Journal of Computational Physics 231 (2012) 2599–2620

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Multilevel coarse graining and nano-pattern discovery in many particle stochastic systems

Evangelia Kalligiannaki [a,*], Markos A. Katsoulakis [b,c], Petr Plecháč [a], Dionisios G. Vlachos [d]

[a] Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA
[b] Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA
[c] Department of Applied Mathematics, University of Crete and Foundation of Research and Technology-Hellas, Greece
[d] Department of Chemical Engineering, University of Delaware, Newark, DE 19716, USA

ARTICLE INFO

Article history:
Received 12 September 2011
Received in revised form 8 December 2011
Accepted 12 December 2011
Available online 20 December 2011

Keywords:
Markov chain Monte Carlo
Coarse graining
Lattice systems
Phase transitions
Pattern formation

ABSTRACT

In this work we propose a hierarchy of Markov chain Monte Carlo methods for sampling equilibrium properties of stochastic lattice systems with competing short and long range interactions. Each Monte Carlo step is composed by two or more sub-steps efficiently coupling coarse and finer state spaces. The method can be designed to sample the exact or controlled-error approximations of the target distribution, providing information on levels of different resolutions, as well as at the microscopic level. In both strategies the method achieves significant reduction of the computational cost compared to conventional Markov chain Monte Carlo methods. Applications in phase transition and pattern formation problems confirm the efficiency of the proposed methods.

1. Introduction

Our primary goal in this work is to develop a systematic mathematical and computational strategy for accelerating microscopic simulation methods with competing short and long range interactions, arising in numerous physical systems for instance in micromagnetics, models of epitaxial growth, etc. We propose the Multilevel Coarse Graining Monte Carlo (ML-CGMC) method, based on a hybrid statistical mechanics and statistics approach. The method introduces a *hierarchy of Markov chain Monte Carlo methods coupling scales and types of interactions* that can sample the exact or controlled error approximations of Gibbs measures $\mu_{N,\beta}(d\sigma) = Z_N^{-1} e^{-\beta H_N(\sigma)} P_N(d\sigma)$ defined on a high dimensional configuration space $\Sigma_N = \{0,1\}^{\Lambda_N}$, with Λ_N a d-dimensional lattice with $N \gg 1$ sites, that can be easily generalized to any probability measure with similar properties. It is a method of constructing efficient *proposal measures* in Metropolis sampling using coarse-graining techniques, aiming at reducing the rejection rate and the computational complexity. The key idea is a decomposition of the sampling distribution to a product measure

$$\mu_{N,\beta}(d\sigma) = \bar{\mu}^{(0)}_{M,\beta}(d\eta) \nu_r(d\sigma|\eta), \tag{1}$$

with $\eta := \mathbf{T}\sigma$ a variable with less degrees of freedom compared to σ, defined by a projection map $\mathbf{T} : \Sigma_N \to \overline{\Sigma}_M, M < N$. $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$ is a measure with a simple explicit representation approximating the marginal $\bar{\mu}_{M,\beta}(d\eta) = \mu_{N,\beta} \circ \mathbf{T}^{-1}(d\eta)$ and $\nu_r(d\sigma|\eta)$ is a uniquely defined (prior) measure, responsible for *reconstructing* variables σ given η [18]. Such a two-level measure decomposition can be

* Corresponding author. Tel.: +1 302 831 0588; fax: +1 302 831 4511.
E-mail addresses: ekalligi@math.udel.edu (E. Kalligiannaki), markos@math.umass.edu (M.A. Katsoulakis), plechac@math.udel.edu (P. Plecháč), vlachos@che.udel.edu (D.G. Vlachos).

0021-9991/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2011.12.011

trivially extended to a multi-level setting where (1) can include different resolution levels interpolating between a coarser level and the microscopic one σ.

We describe a Monte Carlo step of a two-level CGMC method:

1. Sample η from $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$, using Coarse Grained (CGMC) samplers [17,16]. Appropriate coarse-grained measures have been evaluated in earlier work via cluster expansions that can be easily constructed with available analytical error estimates, ensuring that such approximations are controllable, [19].
2. Conditioned on such η, obtained in Step 1, we sample $\nu_r(d\sigma|\eta)$ using an accept/reject method.

A schematic description of this procedure is seen in Fig. 1. Better proposals constructed in Step 1 will lead to fewer rejections in Step 2, furthermore, there is no need to consider all possible microscopic proposals since at the coarse step we do a first screening.

In comparison to two-level rejection-free CGMC previously discussed in [4], the approach here provides a rigorous mathematical framework and employs a rejection-based type algorithm that is computationally easier to implement and, in contrast to the general belief, more efficient than rejection-free methods under certain conditions (e.g. long range interactions and stiff problems), [33]. Instead of further approximating the sampling measure, as is done with cluster expansion keeping the typically computationally expensive multi-body higher order terms [19,2], we use the hybrid statistical and statistics approach that construct $\nu_r(d\sigma|\eta)$. Even when CGMC provides less accurate approximations, the ML-CGMC approach can refine the results by the accept/reject step in the finer space.

A necessary ingredient for applicability of the method is a decomposition of the form (1), which includes a possibly less accurate coarse-grained measure and the correcting accept/reject Step 2 above. This formulation can make the proposed method extentable to off-lattice systems where various coarse-graining schemes are already available [25,12], although without controlled-error approximations. In such off-lattice systems we typically have two main features: a presence of short and long interactions, as well as comparable energy and entropy, hence fluctuations are expected to be important in the modelling and simulation.

Systems with smooth long or intermediate range interactions are well approximated by coarse-graining techniques [16,18,23], and CGMC are reliable simulation methods with controlled error approximations, both for observables and loss of information [20,19]. Furthermore, models where only short-range interactions appear are inexpensive to simulate with conventional methods. However, when both short and long-range interactions are present, the conventional methods become prohibitively expensive, and coarse-graining error estimates are not applicable. The proposed method can handle such systems efficiently by either (a) compressing only the long range interactions for Step 1 and sample with CGMC with low computational cost, including the short range part at the accept/reject Step 2 (potential splitting), or (b) compress all types of interactions for Step 1, and correct appropriately in Step 2 (corrections).

A wide literature exists on sophisticated Markov chain Monte Carlo (MCMC) methods designed to accelerate simulations for large systems, applying for example parallel techniques and/or constructing good first approximations (proposals) in Metropolis sampling [9,31]. In [10] Efendiev et al., the preconditioning MCMC is proposed, a two stage Metropolis method, applied to inverse problems of subsurface characterization. Our algorithm shares the same idea of constructing a proposal density based on meso or macroscopic properties of the model studied and taking advantage of the first stage rejections. Several methods where the trial density is built up sequentially with stage-wise rejection decision appear [3,27]. There are also some similarities with simulated sintering, and transdimensional MCMC, see [28,27] and references therein. However, the novelty of our method lies in the construction of the variable dimensionality (and level of coarse-graining)

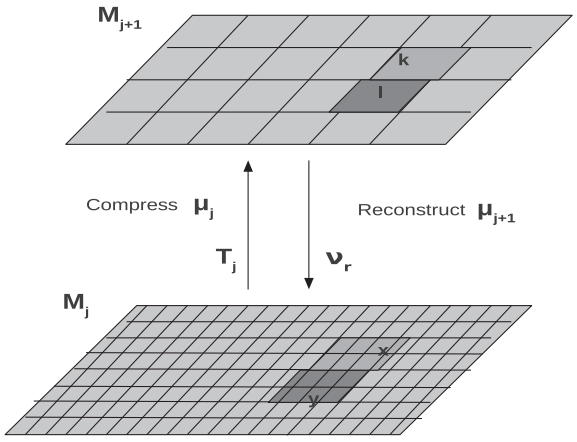

Fig. 1. Schematic of a two-level ML-CGMC. Information exchange between coarser and finer resolutions. Step 1: Sample $\mu_{j+1}(d\eta_{j+1})$, Step 2: Reconstruct with $\nu_r(d\eta_j|\eta_{j+1})$ such that $\mu_j(d\eta_j) = \mu_{j+1}(d\eta_{j+1})\nu_r(d\eta_j|\eta_{j+1})$.

state spaces and the corresponding Gibbs measures relies on statistical mechanics tools that allow a systematic control of the error from one level of coarse-graining to the next. The interplay of different levels of compressed spaces appears also in spatial multigrid methods coupled with CGMC sampling, studied in [6,4], for accelerating lattice kinetic Monte Carlo simulations where, however, the proposed methods are not exact and do not necessarily provide controlled error approximations. Various attempts appear on parallelising Monte Carlo simulations, based on a parallel resolution, such as parallel kinetic CGMC [1], and a combination of CGMC and parallel tempering [36]. In a follow up work we extend our framework for ML-CGMC to kinetic Monte Carlo methods in order to accelerate sampling of evolution processes on large lattice systems, based on the knowledge of explicit coarse level approximating processes [17,16].

In Section 2 we introduce microscopic lattice systems and provide a brief review of coarse graining methods. We also provide the Metropolis–Hastings (MH) MCMC method for numerical simulations and describe microscopic processes. We introduce the ML-CGMC method in Section 3, describing in detail the two-level method, and provide the mathematical analysis that ensures its theoretical validity. Section 4 provides a full comparison of the computational complexity between the conventional MH and the two-level CGMC MH introduced here. More specifically, in Theorem 2 we prove comparative estimates on their mixing times using spectral gap estimates. In turn such spectral estimates are obtained through suitable relative bounds on their respective Dirichlet forms. Concluding, our analysis shows that ML-CGMC has substantial savings over the conventional MH algorithm, generating low cost proposals with small or controllable rejection rates. Sections 5 and 6 give example applications of the two-level CGMC method in canonical and microcanonical sampling. In Section 5.1 a benchmark example is employed in order to demonstrate an explicit application of Theorem 2. In Section 5.2 an order one improvement of the coarse graining error in a phase transition regime is achieved when applying the ML-CGMC method, for a Kac type potential with algebraic decay. Finally in Section 6 we study nanopattern formation in surface diffusion induced by a Morse type potential, and verify that the proposed method can provide correctly microscopic details.

2. Stochastic lattice systems at equilibrium

We consider an Ising-type system on a periodic d-dimensional lattice Λ_N with $N = n^d$ lattice sites. At each site $x \in \Lambda_N$ we define a spin variable $\sigma(x)$ taking values in a finite set. For instance, in a lattice gas model $\sigma(x) \in \{0,1\}$ describes that the site x is vacant or occupied by an atom. The state of the system is described by a configuration $\sigma \in \Sigma_N = \{0,1\}^{\Lambda_N}$, $\sigma = \{\sigma(x) : x \in \Lambda_N\}$. The interaction energy of the system, e.g. interacting particles in the lattice gas model, is defined by the Hamiltonian H_N. We assume systems where the particles interact only through a pair-wise potential and thus the Hamiltonian takes the form

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

where h is an external field. Equilibrium states at the inverse temperature β are described by the (canonical) Gibbs probability measure on the space Σ_N

$$\mu_{N,\beta}(d\sigma) = Z_N^{-1} e^{-\beta H_N(\sigma)} P_N(d\sigma), \tag{3}$$

where Z_N is the normalizing constant (partition function). Furthermore, the product Bernoulli distribution $P_N(d\sigma)$, is the *prior distribution* on Λ_N representing distribution of states in a non-interacting system, or equivalently at $\beta = 0$, when thermal fluctuations-disorder-associated with the product structure of $P_N(d\sigma)$ dominates. By contrast at zero temperature, $\beta = \infty$, interactions and hence order, prevail. Finite temperatures, $0 < \beta < \infty$, describe intermediate states, including possible phase transitions between ordered and disordered states. Specifically, $P_N(d\sigma) = \prod_{x\in\Lambda_N}\rho(d\sigma(x))$ where $\rho(\sigma(x) = 0) = 1/2$, $\rho(\sigma(x) = 1) = 1/2$ is the distribution of a Bernoulli random variable for each $x \in \Lambda_N$. In principle $d\sigma$ denotes the variable increments when the configuration space is a continuous. Since the configuration space Σ_N is discrete, $d\sigma$ is a conventional notation that is used for notation simplicity, denoting

$$\int_{\Sigma_N} f(\sigma)P_N(d\sigma) = \sum_{\sigma_i\in\Sigma_N} f(\sigma_i)\prod_{x\in\Lambda_N}\left(\frac{1}{2}\delta(\sigma(x)) + \frac{1}{2}\delta(\sigma(x)-1)\right),$$

for any function f, where δ denotes the Kronecker delta function.

The coarse-graining techniques have been developed in order to study the behaviour in the regimes when the size of the system $N \to \infty$. In the series of papers [16,17,24] the authors initiated the development of *coarse-graining* (CG) as a computational tool for accelerating Monte Carlo simulations of stochastic lattice dynamics. The coarse-grained model is constructed on a coarse grid $\overline{\Lambda}_M$ by dividing Λ_N into M coarse cells, each of which contains Q (micro-)cells, typically $Q = q^d$ with the coarse-graining scale q in each dimension. Each coarse cell is denoted by C_k, $k \in \overline{\Lambda}_M$. A typical choice for the coarse variable in the context of Ising-type models is the block-spin over each coarse cell C_k,

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

defining the coarse graining map $\mathbf{T}:\Sigma_N\to\overline{\Sigma}_M,\ \mathbf{T}\sigma=\eta$. The exact coarse-grained Gibbs measure is given (with a slight abuse of notation) by $\bar{\mu}_{M,\beta}=\mu_{N,\beta}\circ\mathbf{T}^{-1}$, written in a more convenient form

$$\bar{\mu}_{M,\beta}(d\eta)=\frac{1}{\overline{Z}_M}e^{-\beta\overline{H}_M(\eta)}\overline{P}_M(d\eta), \tag{4}$$

where $\overline{P}_M=P_N\circ\mathbf{T}^{-1}$ defines the coarse-grained prior measure. The *exact coarse-grained Hamiltonian* is defined by the *renormalization group map*, see, e.g. [11],

$$e^{-\beta\overline{H}_M(\eta)}=\mathbb{E}[e^{-\beta H_N}|\eta]=\int e^{-\beta H_N(\sigma)}P_N(d\sigma|\eta). \tag{5}$$

The conditional prior $P_N(d\sigma|\eta)$ is the probability of having a microscopic configuration σ, given a coarse configuration η.

Although typically $\overline{P}_M(d\eta)$ is easy to calculate the exact computation of the coarse-grained Hamiltonian $\overline{H}_M(\eta)$ given by (5) is, in general, impossible even for moderately small values of N. Therefore suitable approximations have to be constructed. An initial approximation can be then proposed by the technique of *cluster expansions* [35] providing improved approximation at suitable phase regimes. The corresponding first order CG Hamiltonian is explicitly given [19],

$$\overline{H}^{(0)}(\eta)=-\frac{1}{2}\sum_{k\in\overline{\Lambda}_M}\sum_{l\neq k}\bar{J}(k,l)\eta(k)\eta(l)-\frac{1}{2}\bar{J}(0,0)\sum_{k\in\overline{\Lambda}_M}\eta(k)(\eta(k)-1)+\sum_{k\in\overline{\Lambda}_M}\bar{h}\eta(k), \tag{6}$$

where the coarse-grained interactions are evaluated explicitly by averaging over the cells $k,l\in\overline{\Lambda}_M$

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

defining the coarse grained Gibbs measure

$$\bar{\mu}^{(0)}_{M,\beta}(d\eta)=\frac{1}{\overline{Z}^{(0)}}e^{-\beta\overline{H}^{(0)}(\eta)}\overline{P}_M(d\eta). \tag{7}$$

The coarse-graining of systems with purely long- or intermediate-range interactions was studied using cluster expansions in [19,2,18,21]. In many applications the long-range potentials exhibit scaling property

$$J(x-y)=L^{-d}V\Big(\frac{n}{L}|x-y|\Big),x,y\in\Lambda_N, \tag{8}$$

where $V\in C^1([0,\infty))$ and it is normalized to ensure that the strength of the potential J is essentially independent of L, i.e., $\sum_{x\neq 0}J(x)\sim\int_0^\infty V(r)dr$. The constant L can be interpreted as a (characteristic) interaction range of the potential. For example, if we have V with properties $V(r)=V(-r)$, $V(r)=0$, $|r|>1$, then a spin at the site x interacts with its neighbours which are at most L lattice points away from x. One of the results therein is on deriving error estimates in terms of the specific relative entropy $\mathcal{R}(\mu|\nu):=N^{-1}\sum_\sigma\log\{\mu(\sigma)/\nu(\sigma)\}\mu(\sigma)$ between the corresponding equilibrium Gibbs measures. Note that 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*,

$$\mathcal{R}\Big(\bar{\mu}^{(0)}_{M,\beta}|\mu_{N,\beta}\circ\mathbf{T}^{-1}\Big)=\mathcal{O}(\epsilon^2),\quad \epsilon\equiv\beta\|\nabla V\|_1\Big(\frac{q}{L}\Big). \tag{9}$$

Systems with short and long range interactions. One of our goals in this work is to study systems where in addition to the long-range potential we have a short range

$$K(x-y)=S^{-d}V_s\Big(\frac{n}{S}|x-y|\Big),\quad x,y\in\Lambda_N, \tag{10}$$

with V_s having similar properties as V in (8) and $S\ll L$ distinguishing the short and long range nature of interactions. A typical case of a short-range potential is encountered in the nearest-neighbour Ising model where $K(x-y)=K=constant$, for $|x-y|=S=1$ and zero otherwise. The new Hamiltonian including the contributing energy from the short range potential $H_s(\sigma)$ and the long range part $H_l(\sigma)$ is

$$H_N(\sigma)=H_s(\sigma)+H_l(\sigma). \tag{11}$$

Study of equilibrium properties of stochastic lattice systems mainly accounts for the evaluation of averages over the coarse-grained $\bar{\mu}_{M,\beta}(d\eta)$ or the microscopic $\mu_{N,\beta}(d\sigma)$ Gibbs measures of observables $\phi:\Sigma_N\to\mathbb{R}$, i.e.,

$$\mathbb{E}[\phi]=\int_{\Sigma_N}\phi(\sigma)\mu_{N,\beta}(d\sigma).$$

Numerical methods evaluating equilibrium averages are the Markov chain Monte Carlo methods, among which the most widely used is the Metropolis–Hastings (MH) method [29,13]. The Metropolis–Hastings algorithm generates proposals σ', for the evolution from the configuration σ to σ', that are defined by the *proposal* probability transition kernel $\rho(\sigma',\sigma)$. The

proposal σ' is accepted with probability $\alpha(\sigma,\sigma')$ or rejected with the probability $1-\alpha(\sigma,\sigma')$. Let $X_0=\sigma_0$ be an arbitrary initial configuration, the nth iteration of the algorithm consists of the following steps

Algorithm 1 (*Metropolis–Hastings algorithm*). Given $X_n=\sigma$

Step 1 Generate $X_n'=\sigma'\sim\rho(\sigma',\sigma)$.
Step 2 Accept–Reject

$$X_{n+1}=\begin{cases} X_n'=\sigma' & \text{with probability } \alpha(\sigma,\sigma'),\\ X_n=\sigma & \text{with probability } 1-\alpha(\sigma,\sigma'),\end{cases}$$

with the acceptance probability depending on the energy difference between configurations σ and σ', $\Delta H_N(\sigma,\sigma')=H_N(\sigma')-H_N(\sigma)$,

$$\alpha(\sigma,\sigma')=\min\left\{1,\exp\{-\beta\Delta H_N(\sigma,\sigma)\}\frac{\rho(\sigma',\sigma)}{\rho(\sigma,\sigma')}\right\}.$$

The algorithm generates an ergodic Markov chain $\{X_n\}$ in the state space Σ_N, with the stationary measure $\mu_{N,\beta}(d\sigma)$. Ergodicity ensures the convergence of empirical averages $\frac{1}{n}\sum_{i=1}^n\phi(X_i)$ to the desired mean $\mathbb{E}[\phi]$, for any $\phi\in L^1(\mu_{N,\beta})$. It is easy to deduce the probability transition kernel associated to MH Algorithm 1

$$\begin{aligned}&\mathcal{K}_c(\sigma,\sigma')=\alpha(\sigma,\sigma')\rho(\sigma,\sigma'),\ \text{for } \sigma'\neq\sigma,\\ &\mathcal{K}_c(\sigma,\sigma)=1-\int_\Sigma\alpha(\sigma,\sigma')\rho(\sigma,\sigma')d\sigma'.\end{aligned}\tag{12}$$

Depending on whether one considers microcanonical or canonical ensemble, the ergodic Markov chain $\{X_n\}$ can be defined by spin-exchange dynamics that preserve the order parameter or spin-flip dynamics, respectively [26,7]. In the spin-exchange the proposed new configuration $\sigma'=\sigma^{(x,y)}$ is obtained from σ by interchanging the spins at x and y, for nearest-neighbour sites x and y,

$$\sigma^{(x,y)}(z)=\begin{cases}\sigma(y), & \text{when } z=x,\\ \sigma(x), & \text{when } z=y,\\ \sigma(z), & \text{otherwise.}\end{cases}$$

Analogously for the spin-flip $\sigma'=\sigma^{(x)}$ is obtained from σ by flipping the spin value at site x,

$$\sigma^{(x)}(z)=\begin{cases}1-\sigma(x), & \text{when } z=x,\\ \sigma(z), & \text{otherwise.}\end{cases}$$

3. The ML-CGMC Metropolis–Hastings method

We present in detail and generalize the Coupled CGMC Metropolis–Hastings method originally proposed in [14]. We introduce the ML-CGMC, a multi-level Metropolis–Hastings method, where each level corresponds to a configuration space resolution level, and provide the associated mathematical analysis that ensures the theoretical validity of the method.

Let $\{\overline{\Sigma}_{M_j}\}_{j=0}^l$ denote a hierarchy of coarse spaces derived from the microscopic space $\Sigma_N=:\overline{\Sigma}_{M_0}$ by a family of mappings $\mathbf{T}_j:\overline{\Sigma}_{M_j}\to\overline{\Sigma}_{M_{j+1}}$, $\mathbf{T}_j\eta_j=\eta_{j+1}$, for $N=M_0>M_1>\cdots>M_l$. The variables η_j denote configurations on spaces $\overline{\Sigma}_{M_j}$, with $\eta_0=:\sigma$ referring to the microscopic variable on Σ_N. The method is composed by a sequence of $l+1$ Metropolis–Hastings steps each one designed to generate samples from $\overline{\Sigma}_{M_j}$ given a coarser sample from $\overline{\Sigma}_{M_{j+1}}$. Properly constructed measures on $\overline{\Sigma}_{M_j}$, $j>0$ form the basis for constructing efficient proposal kernels for Metropolis–Hastings algorithm allowing sampling of large systems. The interplay between different resolution spaces is controlled by the hierarchy of projection mappings $\mathbf{T}_j$ and corresponding inverse mapping procedures (reconstruction). Appropriate reconstruction measures $\nu_{r,j}(d\eta_j|\eta_{j+1})$ are constructed in view of decompositions similar to (13), for each pair of Gibbs measures $\mu_j(d\eta_j)$ and $\mu_{j+1}(d\eta_{j+1})$, see Fig. 1. For the sake of exposition we describe in detail the two-level CGMC method, in which two different scale state spaces $\overline{\Sigma}_{M_{j_1}}=\Sigma_N$ and $\overline{\Sigma}_{M_{j_2}}=\Sigma_M$, $M<N$ are involved. A situation where a ML-CGMC method with three or more levels can be useful is for sampling systems of large size $N\gg1$ with long-range interactions of $O(N)$ that exhibit complex behaviour which cannot be captured when over-coarsening. With a multi-level method one could hierarchically improve the effective approximating potential. For two levels, the proposals in the CG space are uniformly distributed. However, the introduction of a third coarse level can construct better proposal configurations that will increase the acceptance rate of the second level. Consider two coarsening levels Q_1 and Q_2 s.t. $1<Q_1<Q_2$ and $Q_2M_1=Q_2M_2=N$. A large coarsening parameter Q_2 may miss fine details of the potential but still captures information about it, see Fig. 6. A second refinement Q_1 can correct the prediction, in other words we discover all the information hierarchically.

The ML-CGMC method is a hybrid statistical mechanics and statistics approach based on a product measure decomposition of the target distribution

$$\mu_{N,\beta}(d\sigma) = \bar{\mu}^{(0)}_{M,\beta}(d\eta)\nu_r(d\sigma|\eta). \tag{13}$$

The principal idea is that computationally inexpensive CG simulations will reproduce the large scale structure and subsequently microscopic information will be added through the *microscopic reconstruction*. The two levels of the method are described by:

1. *Marginal approximation and CGMC sampling.* $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$, given in (7), is an approximation of marginal $\bar{\mu}_{M,\beta} = \mu_{N,\beta}\circ\mathbf{T}^{-1}$ that is described explicitly with a controllable error [19].
2. *Microscopic reconstruction.* Conditioned on η, obtained by sampling from $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$, we sample $\nu_r(d\sigma|\eta)$ via a simple inverse-mapping distribution $\mu_r(d\sigma|\eta)$ and an accept/reject method. Measure $\nu_r(d\sigma|\eta)$ is responsible for the *reconstruction* procedure of the fine configuration σ constrained on the coarse configuration η.

More specifically, reconstruction is the reverse procedure to coarse-graining, i.e., reproducing *microscopic* properties directly from CGMC simulations. Reconstruction in off-lattice systems e.g. the calculation of diffusion of penetrants through polymer melts, is based in local energy minimization methods [30,37]. These methods suggest reconstruction measures, i.e. $\mu_r(\sigma|\eta)$, showing our method can be generalized to off-lattice systems. A detailed discussion on reconstruction can be found in [18,14]. Note that $\nu_r(d\sigma|\eta)$ is a finite measure, uniquely defined by (13) given $\mu_{N,\beta}(d\sigma)$ and $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$.

Remark. Dependent on the coarse graining mapping $\mathbf{T}$, the complexity of sampling the optimal reconstruction operator may be computationally formidable or even infeasible, suggesting that the proposed method may not be advantageous over a conventional method or even not applicable. On the other hand, our proposed method does not, in principle, require an accurate reconstruction step, since the accept/reject move corrects any irrelevant proposals related to that step, as long as the detailed balance condition is satisfied.

In view of (13) we propose the two-level CGMC algorithm composed of two Metropolis–Hastings steps. The first step samples the measure $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$, on the coarse space $\overline{\Sigma}_M$, using an arbitrary proposal transition kernel $\overline{\rho}(\eta,\eta')$ to produce coarse trial samples η'. The second step, *performed only if the coarse trial sample is accepted*, consists of the simple reconstruction from the coarse state η' with $\mu_r(d\sigma|\eta)$ and an accept/reject method. If a trial coarse sample is rejected the second step is not performed and no computational time is wasted on checking fine trial samples that are most likely to be rejected. The two-level CGMC algorithm is defined as follows: Let $Y_0 = \sigma_0$ be an arbitrary initial configuration, for $n = 0,1,2,\ldots,$

Algorithm 2. Two-level CGMC MH Algorithm

Given $Y_n = \sigma$

Step 1 Compute the coarse configuration $\eta = \mathbf{T}\sigma$.

Step 2 Generate a coarse sample $\eta' \sim \bar{\rho}(\eta,\eta')$.

Step 3 Coarse Level Accept–Reject.

Accept η' with probability:

$$\alpha_{CG}(\eta,\eta') = \min\left\{1, e^{-\beta\Delta\overline{H}^{(0)}(\eta,\eta')}\frac{\bar{\rho}(\eta',\eta)}{\bar{\rho}(\eta,\eta')}\right\}. \tag{14}$$

if η' is accepted then proceed to Step 4,

else generate a new coarse sample, Step 2.

Step 4 Reconstruct σ' given the coarse trial η',

$$\sigma' \sim \mu_r(\cdot|\eta').$$

Step 5 Fine Level Accept–Reject.Accept σ' with probability

$$\alpha_f(\sigma,\sigma') = \min\left\{1, e^{-\beta\left[\Delta H_N(\sigma,\sigma') - \Delta\overline{H}^{(0)}(\eta,\eta')\right]}\frac{\mu_r(\sigma|\eta)}{\mu_r(\sigma'|\eta')}\right\}. \tag{15}$$

where $\Delta H_N(\sigma,\sigma') = H_N(\sigma') - H_N(\sigma)$, and $\Delta\overline{H}^{(0)}(\eta,\eta') = \overline{H}^{(0)}(\eta') - \overline{H}^{(0)}(\eta)$.

In terms of the MH Algorithm 1, the method is generating trial samples σ' from proposal kernel $\rho(\sigma,\sigma')$, with stationary measure $\bar{\mu}^{(0)}_{M,\beta}(d\eta)\mu_r(d\sigma|\eta)$, that depends on the statistical mechanics properties of the system. This fact leads to an increase of the acceptance rate of the MH method, see for example Fig. 4. Indeed, consider the canonical ensemble for $x \in \Lambda_N$ such that $x \in C_k,\ k \in \overline{\Lambda}_M$ with a potential $J(x-y)$ (8), where the following estimate holds [17],

$$\Delta H_N(\sigma,\sigma^{(x)}) - \Delta\overline{H}^{(0)}(\eta,\eta^{(k)}) = \mathcal{O}(Q/(2L+1)^d),$$

with $\eta^{(k)} = \mathbf{T}\sigma^{(x)}$. This estimate shows that the second level acceptance probability $\alpha_f(\sigma,\sigma')$ is controlled by the coarsening parameter, i.e. it is close to 1 for $Q/(2L+1)^d \ll 1$ and most of the coarse samples entering the second level will be accepted. Thus, for fixed L, varying the coarsening parameter we can control the effectiveness of the method as a balance between acceptance rate and computational cost. That is between small values of Q, leading to high acceptance rate, and larger Q, leading to lower computational cost at sampling at the coarse space. We elaborate with this issue in detail in Section 4.

3.1. Examples of multi-level decompositions

In this section we elaborate on how to select the prior measure $\nu_r(d\sigma|\eta)$, that furthermore determines how to perform a multi-level decomposition of the Gibbs measure. The selection of $\nu_r(d\sigma|\eta)$ depends on two features (a) the choice of the coarse grained measure $\bar{\mu}^{(0)}_{M,\beta}(d\eta)$, that is whether we compress the entire interaction potential or split to a short and long range part and (b) on sampling the exact or an approximation of the microscopic measure $\mu_{N,\beta}(d\sigma)$, that suggest the following strategies:

1. *Exact Corrections.* With this strategy all types of interactions are compressed and incorporated into the first level of sampling with $\overline{H}^{(0)}(\eta)$, as is depicted in the algorithmic description of the method Algorithm 2. The reconstructive and corrective characteristics of $\nu_r(d\sigma|\eta)$ appear explicitly in the following formulation, that is reconstruct with the uniform conditional distribution $\mu_r(d\sigma|\eta) = P_N(d\sigma|\eta)$ and correct with $H_N(\sigma) - \overline{H}^{(0)}(\eta)$,

$$\nu_r(d\sigma|\eta) = \overline{Z}^{(0)}_M Z_N^{-1} \exp\{-\beta[H_N(\sigma) - \overline{H}^{(0)}(\eta)]\} P_N(d\sigma|\eta).$$

2. *Potential splitting.* An alternative approach is considered based on splitting the inter particle interactions into *short-* and *long*-range terms. A decomposition of the coarse-graining of the interaction potential can be justified and optimized by known error estimates, see [2]. These estimates suggest a natural way to split the potential into a short-range piece $K(x-y)$ with possible singularities and a locally integrable (or smooth) long-range decaying component $J(x-y)$. Here we suggest to sample on the coarse step according to the effective Hamiltonian $\overline{H}^{(0)}_l(\eta)$, corresponding only to the long-range depended energy $H_l(\sigma)$, as in (11), that suggests a rearrangement of (13) where

$$\nu_r(d\sigma|\eta) = \overline{Z}^{(0)}_{l_M} Z_N^{-1} \exp\left\{-\beta[H_s(\sigma) + H_l(\sigma) - \overline{H}^{(0)}_l(\eta)]\right\} P_N(d\sigma|\eta).$$

In the two-level CGMC sampling the costly long-range part is involved only in the coarse updating where the number of operations to calculate energy differences is reduced and coarsening of the short-range potential is avoided [22].

3. *Approximate CG.* In many applications where meso/macroscopic information is sufficient CGMC sampling is reliable for long-range potentials and the error when neglecting terms $\Delta H_l(\sigma,\sigma') - \Delta\overline{H}^{(0)}_l(\eta,\eta')$ is small. In this strategy we suggest to neglect these terms, despite the encountering of the approximating error, benefiting from a further reduction of the computational complexity, see Section 4. As a result Algorithm 2 is sampling from a probability measure approximating $\mu_{N,\beta}(d\sigma)$,

$$\mu^{(0)}_{N,\beta}(d\sigma) \propto \exp\left\{-\beta H_s(\sigma) - \beta H^{(0)}_l(\eta)\right\} P_N(d\sigma|\eta)\bar{\mu}^{(0)}_{M,\beta}(d\eta).$$

3.2. Reversibility of the ML-CGMC

Mathematical analysis for the two-level CGMC method is carried out in this section for a broad class of probability measures. The method can be generalized to the sampling of any probability measure $\mu(d\sigma)$ on a countable configuration space Σ, by coupling properly configurations between a hierarchy of coarser and finer spaces. Since almost any probability density $\mu(\sigma)$ can be written in the form $e^{-H(\sigma)}$ the method can applied to any model for which one can properly define a function $H(\sigma)$ and the hierarchy of coarse spaces and densities. The following arguments are straightforward but necessary to prove that the algorithm samples processes with the desired stationary measure. Algorithm 2 is defined by the acceptance probabilities

$$\alpha_{CG}(\eta,\eta') = \min\left\{1, \frac{\bar{\mu}^{(0)}(\eta')\bar{\rho}(\eta',\eta)}{\bar{\mu}^{(0)}(\eta)\bar{\rho}(\eta,\eta')}\right\} \tag{16}$$

and

$$\alpha_f(\sigma,\sigma') = \min\left\{1, \frac{\mu(\sigma')\bar{\mu}^{(0)}(\eta)\mu_r(\sigma|\eta)}{\mu(\sigma)\bar{\mu}^{(0)}(\eta')\mu_r(\sigma'|\eta')}\right\}, \tag{17}$$

where with a slight abuse of notation we denote the probability density of measures with the same letter. Note that in the sequel $\eta = \mathbf{T}\sigma$, $\eta' = \mathbf{T}\sigma'$, if not otherwise stated. A trial state σ' is generated by the first level and the simple reconstruction with the transition probability

$$Q(\sigma,\sigma') = \begin{cases} \alpha_{CG}(\eta,\eta')\bar{\rho}(\eta,\eta')\mu_r(\sigma'|\eta'), & \text{for } \sigma \neq \sigma', \\ 1 - \int \alpha_{CG}(\eta,\eta')\bar{\rho}(\eta,\eta')\mu_r(\sigma'|\eta')d\eta', & \text{for } \sigma = \sigma'. \end{cases}$$

It is easy to check that $Q(\sigma,\sigma')$ satisfies the detailed balance condition with $\mu_0(d\sigma) = \bar{\mu}^{(0)}(d\eta)\mu_r(d\sigma|\eta)$, i.e. $Q(\sigma,\sigma')\mu_0(\sigma) = Q(\sigma',\sigma)\mu_0(\sigma')$, see Definition 2. This property led to the simplified formulation of the second level acceptance probability (17) from the one suggested by the Metropolis method $\alpha_f(\sigma,\sigma') = \min\left\{1, \frac{\mu(\sigma')Q(\sigma',\sigma)}{\mu(\sigma)Q(\sigma,\sigma')}\right\}$. The probability of moving from a state σ to a next $\sigma'$is $\alpha_f(\sigma,\sigma')\alpha_{CG}(\eta,\eta')\mu_r(\sigma'|\eta')\bar{\rho}(\eta,\eta')$. Therefore the method generates a Markov chain $\{Y_n\}$, starting from an arbitrary initial state σ_0, with transition kernel

$$\begin{aligned} \mathcal{K}_{tl}(\sigma,\sigma') &= \alpha_f(\sigma,\sigma')\alpha_{CG}(\eta,\eta')\mu_r(\sigma'|\eta')\bar{\rho}(\eta,\eta') \quad \text{for } \sigma \neq \sigma', \\ \mathcal{K}_{tl}(\sigma,\sigma) &= 1 - \int_{\Sigma} \alpha_f(\sigma,\sigma')\alpha_{CG}(\eta,\eta')\mu_r(\sigma'|\eta')\bar{\rho}(\eta,\eta')d\sigma' \end{aligned} \tag{18}$$

With the following Theorem, the proof of which is given in Appendix A, we prove that transition kernel $\mathcal{K}_{tl}(\sigma,\sigma')$ satisfies the detailed balance condition, that ensures the method generates samples from the target measure. Furthermore, irreducibility and aperiodicity properties are satisfied that guarantee ergodicity of $\{Y_n\}$, i.e. $\frac{1}{n}\sum_{j=1}^{n}\phi(Y_j)$ is a convergent approximation of the averages $\int \phi(\sigma)\mu(d\sigma)$ for any $\phi \in L^1(\mu)$.

We denote $E = \{\sigma \in \Sigma : \mu(\sigma) > 0\}$ and $\widetilde{E} = \{\sigma \in \Sigma : \mu_0(\sigma) > 0\}$ the support of the microscopic and the proposal distributions respectively.

Remark. Note that when the (reconstructed) proposed microscopic state σ' is equal to σ, the current state, there is no need of testing the second level acceptance since $\alpha_f(\sigma,\sigma) = 1$. This suggests that if the coarse proposal is rejected, i.e., $\alpha_{CG}(\eta,\eta') \ll 1$ then the probability of rejecting a σ' reconstructed with $\mu_r(\sigma'|\eta')$, is $\alpha_{CG}(\eta,\eta')\,\mu_r(\sigma'|\eta') \ll 1$ and σ' is most likely to be rejected. For this reason we implemented the method described in Algorithm 2 where the rejection decision depends only on $\alpha_{CG}(\eta,\eta')$.

Theorem 1. *For every conditional distribution $\bar{\rho}(\eta,\eta')$on $\overline{\Sigma}$, and $\mu_r(\cdot|\eta)$ on $\{\sigma \in \Sigma : \mathbf{T}\sigma = \eta\}$,*

(i) *The transition kernel $\mathcal{K}_{tl}(\sigma,\sigma')$ satisfies the detailed balance (DB) condition with $\mu(\sigma)$.*
(ii) *$\mu(\sigma)$ is a stationary distribution of the chain.*
(iii) *If $\bar{\rho}(\eta,\eta') > 0, \mu_r(\sigma|\eta) > 0$ for all σ, $\sigma' \in E$ and $E \subset \widetilde{E}$ holds, then $\{Y_n\}$ is μ-irreducible.*
(iv) *$\{Y_n\}$ is aperiodic.*

4. Computational complexity of ML-CGMC

The effectiveness of the ML-CGMC method is a result of the synthesis of the following two arguments. Firstly, the computational cost of a conventional MH method is reduced by a two-level CGMC method. However this is not enough to prove that the method can indeed accelerate conventional methods, and a mathematical spectral analysis emerges as a key algorithmic need in order to compare the speed of convergence of two-level CGMC and a conventional algorithm. As we will see next, Theorem 2 provides the relation of the two methods equilibration times using spectral arguments.

4.1. Computational complexity

An abstract comparison of the computational complexity of a conventional MH and the two-level CGMC is summarized in Table 1, for sampling a Gibbs measure with Hamiltonian (11). By computational complexity here we mean the cost of calculating energy differences involved at the acceptance probabilities. The following analysis is based on the approximate CG strategy with potential splitting as described in Section 3.1. Let us consider the canonical ensemble where energy difference is

$$\Delta H_N(\sigma,\sigma^{(x)}) = (2\sigma(x) - 1)\sum_{y\in\Lambda_N, y\neq x}[K(x-y) + J(x-y)]\sigma(y).$$

For potential $J(x-y)$ with interaction range L each particle interacts with a number of $(2L+1)^d$ neighbours and similarly for $K(x-y)$ with range $S < L$. Therefore the number of operations necessary is $(2L+1)^d + (2S+1)^d$. Similarly the energy differences appearing in the two-level method are

Table 1
Operations count for evaluating energy differences for n MC iterations. The total number of accepted coarse trials $m < n$ is the number of the second level iterations tested.

Method	Operations
Metropolis Hastings	$n \times (2L+1)^d + n \times (2S+1)^d$
Two-level CGMC	$n \times (2L+1)^d/Q + m \times (2S+1)^d$

$$\Delta\overline{H}_l^{(0)}(\eta,\eta^{(k)}) = \sum_{l\in\overline{\Lambda}_M, k\neq l} \bar{J}(k,l)\eta(l) + \bar{J}(0,0)(\eta(k)-1),$$
$$\Delta H_s(\sigma,\sigma^{(x)}) = (2\sigma(x)-1)\sum_{y\in\Lambda_N, y\neq x} K(x-y)\sigma(y),$$

at the first and second level respectively, where $\eta^{(k)} = \mathbf{T}\sigma^{(x)}$. The compressed interaction potential $\bar{J}(k,l)$ has the reduced range L/q and the number of operations for calculating $\Delta\overline{H}_l^{(0)}(\eta,\eta^{(k)})$ is $(2L+1)^d/Q$, while for $\Delta H_s(\sigma,\sigma^{(x)})$ is $(2S+1)^d$. The method, in addition to the reduction of operations due to range suppression, exhibits a computational reduction as a result of the fact that rejected trials at the first level will not be tested at the second and the calculation of differences $\Delta H_s(\sigma,\sigma^{(x)})$ is avoided. A summary of this discussion appears in Table 1. When the correction terms $\Delta H_l(\sigma,\sigma^{(x)}) - \Delta\overline{H}_l^{(0)}(\eta,\eta^{(k)})$ are present the additional computational cost in the second level is small, dependent on the decay of rate of $J(x-y) - \bar{J}(k,l),\ x,y\in\Sigma_N,\ x\in C_k,\ y\in C_l$. Since this term is in principle fast decaying, in implementations we can neglect interactions with distance larger than a cut-off range $L_c < L$ with a small error, that will contribute further in the computational time reduction.

4.2. Comparison of equilibration times

A direct estimation of rate of convergence of a Markov chain generated by a Monte Carlo method is model dependent and in general intractable. Thus for the purpose of this work it is natural to study this property as a comparison of the proposed method with a conventional method. Theorem 2 provides such a comparison of the spectral gap between the conventional MH method and the two-level CGMC method, for sampling a measure $\mu(d\sigma)$ on a countable state space Σ. This comparison is summarised in inequality (20) proving that their relation, in terms the spectral gap, is controlled by the approximation $\bar{\mu}^{(0)}(d\eta)$ of marginals $\bar{\mu}(d\eta) = \mu\circ\mathbf{T}^{-1}(d\eta)$.

For a discrete time Markov chain $\{X_n\}$ with transition kernel $\mathcal{K}$ and stationary distribution μ, the mixing time τ is defined as

$$\tau := min_n\left\{\forall\sigma\in\Sigma : \|\mathcal{K}^n(\sigma,\cdot) - \mu(\cdot)\|_{TV} \leqslant \frac{1}{4}\right\}.$$

For two probability measures μ, ν the total variation norm is $\|\mu-\nu\|_{TV} = \frac{1}{2}\sum_\sigma|\mu(\sigma)-\nu(\sigma)|$. Bounds of the total variation norm appearing can be given in terms of $\mathcal{K}$'s spectral gap $\lambda(\mathcal{K})$ [8], for example for a reversible kernel holds

$$2\|\mathcal{K}^n(\sigma,\cdot) - \mu\|_{TV} \leqslant \frac{1}{\min_{\sigma\in\Sigma}\mu(\sigma)^{1/2}}(1-\lambda(\mathcal{K}))^n. \tag{19}$$

The spectral gap of a kernel $\mathcal{K}$ is defined by $\lambda(\mathcal{K}) = \min\left\{\frac{\mathcal{E}(f,f)}{\mathrm{Var}(f)}; \mathrm{Var}(f)\neq 0\right\}$ with the Dirichlet form $\mathcal{E}(f,f) = \frac{1}{2}\sum_{\sigma,\sigma'}|f(\sigma)-f(\sigma')|^2\mathcal{K}(\sigma,\sigma')\mu(\sigma)$ and the variance $\mathrm{Var}(f) = \frac{1}{2}\sum_{\sigma,\sigma'}|f(\sigma)-f(\sigma')|^2\mu(\sigma')\mu(\sigma)$, where f is a function on Σ square integrable with respect to μ. In view of (19), one can say that between two algorithms producing Markov chains with identical equilibrium distributions *better* in terms of speed of convergence is the one with the *larger spectral gap*. Therefore here we compare the spectral gap corresponding to transition kernels of the proposed method and the conventional MH method. Let $\mathcal{E}[\mathcal{K}_{tl}], \mathcal{E}[\mathcal{K}_c]$ denote the Dirichlet forms and $\lambda(\mathcal{K}_{tl}), \lambda(\mathcal{K}_c)$ the spectral gap corresponding to the two-level CGMC and the Metropolis transition kernels $\mathcal{K}_{tl}(\sigma,\sigma')$ (18) and $\mathcal{K}_c(\sigma,\sigma')$ (12) respectively.

Theorem 2. *Let $\rho(\sigma,\sigma')$ be a symmetric proposal transition probability for the conventional MH algorithm and $\bar{\rho}(\eta,\eta')$ a symmetric proposal transition probability on the coarse space $\overline{\Sigma}$ for the two-level CGMC algorithm, then for any conditional probability $\mu_r(\sigma|\eta)$*

$$\mathcal{A}\underline{\gamma}\lambda(\mathcal{K}_c) \leqslant \lambda(\mathcal{K}_{tl}) \leqslant \bar{\gamma}\lambda(\mathcal{K}_c), \tag{20}$$

where $\mathcal{A} = \inf_{\sigma,\sigma'}\{\mathcal{A}(\sigma,\sigma')\}$ and $\underline{\gamma} > 0, \bar{\gamma} > 0$ such that $\underline{\gamma} \leqslant \mathcal{B}(\sigma,\sigma') \leqslant \bar{\gamma}$, with $\mathcal{A}(\sigma,\sigma')$ and $\mathcal{B}(\sigma,\sigma')$ defined in Lemma 1.

Remark

1. Existence of finite and positive values of $\mathcal{A} = \inf_{\sigma,\sigma'}\{\mathcal{A}(\sigma,\sigma')\}$ is ensured for the models studied in this work. Consider the Gibbs measure $\mu_{N,\beta}(d\sigma)\propto\exp\{-\beta H_N(\sigma)\}$ and $\bar{\mu}^{(0)}_{M,\beta}(d\eta)\propto\exp\{-\beta\overline{H}^{(0)}(\eta)\}$ as defined in Section 2. Let $J(r),\ s.t.\ \sum_r J(r) = J^* < \infty$ with the compressed interactions $\bar{J}$ as defined in (6). All possible values of $\mathcal{A}(\sigma,\sigma')$, (23), are $\mathcal{A}(\sigma,\sigma') = 1$,

$$\mathcal{A}(\sigma,\sigma') = \min\{e^{-\beta\Delta\overline{H}^{(0)}(\eta,\eta')}, e^{-\beta\Delta\overline{H}^{(0)}(\eta',\eta)}\} \geqslant e^{-\beta\bar{J}^*}$$

and

$$\mathcal{A}(\sigma,\sigma') = \min\left\{e^{-\beta(\Delta H_N(\sigma,\sigma')-\Delta\overline{H}^{(0)}(\eta,\eta'))}, e^{-\beta(\Delta H_N(\sigma',\sigma)-\Delta\overline{H}^{(0)}(\eta',\eta))}\right\} = 1+\mathcal{O}\left(\frac{Q}{(2L+1)^d}\right).$$

This is a result of the known estimate, Lemma (2.3) [19],

$$\Delta H_N(\sigma,\sigma') - \Delta\overline{H}^{(0)}(\eta,\eta') = \mathcal{O}\left(\frac{Q}{(2L+1)^d}\right)$$

2. The above estimate proves also that values of $\mathcal{A}(\sigma,\sigma')$ are close to 1 controlled by the approximation parameter Q.
3. Term $\mathcal{B}(\sigma,\sigma')$ depends only on the difference between the proposal kernels in the two methods while the marginal approximation is controlled by $\mathcal{A}$.

Numerically the statement of the Theorem is revealed by a comparison of the average acceptance probabilities in examples, see Figs. 4 and 5. The proof of Theorem 2 is based on the following lemmata. Lemma 1 gives the detailed relation of the two methods transition kernels that will aid in the comparison of their Dirichlet forms and subsequently application of Lemma 2 will lead to the desired result.

Lemma 1. *For symmetric proposal transition kernels $\rho(\sigma,\sigma') = \rho(\sigma',\sigma)$ and $\bar{\rho}(\eta,\eta') = \bar{\rho}(\eta',\eta)$, for all σ, $\sigma' \in \Sigma$ with $\sigma \neq \sigma'$,*

$$\mathcal{K}_{tl}(\sigma,\sigma') = \mathcal{A}(\sigma,\sigma')\mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma'),$$

where

$$\mathcal{B}(\sigma,\sigma') = \frac{\bar{\rho}(\eta,\eta')}{\rho(\sigma,\sigma')}\begin{cases}\mu_r(\sigma'|\eta'), & \text{if } \alpha_f(\sigma,\sigma') = 1\\ \mu_r(\sigma|\eta), & \text{if } \alpha_f(\sigma,\sigma') < 1.\end{cases} \tag{21}$$

Furthermore we define the subsets

$$\begin{aligned}
C_1 &= \left\{(\sigma,\sigma') \in \Sigma\times\Sigma : \{\alpha<1, \alpha_{CG}<1, \alpha_f<1\} \text{ or } \{\alpha=1, \alpha_{CG}=1, \alpha_f=1\}\right\}\\
C_2 &= \left\{(\sigma,\sigma') \in \Sigma\times\Sigma : \{\alpha=1, \alpha_{CG}<1, \alpha_f=1\} \text{ or } \{\alpha<1, \alpha_{CG}=1, \alpha_f<1\}\right\}\\
C_3 &= \left\{(\sigma,\sigma') \in \Sigma\times\Sigma : \{\alpha=1, \alpha_{CG}=1, \alpha_f<1\} \text{ or } \{\alpha<1, \alpha_{CG}<1, \alpha_f=1\}\right\}\\
C_4 &= \left\{(\sigma,\sigma') \in \Sigma\times\Sigma : \{\alpha<1, \alpha_{CG}=1, \alpha_f=1\} \text{ or } \{\alpha=1, \alpha_{CG}<1, \alpha_f<1\}\right\}
\end{aligned} \tag{22}$$

and

$$\mathcal{A}(\sigma,\sigma') = \begin{cases}1, & \text{if } (\sigma,\sigma') \in C_1\\ \min\left\{\frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}, \frac{\bar{\mu}^{(0)}(\eta)}{\bar{\mu}^{(0)}(\eta')}\right\}, & \text{if } (\sigma,\sigma') \in C_2\\ \min\left\{\frac{\mu(\sigma')\bar{\mu}^{(0)}(\eta)}{\mu(\sigma)\bar{\mu}^{(0)}(\eta')}, \frac{\mu(\sigma)\bar{\mu}^{(0)}(\eta')}{\mu(\sigma')\bar{\mu}^{(0)}(\eta)}\right\}, & \text{if } (\sigma,\sigma') \in C_3\\ \min\left\{\frac{\mu(\sigma')}{\mu(\sigma)}, \frac{\mu(\sigma)}{\mu(\sigma')}\right\}, & \text{if } (\sigma,\sigma') \in C_4.\end{cases} \tag{23}$$

Proof. Recall that $\mathcal{K}_c(\sigma,\sigma') = \alpha(\sigma,\sigma')\rho(\sigma,\sigma')$, for $\sigma \neq \sigma'$, which for the conventional MH method and the hypothesis of symmetry of $\rho(\sigma,\sigma')$ becomes $\mathcal{K}_c(\sigma,\sigma') = \min\left\{1, \frac{\mu(\sigma')}{\mu(\sigma)}\right\}\rho(\sigma,\sigma')$. Similarly the two-level CGMC kernel becomes

$$\mathcal{K}_{tl}(\sigma,\sigma') = \min\left\{1, \frac{\mu(\sigma')\bar{\mu}^{(0)}(\eta)\mu_r(\sigma|\eta)}{\mu(\sigma)\bar{\mu}^{(0)}(\eta')\mu_r(\sigma'|\eta')}\right\} \times \min\left\{1, \frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}\right\}\bar{\rho}(\eta,\eta')\mu_r(\sigma|\eta').$$

The proof is based on a case study on the values of the acceptance probabilities $\alpha(\sigma,\sigma') = \min\left\{1, \frac{\mu(\sigma')}{\mu(\sigma)}\right\}$, $\alpha_f(\sigma,\sigma') = \min\left\{1, \frac{\mu(\sigma')\bar{\mu}^{(0)}(\eta)\mu_r(\sigma|\eta)}{\mu(\sigma)\bar{\mu}^{(0)}(\eta')\mu_r(\sigma'|\eta')}\right\}$ and $\alpha_{CG}(\eta,\eta') = \min\left\{1, \frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}\right\}$. All possible combinations on their values are categorized in the sets C_i, $i = 1,\ldots,4$ defined in (22). For $(\sigma,\sigma') \in C_1$ and $\alpha(\sigma,\sigma') < 1$, $\alpha_f(\sigma,\sigma') < 1$ and $\alpha_{CG}(\eta,\eta') < 1$, $\mathcal{K}_c(\sigma,\sigma') = \frac{\mu(\sigma')}{\mu(\sigma)}\rho(\sigma,\sigma')$ and

$$\mathcal{K}_{tl}(\sigma,\sigma') = \frac{\mu(\sigma')\bar{\mu}^{(0)}(\eta)}{\mu(\sigma)\bar{\mu}^{(0)}(\eta')}\frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}\bar{\rho}(\eta,\eta')\mu_r(\sigma|\eta) = \frac{\mu(\sigma')}{\mu(\sigma)}\bar{\rho}(\eta,\eta')\mu_r(\sigma|\eta),$$

that results their relation $\mathcal{K}_{tl}(\sigma,\sigma') = \frac{\bar{\rho}(\eta,\eta')\mu_r(\sigma|\eta)}{\rho(\sigma,\sigma')}\mathcal{K}_c(\sigma,\sigma')$. With similar simple calculations for all sub-cases that are encountered in sets C_i, $i = 1,\ldots,4$ we have, for $(\sigma,\sigma') \in C_1$

$$\mathcal{K}_{tl}(\sigma,\sigma') = \mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma'),$$

with

$$\mathcal{B}(\sigma,\sigma') = \frac{\bar{\rho}(\eta,\eta)}{\rho(\sigma,\sigma')}\begin{cases}\mu_r(\sigma'|\eta'), & \text{if } \alpha_f(\sigma,\sigma') = 1\\ \mu_r(\sigma|\eta), & \text{if } \alpha_f(\sigma,\sigma') < 1\end{cases}$$

For $(\sigma,\sigma')\in C_2$ such that $\alpha(\sigma,\sigma')=1$, $\alpha_{CG}(\eta,\eta')<1$, $\alpha_f(\sigma,\sigma')=1$,

$$\mathcal{K}_{tl}(\sigma,\sigma')=\frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}\frac{\bar{\rho}(\eta,\eta')\mu_r(\sigma'|\eta')}{\rho(\sigma,\sigma')}\mathcal{K}_c(\sigma,\sigma'),$$

and for $\alpha(\sigma,\sigma')<1$, $\alpha_{CG}(\eta,\eta')=1$, $\alpha_f(\sigma,\sigma')<1$,

$$\mathcal{K}_{tl}(\sigma,\sigma')=\frac{\bar{\mu}^{(0)}(\eta)}{\bar{\mu}^{(0)}(\eta')}\frac{\bar{\rho}(\eta,\eta')\mu_r(\sigma|\eta)}{\rho(\sigma,\sigma')}K_c(\sigma,\sigma'),$$

that we can summarize to

$$\mathcal{K}_{tl}(\sigma,\sigma')=\min\left\{\frac{\bar{\mu}^{(0)}(\eta)}{\bar{\mu}^{(0)}(\eta')},\frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}\right\}\mathcal{B}(\sigma,\sigma')K_c(\sigma,\sigma'),$$

since for example in the first case $\frac{\bar{\mu}^{(0)}(\eta')}{\bar{\mu}^{(0)}(\eta)}<1<\frac{\bar{\mu}^{(0)}(\eta)}{\bar{\mu}^{(0)}(\eta')}$ and inversely for the second. Following the same reasoning for $(\sigma,\sigma')\in C_3$,

$$\mathcal{K}_{tl}(\sigma,\sigma')=\min\left\{\frac{\mu(\sigma')\bar{\mu}^{(0)}(\eta)}{\mu(\sigma)\bar{\mu}^{(0)}(\eta')},\frac{\mu(\sigma)\bar{\mu}^{(0)}(\eta')}{\mu(\sigma')\bar{\mu}^{(0)}(\eta)}\right\}\mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma'),$$

and for $(\sigma,\sigma')\in C_4$

$$\mathcal{K}_{tl}(\sigma,\sigma')=\min\left\{\frac{\mu(\sigma')}{\mu(\sigma)},\frac{\mu(\sigma)}{\mu(\sigma')}\right\}\mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma').$$

All these steps prove, the following relation of transition kernels generated by Algorithms 1 and 2,

$$\mathcal{K}_{tl}(\sigma,\sigma')=\mathcal{A}(\sigma,\sigma')\mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma'),\ \text{for all } \sigma,\sigma'\in\Sigma,$$

with $\mathcal{A}(\sigma,\sigma')$ and $\mathcal{B}(\sigma,\sigma')$ defined in (23) and (21). □

The proof of Theorem 2 is based on the application of Lemma 1 and Lemma 3.3 in the work of Diaconis and Saloff-Coste [8] that is stated next for completeness.

Lemma 2. *Let $(\mathcal{K},\mu)$ and $(\mathcal{K}',\mu')$ be two Markov chains on the same finite set X. Assume that there exist $A,a>0$ such that*

$$\mathcal{E}[\mathcal{K}']\leqslant A\mathcal{E}[\mathcal{K}],\quad a\mu\leqslant\mu',$$

then $\lambda'\leqslant\frac{A}{a}\lambda$.

We conclude with the proof of Theorem 2.

Proof. We compare the Dirichlet forms $\mathcal{E}[\mathcal{K}_{tl}],\mathcal{E}[\mathcal{K}_c]$ using the definition of Dirichlet form and Lemma 1. By the definition of $\mathcal{A}(\sigma,\sigma')$ for all $\sigma,\sigma'\in\Sigma$ holds $0<\mathcal{A}(\sigma,\sigma')\leqslant 1$. Then by Lemma 1

$$\inf_{\sigma,\sigma'}\{\mathcal{A}(\sigma,\sigma')\}\mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma')\leqslant\mathcal{K}_{tl}(\sigma,\sigma')\leqslant\mathcal{B}(\sigma,\sigma')\mathcal{K}_c(\sigma,\sigma').$$

Let $\underline{\gamma}>0$ and $\bar{\gamma}>0$ such that $\underline{\gamma}\leqslant\mathcal{B}(\sigma,\sigma')\leqslant\bar{\gamma}$ for all $\sigma,\sigma'\in\Sigma$. Then

$$\inf_{\sigma,\sigma'}\{\mathcal{A}(\sigma,\sigma')\}\underline{\gamma}\mathcal{K}_c(\sigma,\sigma')\leqslant\mathcal{K}_{tl}(\sigma,\sigma')\leqslant\bar{\gamma}\mathcal{K}_c(\sigma,\sigma'),\ \text{for all } \sigma,\sigma'\in\Sigma.$$

Recalling the definition of Dirichlet form for kernel $\mathcal{K}_{tl}(\sigma,\sigma')$, $\mathcal{E}[\mathcal{K}_{tl}](f,f)=\frac{1}{2}\sum_{\sigma,\sigma'}|f(\sigma)-f(\sigma')|^2\mathcal{K}_{tl}(\sigma,\sigma')\mu(\sigma)$, and the above relation we can write

$$\inf\{\mathcal{A}(\sigma,\sigma')\}\underline{\gamma}\mathcal{E}[\mathcal{K}_c](f,f)\leqslant\mathcal{E}[\mathcal{K}_{tl}](f,f)\leqslant\bar{\gamma}\mathcal{E}[\mathcal{K}_c](f,f),$$

for any function f square integrable with respect to μ. Application of Lemma 2, for which here $\mu'\equiv\mu$, thus $a=1$ and $A=\inf_{\sigma,\sigma'}\{\mathcal{A}(\sigma,\sigma')\}\underline{\gamma}$ for the left hand side inequality, and $A=\bar{\gamma}$ for the right hand side, gives the relation of spectral gaps

$$\inf_{\sigma,\sigma'}\{\mathcal{A}(\sigma,\sigma')\}\underline{\gamma}\lambda(\mathcal{K}_c)\leqslant\lambda(\mathcal{K}_{tl})\leqslant\bar{\gamma}\lambda(\mathcal{K}_c).\qquad\square$$

5. Benchmark examples for the canonical ensemble

5.1. Combined Ising and Curie Weiss model

We consider a benchmark example of competing short- and long-range interactions, which exhibits critical behaviour. Furthermore analytical expressions for the free energy in the thermodynamic limit, $N\to\infty$, are known for the one dimen-

sional model as well as representative phase diagrams for higher dimensions [15]. Application of Theorem 2 is described for this model with an explicit calculation of constants $\mathcal{A}$ and $\underline{\gamma}, \bar{\gamma}$ appearing in inequality (20), for the one dimensional model. Numerical tests are presented for a two-dimensional model and in [14] one can find a detailed numerical study for the one dimensional case. The energy of the system at configuration $\sigma \in \Sigma_N = \{0,1\}^{\Lambda_N}$ is

$$H_N(\sigma) = H_s(\sigma) + H_l(\sigma) - h\sum_{x\in\Lambda_N}\sigma(x), \tag{24}$$
$$H_s(\sigma) = -\frac{K_0}{2}\sum_{x\in\Lambda_N}\sum_{|x-y|=1}\sigma(x)\sigma(y), \quad H_l(\sigma) = -\frac{J_0}{2N}\sum_{x\in\Lambda_N}\sum_{y\neq x}\sigma(x)\sigma(y).$$

The interactions involved in $H_s(\sigma)$ are of the nearest-neighbour type with strength K_0. Equivalent neighbour, distance independent, interactions $J_0/2N$ define $H_l(\sigma)$, that represent a mean-field approximation of a potential $J(x-y)$, as in (8) for example, averaged over all lattice sites. The coarse grained Hamiltonian $\overline{H}_l(\eta)$ is exact, equal to the microscopic energy $H_l(\sigma)$, and no coarsening error is involved from compressing long range interactions that allows us to study the effect of the splitting of short- and long-range interactions. Indeed, for any coarsening level q

$$H_l(\sigma) = \overline{H}_l(\eta) = -\frac{J_0}{2N}\sum_{k\in\overline{\Lambda}_M}\sum_{l\in\overline{\Lambda}_M l\neq k}\eta(k)\eta(l) - \frac{J_0}{2N}\sum_{k\in\overline{\Lambda}_M}\eta(k)(\eta(k)-1),$$

where $\eta(k) = \sum_{x\in C_k}\sigma(x),\ k\in\overline{\Lambda}_M$. We consider the canonical ensemble with the spin-flip dynamics, that is for the conventional MH each trial at the nth MC iteration with configuration σ is $\sigma^{(x)}$. Similarly sampling at the coarse space, in the first step of the two-level method, is achieved with a birth–death process on a coarse cell, i.e., a trial at the nth MC iteration with coarse configuration η is $\eta^{(k)} = \eta \pm \delta_k$ with $\delta_k(k) = 1$, $\delta_k(v) = 0$, $v \neq k$, more details on implementation of the method are given in Appendix B.

The detailed calculations on the application of Theorem 2 are given in Appendix B, where we prove that the constants appearing in inequality (20) are

$$\underline{\gamma} = \bar{\gamma} = 1, \text{ and}$$
$$\mathcal{A} = \begin{cases} \min\{e^{-|J_0|}, e^{-2|K_0|}\} & \text{for } K_0 \neq 0, \\ 1 & \text{for } K_0 = 0. \end{cases} \tag{25}$$

For $K_0 = 0$ the energy associated to the second step of the method is zero, and according to (15) the acceptance probability $\alpha_f(\sigma,\sigma') = 1$, for all $\sigma,\sigma' \in \Sigma_N$, i.e., all proposed states from the first step are accepted. Therefore, in view of the exact coarsening of $H_l(\sigma)$, the two-level method and the conventional MH are equivalent. This is verified also by Theorem 2 since inequality (20) for this trivial case becomes $\lambda(\mathcal{K}_{tl}) = \lambda(\mathcal{K}_c)$. For $K_0 \neq 0$ we have

$$\min\{e^{-|J_0|}, e^{-2|K_0|}\}\lambda(\mathcal{K}_c) \leqslant \lambda(\mathcal{K}_{tl}) \leqslant \lambda(\mathcal{K}_c), \tag{26}$$

showing that the parameter controlling the relation of the convergence rates is the relative strength of the short and long range interactions J_0 and K_0. Table 2 presents a comparison of the computational time between the MH and the two-level CGMC method for a two-dimensional lattice $N = 16\times 16$ and coarsening parameter $Q = q\times q$, for $q = 4$ and $q = 8$. The results demonstrate a reduction of computational effort at a rate close to $\mathcal{O}(Q)$. As is expected this rate is not exactly $\mathcal{O}(Q)$because of the additional computational effort necessary for implementing the local reconstruction. The hysteresis diagram is indicated in Fig. 2, for the average total coverage $\langle c\rangle$ first upon increasing the field h from low values and then decreasing it from high values. The total coverage is $c(\sigma) = N^{-1}\sum_x\sigma(x)$ and average quantities $\langle c\rangle$ are computed after approximate equilibration in the Monte Carlo sampling, using 1000 samples. The tests demonstrate that the two-level CGMC method predicts correctly the phase transition regime, compared to the conventional MH, with a reduced computational cost. Furthermore the hysteresis diagrams obtained with the two-level method agree with the conventional method outcome, for a wide range (not shown) of the relative strengths of interactions.

Table 2
CPU cost comparisons for different resolutions in hysteresis simulations with parameters $K_0 = 1$, $J_0 = 5$, $h\in[0,6]$, $N = 16\times 6$.

Method	CPU (min)
MH	327
Two-level CGMC $q = 4$	28
Two-level CGMC $q = 8$	9

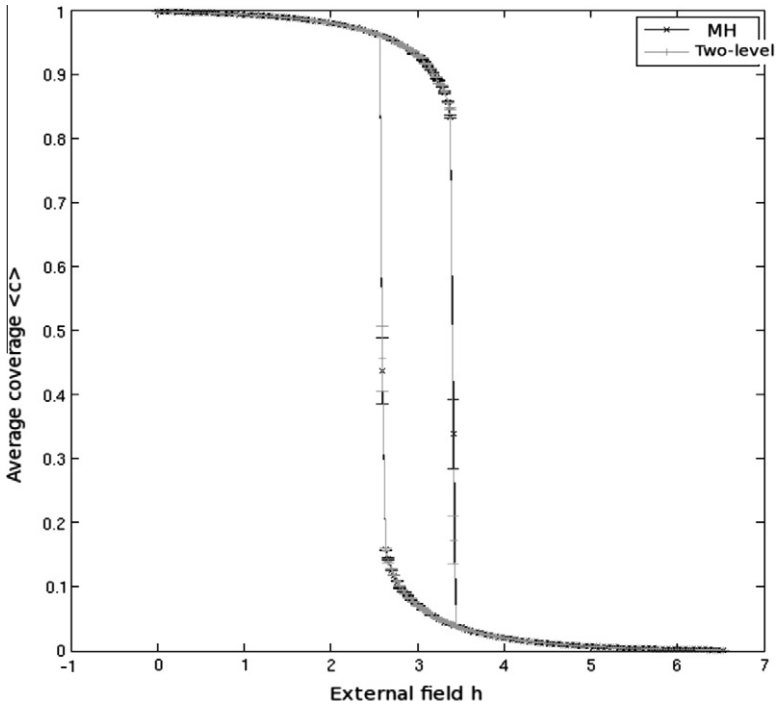

Fig. 2. Comparison of average coverage in hysteresis simulations obtained by microscopic MH and two-level CGMC methods for model (24) with parameters $N = 64 \times 64$, $q = 4$, $K_0 = 1$, $J_0 = 5$.

5.2. Kac type interactions

Known error estimates indicate that the potential decay is one of the parameters controlling the approximation error in coarse graining techniques. The effect of potential singularities on the coarsening error and improving strategies encountering multi-body interactions and/or potential splitting has been studied in [2]. The example application studied in this section exhibits an algebraic decay for which direct CGMC error can be significant. The two-level CGMC method with potential splitting strategy is applied here and is shown to improve the coarsening error. We consider a long range Kac type interaction potential $J(r) = N^{-1}V(r/N)$, $r = |x - y| > 0$, $x,y \in \Lambda_N$

$$V(r) = \frac{v}{r^{3/2}}, \text{ if } r > 0, \tag{27}$$

where constant v is chosen to ensure the conservation of the total mass $J_0 = \int J(r)dr$. The potential splitting strategy is applied by decomposing the interacting potential into a short-range $J_s(r)$, as well as a long-range $J_l(r)$, defined by

$$J_s(r) = \begin{cases} J(r) & \text{for } 0 < r \leqslant S, \\ 0 & \text{for } S < r \end{cases}$$

and

$$J_l(r) = J(r) - J_s(r).$$

This splitting defines the energy of the system in the form $H_N(\sigma) = H_s(\sigma) + H_l(\sigma)$ as in (11) with $K(x - y) = J_s(|x - y|)$ and $J(x - y) = J_l(|x - y|)$.

In Fig. 3 and Table 3 we present simulation results, for a one dimensional system, in the canonical ensemble obtained by the conventional MH with interaction potential $J(|x - y|)$, $x,y \in \Lambda_N$, the two-level CGMC with interactions $\bar{J}_l(k,l),\ k,l \in \overline{\Lambda}_M$ and $J_s(|x - y|)$, both sampling on the microscopic space $\Sigma_N = \{0,1\}^{\Lambda_N}$, and the CGMC with interaction potential $\bar{J}(k,l),\ k,l \in \overline{\Lambda}_M$ sampling on the coarse space $\overline{\Sigma}_M$. A reduction of the computational time of the conventional sampling is achieved by the two-level method as the numerical tests summarized in Table 3 confirm. Furthermore the proposed method reduces the coarsening error of CGMC simulations since the most singular part of the interaction potential, $J_s(r)$, is treated in the microscopic space and coarsening is applied only to its fast decaying part $J_l(r)$, see Fig. 3 and Table 3. The error appearing in Table 3 is the l^2 distance of the average total coverage $\langle c \rangle$, of the indicated method, from the conventional MH result, as a function of the external field h.

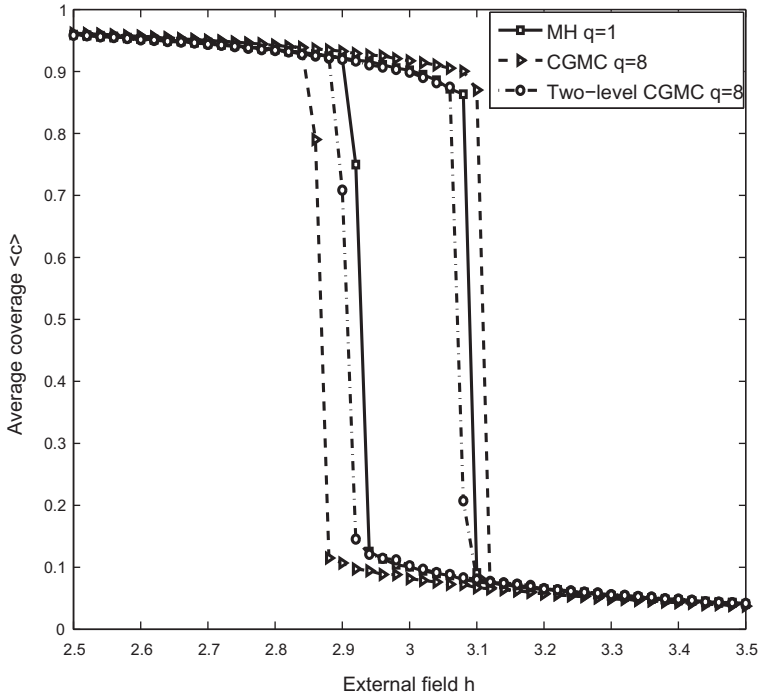

Fig. 3. Comparison of average coverage in hysteresis simulations obtained by microscopic MH, CGMC and two-level CGMC methods for model (27) with parameters $\beta J_0 = 6$, $N = 512$, $q = 8$, $S = 1$.

Table 3
CPU cost and error comparisons for different resolutions and methods. $N = 512$, $\beta J_0 = 6$, $S = 1$.

Method	CPU time	Error
MH $q = 1$	252	0
CGMC $q = 8$	47	0.78
Two-level CGMC $q = 8$	57	0.65
CGMC $q = 64$	5	2.10
Two-level CGMC $q = 64$	6	0.81

The two-level CGMC is an efficient method of generating proposals in Metropolis sampling, verified in Fig. 4 where we observe a significant increase of the average acceptance probability of the conventional method. The average acceptance probability is defined as the average over the number of MC iterations of acceptance probabilities $\alpha(\sigma,\sigma')$ for the conventional MH method, Algorithm 1, and $\alpha_f(\sigma,\sigma')$ for the two-level CGMC method, Algorithm 2. Overall the total average acceptance probability of the two-level method approaches the one of the conventional method, which confirms the statement of Theorem 2, see Fig. 5. The total average acceptance probability is defined similarly as the average of the product of the two levels acceptance probabilities $\alpha_{CG}(\eta,\eta')\,\alpha_f(\sigma,\sigma')$. Thus we confirm the outcome of Theorem 2 that clearly suggest that the CPU savings reported in Table 3 are entirely due to the inexpensive creation of suitable proposals using the CGMC method.

6. Nanopattern formation in heteroepitaxy – Microcanonical ensemble

Heteroepitaxy describes the procedure of deposition of a crystalline material on a substrate of a different material, a procedure that in general exhibits a variety of patterns. Models describing these procedures are characterized by the interplay of short range attraction and long range repulsion interactions between particles. Such an interplay can lead to the formation of patterns, such as discs and stripes [5]. With this application we show how the proposed method can provide microscopic information, benefiting from the low computational cost of the coarse graining technique.

The energy of the system is given by $H_N(\sigma)$, as in (2), with isotropic interaction potential

$$J(r) = J_0\left(e^{-(r/r_a)^2} - \chi e^{-(r/r_r)^2}\right), \quad r = |x-y| > 0, \quad x,y \in \Lambda_N. \tag{28}$$

J_0 is the strength of the potential, r_a and r_r are the dimensionless length scales of attraction and repulsion, respectively and χ is the repulsion strength parameter. In general $r_a < r_r$ determining the short range nature of attractive interactions. A study of

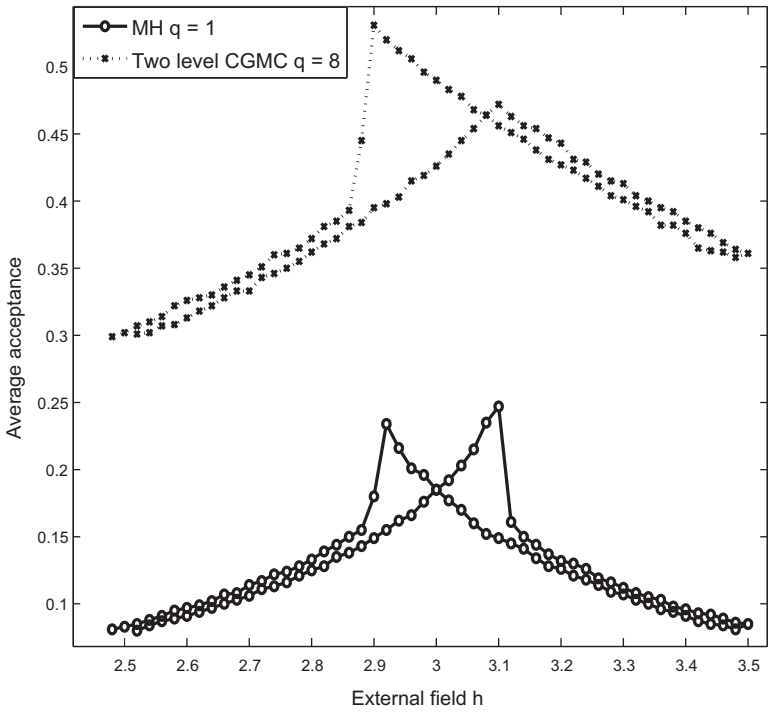

Fig. 4. The comparison of the average acceptance probabilities between the second step of ML-CGMC and the conventional MH, in hysteresis simulations depicted in Fig. 3, demonstrates increased acceptance rate due to better proposals in ML-CGMC.

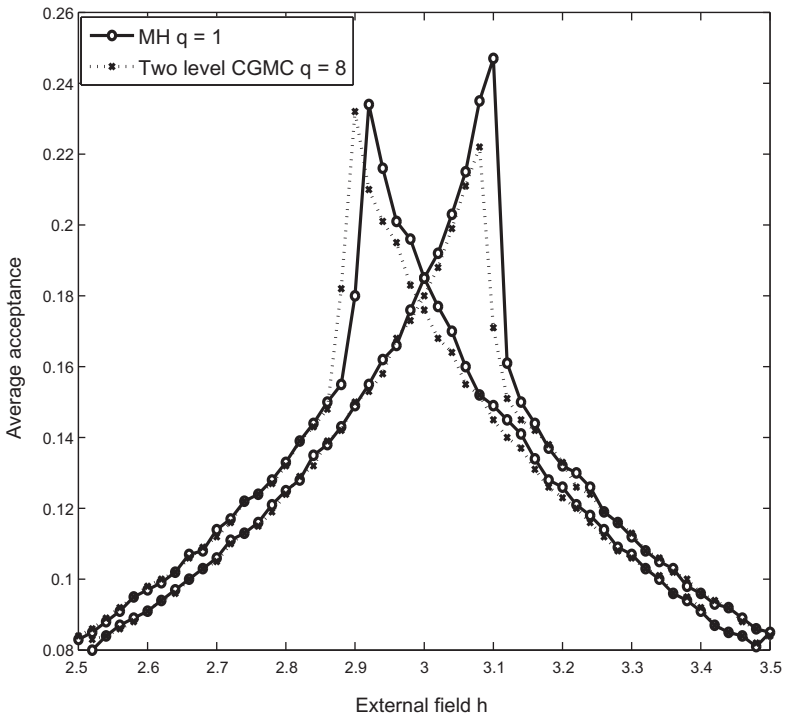

Fig. 5. Comparison of the total average acceptance of the ML-CGMC and the conventional MH confirms equivalent equilibration rates.

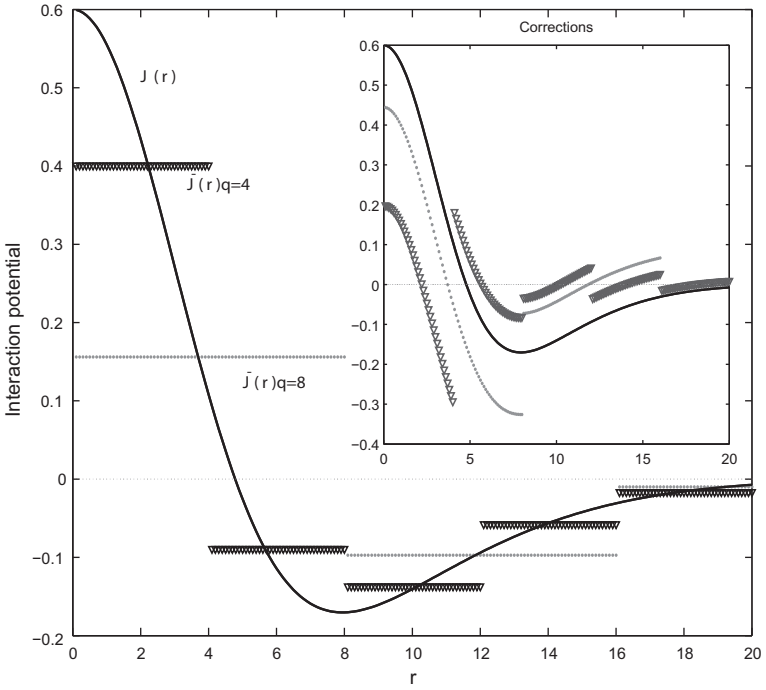

Fig. 6. Potential functions $J(r)$, coarse grained $\bar{J}(r)$ and $J_c(r)$ for different coarsening levels q.

the kinetic phase diagrams and application of CGMC methods for this system is given in [5]. The underlying dynamics describing the surface diffusion of particles consideredare spin-exchange, obeying the exclusion principle and conserving the order parameter microcanonically. The order parameter here is the total coverage $c_0 = N^{-1}\sum_{x\in\Lambda_N}\sigma(x)$ representing the number of occupied sites on the lattice, where $\sigma(x)\in\{0,1\}$, $x\in\Lambda_N$.

In the tables and figures following we provide simulation results from the conventional MH, the CGMC and the two-level CGMC method. For the two-level CGMC method, two sampling strategies were tested following the description in Section 3.1, the exact sampling with correction terms, strategy 1., (Figs. 7, 8 and Tables 4, 5), and the approximate CG with potential splitting, strategy 3., (Fig. 9). For the strategy with corrections the potential functions appearing in all methods are depicted in Fig. 6. The compressed interaction potential is $\bar{J}(k,l),\ k,l\in\overline{\Lambda}_M$, as in (6), used for the coarse space simulations and the first step of the two-level method. In the second step of the latest method the potential function is $J_c(|x-y|)=J(|x-y|)-\bar{J}(|x-y|)$ where $\bar{J}(|x-y|)=\bar{J}(k,l),\ x,y\in\Lambda_N$ with $x\in C_k$, $y\in C_l$, representing the correction of compressing interactions at the first step of the method.

In all simulations presented in the sequel the range of pure attractive and repulsive forces are $r_a = 4.47$, $r_r = 10$ respectively and the repulsion strength is $\chi = 0.1$. Since the potential function (28) follows an exponential decay, in implementations we use a cut-off range L for the potential $J(r)$ such that $J(r) < 10^{-6}$ for $r > L$, which accounts on $(2L)^2$ interactions for each site. For all sampling methods the number of MC iterations used is 5×10^7.

The numerical results verify that the proposed method provides the expected behaviour of the system for a finite lattice as its size increases, that is feature properties converge, as is evident in Table 4. For the total coverage $c_0 = 0.9$ that leads to a pattern of periodic inverted discs Table 4 presents the average diameter of discs and the computational time when varying the lattice size N. Feature statistics are calculated with the use of edge detection techniques of image processing [32]. In order to demonstrate the effect of the statistical errors we present confidence interval estimates. Such intervals are obtained after transforming the simulations data to follow an approximately normal distribution. For the sake of comparison in the

Table 4
Finite system size behaviour of the two-level CGMC method. Average discs diameter $\langle d\rangle$ and computational time. $c_0 = 0.9$, $\beta J_0 = 0.6$, $q = 8$, $L = 24$.

Lattice	Diameter	Std	Confidence interval	CPU time (min)
128×128	10.2	3.6	[5.7,13.3]	30
256×256	9.5	2.5	[8.7,9.9]	43
512×512	8.8	3.4	[8.5,9.1]	66
1024×1024	8.6	3.6	[8.5,8.7]	110
MH method				
512×512	8.27	1.2	[8.1,8.3]	805

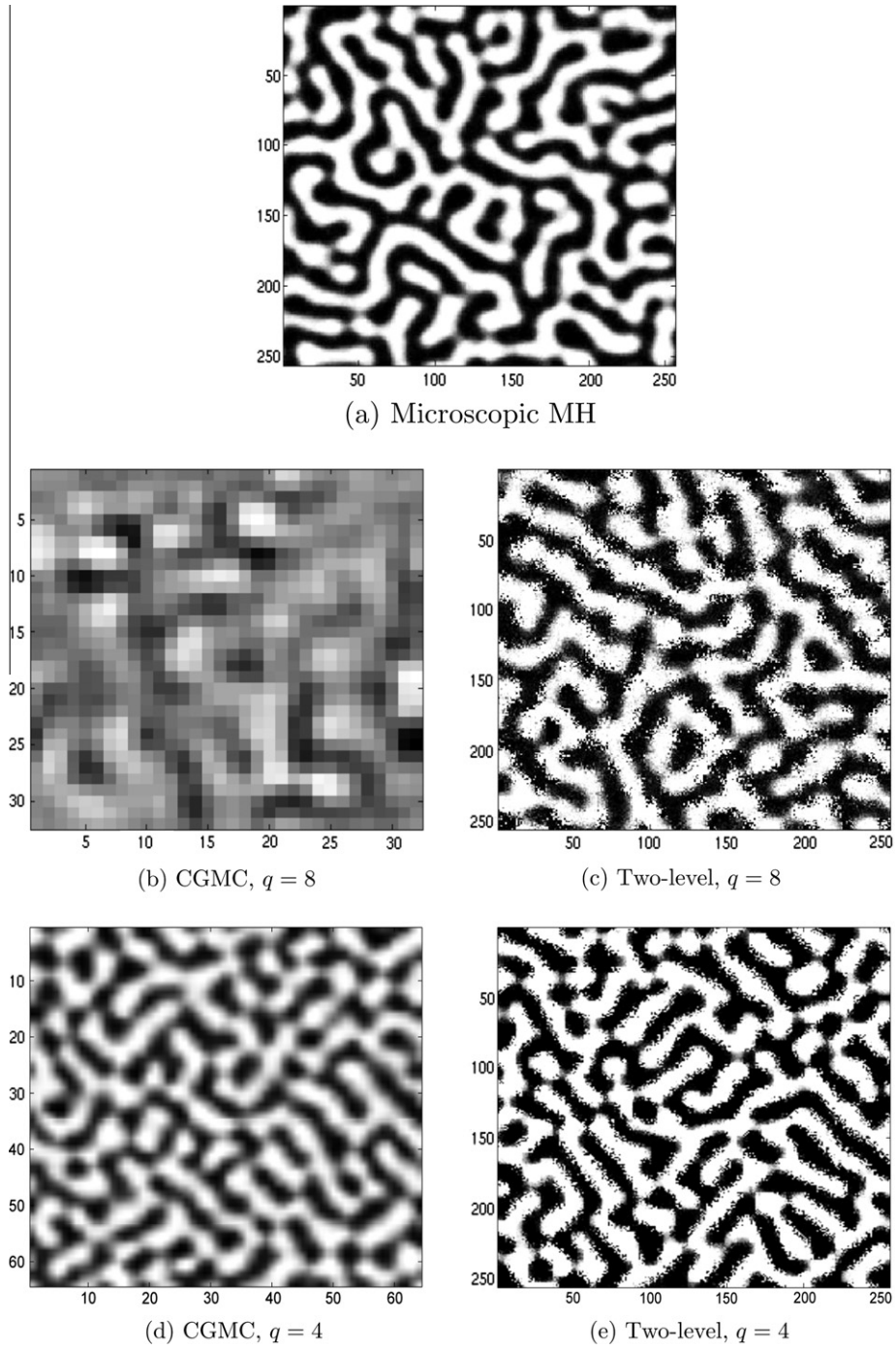

Fig. 7. The two-level CGMC method captures the correct feature scaling even when CGMC is inaccurate. Simulations for model (28) with parameters $c_0 = 0.5$, $\beta J_0 = 0.2$, $\chi = 0.1$ and $N = 256 \times 256$.

same table we present results obtained with the MH method for a 512×512 lattice, that indicates also the high computational cost of a conventional microscopic simulation.

For relatively large coarsening parameter CGMC sampling fails to predict the type of the pattern while the two-level CGMC refines the error introduced and can provide correct feature scales, see Fig. 7(b) and (c). Figs. 7 and 8 show the averaged equilibrium conformations for two total coverage values, $c_0 = 0.9$ and $c_0 = 0.5$ leading to the formation of inverted discs and stripes respectively. Black and white dots indicate occupied and vacant sites respectively. As the coarsening parameter

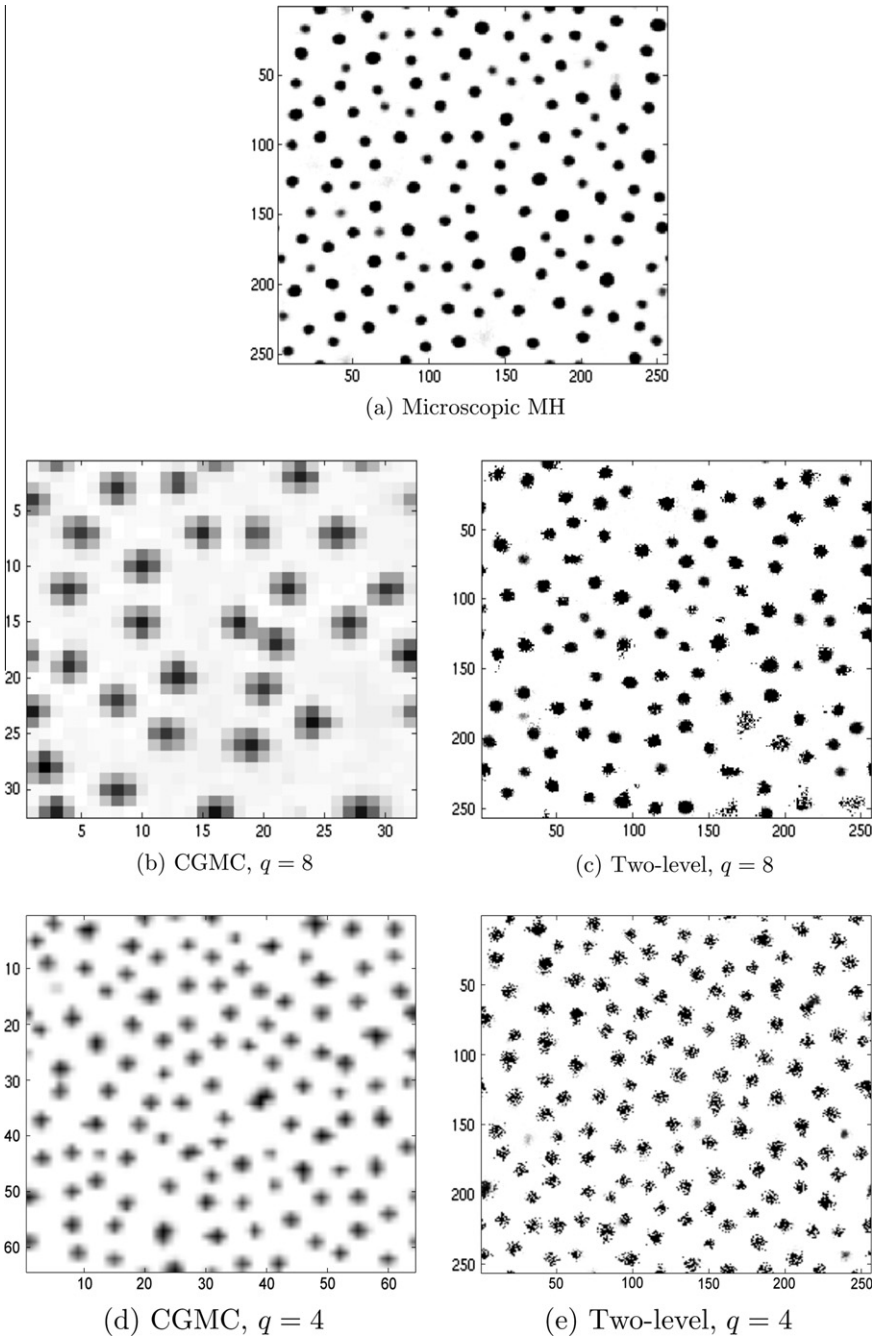

Fig. 8. Inverted discs pattern. Simulations for model (28) with parameters $c_0 = 0.9$, $\beta J_0 = 0.6$, $\chi = 0.1$ and $N = 256 \times 256$.

decreases CGMC prediction becomes better albeit at a higher computational cost, i.e., for $q = 4$ in Fig. 7(d). In such cases the proposed method can be employed as an effective reconstruction method.

A substantial reduction of the computational time compared to the MH method is achieved with the two-level method as is demonstrated in Table 5. Furthermore even for high values of the parameter q the two-level method estimates approximate well the conventional method results.

We also performed numerical tests with the approximate splitting approach, where we split the interactions up to the range S, as was described for the application in Section 5.2, and neglect the correction terms at the second step, see Section 3.1. This strategy captures the qualitative picture but misses the characteristic length scale as is evident in Fig. 9. Model parameters used to generate these figures are the same as in Fig. 7. From the simulations tested we can conclude that this

Table 5
Different resolutions for the two-level CGMC with corrections capture correctly the feature statistics, $N = 256 \times 256$, $c_0 = 0.9$.

Method	Diameter $\langle d \rangle$	Std	CPU time (min)
MH	8.7	1.3	252
Two-level $q = 2$	8.4	1.6	66
Two-level $q = 4$	7.9	4.6	22
Two-level $q = 8$	8.9	3.2	23

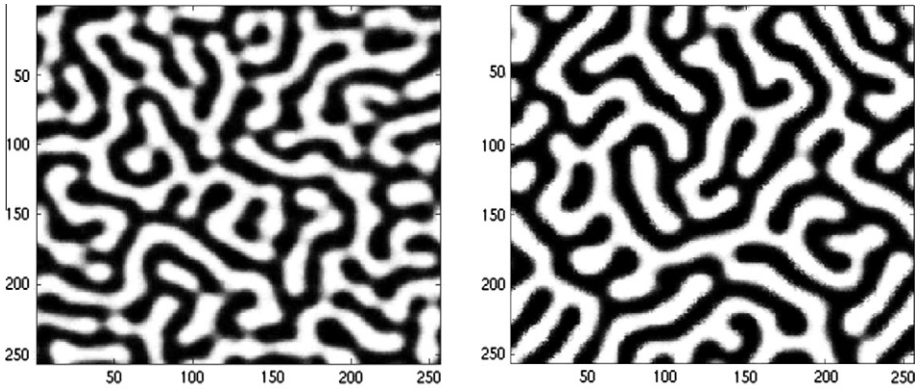

Fig. 9. Splitting interaction potential strategy captures the qualitative picture but misses the characteristic scale.

approach is not recommended since the splitting disturbs significantly the relative strength of attractive and repulsive forces that has an impact on characteristic length scales of the features.

7. Conclusion

In this work we have presented the ML-CGMC method and studied its efficiency theoretically and numerically. The hierarchy of CGMC methods was constructed with introducing multilevel decompositions of the Gibbs measure. The interplay between different resolution levels is achieved with the accept/reject method that incorporates correctly the information lost during compression. The potential splitting strategy for the Kac type model provides good approximations to the conventional simulations correcting the coarsening error, while reduces significantly the computational cost. We successfully apply the method in a nano-pattern discovery problem, verifying that the ML-CGMC method provides the expected behaviour of the system even when CGMC is inaccurate. Numerical tests have shown that a potential splitting in models with competing interactions can disturb the system characteristics, therefore the appropriate strategy is to sample with fully compressed interactions on the coarse space and include a corrective potential at the reconstructing step.

Acknowledgements

The research was supported by the National Science Foundation, E.K. and P.P. under the Grant NSF-CMMI-0835582, M.A.K. under the Grant NSF-CMMI-0835673 and DMS-715125 and D.G.V. under the Grant NSF-CMMI-0835548.

Appendix A. Proofs

Before studying the proposed method's mathematical properties we need to introduce some definitions and theoretical facts [34]. Let $\{X_n\}$ be a Markov chain on space Σ with transition kernel $\mathcal{K}$.

Definition 1

(i) A transition kernel $\mathcal{K}$ has the *stationary measure* μ if

$$\int_{\Sigma} \mathcal{K}(\sigma, \sigma')\mu(d\sigma') = \mu(\sigma), \text{for all } \sigma \in \Sigma.$$

(ii) $\mathcal{K}$ is called reversible with respect to μ if

$$(g,\mathcal{K}f)_\mu = (\mathcal{K}g,f)_\mu, \text{for all } g,f \in L^2(\mu).$$

where $(g,f)_\mu = \int_\Sigma \overline{g(\sigma)} f(\sigma)\mu(d\sigma)$, $\bar{g}$ denoting the complex conjugate of g and $\mathcal{K}g(\sigma) = \int_\Sigma \mathcal{K}(\sigma,d\sigma')g(\sigma'), \forall \sigma \in \Sigma$. A sufficient condition for μ being a stationary measure of $\mathcal{K}$ that is often easy to check is the detailed balance (DB) condition.

Definition 2 (*Detailed Balance*). A Markov chain with transition kernel $\mathcal{K}$ satisfies the detailed balance condition if there exists a function f satisfying

$$\mathcal{K}(\sigma,\sigma')f(\sigma) = \mathcal{K}(\sigma',\sigma)f(\sigma'), \text{for all } \sigma,\sigma' \in \Sigma. \tag{29}$$

Here we continue with the proof of Theorem 1.

Proof

(i) Let $\sigma \neq \sigma'$, recalling the definition of transition kernel $\mathcal{K}_{tl}(\sigma,\sigma')$ (18) we have

$$\begin{aligned}
\mathcal{K}_{tl}(\sigma,\sigma')\mu(\sigma) &= \alpha_f(\sigma,\sigma')\alpha_{CG}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu(\sigma)\\
&= \min\left\{1,\frac{\mu(\sigma')\bar{\mu}^{(0)}(\mathbf{T}\sigma)\mu_r(\sigma|\mathbf{T}\sigma)}{\mu(\sigma)\bar{\mu}^{(0)}(\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')}\right\} \times \min\left\{1,\frac{\bar{\mu}^{(0)}(\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma',\mathbf{T}\sigma)}{\bar{\mu}^{(0)}(\mathbf{T}\sigma)\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')}\right\}\mu_r(\sigma'|\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu(\sigma)\\
&= \min\left\{1,\frac{\mu(\sigma)\bar{\mu}^{(0)}(\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')}{\mu(\sigma')\bar{\mu}^{(0)}(\mathbf{T}\sigma)\mu_r(\sigma|\mathbf{T}\sigma)}\right\} \times \min\left\{1,\frac{\bar{\mu}^{(0)}(\mathbf{T}\sigma)\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')}{\bar{\mu}^{(0)}(\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma',\mathbf{T}\sigma)}\right\}\mu_r(\sigma|\mathbf{T}\sigma)\bar{\rho}(\mathbf{T}\sigma',\mathbf{T}\sigma)\mu(\sigma')\\
&= \mathcal{K}_{tl}(\sigma',\sigma)\mu(\sigma').
\end{aligned}$$

(ii) follows from (i). Detailed balance with $\mu(\sigma)$ is sufficient to guarantee that $\mu(\sigma)$ is the stationary distribution of kernel $\mathcal{K}_{tl}(\sigma,\sigma')$, (Theorem 6.46 in [34]).

(iii) To prove that chain $\{Y_n\}$ is μ-irreducible we need to prove $\mathcal{K}_{tl}(\sigma,A) > 0$, for all $\sigma \in E$ and A measurable such that $\mu(A) > 0$. We have

$$\mathcal{K}_{tl}(\sigma,A) = \int_A \mathcal{K}_{tl}(\sigma,\sigma')d\sigma' \geqslant \int_{A-\{\sigma\}} \mathcal{K}_{tl}(\sigma,\sigma')d\sigma' = \int_{A-\{\sigma\}} \alpha_f(\sigma,\sigma')\alpha_{CG}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')d\sigma'.$$

From assumptions on $\bar{\rho}(\eta,\eta')$ and $\mu_r(\sigma|\eta)$ term $\mu_r(\sigma'|\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')$ is positive for all $\sigma,\sigma' \in E$. Also since $A \subset E$ and $E \subset \widetilde{E} = supp(\mu_0)$, $\alpha_f(\sigma,\sigma')$ and $\alpha_{CG}(\mathbf{T}\sigma,\mathbf{T}\sigma')$ are positive. These ensure that $\mathcal{K}_{tl}(\sigma,A) > 0$.

(iv) A sufficient condition ensuring that $\{Y_n\}$ is aperiodic is that $\mathcal{K}(\sigma,\{\sigma\}) > 0$ for some $\sigma \in E$, that means the event $Y_{n+1} = Y_n$ happens with positive probability. We have

$$\mathcal{K}_{tl}(\sigma,\{\sigma\}) = 1 - \int_{\{\sigma'\neq\sigma\}} \alpha_f(\sigma,\sigma')\alpha_{CG}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')d\sigma'.$$

If for all $\sigma \in \Sigma$, $\mathcal{K}_{tl}(\sigma,\{\sigma\}) = 0$ then

$$\int_{\{\sigma'\neq\sigma\}} \alpha_f(\sigma,\sigma')\alpha_{CG}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')d\sigma' = 1,$$

which means that $\alpha_f(\sigma,\sigma') = 1$ and $\alpha_{CG}(\mathbf{T}\sigma,\mathbf{T}\sigma') = 1$ for almost all $\sigma \in \{\sigma \in E : \bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma') > 0, \text{for some } \sigma' \in E\}$. This would mean that the reconstructed proposal kernel $\bar{\rho}(\mathbf{T}\sigma,\mathbf{T}\sigma')\mu_r(\sigma'|\mathbf{T}\sigma')$ is sampling from the exact target measure μ which in general is not true. Therefore there exists $\sigma \in E$ such that $\mathcal{K}_{tl}(\sigma,\{\sigma\}) = 0$. □

Appendix B. Benchmark example calculations

The detailed calculations of the application of Theorem 2 for the benchmark example described in Section 5.1 are provided here. The conventional MH algorithm is described by a transition kernel proposing a spin-flip at the site $x \in \Lambda_N$ with the probability $1/N$, that is $\rho(\sigma,\sigma') = N^{-1}\sum_{x\in\Lambda_N}\delta(\sigma' - \sigma^{(x)})$, and acceptance probability

$$\alpha(\sigma,\sigma^{(x)}) = \min\left\{1, e^{(1-2\sigma(x))[K_0(\sigma(x-1)+\sigma(x+1))+\frac{J_0}{N}\sum_{y\neq x}^{N}\sigma(y)]}\right\},$$

where, for simplicity, we consider the case of zero external field h. In the two-level CGMC algorithm the potential splitting approach (Section 3.1) determines the energy terms of each step being $\overline{H}_l(\eta) = H_l(\sigma)$ and $H_N(\sigma) - \overline{H}_l(\eta) = H_s(\sigma)$. The pro-

posal probability distribution is $\bar{\rho}(\eta,\eta') = M^{-1}\sum_{k\in\overline{\Lambda}_M}\delta(\eta'-\eta^{(k)})$ where $\eta^{(k)}(l)=\eta(l)$, $l\neq k$ and $\eta^{(k)}(k)=\eta(k)\pm 1$, that implements the adsorption (birth +) or desorption (death −) in the cell k. The acceptance probability of the first level is

$$\alpha_{CG}(\eta,\eta^{(k)}) = \frac{Q-\eta(k)}{Q}\min\left\{1,e^{-\Delta_{k+}\overline{H}_l(\eta)}\right\} + \frac{\eta(k)}{Q}\min\left\{1,e^{-\Delta_{k-}\overline{H}_l(\eta)}\right\}.$$

where $\Delta_{k+}\overline{H}_l(\eta) = -\frac{J_0}{N}\sum_{i\in\overline{\Lambda}_M}\eta(i)$ and $\Delta_{k-}\overline{H}_l(\eta) = -\frac{J_0}{N}\left[1-\sum_{i\in\overline{\Lambda}_M}\eta(i)\right]$. The reconstruction probability distribution is a product of uniform distributions in each cell, that is the probability of constructing $\sigma^{(x)}$ given $\eta^{(k)}$ is $\mu_r(d\sigma^{(x)}|\eta^{(k)}) = \prod_{i\in\overline{\Lambda}_M}\mu_r^{(i)}\left(d\sigma_{C_i}^{(x)}|\eta^{(k)}(i)\right)$, with $\sigma_{C_i}^{(x)} = \{\sigma^{(x)}(y), y\in C_i\}$ and

$$\mu_r^{(i)}\left(d\sigma_{C_i}^{(x)} = \sigma_{C_k}^{(x)}|\eta^{(k)}(k)\right) = \frac{1}{Q},$$

describing the probability of finding an arbitrary $x\in C_k$. Note that at each MC iteration the change in the new state $\eta^{(k)}$ happens in cell k, thus we need to reconstruct $\sigma^{(x)}$, from $\eta^{(k)}$, only in one cell. The second level acceptance probability is

$$\alpha_f(\sigma,\sigma^{(x)}) = \min\{1,e^{-\Delta_x H_s(\sigma)}\} = \min\left\{1,e^{K_0(1-2\sigma(x))(\sigma(x-1)+\sigma(x+1))}\right\}.$$

Term $\mathcal{B}(\sigma,\sigma^{(x)})$, (21), is identically equal to one for all $\sigma\in\Sigma_N$ and $x\in\Lambda_N$. Indeed let $\sigma\in\Sigma_N$ then for all $x\in\Lambda_N$ such that $x\in C_k$ with $\eta = \mathbf{T}\sigma$ and $\eta^{(k)} = \mathbf{T}\sigma^{(x)}$

$$\frac{\bar{\rho}(\eta,\eta^{(k)})\mu_r(\sigma^{(x)}|\eta^{(k)})}{\rho(\sigma,\sigma^{(x)})} = \frac{\frac{1}{M}\frac{1}{Q}}{\frac{1}{N}} = 1.$$

Therefore we have $\underline{\gamma} = \bar{\gamma} = 1$. For any $x\in\Lambda_N$ such that $x\in C_k,\ k\in\overline{\Lambda}_M$ we have from (23)

$$\mathcal{A}(\sigma,\sigma^{(x)}) = \begin{cases} 1, & \text{if}(\sigma,\sigma^{(x)})\in C_1 \\ \min\{e^{-\beta\Delta_k\overline{H}_l(\eta)}, e^{\beta\Delta_k\overline{H}_l(\eta)}\}, & \text{if}(\sigma,\sigma^{(x)})\in C_2 \\ \min\{e^{-\beta\Delta_x H_s(\sigma)}, e^{\beta\Delta_x H_s(\sigma)}\}, & \text{if}(\sigma,\sigma^{(x)})\in C_3 \\ \min\{e^{-\beta\Delta_x H_N(\sigma)}, e^{\beta\Delta_x H_N(\sigma)}\}, & \text{if}(\sigma,\sigma^{(x)})\in C_4 \end{cases}$$

Set $C_4=\emptyset$ according to the following argument. Let $(\sigma,\sigma^{(x)})\in C_4$ such that $\alpha_{CG}(\eta,\eta^{(k)})=1$, $\alpha_f(\sigma,\sigma^{(x)})=1$ and $\alpha(\sigma,\sigma^{(x)})<1$. The first two relations are equivalent to $\Delta_k\overline{H}_l(\eta)\leqslant 0$ and $\Delta_x H_s(\sigma)\leqslant 0$ that imply $\Delta_x H_N(\sigma)\leqslant 0$, since $H_N(\sigma) = H_s(\sigma)+\overline{H}_l(\eta)$, thus $\alpha(\sigma,\sigma^{(x)})=1$, a contradiction. Analogous argument holds for the case $\alpha(\sigma,\sigma^{(x)})=1$, $\alpha_{CG}(\eta,\eta^{(k)})<1$, $\alpha_f(\sigma,\sigma^{(x)})<1$ that proves $C_4=\emptyset$.

Consider $K=0$ then set $C_2=\emptyset$ that can be proved with a simple argument. $\mathcal{A}(\sigma,\sigma^{(x)})=1$ for all $(\sigma,\sigma^{(x)})\in C_3$ which is also easy to prove based on the fact that $\Delta_x H_s(\sigma)=0$. Therefore when $K_0=0$, $\mathcal{A}(\sigma,\sigma^{(x)})=1$, for all $\sigma\in\Sigma_N$ and all $x\in\Lambda_N$.

For $K_0\neq 0$, using the analytic expression of $\Delta_x H_s(\sigma)$ and $\Delta_k\overline{H}_l(\eta)$ we have, for $(\sigma,\sigma^{(x)})\in C_2$

$$\min\left\{e^{-\beta\Delta_{k+}\overline{H}_l(\eta)}, e^{\beta\Delta_{k+}\overline{H}_l(\eta)}\right\} \geqslant e^{-|\beta\Delta_{k+}\overline{H}_l(\eta)|}\big|_{\{\eta(l)=Q,\forall l\in\overline{\Lambda}_M\}} = e^{-|\frac{J_0}{N}N|} = e^{-|J_0|},$$

and

$$\min\left\{e^{-\beta\Delta_{k-}\overline{H}_l(\eta)}, e^{\beta\Delta_{k-}\overline{H}_l(\eta)}\right\} \geqslant e^{-|\frac{J_0}{N}(N-1)|} > e^{-|J_0|}.$$

Similarly for $(\sigma,\sigma^{(x)})\in C_3$

$$\min\left\{e^{-\beta\Delta_x\overline{H}_s(\sigma)}, e^{\beta\Delta_x\overline{H}_s(\sigma)}\right\} \geqslant e^{-2|K_0|}.$$

Therefore $\inf_{\sigma,x}\mathcal{A}(\sigma,\sigma^{(x)}) = \mathcal{A}(\sigma,\sigma^{(x)})|_{\{\sigma:\sigma(x)=1,\forall x\in\Lambda_N\}} = \min\{e^{-|J_0|}, e^{-2|K_0|}\}$.

Remark. We should note that the reconstruction used in implementations is slightly different from the one just described, taking advantage of the knowledge from the coarse step whether an adsorption or desorption is proposed,

$$\mu_r\left(d\sigma_{C_l}^{(x)} = \sigma_{C_k}^{(x)}|\eta^{(k)}(k)\right) = \frac{1}{Q-\eta(k)}\delta(k-k+) + \frac{1}{\eta(k)}\delta(k-k-),$$

describing the probability of finding $x\in C_k$ with $\sigma(x)=0$ for $k=k+$ and $\sigma(x)=1$ for $k=k-$.

References

[1] G. Arampatzis, M.A. Katsoulakis, P. Plecháč, M. Taufer, L. Xu, Hierarchical fractional-step approximations and parallel kinetic Monte Carlo algorithms, submitted for publication, arXiv:1105.4673.
[2] S. Are, M.A. Katsoulakis, P. Plecháč, L. Rey-Bellet, Multibody interactions in coarse-graining schemes for extended systems, SIAM J. Sci. Comput. 31 (2008) 987–1015.

[3] D.M. Ceperley, Path integrals in the theory of condensed helium, Rev. Mod. Phys. 67 (1995) 279–355.
[4] A. Chatterjee, D. Vlachos, Multiscale spatial Monte Carlo simulations: Multigriding, computational singular perturbation, and hierarchical stochastic closures, J. Chem. Phys. 124 (2006) 064110.
[5] A. Chatterjee, D. Vlachos, Systems tasks in nanotechnology via hierarchical multiscale modeling: Nanopattern formation in heteroepitaxy, Chem. Eng. Sci. 62 (18–20) (2007) 4852–4863.
[6] J. Dai, W.D. Seider, T. Sinno, Coarse-grained lattice kinetic Monte Carlo simulation of systems of strongly interacting particles, J. Chem. Phys. 128 (2008) 194705.
[7] A. DeMassi, E. Presutti, Mathematical Methods for Hydrodynamic Limits, Lecture Notes in Mathematics, Springer, New York Berlin Heidelberg, 1991.
[8] P. Diaconis, L. Saloff-Coste, Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Prob. 6 (1996) 695–750.
[9] S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, Hybrid Monte Carlo, Phys. Lett. B 195 (1987) 216–222.
[10] Y. Efendiev, T. Hou, W. Luo, Preconditioning Markov chain Monte Carlo simulations using coarse-scale models, SIAM J. Sci. Comput. 28 (2006) 776–803.
[11] N. Goldenfeld, Lectures on Phase Transitions and the Renormalization Group, vol. 85, Addison-Wesley, New York, 1992.
[12] V.A. Harmandaris, N.P. Adhikari, N.F.A. van der Vegt, K. Kremer, Hierarchical modeling of polystyrene: From atomistic to coarse-grained simulations, Macromolecules 39 (2006) 6708.
[13] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970) 97–109.
[14] E. Kalligiannaki, M.A. Katsoulakis, P. Plechac, Coupled coarse graining and Markov chain Monte Carlo for lattice systems, Numerical Analysis and Multiscale Computations, Lect. Notes Comput. Sci. Eng. 82 (2011).
[15] M. Kardar, Crossover to equivalent-neighbor multicritical behavior in arbitrary dimensions, Phys. Rev. B 28 (1983) 244–246.
[16] M. Katsoulakis, A. Majda, D. Vlachos, Coarse-grained stochastic processes for microscopic lattice systems, Proc. Natl. Acad. Sci. 100 (2003) 782–787.
[17] M.A. Katsoulakis, A.J. Majda, D.G. Vlachos, Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems, J. Comput. Phys. 186 (2003) 250–278.
[18] M.A. Katsoulakis, P. Plecháč, L. Rey-Bellet, Numerical and statistical methods for the coarse-graining of many-particle stochastic systems, J. Sci. Comput. 37 (2008) 43–71.
[19] M.A. Katsoulakis, P. Plecháč, L. Rey-Bellet, D.K. Tsagkarogiannis, Coarse-graining schemes and a posteriori error estimates for stochastic lattice systems, ESAIM-Math. Model. Numer. Anal. 41 (2007) 627–660.
[20] M.A. Katsoulakis, P. Plecháč, A. Sopasakis, Error analysis of coarse-graining for stochastic lattice dynamics, SIAM J. Numer. Anal. 44 (2006) 2270–2296.
[21] M.A. Katsoulakis, L. Rey-Bellet, P. Plecháč, D.K. Tsagkarogiannis, Mathematical strategies in the coarse-graining of extensive systems: error quantification and adaptivity, J. Non Newt. Fluid Mech. (2008).
[22] M.A. Katsoulakis, L. Rey-Bellet, P. Plecháč, D.K. Tsagkarogiannis, Coarse-graining schemes for stochastic lattice systems with short and long range interactions, submitted for publication, arXiv:1003.1506.
[23] M.A. Katsoulakis, J. Trashorras, Information loss in coarse-graining of stochastic particle dynamics, J. Stat. Phys. 122 (2006) 115–135.
[24] M.A. Katsoulakis, D.G. Vlachos, Coarse-grained stochastic processes and kinetic Monte Carlo simulators for the diffusion of interacting particles, J. Chem. Phys. 119 (2003) 9412–9427.
[25] K. Kremer, F. Müller-Plathe, Multiscale problems in polymer science: Simulation approaches, MRS Bull. (March 2001) 205.
[26] T.M. Liggett, Interacting Particle Systems, Springer, New York Berlin Heidelberg, 1985.
[27] J.S. Liu, Monte Carlo Strategies in Scientific Computing, Springer-Verlag, New York Berlin Heidelberg, 2001.
[28] J. Liu, C. Sabatti, Simulated Sintering: Markov chain Monte Carlo with spaces of varying dimensions, Bayesian Statistics, J.M. Bernardo, J.O. Berger, A.P. Dawid, A.F.M. Smith (Eds.), vol. 6, 1998, pp. 402–405.
[29] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21 (1953) 1087–1092.
[30] F. Müller-Plathe, Coarse-graining in polymer simulation: From the atomistic to the mesoscale and back, Chem. Phys. Chem. 3 (2002) 754.
[31] R.M. Neal, An improved acceptance procedure for the hybrid Monte Carlo algorithm, J. Comput. Phys. 111 (1994) 194–203.
[32] J. Parker, Algorithms for Image Processing and Computer Vision, Wiley, 1996.
[33] J.S. Reese, S. Raimondeau, D.G. Vlachos, Monte Carlo algorithms for complex surface reaction mechanisms: Efficiency and accuracy, J. Comput. Phys. 173 (1) (2001) 302–321.
[34] C.P. Robert, G. Casella, Monte Carlo Statistical Methods, Springer-Verlag, New York, 2004.
[35] B. Simon, The statistical mechanics of lattice gases, Princeton series in Physics, vol. I, 1993.
[36] J.M. Taylor, P. Plecháč, Multi-level coarse graining methods for sampling stochastic particle systems.
[37] W. Tschöp, K. Kremer, O. Hahn, J. Batoulis, T. Bürger, Simulation of polymer melts. II. From coarse-grained models back to atomistic description, Acta Polym. 49 (1998) 75.