Mathematical strategies and error quantification in the coarse-graining of many-body stochastic systems

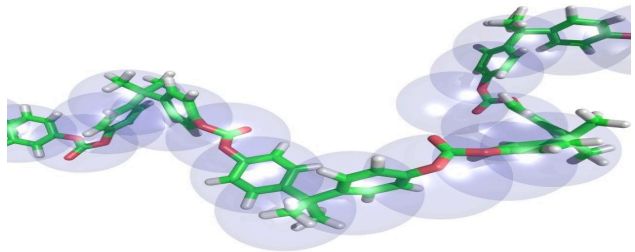
Markos Katsoulakis

University of Massachusetts, Amherst, USA

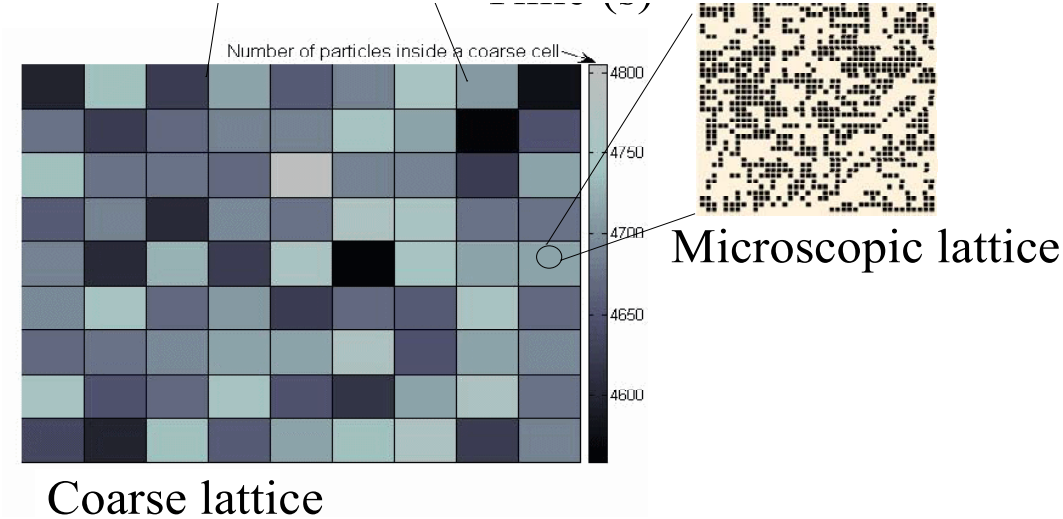
and

University of Crete, Greece

1. Coarse-graining of polymers; DPD methods

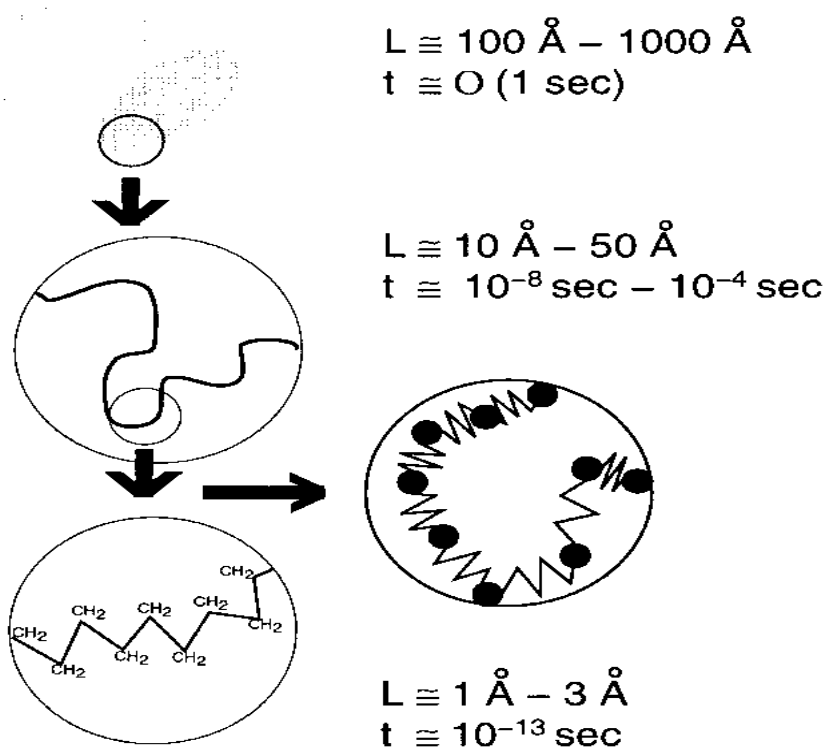


2. Stochastic lattice dynamics/ KMC



Microscopics $\mapsto$ CG system $\mapsto$ Reconstructed Microscopics

Tschöp, Kremer, Batoulis, Bürger and Hahn *Acta Polymer.* '98.



Microscopics: United Atom (UA) Model

• Continuum model: $X \in (\mathbb{R}^3)^N$ – positions of n atoms on one macromolecule; m macromolecules; $N = nm$.

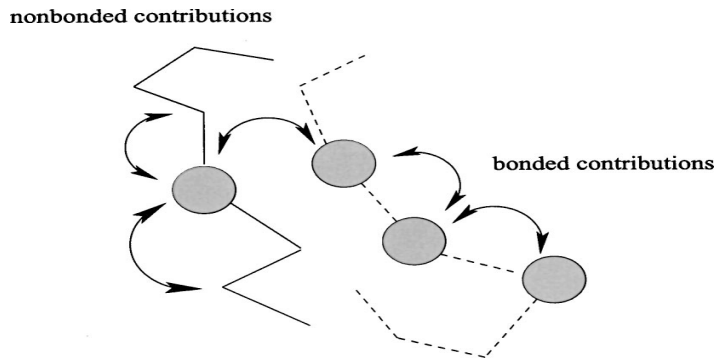
• Hamiltonian: $H_N(X) = H_{\mathsf{b}}(X) + H_{\mathsf{nb}}(X) + H_{\mathsf{Coul}}(X) + H_{\mathsf{wall}} + H_{\mathsf{kin}}$

Bonded Interactions: Gaussian, FENE, etc. short-range

$$H_{\mathsf{b}}(X) = \sum_i U_{\mathsf{b}}(\theta_i, \phi_i, r_i) \qquad \text{short-range}$$

Non-Bonded Interactions: 12-6 Lennard Jones long-range

$$H_{\mathsf{nb}}(X) = \sum_{i,j} U_{nb}^{LJ}(|x_i - x_j|)$$



Equilibrium Gibbs measure at $\beta = \frac{1}{kT}$.

$$\mu(dX) = \frac{1}{Z} e^{-\beta H(X)} \prod dx_i$$

Molecular Dynamics (via Langevin thermostat)

• UA is a typical set-up for CG in polymer science literature:

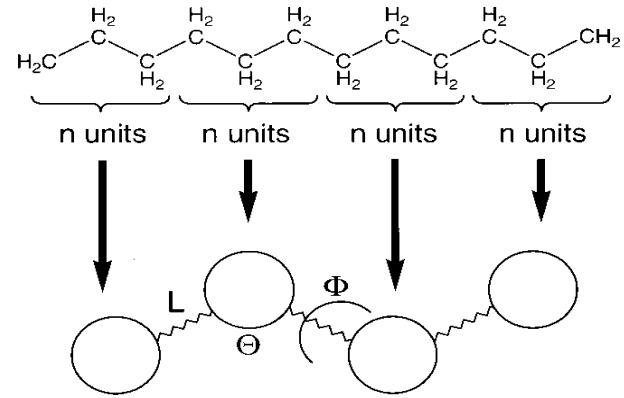
Briels, et. al. *J.Chem.Phys.* '01;

Doi et. al. *J.Chem.Phys.* '02;

Kremer et. al. *Macromolecules* '06, etc.

Also the parametric statistics approach:

Müller-Plathe *Chem.Phys.Chem* '00.



CG procedure, "blobs": for instance Doi, et al. '02

$\mathbf{T}X = Q = (q_1, \ldots, q_m) \in \mathbf{Q}$, where $q_i \in \mathbb{R}^3$.

Exact CG Hamiltonian $\bar{H}(Q)$ via **Renormalization map**:

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

Break-up of computational task: Simplifying assumptions

(i) $\bar{H}$ decouples:

$$\bar{H}(Q) = \bar{H}_{\text{b}} + \bar{H}_{\text{nb}} = \sum_{\text{CG var.}} \bar{U}_{\text{b}} + \bar{U}_{\text{nb}}$$

(ii) $\bar{U}_{\text{b}} = \bar{U}^{\theta}_{\text{b}} + \bar{U}^{\phi}_{\text{b}} + \bar{U}^{r}_{\text{b}}$ where each term depends *only* on torsion angle ϕ, rotation angle θ and distance r respectively between successive CG particles.

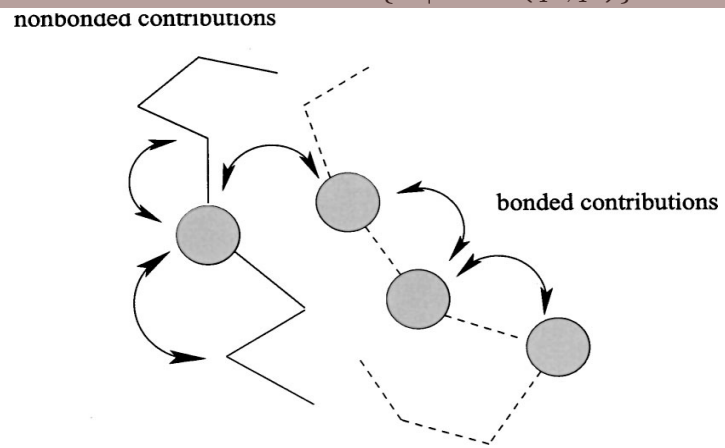
(iii) $\bar{U}_{\text{nb}}$ depends *only* on two-body interactions between CG particles; no multi-body interactions included.

How to calculate the CG non-bonded interactions $\bar{U}_{\text{nb}}$:

McCoy-Curro scheme, *Macromolecules* '98.

For two isolated small molecules with centers of mass at q_1, q_2:

$$U_{\text{nb}}(|q_1 - q_2|) = -\frac{1}{\beta} \log \int_{\{X|\mathbf{T}X=(q_1,q_2)\}} e^{-\beta H(X)}\, dX$$



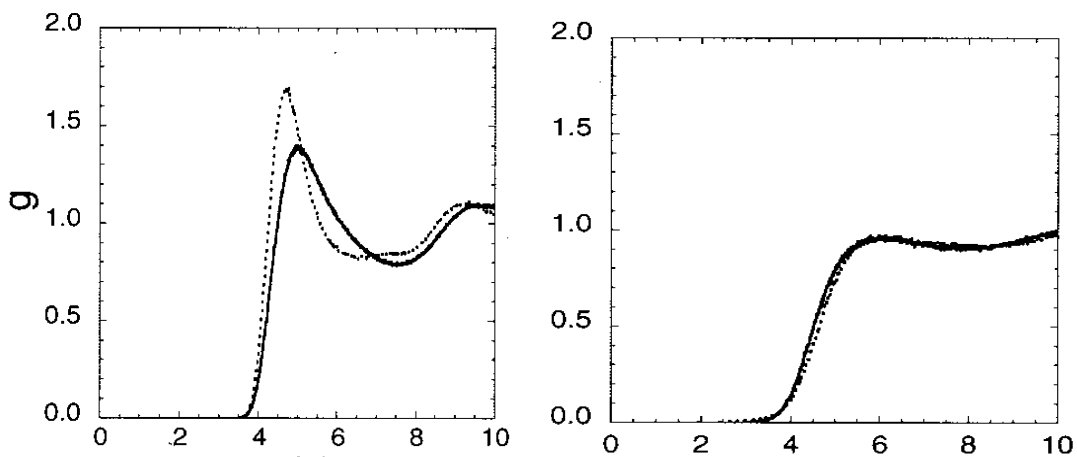
- The calculation is computationally feasible but disregards multi-body interactions.
- Extension to long chains: Doi et al. *J. Chem. Phys.* '02.

Challenges in coarse-graining methods

Often: wrong predictions in dynamics, phase transitions, melt structure, crystallization, etc. See for instance:

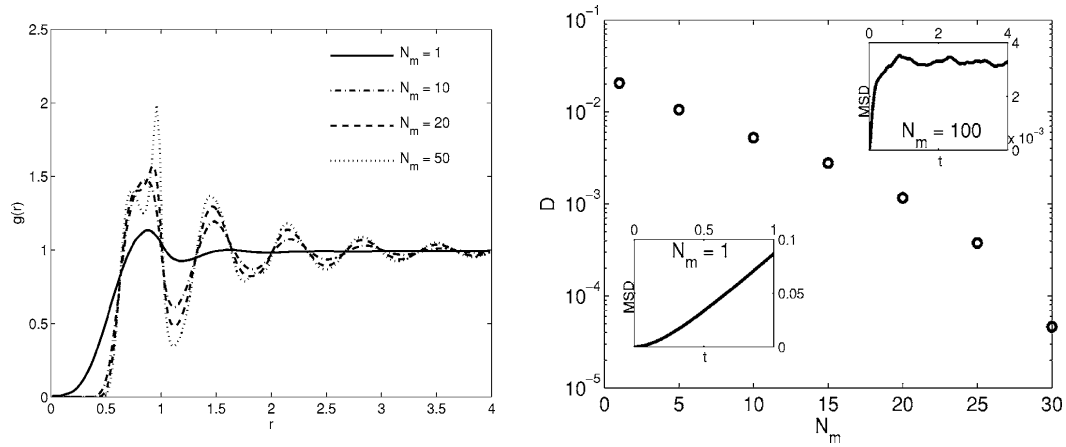
- CG in polymers: sensitive dependence to temperature

low vs. high



Doi et al. *J.Chem.Phys.* '02

• DPD: Pivkin, Karniadakis *J. Chem. Phys.* (2006): artificial crystallization



• "classical" example: 1-D nearest neigbor Ising vs. Curie-Weiss (or Mean Field)

Mathematics and Numerics of CG

1. Error Quantification and numerical accuracy of CG methods.

2. The role of randomness: need to approximate the measure rather than just $H = H(X)$:

$$e^{-\beta H(X)} dX \sim \mu_{\mathbf{micro}}(\mathbf{dX}) \quad \mapsto \quad \mathbf{T}^* \mu_{\mathbf{micro}}(\mathbf{dX}) \approx \mu_{\mathbf{cg}}(\mathbf{dQ}) \sim e^{-\beta \bar{H}(Q)} dQ$$

3. The role of multi-body CG interaction terms.

4. "Reverse map"-reconstruct microscopic info from CG:

Mathematical formulation in terms of relative entropy; loss of information during CG–information re-insertion in reverse map.

joint work with: P. Plecháč (U of TN, ORNL), V. Harmandaris (Max Planck Inst. Polymers, Mainz)

2. Stochastic lattice dynamics–Ising Systems

$\sigma(x) = 0$ or 1: site x is resp. empty or occupied.

Hamiltonian: $H_N(\sigma) = -\frac{1}{2}\sum_{x\neq y} J(x,y)\sigma(x)\sigma(y) + h\sum_x \sigma(x)$

- J: potential with interaction range L,

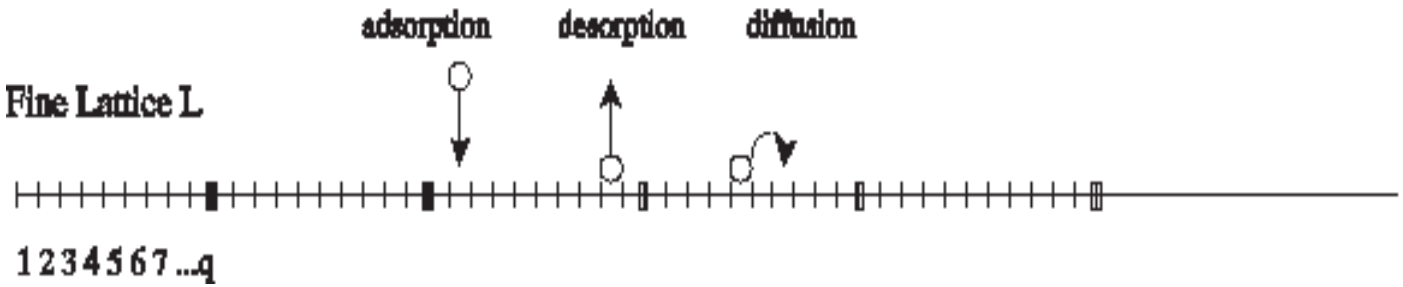
$$J(x-y) = \frac{1}{L}V\Big(\frac{i-j}{L}\Big), x = i/N, y = j/N$$

possibly short-/long- range interactions.

Canonical Gibbs measure: at the inverse temperature $\beta = \frac{1}{kT}$,

$$\mu_{\Lambda,\beta}(\sigma = \sigma_0) = \frac{1}{Z_{\Lambda,\beta}} \exp\Big\{-\beta H_N(\sigma_0)\Big\} P_N(\sigma = \sigma_0)$$

Arrhenius adsoprtion/desorption dynamics:



$\sigma(x) = 0$ or 1: site x is resp. empty or occupied.

Generator: $L_X f(\sigma) = \sum_x c(x, \sigma, X)[f(\sigma^x) - f(\sigma)]$

Transition rate: $c(x, \sigma, X) = c_0 \exp\Big[-\beta U(x)\Big]$

$U(x)$: Energy barrier a particle has to overcome in jumping from a lattice site to the gas phase.

- Detailed Balance
- $U(x) = U(x, \sigma, X) = \sum_{z \neq x} J(x - z)\sigma(z) - h(X)$.
- strong interactions/low temperature $\rightarrow$ clustering/phase transitions

Why study this system?

0. Many-particle system, related to realistic models, KMC, etc.

1. Strong interactions/low temperature $\rightarrow$ clustering/phase transitions. "Complex" landscape: metastability of islands.

How CG performs in predicting phase transitions and various rare events?

2. Equilibrium/ Detailed Balance.

How CG performs in transient and long time regimes?

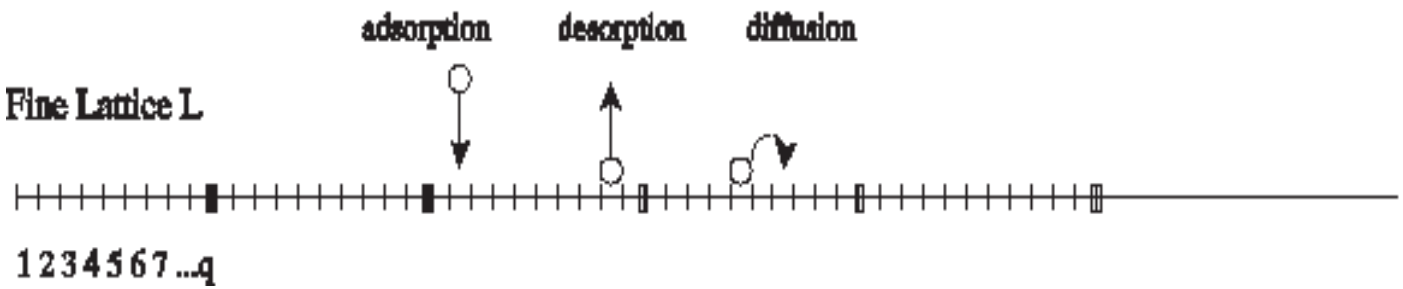
3. Numerous analytic benchmark solutions; a variety of mathematical physics tools.

Hierarchical coarse-graining of stochastic lattice dynamics

K., Majda, Vlachos,*Proc. Nat. Acad. Sci.*'03, *JCompPhys*'03;

K., Vlachos *J.Chem.Phys.*'03

Construct a **stochastic process** for a hierarchy of "mesoscopic" length or time scales. Coarse-grained Monte Carlo algorithm (CGMC).



Coarse observable at resolution q:

$$\eta_t(k) = \mathbf{T}\sigma_t(k) := \sum_{y \in D_k} \sigma_t(y)$$

In general it is **non-markovian**

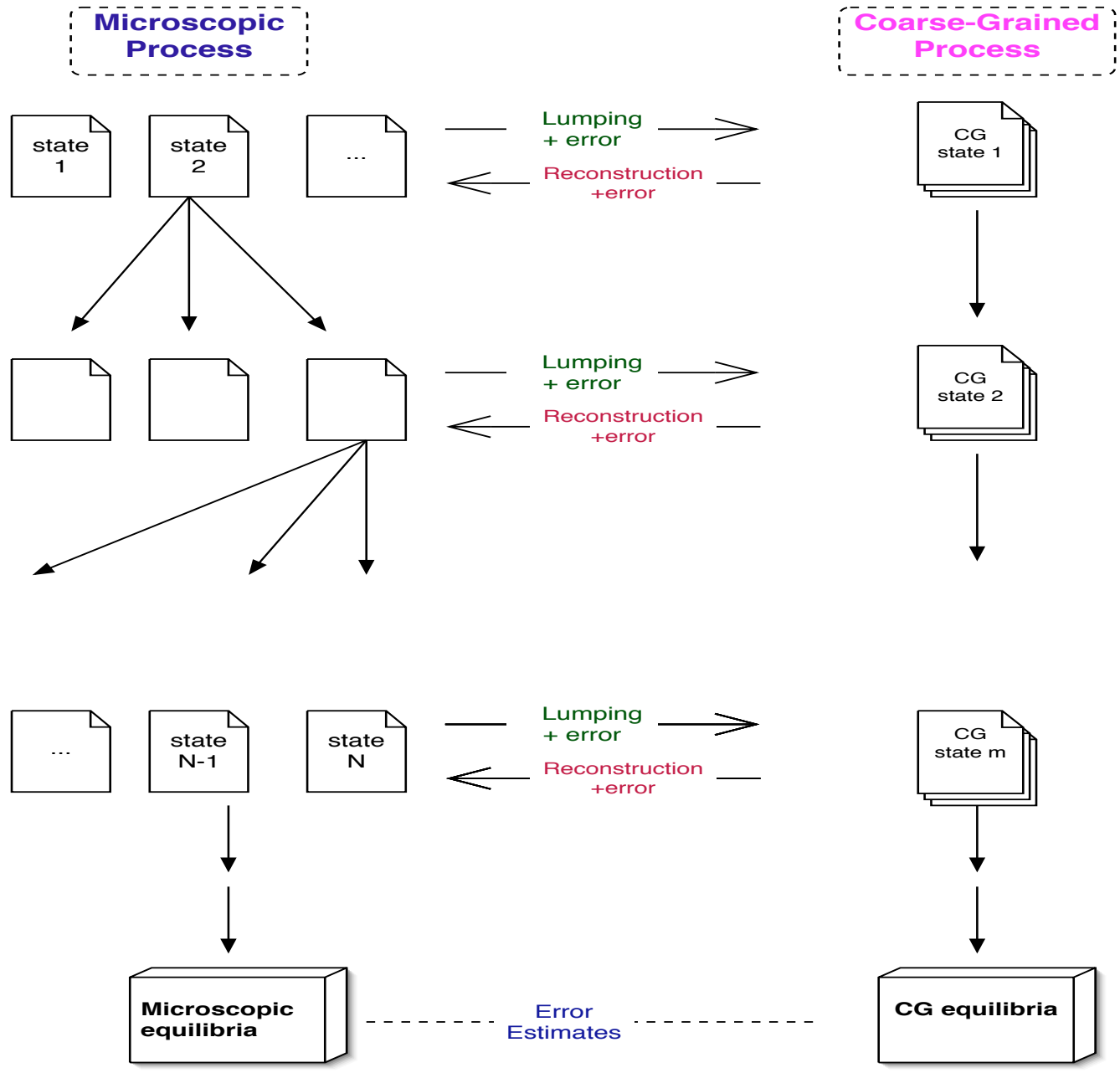
Stochastic closures: can we write a new approximating Markov process for η_t?

• "projective dynamics": Koleshik, Novotny, Rikvold, *PRL* '98; coarse rates for total coverage calculated by sampling;

Ergodicity: Are the long-time dynamics reproduced?

• Errors can contaminate the simulation at long times; wrong switching times in bistable systems: Hanggi et al *PRA* '84 (well-mixed systems).

• Connections to *lumpable* Markov processes


Microscopic Process
Coarse-Grained Process
state 1
state 2
...
Lumping + error
Reconstruction +error
CG state 1
CG state 2
state N-1
state N
CG state m
Microscopic equilibria
Error Estimates
CG equilibria

1. CG Schemes at Equilibrium

K., Plechac, Rey-Bellet, Tsagkarogiannis, [*M^2AN, '07, J. Non. Newt. Fluid Mech. '08, preprint*]

- **CG Hamiltonian–Renormalization Group Map**: $N = mq$

$$e^{-\beta \bar{H}_m(\eta)} = \int e^{-\beta H_N(\sigma)} \, P_N(d\sigma \,|\, \eta) \equiv \mathbb{E}[e^{-\beta H_N} \,|\, \eta]$$

- **Correction terms** around a first "good guess" $\bar{H}_m^{(0)}$:

$$\bar{H}_m(\eta) = \bar{H}_m^{(0)}(\eta) - \frac{1}{\beta} \log \mathbb{E}[e^{-\beta(H_N - \bar{H}_m^{(0)})} \,|\, \eta] \,, \quad m = N, N-1, ...$$

- **Heuristics:** Expansion of $e^{\Delta H}$ and log:

$$= \mathbb{E}\left[\Delta H \,|\, \eta\right] + \mathbb{E}\left[(\Delta H)^2 \,|\, \eta\right] - \mathbb{E}\left[\Delta H \,|\, \eta\right]^2 + O((\Delta H)^3)$$

formal calculations inadequate since:

$$\Delta H \equiv H_N - \bar{H}_m^{(0)} = N \cdot O(\epsilon)$$

- **Rigorous analysis – Cluster expansion**: around $\bar{H}_m^{(0)}$

Systems with short+long-range interactions

$$H_N(\sigma) = H_N^l(\sigma) + H_N^s(\sigma)\,;$$

J: long range potential $\sim H_N^{(l)}$ radius L. K: short range potential $\sim H_N^{(s)}$ with radius $S << L$.

Examples: Surface processes, epitaxial growth, polymers, etc.

CG: approximation of the free energy-landscape.

CG prior: $\bar{P}_m(\eta) = P_N(\{\sigma : \mathbf{T}\sigma = \eta\})$

- Splitting strategy:

$$e^{-\beta H_N(\sigma)} P_N(d\sigma) \quad = \quad e^{-\beta H_N^s(\sigma)} e^{-\beta\left(H_N^l(\sigma) - \bar{H}_m^l(\eta)\right)} P_N(d\sigma|\eta) e^{-\beta \bar{H}_m^l(\eta)} \bar{P}_m(\eta)$$

Case 1: Long- and intermediate-range interactions

Approximate CG Hamiltonian:

$$\bar{H}^{(0)}(\eta) = -\frac{1}{2}\sum_{l\in\Lambda_M^c}\sum_{k\neq l}\bar{J}(k,l)\eta(k)\eta(l) - \frac{1}{2}\bar{J}(0,0)\sum_{l\in\Lambda_M^c}\eta(l)(\eta(l)-1) + \sum_{l\in\Lambda_M^c}\bar{h}(l)\eta(l)$$

- $\mathbb{E}\left[H_N - \bar{H}^{(0)} \,|\, \eta\right] = 0$

Involves two-body CG interaction only:

$$\bar{J}(k,l)\eta(k)\eta(l) = \int \sum_{x\in C_k, y\in C_l} J(x-y)\sigma(x)\sigma(y)P_N(d\sigma \,|\, \eta_k, \eta_l)$$

Where

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

- Analytical version of McCoy-Curro scheme in polymers:

$$\bar{U}^{mcc}(\eta_k, \eta_l; k-l) = -\frac{1}{\beta}\log\int e^{-\beta H_N(\sigma)}P_N(d\sigma \,|\, \eta_k, \eta_l)$$

Corrections to the Hamiltonian $\bar{H}^{(0)} \mapsto$ Multi-body terms

$$\bar{H}_m(\eta) = \bar{H}_m^{(0)}(\eta) + \bar{H}_m^{(1)}(\eta) + ...$$

$$\bar{H}^{(1)}(\eta) = \beta \sum_{k_1} \sum_{k_2 > k_1} \sum_{k_3 > k_2} [j^2_{k_1 k_2 k_3}(-E_1(k_1)E_2(k_2)E_1(k_3) + ...$$

- $E_r(k) \equiv E_r(\eta(k)) = (2\eta(k)/q - 1)^r + o_q(1)$
- "Moments" of interaction potential J:

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

Computational complexity-Compression of $\bar{H}^{(1)}$

- Evaluation of the Hamiltonian:

	Count	Speed-up
Microscopic: $H_N(\sigma)$	$O(NL^d)$	1
CG0: $\bar{H}^{(0)}$	$O(ML^d/q^d)$	$O(q^{2d})$
CG1: $\bar{H}^{(0)} + \bar{H}^{(1)}$	$O(ML^{2d}/q^{2d})$	$O(q^{3d}/L^d)$

- Decay of J (e.g. Coulomb) $\mapsto$ $J - \bar{J}$ decays faster.

Rigorous analysis – Cluster expansion

Idea: Identify clusters that do not "communicate"–factorize– then Taylor expand.

Step 1: Rewrite

$$\mathbb{E}\left[e^{-\beta(H_N-\bar{H}^{(0)})}\,|\,\eta\right] = \int \prod_{k\le l}(1+(e^{-\beta\Delta_{kl}J(\sigma)}-1))\,P_N(d\sigma\,|\,\eta)$$

where

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

Step 2: Assume $e^{\cdots}-1$ small and expand

$$\prod_{k\le l}(1+(e^{-\beta\Delta_{kl}J(\sigma)}-1)) = \sum_{G\in\mathcal{G}_M}\prod_{\{k,l\}\in G}(e^{-\beta\Delta_{kl}J(\sigma)}-1))$$

Convergence criterion for the resulting series (Kotecký-Preiss-Dobrushin)

Error Quantification in CG Schemes

Theorem 1:

Define the "small" parameter $\epsilon \equiv \beta \frac{q}{L} \|\nabla J\|_1$

1. *Approximation of the CG free-energy landscapes*

$$\bar{H}_m(\eta) = \bar{H}_m^{(0)}(\eta) - \frac{1}{\beta} \log \mathbb{E}[e^{-\beta(H_N - \bar{H}_m^{(0)})} \,|\, \eta] = \bar{H}_m^{(0)}(\eta) + \bar{H}_m^{(1)}(\eta) + NO(\epsilon^3)\,.$$

2. *Loss of information during coarse-graining*

- Specific relative entropy:

$$\mathcal{R}\,(\mu \,|\, \nu) := \frac{1}{N} \sum_{\sigma} \log \left\{ \frac{\mu(\sigma)}{\nu(\sigma)} \right\} \mu(\sigma) \qquad .\diamond$$

$$\mathcal{R}\left(\bar{\mu}_{M,q,\beta}^{(\alpha)} \,|\, \mu_{N,\beta} o \mathbf{T}^{-1}\right) = O\left(\epsilon^{\alpha+2}\right).$$

- $\mathbf{T}\sigma =$ **Projection** on coarse variables$= \sum_{y \in D_k} \sigma_{(}y)$.

Remarks:

- Information Theory interpretation: The relative entropy describes the increase in descriptive complexity of a random variable due to "wrong information".

- Controlling the expansion: "high-temperature" cluster expansion techniques (Cammarota CMP 82, Procacci, De-Lima, Scoppola LMP 98)

- Related work: M. Suzuki et. al.'95, Cassandro/Presutti '96, Bovier/Zahradnik '97; cluster expansions around mean-field; focus on criticality.

General Case: combined short+long range interactions:

K., Plechac, Rey-Bellet, Tsagkarogiannis, [preprint '08]

Results on the long range interactions suggest a separation into:

- smooth, long-range interactions (expensive with KMC-very efficient with CGMC)
- separately handle short range interactions*

$$e^{-\beta H_N(\sigma)} P_N(d\sigma) \quad = \quad e^{-\left(\beta H_N^l(\sigma) - \bar{H}_m^l(\eta)\right)} \left[e^{-\beta H_N^s(\sigma)} P_N(d\sigma|\eta) \right] e^{-\bar{H}_m^l(\eta)} \bar{P}_m(\eta)$$

***Related Cluster Expansion:** Bertini, Cirillo, Olivieri, *J. Stat. Phys.* '99.

Double/Triple terms in CG short range interactions:

$$\bar{H}^{(1)}_{k-1,k,k+1}(\eta(k-1),\eta(k),\eta(k+1)) = -\frac{1}{\beta}\log\Big(1-\lambda\Phi^1_{k-1}(\eta(k-1))\Phi^1_k(\eta(k)) - \lambda\Phi^1_k(\eta(k))\Phi^1_{k+1}(\eta(k+1)) + \lambda^2\Phi^1_{k-1}(\eta(k-1))\Phi^2_k(\eta(k))\Phi^1_{k+1}(\eta(k+1))\Big)$$

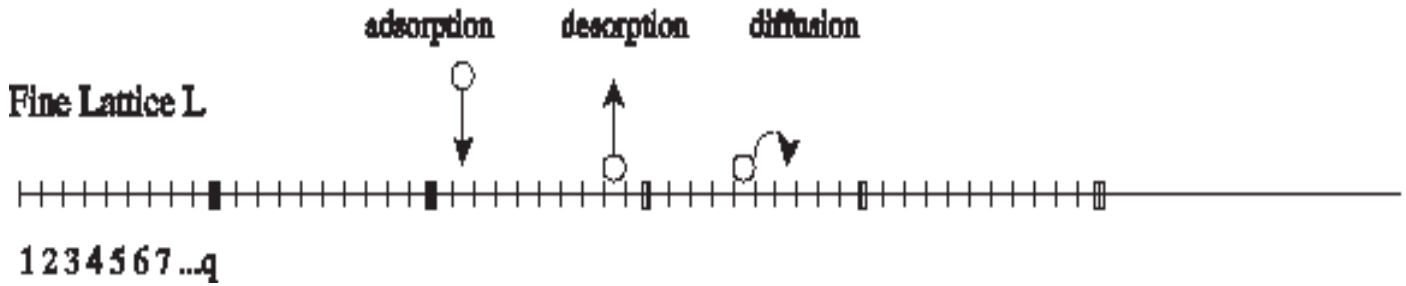
where $\lambda = \tanh(\beta K)$,

$$\Phi^1_k(\eta) := \int \sigma(x)\hat{\rho}_k \quad \text{and} \quad \Phi^2_k(\eta) := \int \sigma(x)\sigma(y)\hat{\rho}_k$$

- **Semi-analytical splitting method:** Fine scales are simulated (cheaply) in the Φ-terms, then a CGMC step is performed.

- Triple terms are important only at lower temperatures.

CG Markovian Dynamics



Birth-Death type process, with interactions.

$$L_c g(\eta) = \sum_{k \in \Lambda_c} c_a(k,\eta)\Big[g(\eta+\delta_k) - g(\eta)\Big] + c_d(k,\eta)\Big[g(\eta-\delta_k) - g(\eta)\Big].$$

- **Coarse-grained** rates: Detailed Balance

Adsorption rate of a single particle in the k-coarse cell

$$c_a(k,\eta) = q - \eta(k)$$

Desorption rate

$$c_d(k,\eta) = \eta(k) \exp\Big[-\beta\Big(U_0 + \bar{U}(k)\Big)\Big]$$

with or w/o higher order terms.

Formal derivation

Step 1: From the microscopic generator:

$$\frac{d}{dt}Eg(\eta) = \quad E\sum_{k\in\Lambda_c}\left\{\sum_{x\in D_k} c(x,\sigma)\big(1-\sigma(x)\big)\right\}\times \Big[g(\eta+\delta_k)-g(\eta)\Big] +$$

$$E\sum_{k\in\Lambda_c}\left\{\sum_{x\in D_k} c(x,\sigma)\sigma(x)\right\}\times \Big[g(\eta-\delta_k)-g(\eta)\Big].$$

"Closure" argument: Express as a function of the coarse variables the terms

$$\left\{\sum_{x\in D_k} c(x,\sigma)\,\ldots\right\},\quad \left\{\sum_{x\in D_k} c(x,\sigma)\,\ldots\right\}$$

- $\sum_{x\in D_k} c(x,\sigma)\Big(1-\sigma(x)\Big) = \Big(q-\eta(k)\Big) := c_a(k,\eta)$

- $\sum_{\mathbf{D_k}} \mathbf{c(x,\sigma)\sigma(x)} = \sum_{D_k} \sigma(x) \exp\Big[-\beta\Big(U_0+U(x)\Big)\Big] \overset{??}{=} \mathbf{c_d(k,\eta)}$

One possibility: $c(x,\sigma) \approx$ const. on coarse cell D_k, e.g.

1. high temperature/external field, or
2. **q << L** q: level of coarse-graining, L: interaction range

We have

$$c_d(k,\eta) \approx \eta(k) \exp\Big[-\beta\Big(U_0+\bar{U}(k)\Big)\Big]$$

where $U(x) = \bar{U}(l) + O\Big(\frac{q}{L}\Big)$, and

$$\bar{U}(l) = \sum_{\substack{k\in\Lambda_c \\ k\neq l}} \bar{J}(l,k)\eta(k) + \bar{J}(0,0)\Big(\eta(l)-1\Big) - \bar{h}\,.$$

I. Error Estimates for observables – Dynamics

[K., P. Plechac, A. Sopasakis, *SIAM Num. Anal.* '06]

Theorem 1:

q: level of coarse-graining

L: # of interacting neighbors

coarse grained observables/quantity of interest: ψ,

microscopic dynamics: σ_t,

coarse-grained dynamics: η_t

Then for any fixed time $0 < T < \infty$

$$|E\psi(\mathbf{T}\sigma_T) - E\psi(\eta_T)| \leq C_T\epsilon^2$$

- $\mathbf{T}\sigma_t =$ **Projection** on coarse variables$= \sum_{y \in D_k} \sigma_t(y)$.
- Error accumulation as $T \to \infty$? 2nd order error estimates at **equilibrium**

Difficulty: $\mathbf{T}\sigma_t(k) = \sum_{y \in D_k} \sigma_t(y)$. is not a Markov process.

Elements of the proof:

1. γ_t: Markovian reconstruction of the microscopic process σ_t from the coarse process η_t with controlled error:
   - $\mathbf{T}(\gamma_t)_{t \geq 0}$ and $(\eta_t)_{t \geq 0}$ have the same distribution
   - $|E\phi(\sigma_T) - E\phi(\gamma_T)| \leq C_T \epsilon^2$,
2. Stochastic averaging $\rightarrow$ cancellations and 2nd order accuracy.
3. Bernstein-type estimates to control discrete derivatives– here related to the number of jumps-extended system!
4. Weak topology estimates for SDE: Talay-Tubaro (1990), Szepessy, Tempone, Zouraris (2001),..., K., Szepessy (2006).

Error II–Loss of information during coarse-graining

[with José Trashorras (Paris IX), J. Stat. Phys. (2006)]

- $\mu_{m,q,\beta}(t)$: Coarse-grained PDF at time t.
- $\mu_{N,\beta} o \mathbf{T}(t)$: Projection of the microscopic PDF at time t on the coarse observables.

Theorem 2:

$$\mathcal{R}\left(\mu_{m,q,\beta}(t) \,|\, \mu_{N,\beta} o \mathbf{T}(t)\right) = O_T(\frac{q}{L}), \quad t \in [0,T]$$

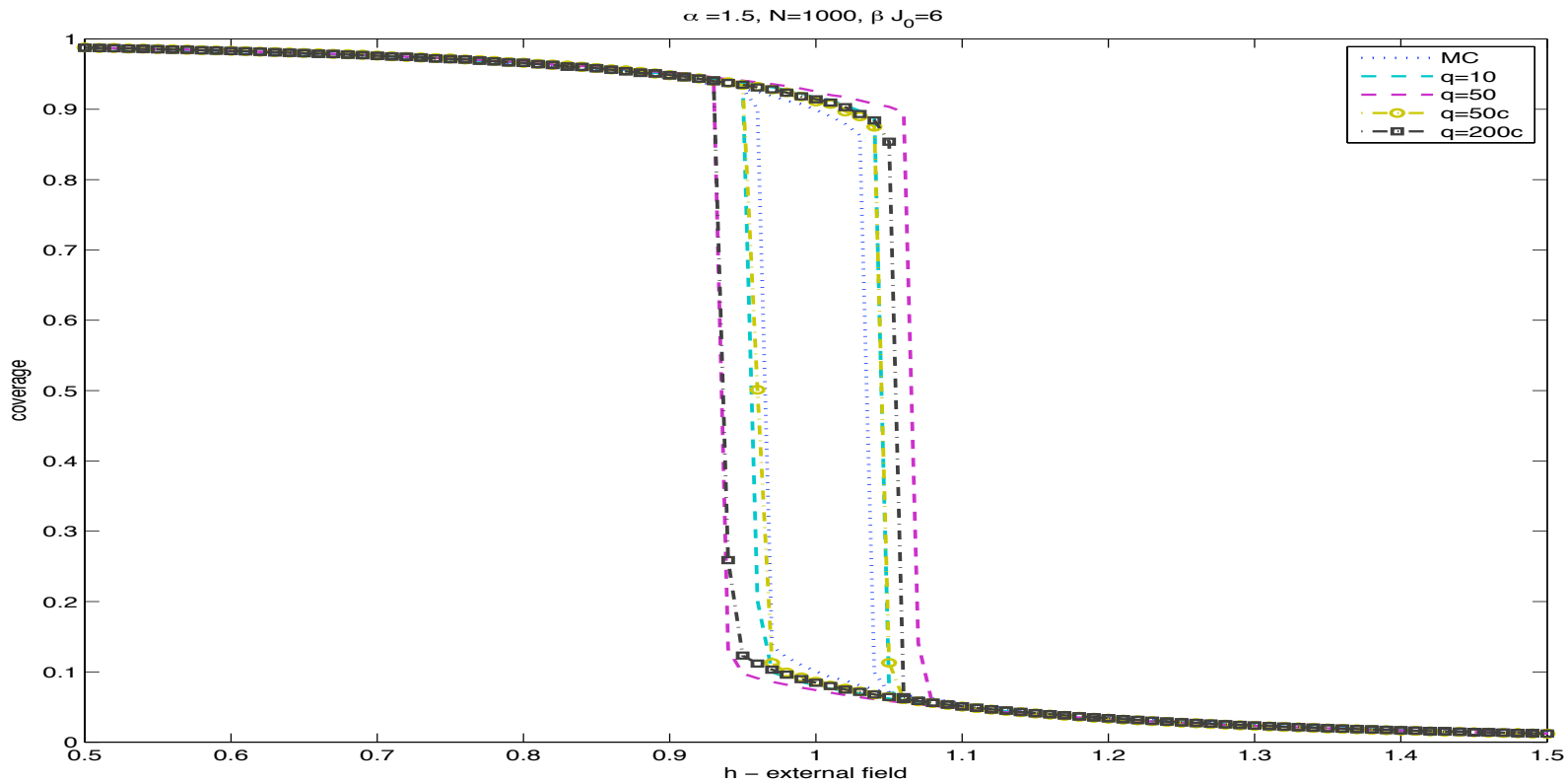
where

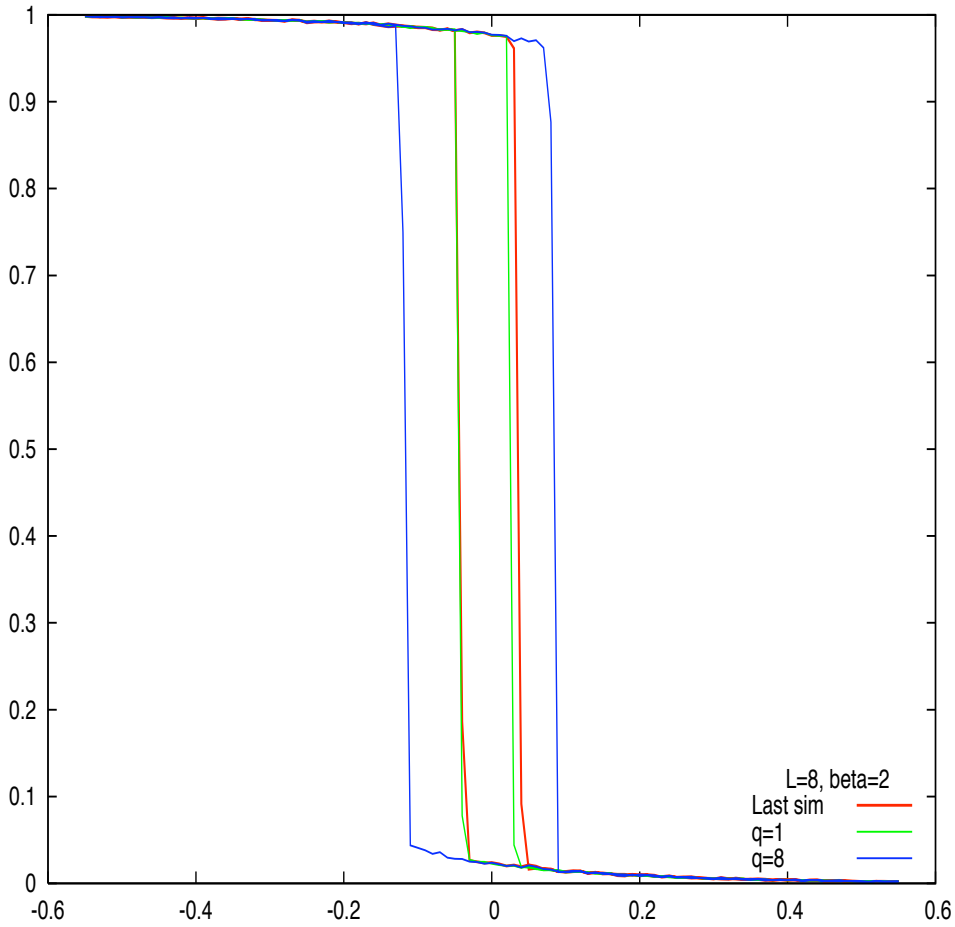
$$\mathcal{R}\left(\mu \,|\, \nu\right) := \frac{1}{N} \sum_{\sigma} \log\left\{\frac{\mu(\sigma)}{\nu(\sigma)}\right\} \mu(\sigma) \quad .\diamond$$

Some computational tests

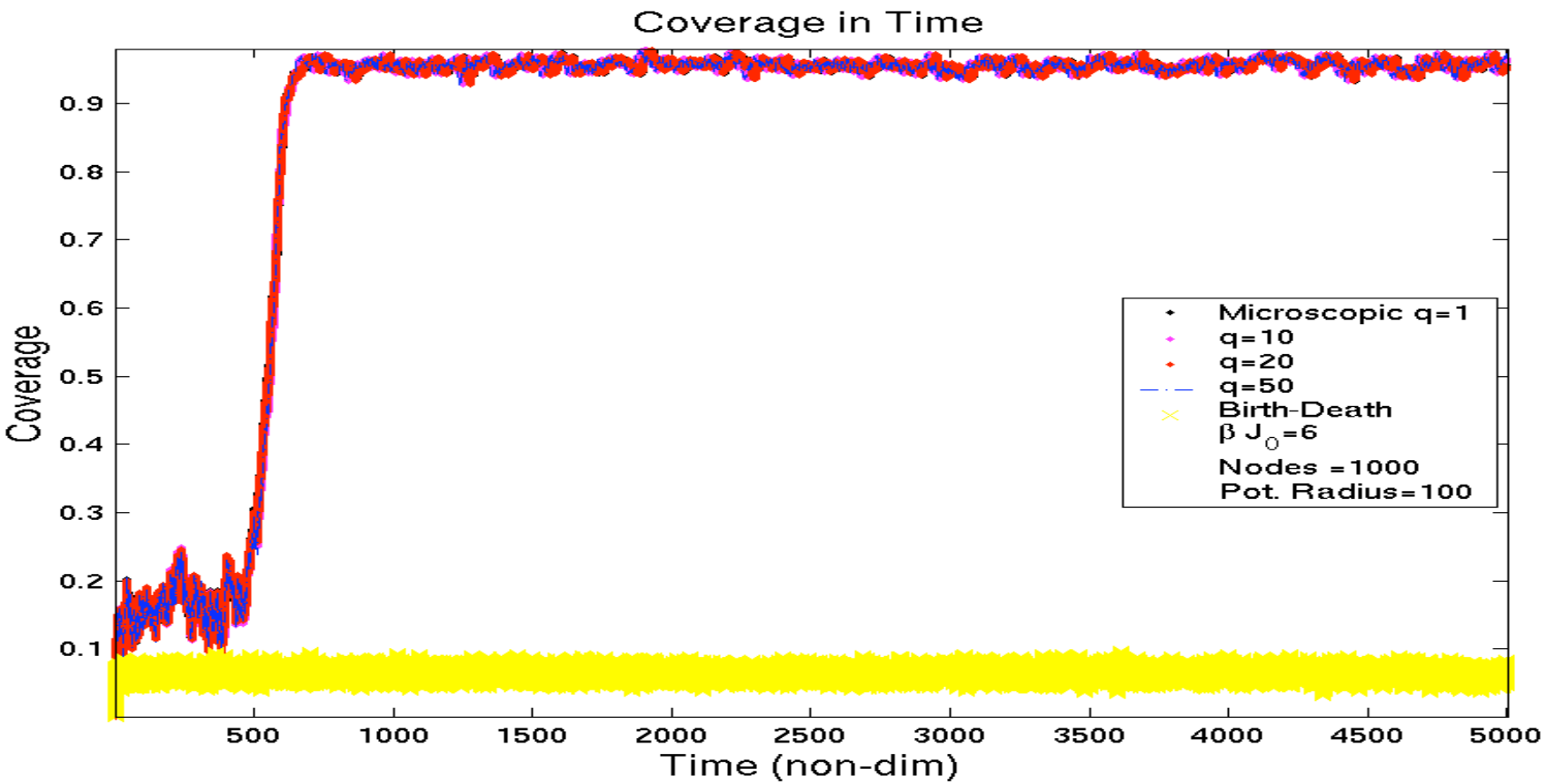
CG Arrhenius lattice dynamics Metastable regime

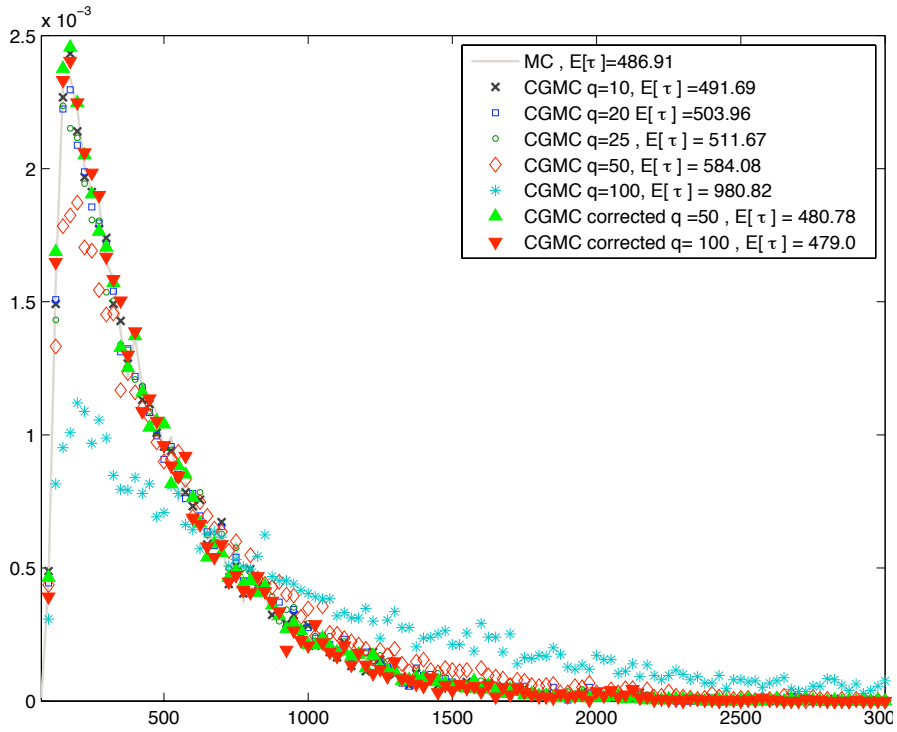
1. Power law interactions: $J(r) = r^{-\alpha}$.

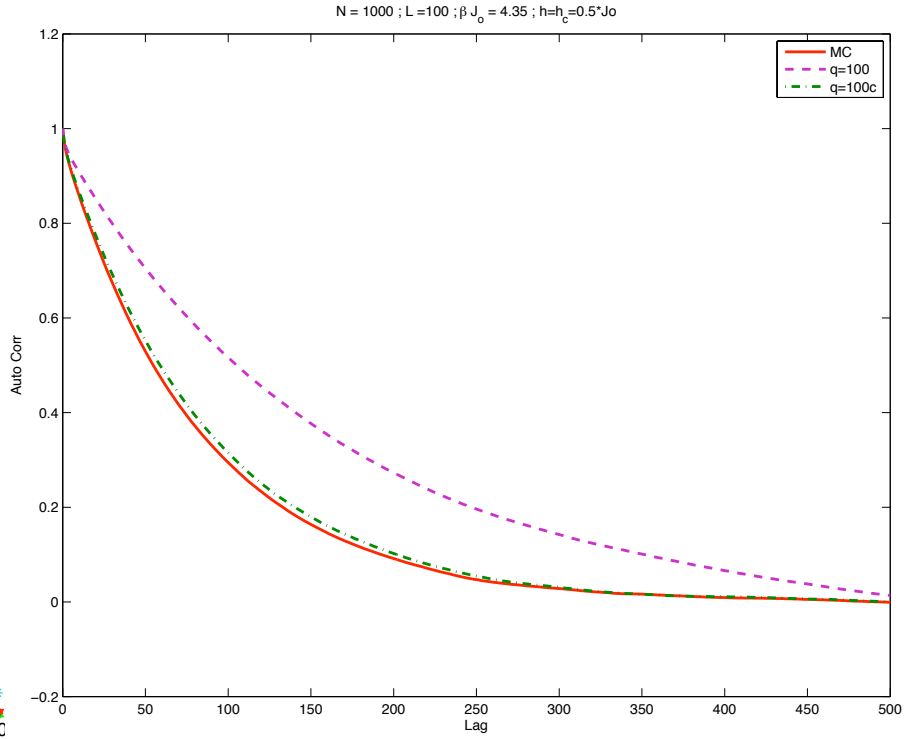



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
L=8, beta=2
Last sim
q=1
q=8

2. Switching Time PDFs/Autocorrelations-corrections

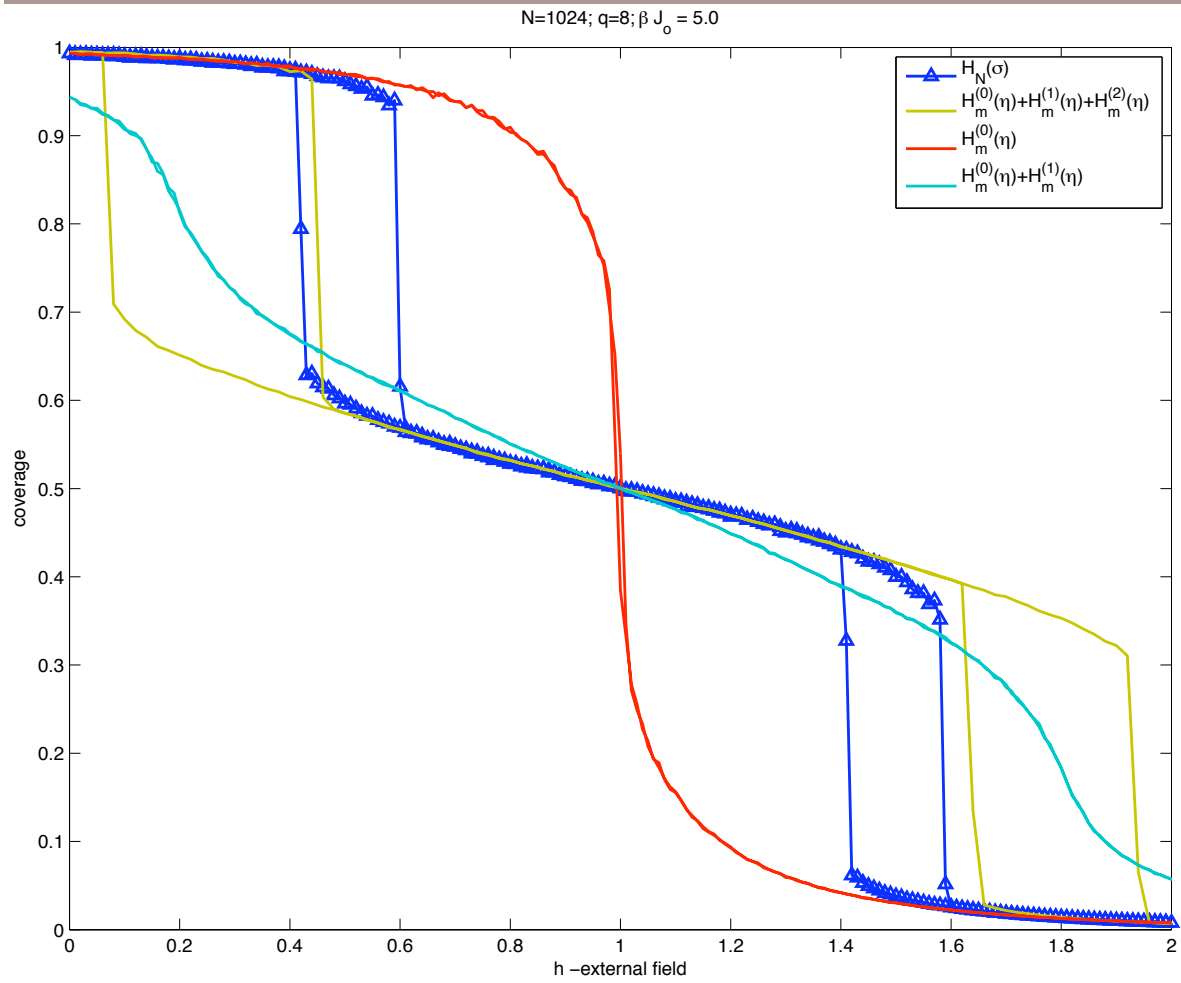






joint work with Sasanka Are (UMass)

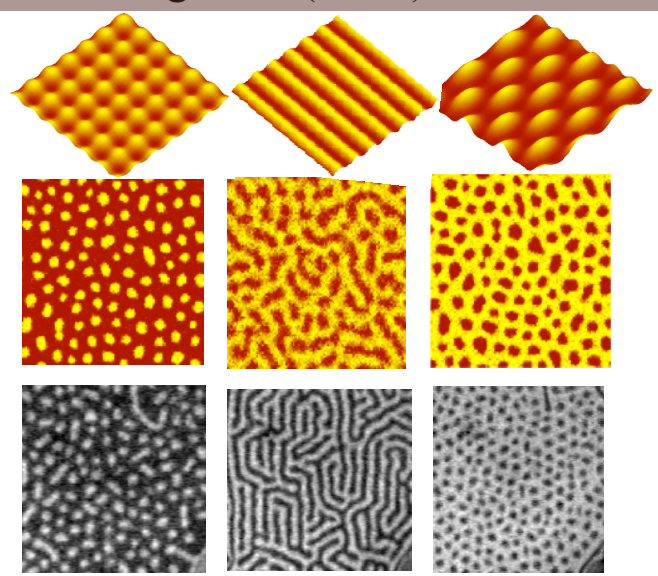
Hysteresis Diagram for system with power law potentials exhibiting short and long range interactions: comparisons of KMC with CGMC. Sasanka Are (UMass)



Example: Hetero-epitaxy in a Pb/Cu system

Plass, Last, et al., Nature (2001)

Simulation with CGMC at mesoscopic length scales:
Chaterjee, Vlachos, Chem Eng. Sci. (2007)



Comparison of deterministic (top row), CGMC (middle row), and experimental patterns (bottom) of Pb/Cu(111) system as the Pb concentration increases from low (left) to high (right).

Reverse CG map–Microscopic Reconstruction

[Tschöp et al *Acta Polymer.* '98], [K., Trashorras, *J. Stat. Phys.* '06], [K., Plechac, Sopasakis, *SIAM Num. Anal.* '06]

[Trashorras, Tsagkarogiannis '08]: systematic equilibrium study

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

We can think of the conditional probability $\mu_N(d\sigma|\eta)$ as reconstructing (perfectly) $\mu_N(d\sigma)$ from the (exactly) CG measure $\bar{\mu}_M(d\eta)$.

Mathematical formulation:

1. *CG Scheme*: $\bar{\mu}_M^{\text{app}}(d\eta) \approx \bar{\mu}_M(d\eta)$

2. *Reconstruction*: Construct a "suitable" conditional probability $\nu_N(d\sigma|\eta)$ and define the approximate microscopic measure

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

Efficiency of the reconstruction:

$$\mathcal{R}\left(\mu_N^{\text{app}}|\mu_N\right) = \mathcal{R}\left(\bar{\mu}_M^{\text{app}}|\bar{\mu}_M\right) + \int \mathcal{R}\left(\nu_N(\cdot|\eta)\,|\,\mu_N(\cdot|\eta)\right)\bar{\mu}_M^{\text{app}}(d\eta)$$

Example: $\bar{\mu}_M^{\text{app}}(d\eta) = \bar{\mu}_M^{(0)}(d\eta)\,, \quad \nu_N(d\sigma\,|\,\eta) = P_N(d\sigma\,|\,\eta)\,,$

a. $P_N(\sigma|\eta)$ is a product measure $\Longrightarrow$ "local" reconstruction at each coarse-cell;

b. Reconstruction for equilibrium and dynamics;

c. Numerical error estimate for reconstructed microscopic dynamics γ_t:

$$|E\phi(\sigma_T) - E\phi(\gamma_T)| \leq C_T\epsilon^2\,,$$

Thus far: **Applied math/statistical mechanics perspective** of expanding (using cluster expansions) around a "carefully" chosen first CG guess,

A statistics perspective:

$$e^{-\beta H_N(\sigma)} P_N(d\sigma) \quad = \quad e^{-\beta\left(H_N(\sigma) - \bar{H}^s_m(\eta) - H^l_m(\eta)\right)} P_N(d\sigma|\eta) e^{-\beta\left(\bar{H}^s_m(\eta) + \bar{H}^l_m(\eta)\right)} \bar{P}_m(\eta)$$

Importance Sampling based on proposals sampled from CG distibutions

CG approximating measure (or an "easy" part of it); local reconstruction.

K., Plechac, Rey-Bellet [J. Sci. Comp.], to appear (2008).

CG diagnostics, a posteriori error–Adaptive CG

[K., Plechac, Rey-Bellet, Tsagkarogiannis, J.Non-Newt. Fluid Mech. to appear, '08]

1. Cluster expansions $\rightarrow$ a posteriori expansion for the relative entropy.

$$\bar{H}_m(\eta) = \bar{H}_m^{(0)}(\eta) - \frac{1}{\beta} \log \mathbb{E}[e^{-\beta(H_N - \bar{H}_m^{(0)})} \mid \eta] = \bar{H}_m^{(0)}(\eta) + \bar{H}_m^{(1)}(\eta) + \ldots$$

The error indicator $R(.)$ is given by the terms $\bar{H}^{(1)}, \bar{H}^{(2)}$ and depends only on the coarse variable η:

$$\mathcal{R}\left(\mu_{m,q}^{(0)} \mid \mu_N o \mathbf{T}\right) = E_{\bar{G}^{(0)}}[R(\eta)] + \log\left(E_{\mu_{m,q}^{(0)}}[e^{R(\eta)}]\right) + \mathbf{O}(\epsilon^3)$$

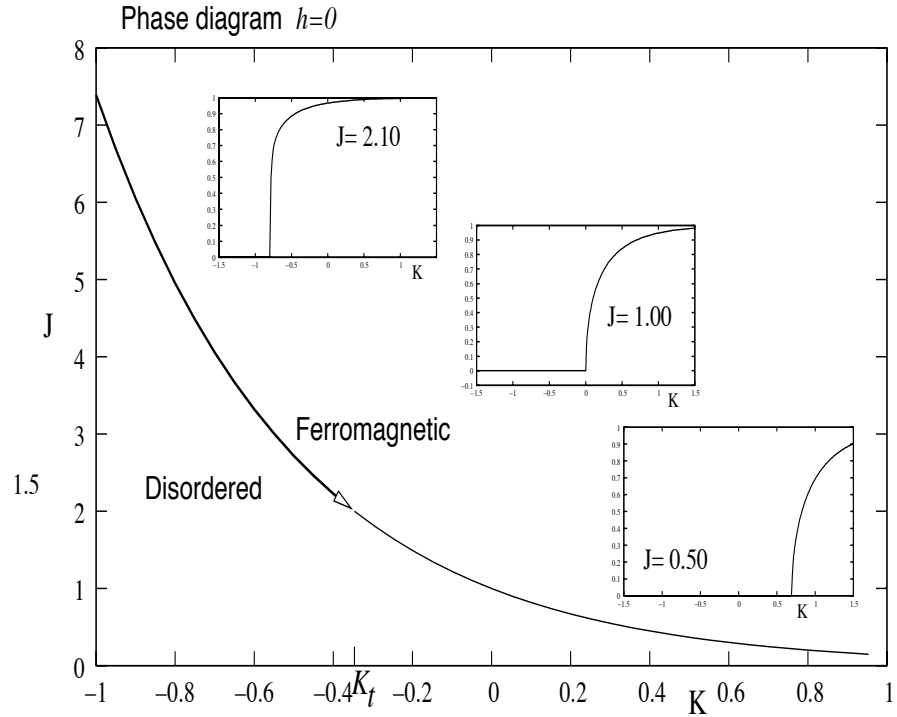
2. "Goal-oriented" a posteriori estimates and adaptivity?

Typical observables: spatial correlation functions of coarse observables

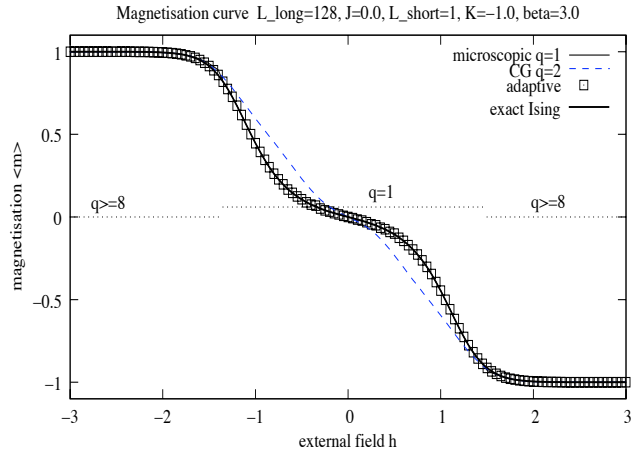
A mathematical prototype: *Competing short ($L_K = 1$) and long ($L_J = 64$) range*

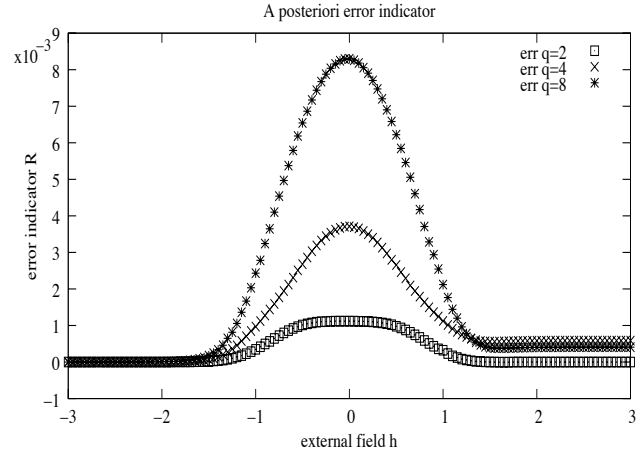
$$H_N = -K \sum_{|x-y|=1} \sigma(x)\sigma(y) - \frac{J}{2N} \sum_{x,y} \sigma(x)\sigma(y) + h \sum_{x} \sigma(x)$$

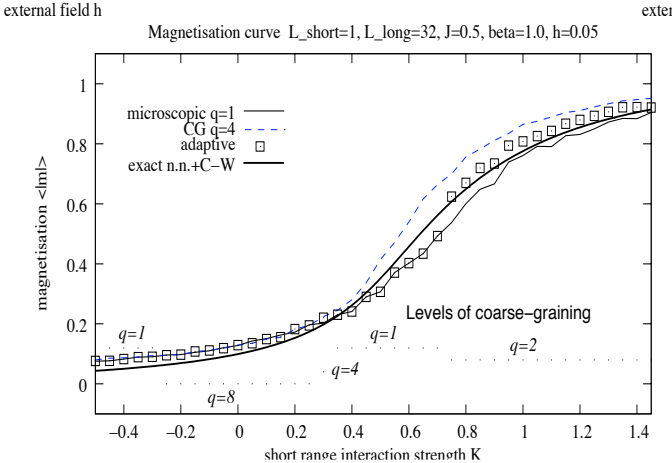
Exact solution in 1D/2D (M. Kardar, PRB '83)



Example: Adaptive computation of phase diagrams







Spatial adaptivity:
[Chaterjee, K., Vlachos, *Phys. Rev. E*'05; *J. Chem. Phys.* '05]

Concluding Remarks

1. Error Quantification and numerical accuracy of CG methods.

Information Theory and Quantity of Interest approaches.

2. Compression of the measure rather than just $H = H(X)$:

$$e^{-\beta H(X)} dX \sim \mu_{\mathbf{micro}}(\mathbf{dX}) \quad \mapsto \quad \mu_{\mathbf{cg}}(\mathbf{dQ}) \sim e^{-\beta \bar{H}(Q)} dQ$$

3. The role of multi-body CG interaction terms in the two-body CG interactions.

4. Adaptive CG schemes

5. "Reverse map"-reconstruct microscopic info from CG simulations.

Acknowledgments

S. Are (JPMorgan-Chase)

V. Harmandaris (Max Planck Inst. for Polymers, Mainz)

A. Majda (NYU),

P. Plecháč (Univ. of TN and ORNL),

L. Rey-Bellet (UMass)

A. Sopasakis (Univ. N. Carolina-Charlotte)

A. Szepessy (KTH, Sweden)

J.Trashorras (Paris IX, France)

D.Tsagkarogiannis (Max Planck Math. Sci. Inst., Leipzig)

D.G. Vlachos (Chem. Eng., Univ. of Delaware)

Supported in part by:

-National Science Foundation

-U. S. Department of Energy

-Marie-Curie IRG Programme

REFERENCES

Reviews

1. Chatterjie, Vlachos, J. Comput-Aided Mater. Des. (2007).
2. K., Plech´aˇc, Rey-Bellet, accepted, J. Sci. Comp. (2008).

Coarse-grained and Hybrid models

1. K., Majda, Vlachos, J. Comp. Phys. (2003)
2. K., Majda, Vlachos, Proc. Nat. Acad. Sci. (2003).
3. K., Vlachos, J. Chem. Phys. (2003).
4. Khouider, Majda, K., PNAS (2003).
5. K., Majda, Sopasakis, Comm. Math. Sci. (2004)/(2005).
6. K., Majda, Sopasakis, Nonlinearity (2006).
7. Chatterjee and D. G. Vlachos. J. Chem. Phys. (2006).
8. Chatterjie, Vlachos, Chem. Eng. Sci. (2007).

Error analysis and adaptivity

1. Chatterjee, K., Vlachos, Phys. Rev. E (2005).
2. Chatterjee, Vlachos, K., J. Chem. Phys. (2005).
3. K., Trashorras, J. Stat. Phys., (2006).
4. K., Plech´aˇc, Sopasakis, SIAM Num. Analysis, (2006).
5. K., Rey-Bellet, Plech´aˇc, Tsagkarogiannis, M2AN (2007).
6. K., Rey-Bellet, Plech´aˇc, Tsagkarogiannis, Jour. Non-Newt.Fluid Mech. (2008).

Temporal CG

1. Chatterjee, Vlachos, K., J. Chem. Phys. (2005).
2. K., Szepessy, Comm. Math. Sci., (2006).
3. Are, K., Szepessy, preprint (2008)