

Hierarchical Pattern Discovery in Stochastic Lattice Systems

Petr Plecháč

Joint Institute for Computational Sciences
University of Tennessee and Oak Ridge National Laboratory

Markos A. Katsoulakis

Department of Mathematics & Statistics
University of Massachusetts

Dionisios G. Vlachos

School of Chemical Engineering
University of Delaware

Abstract. We develop a hierarchical approach for pattern discovery in many-body stochastic systems, motivated by challenges in guiding engineering tasks for nanopattern formation in heteroepitaxial processes. Patterns in such systems have rich morphologies at mesoscales that change dramatically as control parameters vary; typically they form as a result of microscopic particle dynamics in a complex energy landscape, in the presence of stochastic fluctuations.

Studying pattern formation mechanisms as functions of the control parameters of the system poses a significant computational challenge which is currently intractable with conventional Kinetic Monte Carlo (kMC) methods. We present hierarchical strategies towards this *systems' task* goal by combining mesoscopic PDE and Coarse-Grained Monte Carlo (CGMC) approximations of kMC algorithms that we have developed in our earlier work. More precisely, (i) we employ deterministic mesoscopic PDE as means to obtain an approximate (and in principle rather crude) phase diagram of the system; subsequently, (ii) we employ adaptive CGMC at selected regions of the approximate phase diagram in order to refine it by including interactions and fluctuations properly. Our adaptivity framework allows us to obtain accurate and near-optimal coarse-grainings for each parameter regime, ensuring proper “knowledge representation”– in the information theory sense – of the complex system for the desired observables, e.g., spatial correlations, power spectra or scaling laws. In turn such tools can be also used in model reduction and control of the underlying complex systems.

1. Introduction. The recent quest for forming periodic arrays of metal clusters and quantum dots of a material on crystal surfaces of a different metal or semiconductor, i.e., heteroepitaxy, has motivated the development of models that can ultimately aid with the design and control of nanotechnology-related processes. To achieve these goals we focus on developing models and simulation techniques based on kinetic Monte Carlo (kMC) that correctly capture a rich variety of experimentally observed pattern shapes, sizes, orientations and densities. These observables depend on a number of experimentally controllable factors such as temperature, film thickness and material properties ([17]). While phenomenological, equilibrium free-energy-based models ([4, 15]) can predict the most energetically stable patterns at zero temperature incorporating correctly the microscopic physics, entropic effects and thermal fluctuations is essential for predicting nucleation and growth of nanopatterns. Simulation at large spatio-temporal scales, let alone generation of phase diagrams and detection of metastable patterns is limited in simulations by large separation of lengths (1 nm–0.1 μm) and temporal scales (1 ps–1 min). In the last few years we have been developing a bottom-up hierarchy of coarse-grained approximations that leads to a class of efficient Monte Carlo simulations (CGMC), that allows us to simulate at large spatio-temporal scales where morphological features are observed, and are determined by the interplay between thermal fluctuations and molecular interactions. We refer to the review articles [2, 14]. However, in order to obtain consistent and accurate predictions of phase diagrams from the CGMC models, the level of coarse-graining has to be linked to a reliable error estimation.

2. Coarse-graining in pattern formation problems.. In spite of the tremendous speed-up and significant improvements of the CGMC algorithms, it is still computationally costly to use the CGMC method for *systems' tasks*, such as generating a phase diagram for complex systems and nano-structure control (e.g., a simulation study depicted in Fig. 4.2 takes about a day to perform). To overcome this computational challenge, the authors proposed in the recent work [3] a systematic top-down modeling approach in order to efficiently study pattern formation by using the hierarchy of our previously developed coarse-grained models, [13], combining mesoscopic equations and CGMC: a first (less accurate) approximation of the phase diagram of nanopatterns is generated via linear and nonlinear analysis of

mesoscopic equations, Fig. 4.1; “educated” CGMC simulations, which include thermal fluctuations, are subsequently performed at selected regions of the phase diagram to refine it and provide additional information, such as the nanoparticle size, shape, and spacing distributions. This top-down approach is feasible because the coarse-grained models are derived from the same kMC model, and describe the same essential physics at different levels of CG.

On the other hand in [10] we developed, for the first time in the literature, an automatic (on-the-fly) CG/refinement method that recovers accurately phase-diagrams for systems with competing complex interactions, by employing a posteriori error estimates for the loss of information during CG: the estimates allow us to *change adaptively* the CG level within the CG hierarchy, once suitably large or small errors are detected, and dramatically speed up the calculations of phase diagrams, see Fig. 3.1.

These two recent results on performing certain systems tasks using CGMC and mesoscopic PDE form the basis of our proposed approach for *pattern discovery* in complex stochastic systems:

- (i) we employ deterministic mesoscopic PDE as means to obtain an approximate, and in principle rather crude, phase diagram of the system;
- (ii) subsequently, we use Adaptive CGMC (AdCGMC) simulations at selected regions of the approximate phase diagram in order to refine it by including properly coarse-grained interactions and fluctuations.

Our adaptivity framework allows us to obtain accurate and near-optimal coarse-grainings for each parameter regime, ensuring proper “knowledge representation” – in the information theory sense – of the complex system for the desired observables, e.g., spatial correlations, power spectra or scaling laws. In turn such tools can be also used in model reduction and control of the underlying complex systems.

2.1. Coarse-grained Hamiltonians and error estimates. Interactions in particle systems modelling the heteroepitaxial growth are described by the microscopic Hamiltonian $H(\sigma)$ that defines total interaction energy of the system associated with the configuration σ . The simplest models are formulated in terms of Ising-type spin variables $\sigma(x) \in \{0, 1\}$ which describes whether a particle is present at the site $x \in \Lambda_N$ of the lattice Λ_N with N sites. For the sake of simplicity we demonstrate the ideas based on real-space coarse-graining, i.e., by defining coarse-variables (observables) derived from coarsening the original microscopic lattice (i.e., defining the block-spin variables $\eta(k)$ on a coarser lattice Λ_M). In our earlier work (e.g., [7, 11, 13]) we derived the coarse-grained Hamiltonian

$$\begin{aligned} \bar{H}^{(0)}(\eta) = & -\frac{1}{2} \sum_{l \in \bar{\Lambda}_M} \sum_{k \neq l, k \in \bar{\Lambda}_M} \bar{J}(k, l) \eta(k) \eta(l) - \\ & \frac{1}{2} \bar{J}(0, 0) \sum_{l \in \bar{\Lambda}_M} \eta(l) (\eta(l) - 1) + \sum_{k \in \bar{\Lambda}_M} \bar{h} \eta(k). \end{aligned} \quad (2.1)$$

where the two-body CG potential \bar{J} is defined by the local average or equivalently as an interaction between two block-spins

$$\begin{aligned} \bar{J}(k, l) = & \frac{1}{q^2} \sum_{x \in D_k} \sum_{y \in D_l} J(|x - y|) \quad \text{and} \\ \bar{J}(k, k) = & \frac{1}{q(q-1)} \sum_{x \in D_k} \sum_{y \in D_k} J(|x - y|). \end{aligned} \quad (2.2)$$

In [11] our goal was to develop more accurate coarse-graining schemes than those proposed in [7, 13], and quantify their effectiveness in terms of a priori and a posteriori error analysis. The work [11] can also be viewed as a blueprint for tackling more complex problems described in the subsequent sections.

The principle idea is motivated by the renormalization group map, [5, 6]

$$e^{-\beta \bar{H}_M(\eta)} = \int e^{-\beta H_N(\sigma)} P_N(d\sigma | \eta), \quad (2.3)$$

where $\bar{H}_M(\eta)$ is, by definition, the *exactly coarse-grained* Hamiltonian and $P_N(d\sigma | \eta)$ is the conditional probability (with respect to the prior distribution P_N) of having a microscopic configuration σ given a CG configuration η . Furthermore, we denote the coarse-graining operator $\eta = \mathbf{T}\sigma = \sum_{x \in D_k} \sigma(x)$. Note that, due to the high-dimensional

integration, $\bar{H}_M(\eta)$ cannot be calculated explicitly and used in numerical simulations. Our approach is to approximate it by viewing it as a perturbation of $\bar{H}^{(0)}$ in (2.1). Using this first approximation we have

$$\begin{aligned}\bar{H}_M(\eta) &= \bar{H}_M^{(0)}(\eta) - \\ &\frac{1}{\beta} \log \int e^{-\beta(H_N(\sigma) - \bar{H}_M^{(0)}(\eta))} P_N(d\sigma|\eta).\end{aligned}\quad (2.4)$$

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

$$\bar{H}_M(\eta) = \bar{H}_M^{(0)}(\eta) + \bar{H}_M^{(1)}(\eta) \cdots + \bar{H}_M^{(p)}(\eta) + N \mathcal{O}(\epsilon^{p+1}), \quad (2.5)$$

uniformly in η ; also recall that Hamiltonian scales linearly with N , hence the $N \times \mathcal{O}(\epsilon^{p+1})$ term. The small parameter ϵ is given explicitly and encapsulates the dependence of the error on parameters such as temperature ($\beta = 1/kT$), regularity and range of the inter-particle potential, and the level of coarse-graining (the size of the block spins), [14].

While (2.5) gives error bounds at the level of the Hamiltonian, it is important to have error bounds for the corresponding canonical Gibbs measure, since the latter determines the most probable states η at equilibrium. Truncating the expansion we obtained the following Gibbs measures

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

At the level of Gibbs distributions coarse-graining can be also viewed as an information compression and therefore it is natural to measure errors in this context in terms of the relative entropy which quantifies the information loss. Recall that for two probability distribution π_1 and π_2 defined on a common finite state space \mathcal{X} , the relative entropy of π_1 with respect to π_2 is defined as

$$\mathcal{R}(\pi_1 | \pi_2) = \sum_{\sigma \in \mathcal{X}} \pi_1(\sigma) \log \frac{\pi_1(\sigma)}{\pi_2(\sigma)}.$$

Furthermore, since we are dealing with extended systems and compressing local interactions, the errors will be extensive quantities and it is thus natural to measure the error per unit volume, i.e., in terms of the relative entropy per unit volume. Using (2.5), one can prove the following estimates for the loss of information during coarse-graining [11]

THEOREM 2.1. (Relative entropy error bounds)

$$\begin{aligned}\frac{1}{N} \mathcal{R}(\bar{\mu}_{M,\beta}^{(0)} | \mu_{N,\beta} \circ \mathbf{T}^{-1}) &= \mathcal{O}(\epsilon^2), \\ \frac{1}{N} \mathcal{R}(\bar{\mu}_{M,\beta}^{(p)} | \mu_{N,\beta} \circ \mathbf{T}^{-1}) &= \mathcal{O}(\epsilon^{p+2}),\end{aligned}$$

where $p = 1, \dots$ and ϵ is given explicitly from the expansion of the Hamiltonian.

2.2. Coarse-graining schemes. The first scheme is based on the approximation that uses $H_M^{(0)}$ and it has been extensively studied in [7, 8, 9, 12]. The first result concerns the approximation of long-time behavior in the kMC simulations that use the CG Hamiltonian $\bar{H}_M^{(0)}$.

SCHEME 2.1 (2^{nd} -order coarse-graining). *The 2nd-order coarse-graining algorithm has the following characteristics*

1. *Hamiltonian:* $\bar{H}_M^{(0)}$, with inter-particle interactions defined by (2.2).
2. *Gibbs measure:*

$$\bar{\mu}_{M,\beta}^{(0)}(d\eta) = \frac{1}{\bar{Z}_M^{(0)}} e^{-\beta \bar{H}_M^{(0)}(\eta)} \bar{P}_M(d\eta).$$
3. *Relative entropy error:*

$$\frac{1}{N} \mathcal{R}(\bar{\mu}_{M,\beta}^{(0)} | \mu_{N,\beta} \circ \mathbf{T}^{-1}) = \mathcal{O}(\epsilon^2).$$

This scheme is second-order accurate.

The correction terms $\bar{H}_M^{(1)}(\eta)$, $\bar{H}_M^{(2)}(\eta)$ etc. include multi-body interactions and are calculated explicitly in [11]. The term $\bar{H}_M^{(0)}(\eta)$ consists of only two-body interactions whereas $\bar{H}_M^{(1)}(\eta)$ includes a correction to the two-body interactions and a new term that consists of three-body interactions. All detailed formulas for the CG interactions can be found in [11].

As an example we define the improved CG scheme with $p = 1$.

SCHEME 2.2 (3^{rd} -order accurate). *Hamiltonian:* $\bar{H}_M^{(0)} + \bar{H}_M^{(1)}$, where the corrections are

$$\begin{aligned} \bar{H}_M^{(1,1)}(\eta) = & \sum_{k < l} \Lambda_2^{(1)}(k, l; \eta(k), \eta(l)) + \\ & \sum_k \Lambda_1^{(1)}(k; \eta(k)), \end{aligned} \quad (2.6)$$

and

$$\begin{aligned} \bar{H}_M^{(1,2)}(\eta) = & \sum_{k \neq l} \Lambda_2^{(2)}(k, l, k; \eta(k), \eta(l), \eta(k)) \\ & + \sum_{k < l < m} \Lambda_3^{(2)}(k, l, m; \eta(k), \eta(l), \eta(m)). \end{aligned} \quad (2.7)$$

The approximation order of the scheme is given by the error estimate in terms of the specific relative entropy

$$\mathcal{R}(\bar{\mu}_{M,\beta}^{(2)} | \mu_{N,\beta} \circ \mathbf{T}^{-1}) = \mathcal{O}(\epsilon^3), \quad (2.8)$$

2.3. Compression of multi-body interactions. The multi-body interactions such as the three-body term in $\bar{H}_M^{(1)}$ of Scheme 2.2 can be computationally expensive due to the potentially large number of three-body terms. Hence it is important to understand when the multi-body interactions are necessary in a CG scheme in order to achieve an error for a given tolerance in (2.6). When multi-body corrections are indeed required, we study how they can be compressed via suitable truncations. For example, the expression (2.6) and the explicit dependence of ϵ on properties of the interaction potential suggest that in coarse-graining smooth long-range potentials the scheme with only two-body interactions is an accurate approximation and does not require the computationally expensive multi-body interactions of Scheme 2.2. On the other hand, for singular potentials we can make use of decay properties of the interactions and show that the multi-body interactions can be compressed by truncating them to a given tolerance. see [1]. For example, truncations of two- and three-body terms can be essentially restricted to the nearest neighbor triplets only, in all dimensions, due to the decay properties of typical inter-particle potentials J . For example, we can truncate the correction term $\bar{H}_M^{(1,2)}$ as

$$\begin{aligned} \bar{H}_M^{(1,2)}(\eta) = & \sum_k \Lambda_2^{(2)}(k, k+1, k; \eta(k), \eta(k+1), \eta(k)) \\ & + \sum_k \Lambda_3^{(2)}(k-1, k, k+1; \eta(k-1), \eta(k), \eta(k+1)). \end{aligned}$$

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

Process	CPU (secs)
q=1 (no coarse-graining)	322192.06
q=8 ($H^{(0)}$ only)	5232.62
q=8c (no compression)	69473.09
q=8c (compression)	6900.72

TABLE 2.1

CPU cost comparisons of different CG algorithms for potentials with a singularity, cite AKPR. The cases $q = 8c$ include higher order corrections (i.e., multibody interactions) and they are treated with and without compression. The CPU time is for calculating the phase diagram using the same number of samples. Note that in this simulation the case $q = 8$ failed to approximate the true phase curve.

Further strategies to decrease the computational cost of implementing the multi-body interactions are discussed in [1]. We emphasize that the CG approximation may fail to capture the phenomena we want to observe in the simulation. This is demonstrated in Fig 2.1 where we compare probability distribution functions for the transition time between two metastable states (low and high total coverage) in an Ising-type lattice model with a piece-wise constant interaction potential. While the potential is long-range the CG scheme that uses $H^{(0)}$ only and $q = 100$ predicts a wrong mean transition time. The higher order scheme gives a reasonable approximation even in this case ($q = 100c$). Therefore it is important to find a posteriori error indicators which will alert us that the CG simulation misses important features of the full microscopic model. We discuss this aspect of our work in the next section.

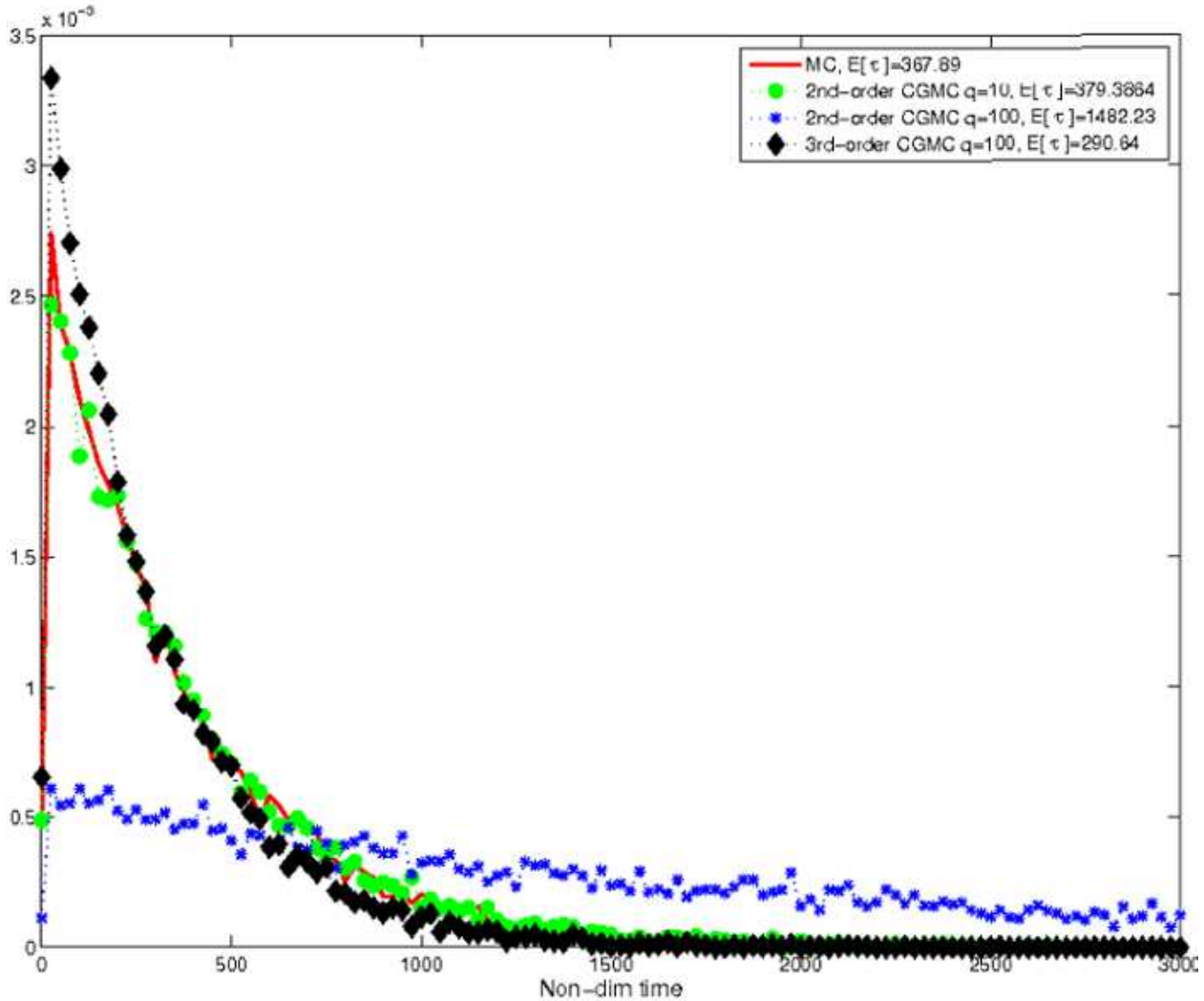


FIG. 2.1. Probability density function (PDF): comparisons between different coarse-graining levels q . The estimated mean times for each PDF are shown in the legend.

3. A posteriori estimation and on-the-fly error estimation. The error estimate in Theorem 2.1, along with the cluster expansion provide us with an explicit representation of the error in the coarse-grained numerical approximation. For instance, in [11] we showed the following a posteriori error for Scheme 2.1

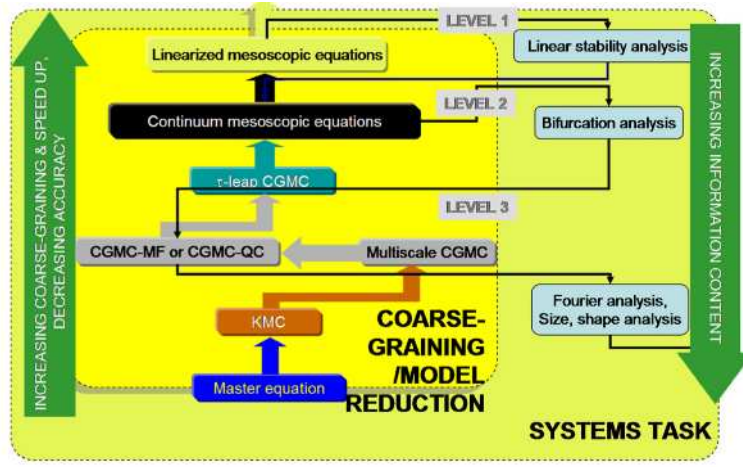


FIG. 2.2. Systems tasks tools (outer box) are used in conjunction with hierarchical coarse-grained models (inner box) derived from kinetic Monte Carlo models for studying pattern formation and ultimately a kinetic phase diagram. ([3])

THEOREM 3.1. (A posteriori error) 1. *Loss of information estimate:*

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

where the residuum operator is $R(\eta) = \bar{H}_M^{(1,1)}(\eta) + \bar{H}_M^{(1,2)}(\eta)$.

2. By formally expanding in the Taylor series we get the (practical) error indicator

$$\text{err} \sim \mathbb{E}_{\bar{\mu}_{M,\beta}^{(0)}} \left[\frac{R(\eta) - \mathbb{E}_{\bar{\mu}_{M,\beta}^{(0)}} R(\eta)}{\sqrt{N}} \right]^2. \quad (3.1)$$

3. A posteriori estimates for a given observable $\phi(\eta)$:

$$\mathbb{E}\phi(\mathbf{T}\sigma) - \mathbb{E}\phi(\eta) \sim \mathbb{E}\phi(\eta) \left[1 - \frac{e^{-R(\eta)}}{\mathbb{E}e^{-R(\eta)}} \right], \quad (3.2)$$

Observables relevant to patterning: Typical *target quantities* that we would like to eventually control depend only on *coarse observables*, for example, (i) nanoparticle size, (ii) shape, (iii) spacing (or density of nanoparticles), and (iv) spatial organization (e.g., the type of pattern, such as stripes, hexagons, squares, and the degree of perfection in an array). Mathematically, the morphology of a patterned surface can be characterized by a spatial average of the 2-point correlation function of the local coverage $\eta_t = \mathbf{T}\sigma_t$

$$G(k, t) = \frac{1}{M^2} \sum_{i \in \Lambda_N} \eta_t(i) \eta_t(i+k) - \mathbb{E}[\eta_t(i)] \mathbb{E}[\eta_t(i+k)].$$

Regular patterns are then more easily characterized by the spectral density function defined as the Fourier image $S \equiv \hat{G}$.

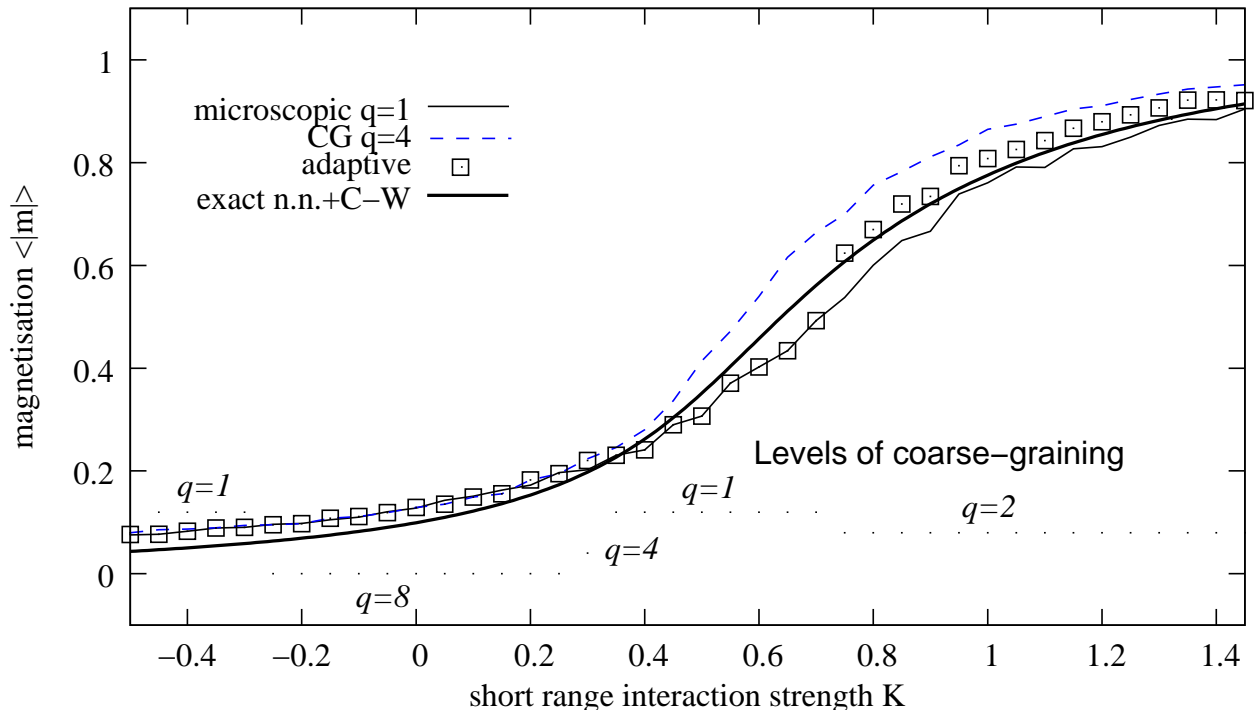


FIG. 3.1. Demonstration of computing a phase diagram for a system with competing short and long-range interactions using adaptive CGMC. The figure compares the adaptive strategy (the levels of coarse-graining, i.e., the size of block-spins is noted in the figure) with fully resolved (microscopic) simulation ($q = 1$), with a fixed coarse-graining and with the exact solution which is available in this example. For more details see [10]

In [10] we demonstrated the use of adaptive diagnostics and the ensuing adaptive coarse-grainings in the numerical calculation of phase diagrams in systems with combined short and long-range interactions, see for instance Fig. 3.1. We discuss further this perspective in the following section.

4. Hierarchical strategies for pattern discovery. To overcome the computational challenge, the authors proposed in [3] a systematic top-down modeling approach in order to efficiently study pattern formation by using hierarchically our coarse-grained models (in particular, combination of mesoscopic equations and CGMC): a less accurate phase diagram of nanopatterns is generated via linear and nonlinear analysis of mesoscopic equations, Fig. 2.2; “educated” CGMC simulations, which include thermal fluctuations, are subsequently performed at selected regions of the phase diagram to refine it and provide additional information, such as the nanoparticle size, shape, and spacing distributions.

The approach was applied in [3] to simulating an ensemble of N_p atoms chemisorbed on N_s lattice sites of a (100) surface. Only one adsorbed atom per lattice site is allowed and thus the model can be simulated by Ising-type spins σ . An adatom can interact with another adatom via a short-range attractive and a long-range repulsive two-body potential. The form of potentials was chosen to be a combination of two Gaussians

$$\beta J(r) = \beta J_0 (e^{-(r/r_a)^2} - \chi e^{-(r/r_r)^2}).$$

The selected form of the potential has an explicit Fourier transform, which in turn enables derivation of explicit formulas for the wavelength of emerging patterns and for the conditions leading to the onset of patterns providing insights into the patterning process. The range of repulsion has been estimated using linear elasticity theory. The model also includes microscopic surface diffusion with a diffusion transition rate

$$\Gamma = \frac{1}{4} \Gamma_m e^{-\beta U(x)} \sigma(x) (1 - \sigma(y)).$$

This process describes a random hop of an adatom at the site x of the lattice to a vacant site y . Arrhenius dynamics is assumed for surface diffusion and the molecular potential $U(x)$ is linked to the interaction potential by $U(x) = \sum_{y \neq x} J(x-y)\sigma(y)$. In the CGMC simulation U is replaced by its coarse-grained approximation \bar{U} that involves \bar{J} . In the limit of long-range interactions we can employ a mean-field approximation in which the site occupancy is replaced by the local coverage c . The ensemble-averaged master equation yields the continuum mesoscopic equation describing the evolution of the local coverage c

$$\partial_t c = -\nabla \cdot (e^{-\beta V} (\nabla c - c(1-c)\nabla V)),$$

where $V(x) = \int J(x-y)c(y) dy$. This mesoscopic equation was derived and studied in [18]. In the present context the mesoscopic equation is used to obtain a first approximation of the phase diagram. Due to the choice of the potential J the linear stability and bifurcation analysis can be carried explicitly as the convolution $J * c$ can be handled in the Fourier space, see [16]. Such analysis produces the phase diagram in Fig 4.1.

On the other hand we have developed for the first time in the literature, an automatic (on-the-fly) CG/refinement method that recovers accurately phase-diagrams for systems with competing complex interactions, by employing a posteriori error estimates for the loss of information during CG: the estimates allow us to *change adaptively* the CG level within in the CG hierarchy, once suitably large or small errors are detected, and dramatically speed up the calculations of phase diagrams, [10]. In the cases presented in [10] it turned out that most of the phase diagram is constructed using coarse levels and hence inexpensive CG simulations are used, while the relatively fewer regimes where critical phenomena occur, require finer, or even fully resolved simulations. The transitions from finer to coarser scales and back are done on-the-fly, based on the a posteriori error computation. The refinement or coarsening in [10] was governed by the error indicator in (3.1), although this indicator does not easily relate to the absolute error of a given observable (e.g., local coverage, correlations, etc.). In the presented simulations in Fig. 3.1 a simple strategy has been adopted: the change of the level of coarse-graining is controlled by the relative value of the indicator with respect to its maximal value along the simulation path. For controlling error in observables such as the 2-point spatial correlation function, which is a more appropriate choice for the patterning applications considered here, we can employ the new estimator in (3.2).

These two recent studies on performing (certain) systems tasks using Adaptive CGMC and mesoscopic PDE form the basis of our proposed approach for *pattern discovery* in complex stochastic systems:

Step 1: First, we obtain a rather "crude" phase diagram by employing deterministic mesoscopic PDE, e.g., [13, 3], derived from the microscopic model under the assumption of a mean-field approximation.

Step 2: Next, we perform Adaptive CGMC (AdCGMC) at selected regions of the phase diagram obtained in Step 1 in order to refine it and include fluctuations. A (near) optimal level of CG can be obtained for each parameter regime, ensuring proper representation of the complex system on the desired observables, e.g., spatial correlations, power spectra, scaling laws, etc.

Single-level CGMC can provide inaccurate phase diagrams, e.g., $q = 4$ in Fig. 3.1 (see also examples in [1]), especially in parameter regimes where we may have over-coarse-grained the typically unknown mesoscopic morphologies. On the other hand, Adaptive CGMC is essentially a *multi-level* algorithm, at least as applied in the evaluation of phase diagrams, exchanging information between scales in an error-controlled manner, as shown in (3.1); thus, through the minimization of the a posteriori error our approach will yield a (near) optimal CGMC, providing reliable fast simulators for pattern formation problems.

REFERENCES

- [1] Sasanka Are, Markos A. Katsoulakis, Petr Plecháč, and Luc Rey-Bellet. Multibody interactions in coarse-graining schemes for extended systems. *SIAM J. Sci. Comput.*, 31(2):987–1015, 2008.
- [2] A. Chatterjee and D. G. Vlachos. An overview of spatial microscopic and accelerated kinetic Monte Carlo methods for materials' simulation. *J. Computer-Aided Materials Design*, 14(2):253, 2007.
- [3] A. Chatterjee and D.G. Vlachos. Systems tasks in nanotechnology via hierarchical multiscale modeling: Nanopattern formation in heteroepitaxy. *Chem. Eng. Sci.*, 62(18-20):4852–4863, 2007.
- [4] I. Daruka and A. L. Barabasi. equilibrium phase diagrams for dislocation free self-assembled quantum dots. *Appl. Physics Lett.*, 72:2102–2104, 1998.
- [5] N. Goldenfeld. *Lectures on Phase Transitions and the Renormalization Group*, volume 85. Addison-Wesley, New York, 1992.

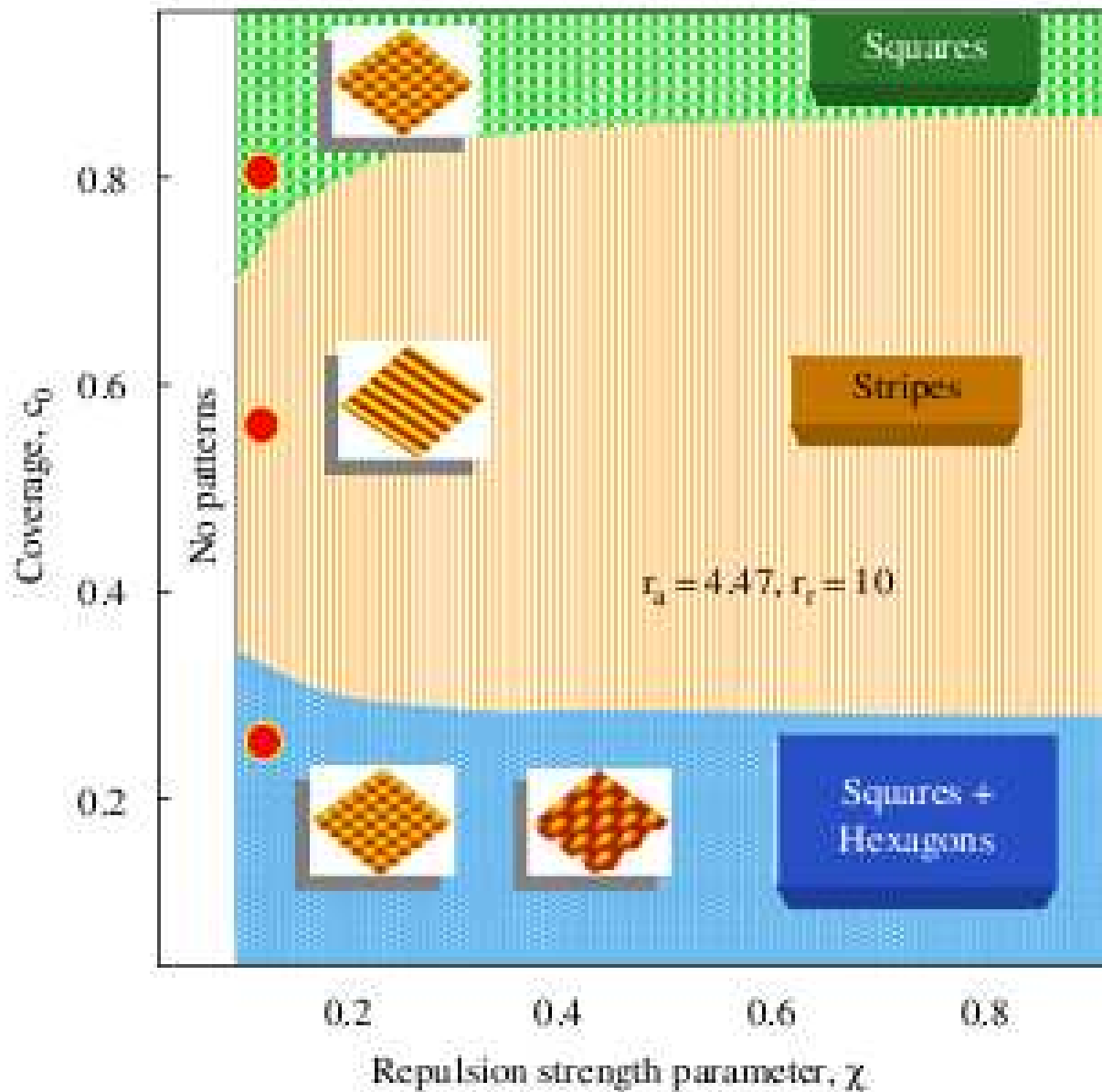


FIG. 4.1. Kinetic phase diagram for three different pattern shapes (schematics shown) obtained from bifurcation analysis of the mesoscopic equations around the marginally stable solution. ([3])

- [6] L. Kadanoff. Scaling laws for ising models near T_c . *Physics*, 2:263, 1966.
- [7] M. Katsoulakis, A. Majda, and D. Vlachos. Coarse-grained stochastic processes for microscopic lattice systems. *Proc. Natl. Acad. Sci.*, 100(3):782–782, 2003.
- [8] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos. Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems. *J. Comp. Phys.*, 112:250–278, 2003.
- [9] M. A. Katsoulakis, P. Plecháč, and A. Sopsakis. Error analysis of coarse-graining for stochastic lattice dynamics. *SIAM J. Numer. Anal.*, 44(6):2270–2296, 2006.
- [10] M. A. Katsoulakis, L. Rey-Bellet, P. Plecháč, and D. K. Tsagkarogiannis. Mathematical strategies in the coarse-graining of extensive systems: error quantification and adaptivity. *J. Non Newt. Fluid Mech.*, 2008.

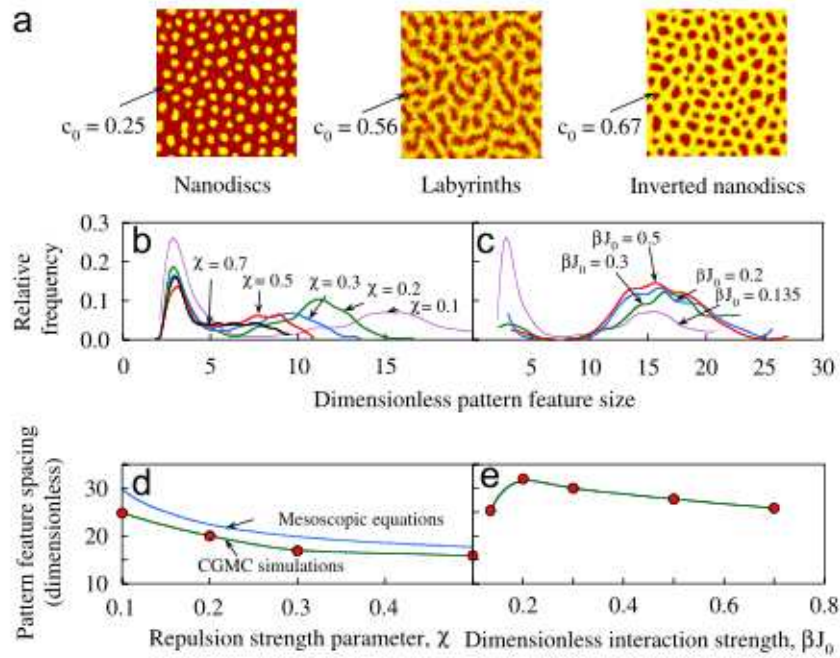


FIG. 4.2. (a) Pattern shapes obtained from CGMC simulations. From left to right: nanodiscs at low coverage, labyrinth at intermediate coverage, and inverted (vacancy) nanodiscs are shown. (b) the particle size distribution at a fixed coverage $c_0 = 0.67$ and temperature for varying repulsion strength parameter χ and (c) for varying temperature with $\xi = 0.1$ fixed. (d) The pattern feature spacing obtained from mesoscopic equation (lines) around the marginally stable solution is in qualitative agreement with CGMC simulations at coverage $c_0 = 0.67$. (e) The pattern feature spacing is nearly constant for the wide range of βJ_0 shown at coverage $c_0 = 0.67$. (for more details about simulations see [3])

- [11] M. A. Katsoulakis, L. Rey-Bellet, P. Plecháč, and D. K. Tsagkarogiannis. Coarse-graining schemes and a posteriori error estimates for stochastic lattice systems. *ESAIM-Math. Model. Num. Analysis*, 41(3):627–660, 2007.
- [12] M. A. Katsoulakis and J. Trashorras. Information loss in coarse-graining of stochastic particle dynamics. *J. Stat. Phys.*, 122(1):115–135, 2006.
- [13] M. A. Katsoulakis and D. G. Vlachos. Hierarchical kinetic Monte Carlo simulations for diffusion of interacting molecules. *J. Chem. Phys.*, 118:9412, 2003.
- [14] Markos A. Katsoulakis, Petr Plecháč, and Luc Rey-Bellet. Numerical and statistical methods for the coarse-graining of many-particle stochastic systems. *J. Sci. Comput.*, 37(1):43–71, 2008.
- [15] K. O. Ng and D. Vanderbilt. Stability of periodic domain structures in a two-dimensional dipolar model. *Phys. Review B*, 52:2177–2183, 1999.
- [16] R. Plass, N.C. Bartelt, and G.L. Kellogg. Dynamic observations of nanoscale self-assembly on solid surfaces. *J. Phys.: Condens. Matter*, 14:4227–4240, 2002.
- [17] E. Vasco. Theoretical optimization of self-organized growth of nanoscale arrays through a figure of merit. *Appl. Physics Lett.*, 85:3714–3716, 2004.
- [18] D.G. Vlachos and M.A. Katsoulakis. Derivation and validation of mesoscopic theories for diffusion of interacting molecules. *Phys. Rev. Letters*, 85(18):3898–3901, 2000.