296.
[20] 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., (2008).
[21] ———, *Coarse-graining schemes for stochastic lattice systems with short and long range interactions*, submitted to Math. Comp., (2010).
[22] M. A. Katsoulakis and J. Trashorras, *Information loss in coarse-graining of stochastic particle dynamics*, J. Stat. Phys., 122 (2006), pp. 115–135.
[23] M. A. Katsoulakis and D. G. Vlachos, *Coarse-grained stochastic processes and kinetic monte carlo simulators for the diffusion of interacting particles*, J. Chem. Phys., 119 (2003), pp. 9412–27.
[24] C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Springer, Berlin, 1999.
[25] D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, 2000.
[26] T. M. Liggett, *Interacting particle systems*, Springer, New York Berlin Heidelberg, 1985.
[27] U. Löw, V. J. Emery, K. Fabricius, and S. A. Kivelson, *Study of an Ising model with competing long- and short-range interactions*, Phys. Rev. Lett., 72 (1994), pp. 1918–1921.
[28] A. J. Majda and B. Gershgorin, *Quantifying uncertainty in climate change science through empirical information theory*, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 14958–14963.
[29] R. Plass, J. A. Last, N. C. Bartelt, and G. L. Kellogg, *Nanostructures: Self-assembled domain patterns*, Nature, 412 (2001), pp. 875–875.
[30] S. J. Plimpton, C. C. Battaile, M. Chandross, L. Holm, A. P. Thompson, V. Tikare, G. Wagner, E. Webb, X. Zhou, C. Garcia Cardona, and A. Slepoy, *Crossing the Mesoscale No-Man's Land via Parallel Kinetic Monte Carlo*, tech. rep., Sandia National Lab., October 2009.
[31] T. P. Schulze, *Efficient kinetic Monte Carlo simulation*, J. Comp. Phys., 227 (2008), pp. 2455–2462.

[32] M. Stamatakis and D. G. Vlachos, *A graph-theoretical kinetic Monte Carlo framework for on-lattice chemical kinetics*, J. Chem. Phys., 134 (2011).
[33] A. F. Voter, *Classicaly exact overlayer dynamics - Diffusion of rhodium clusters on RH(100)*, Physical Review B, 34 (1986), pp. 6819–6829.