Mathematical strategies for coarse-graining and sensitivity analysis of high-dimensional stochastic systems.

Petr Plecháč

Dept. Mathematical Sciences, University of Delaware
plechac@math.udel.edu
http://www.math.udel.edu/~plechac

Stony Brook University, Apr 10, 2014

Department of Energy (OASCR) funded project:
Mathematical Foundations for Uncertainty Quantification in Materials Design

- P. Dupuis (Applied Mathematics, Brown University)
- G. Arampatzis, M.A. Katsoulakis, Y. Pantazis, L. Rey-Bellet (Mathematics, University of Massachusetts, Amherst)
- P. Plechac (Mathematics, University of Delaware)
- D. G. Vlachos (Chemical Engineering, University of Delaware)

NSF-CMMI: *CDI-Type II: Hierarchical Stochastic Algorithms for Materials engineering.*

DOE: *Multiscale Mathematics for biomass conversion to renewable hydrogen*

References:
http://www.math.udel.edu/~plechac
http://www.math.udel.edu/~plechac/CDIsite-php

Goals for this talk

(a) to present information theory-based tools for quantifying error, and model sensitivity,

Goals for this talk

(a) to present information theory-based tools for quantifying error, and model sensitivity,

(b) to derive coarse-graining schemes in a systematic manner,

Goals for this talk

(a) to present information theory-based tools for quantifying error, and model sensitivity,

(b) to derive coarse-graining schemes in a systematic manner,

(c) to understand the validity regimes of existing coarse-graining methods by developing a mathematical and statistical error analysis,

Goals for this talk

(a) to present information theory-based tools for quantifying error, and model sensitivity,

(b) to derive coarse-graining schemes in a systematic manner,

(c) to understand the validity regimes of existing coarse-graining methods by developing a mathematical and statistical error analysis,

(d) to explain parameterized effective dynamics for non-equilibrium systems,

Goals for this talk

(a) to present information theory-based tools for quantifying error, and model sensitivity,

(b) to derive coarse-graining schemes in a systematic manner,

(c) to understand the validity regimes of existing coarse-graining methods by developing a mathematical and statistical error analysis,

(d) to explain parameterized effective dynamics for non-equilibrium systems,

(f) to present coarse-graining as a computational tool.

Outline

- Model errors and sensitivity
- Role of relative entropy and relative entropy rate
- Error and sensitivity bounds
- Coarse-graining stochastic particle systems: Equilibrium
- Non-equilibrium steady states.
- Minimizing the error and parametrization CG models
- Coarse-graining and acceleration of MC simulations
- Benchmarks

Coarse-graining of Stochastic Processes

General task

1. Large-dimensional configuration/phase space $x \in \Sigma$
2. Stochastic process $\{X_t\}_{t>0}$, i.e., probability measure $\mu_t(dX)$
3. Observable $\Phi : \Sigma \to \mathbb{R}$
4. Compute/estimate $\mathbb{E}_{\mu_t}[\Phi(X_t)]$

Coarse-graining of Stochastic Processes

General task

1. Large-dimensional configuration/phase space $x \in \Sigma$
2. Stochastic process $\{X_t\}_{t>0}$, i.e., probability measure $\mu_t(dX)$
3. Observable $\Phi : \Sigma \to \mathbb{R}$
4. Compute/estimate $\mathbb{E}_{\mu_t}[\Phi(X_t)]$

Reduce number of degrees of freedom (DOFs)

1. Projection on a smaller space: $\mathbf{T} : \Sigma \to \bar{\Sigma}$, $\Sigma = \bar{\Sigma} \oplus \hat{\Sigma}$
2. Coarse-grained Stochastic process: $\{Y_t\}_{t>0}$, i.e., $\bar{\mu}_t(dy)$
3. Coarse observable: $\bar{\Phi} : \bar{\Sigma} \to \mathbb{R}$
4. Compute: $\mathbb{E}_{\bar{\mu}_t}[\bar{\Phi}(Y_t)]$
5. Estimate the error: $\left|\mathbb{E}[\bar{\Phi}(\mathbf{T}X_t)] - \mathbb{E}_{\bar{\mu}_t}[\bar{\Phi}(Y_t)]\right|$, $\mathcal{R}(\bar{\mu}_t \,||\, \mu_t)$

Examples

- Surface chemistry: microscopically active interface of boundary layer interacting with bulk (fluid) phase, pattern formation on surfaces.
- Magnetic elements: efficient simulation of mesoscopic inhomogenities in the presence of noise, external field varying on micron to cm scales, self-assembly, magnetic domains in thin films, nucleation and reversal processes in magnetic particles.
- Polymeric fluids: constitutive relations from *microscopic* models (e.g., FENE-type) coupled with fluid dynamics at the macroscopic level (continuum mechanics PDEs).
- Stochastic phase-field models: solidification, dendritic growth in alloys, phase transformations in solids.
- Atmosphere/Ocean interactions: tropical convection, subgrid phenomena
- Cell biology: epidermal growth factor binding/dimerization

Computational challenges and mathematical questions

- Disparity in scales *and* models: DNS require averaging of large systems

Computational challenges and mathematical questions

- Disparity in scales *and* models: DNS require averaging of large systems
- Model reduction: no clear scale separation lead to hierarchical *coarse-graining*

Computational challenges and mathematical questions

- Disparity in scales *and* models: DNS require averaging of large systems
- Model reduction: no clear scale separation lead to hierarchical *coarse-graining*
- Closures – stochastic vs deterministic: when is *randomness* important ?

Computational challenges and mathematical questions

- Disparity in scales *and* models: DNS require averaging of large systems
- Model reduction: no clear scale separation lead to hierarchical *coarse-graining*
- Closures – stochastic vs deterministic: when is *randomness* important ?
- Construction of effective potentials, forces.
- Numerical analysis:
  1. *error control*, stability, consistency
  2. allocation of computational resources: adaptive grids, model refinement.
  3. parallel computing

Desired properties of coarse-graining algorithms

- Compress long-range interactions, decrease number of DOFs, fast evaluation of interactions.
- Coupling with meanfield models when fluctuations are not important.
- Correct energy transport between different scales.
- Larger time-steps, simulations over longer time scales.
- Allow for large length-scale simulations and long time scales
- Correct statistical mechanics limits.
- Retain the correct noise of microscopic models (nucleation, phase transitions, switching etc are properly modelled at larger scales)
- Mathematical analysis is possible in order to assess and control errors.

Model error and sensitivity

Two probabilistic models P and Q on the common measurable space $(\Omega, \mathcal{B})$

Our applications: reaction networks, spatially heterogeneous chemical kinetics, molecular systems at equilibrium or with non-equilibrium steady states

Inspired by works:
D. Giannakis, A. J. Majda I. Horenko,*Physica D* (2012)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2011)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2010)
extend empirical information theory techniques for path-space application in

- Discrimination between the two models – distance

Model error and sensitivity

Two probabilistic models P and Q on the common measurable space $(\Omega, \mathcal{B})$

Our applications: reaction networks, spatially heterogeneous chemical kinetics, molecular systems at equilibrium or with non-equilibrium steady states

Inspired by works:
D. Giannakis, A. J. Majda I. Horenko,*Physica D* (2012)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2011)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2010)
extend empirical information theory techniques for path-space application in

- Discrimination between the two models – distance
- Error for observables: $|\mathbb{E}_P[f] - \mathbb{E}_Q[f]|$

Model error and sensitivity

Two probabilistic models P and Q on the common measurable space $(\Omega, \mathcal{B})$

Our applications: reaction networks, spatially heterogeneous chemical kinetics, molecular systems at equilibrium or with non-equilibrium steady states

Inspired by works:
D. Giannakis, A. J. Majda I. Horenko,*Physica D* (2012)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2011)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2010)
extend empirical information theory techniques for path-space application in

- Discrimination between the two models – distance
- Error for observables: $|\mathbb{E}_P[f] - \mathbb{E}_Q[f]|$
- Sensitivity under perturbations $P^\theta \to Q \equiv P^{\theta+\epsilon}$

Model error and sensitivity

Two probabilistic models P and Q on the common measurable space $(\Omega, \mathcal{B})$

Our applications: reaction networks, spatially heterogeneous chemical kinetics, molecular systems at equilibrium or with non-equilibrium steady states

Inspired by works:
D. Giannakis, A. J. Majda I. Horenko,*Physica D* (2012)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2011)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2010)
extend empirical information theory techniques for path-space application in

- Discrimination between the two models – distance
- Error for observables: $|\mathbb{E}_P[f] - \mathbb{E}_Q[f]|$
- Sensitivity under perturbations $P^\theta \to Q \equiv P^{\theta+\epsilon}$
- Parameter identifiability in parameterized models P^θ

Model error and sensitivity

Two probabilistic models P and Q on the common measurable space $(\Omega, \mathcal{B})$

Our applications: reaction networks, spatially heterogeneous chemical kinetics, molecular systems at equilibrium or with non-equilibrium steady states

Inspired by works:
D. Giannakis, A. J. Majda I. Horenko, *Physica D* (2012)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2011)
A. J. Majda, B. Gershgorin, *Proc. Natl. Acad. Sci.* (2010)
extend empirical information theory techniques for path-space application in

- Discrimination between the two models – distance
- Error for observables: $|\mathbb{E}_P[f] - \mathbb{E}_Q[f]|$
- Sensitivity under perturbations $P^\theta \to Q \equiv P^{\theta+\epsilon}$
- Parameter identifiability in parameterized models P^θ
- "Best-fit" for reduced models.

Relative Entropy and $\mathcal{R}$-projections

- Pseudo-distance (Kullback-Leibler divergence)

$$\mathcal{R}(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) \, dP$$

for $P \ll R$, $Q \ll R$ $\quad \mathcal{R}(P \,||\, Q) = \int p_R \log\left(\frac{p_R}{q_R}\right) \, dR$

Relative Entropy and $\mathcal{R}$-projections

- Pseudo-distance (Kullback-Leibler divergence)

$$\mathcal{R}(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) \, dP$$

for $P \ll R$, $Q \ll R$ $\quad \mathcal{R}(P \,||\, Q) = \int p_R \log\left(\frac{p_R}{q_R}\right) \, dR$

- Properties: (i) $\mathcal{R}(P \,||\, Q) \geq 0$ and
  (ii) $\mathcal{R}(P \,||\, Q) = 0$ iff $P = Q$ a.e.

Relative Entropy and $\mathcal{R}$-projections

- Pseudo-distance (Kullback-Leibler divergence)

$$\mathcal{R}(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) \, dP$$

for $P \ll R$, $Q \ll R$ $\quad \mathcal{R}(P \,||\, Q) = \int p_R \log\left(\frac{p_R}{q_R}\right) \, dR$

- Properties: (i) $\mathcal{R}(P \,||\, Q) \geq 0$ and
  (ii) $\mathcal{R}(P \,||\, Q) = 0$ iff $P = Q$ a.e.

- $\mathcal{R}$-geometry of probability distributions $\mathcal{B}(R, \rho) = \{P \,||\, \mathcal{R}(P \,||\, R) < \rho\}$
  $\mathcal{R}$-projection on $\mathcal{A}$ convex, TV closed, $\mathcal{A} \cap \mathcal{B}(R, \rho) \neq \emptyset$ (Kullback, Csiszár)

$$\mathcal{R}(Q \,||\, R) = \min_{P \in \mathcal{A}} \mathcal{R}(P \,||\, R)$$

Relative Entropy and $\mathcal{R}$-projections

- Pseudo-distance (Kullback-Leibler divergence)

$$\mathcal{R}(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) dP$$

for $P \ll R$, $Q \ll R$ $\quad \mathcal{R}(P \,||\, Q) = \int p_R \log\left(\frac{p_R}{q_R}\right) dR$

- Properties: (i) $\mathcal{R}(P \,||\, Q) \geq 0$ and
  (ii) $\mathcal{R}(P \,||\, Q) = 0$ iff $P = Q$ a.e.

- $\mathcal{R}$-geometry of probability distributions $\mathcal{B}(R, \rho) = \{P \,||\, \mathcal{R}(P \,||\, R) < \rho\}$
  $\mathcal{R}$-projection on $\mathcal{A}$ convex, TV closed, $\mathcal{A} \cap \mathcal{B}(R, \rho) \neq \emptyset$ (Kullback, Csiszár)

$$\mathcal{R}(Q \,||\, R) = \min_{P \in \mathcal{A}} \mathcal{R}(P \,||\, R)$$

- "Geometry": tangent hyperplane to $\mathcal{B}(R, \rho)$ at Q, $\rho = \mathcal{R}(Q \,||\, R)$

$$P \text{ s.t. } \int \log \frac{dQ}{dR}\, dP = \rho\,, \quad \mathcal{R}(P \,||\, R) = \mathcal{R}(P \,||\, Q) + \mathcal{R}(Q \,||\, R)$$

Relative Entropy and $\mathcal{R}$-projections

- Pseudo-distance (Kullback-Leibler divergence)

$$\mathcal{R}(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) \, dP$$

for $P \ll R$, $Q \ll R$ $\quad \mathcal{R}(P \,||\, Q) = \int p_R \log\left(\frac{p_R}{q_R}\right) \, dR$

- Properties: (i) $\mathcal{R}(P \,||\, Q) \geq 0$ and
  (ii) $\mathcal{R}(P \,||\, Q) = 0$ iff $P = Q$ a.e.

- The "best fit" in relative entropy: $\min_{R \in \mathcal{A}} \mathcal{R}(R \,||\, Q)$
  modeling error + numerical error + statistical error
  Modelling error$\sim \mathcal{R}(P \,||\, Q) \sim \epsilon^{\alpha}$
  Bounds on the weak error:

$$|\mathbb{E}_P[f] - \mathbb{E}_Q[f]| \leq C_f \Phi(\mathcal{R}(P \,||\, Q))$$

Bounding the approximation error

modeling error

Bounding the approximation error

modeling error + numerical error

Bounding the approximation error

modeling error + numerical error + statistical error

Csiszar-Kullback-Pinsker inequality:

$$\|P - Q\|_{\mathrm{TV}} \leq \sqrt{2\mathcal{R}\left(P \,||\, Q\right)}$$

Bounding the approximation error

modeling error + numerical error + statistical error

Csiszar-Kullback-Pinsker inequality:

$$\|P - Q\|_{\mathrm{TV}} \leq \sqrt{2\mathcal{R}\left(P \,||\, Q\right)}$$

χ^2-divergence:

$$\chi^2\left(P \,||\, Q\right) = \int \left(\frac{dP}{dQ} - 1\right)^2 dQ\,, \quad \text{if } P \ll Q,$$

Property:

$$\mathcal{R}\left(P \,||\, Q\right) \leq \chi^2\left(P \,||\, Q\right)\,.$$

CG: Error Quantification and Parameterization using RE in molecular simulations: Katsoulakis, P.P. Sopasakis (2006), M.S. Shell (2008), Katsoulakis, P.P., Rey-Bellet, Tsagkarogiannis (07, 08, 09, 13), M.S. Shell (08,12), Bilionis et al (2012), Zabaras et al (2013) ...

Two bounds:
χ^2 bound:

$$|\mathbb{E}_P[f] - \mathbb{E}_Q[f]| \leq \sqrt{\mathrm{Var}_Q[f]}\sqrt{\chi^2(P \,||\, Q)}$$

$$\begin{aligned}
|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| &= \Big| \int f \Big(1 - \frac{dP}{dQ}\Big)\, dQ \Big| = \Big| \int f \Big(1 - \frac{dP}{dQ}\Big)\, dQ - \mathbb{E}_Q[f] \int \Big(1 - \frac{dP}{dQ}\Big)\, dQ \Big| \\
&= \Big| \int (f - \mathbb{E}_Q[f]) \Big(1 - \frac{dP}{dQ}\Big)\, dQ \\
&\leq \left(\int (f - \mathbb{E}_Q[f])^2\, dQ \right)^{1/2} \left(\Big(1 - \frac{dP}{dQ}\Big)^2\, dQ \right)^{1/2} = \sqrt{\mathrm{Var}_Q[f]}\sqrt{\chi^2(P \,||\, Q)}\,.
\end{aligned}$$

Two bounds:
χ^2 bound:

$$|\mathbb{E}_P[f] - \mathbb{E}_Q[f]| \leq \sqrt{\mathrm{Var}_Q[f]}\sqrt{\chi^2(P\,||\,Q)}$$

$$\begin{aligned}
|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| &= \Big|\int f\Big(1 - \frac{dP}{dQ}\Big)\,dQ\Big| = \Big|\int f\Big(1 - \frac{dP}{dQ}\Big)\,dQ - \mathbb{E}_Q[f]\int\Big(1 - \frac{dP}{dQ}\Big)\,dQ\Big| \\
&= \Big|\int (f - \mathbb{E}_Q[f])\Big(1 - \frac{dP}{dQ}\Big)\,dQ \\
&\leq \left(\int (f - \mathbb{E}_Q[f])^2\,dQ\right)^{1/2}\left(\Big(1 - \frac{dP}{dQ}\Big)^2\,dQ\right)^{1/2} = \sqrt{\mathrm{Var}_Q[f]}\sqrt{\chi^2(P\,||\,Q)}.
\end{aligned}$$

Csiszar-Kullback-Pinsker bound: $\|P - Q\|_{\mathrm{TV}} \leq \sqrt{2\mathcal{R}(P\,||\,Q)}$

$$|\mathbb{E}_P[\phi] - \mathbb{E}_Q[\phi]| \leq \|f\|_\infty\sqrt{2\mathcal{R}(P\,||\,Q)}$$

Variational bounds

Variational characterization of the logarithmic moment generating function

Theorem

Let $(\Omega, \mathcal{B})$ be probability space, f bounded measurable function on Ω and $P \in \mathcal{P}(\Omega)$

$$\frac{1}{c} \log \mathbb{E}_P[e^{cf}] = \sup_{Q \in \mathcal{P}(\Omega)} \left\{ \mathbb{E}_Q[f] - \frac{1}{c} \mathcal{R}(Q \,||\, P) \right\}$$

Corollary: For $f - \mathbb{E}_P[f]$

$$\mathbb{E}_Q[f] - \mathbb{E}_P[f] \leq \frac{1}{c} \log \mathbb{E}_P[e^{c(f - \mathbb{E}_P[f])}] + \frac{1}{c} \mathcal{R}(Q \,||\, P)$$

$$\mathbb{E}_Q[f] - \mathbb{E}_P[f] \geq -\frac{1}{c} \log \mathbb{E}_P[e^{-c(f - \mathbb{E}_P[f])}] - \frac{1}{c} \mathcal{R}(Q \,||\, P)$$

$$\tilde{\Lambda}_{P,f}(c) \equiv \log \mathbb{E}_P[e^{c(f-\mathbb{E}_P[f])}] = \sup_{Q \ll P} \{c(\mathbb{E}_Q[f] - \mathbb{E}_P[f]) - \mathcal{R}(Q \,||\, P)\} \,.$$

Tight variational bounds [Chowdhary& Dupuis 2009]

$$\sup_{c>0} \left\{ -\frac{1}{c}\tilde{\Lambda}_{P,f}(-c) - \frac{1}{c}\mathcal{R}(Q \,||\, P) \right\} \leq \mathbb{E}_Q[f] - \mathbb{E}_P[f] \leq$$

$$\leq \inf_{c>0} \left\{ \frac{1}{c}\tilde{\Lambda}_{P,f}(c) + \frac{1}{c}\mathcal{R}(Q \,||\, P) \right\}$$

Side remark: Let $\Psi : \mathbb{R} \to \mathbb{R}$ be a convex and such that $\Psi(0) = \Psi'(0) = 0$ and

$$\tilde{\Lambda}_{P,f}(c) \equiv \log \mathbb{E}_P[e^{c(f-\mathbb{E}_P[f])}] \leq \Psi(c) \,,$$

and define $\Psi^{\sharp}_{+}(t) = \inf_{c>0} \left\{ \frac{1}{c}(t + \Psi(c)) \right\}$ then

$$\mathbb{E}_Q[f] - \mathbb{E}_P[f] \leq \Psi^{\sharp}_{+}(\mathcal{R}(Q \,||\, P)) \,.$$

Example: $f = \mathbb{1}_{A^*}$ then Csiszar-Kullback-Pinsker inequality.

Asymptotics

Lemma

Assume the cumulant generating function $\tilde{\Lambda}_{P,f}(c) \equiv \log \mathbb{E}_P[e^{c(f-\mathbb{E}_P[f])}]$ exists in a neighborhood of the origin and write $\rho^2 = \mathcal{R}(Q \,||\, P)$. Unique solution $c^*(\rho)$ of

$$(P_+) \qquad \inf_{c>0} \left\{ \frac{1}{c}\tilde{\Lambda}_{P,f}(c) + \frac{1}{c}\mathcal{R}(Q \,||\, P) \right\},$$

$$(P_-) \qquad \sup_{c>0} \left\{ -\frac{1}{c}\tilde{\Lambda}_{P,f}(-c) - \frac{1}{c}\mathcal{R}(Q \,||\, P) \right\}.$$

Furthermore, $c^*(\rho)$ is C^∞ in a neighborhood of $\rho = 0$ and admits the expansion

$$c^*(\rho) = c_1^* \rho + \mathcal{O}(\rho^2), \quad c_1^* = \sqrt{\frac{2}{\mathrm{Var}_P[f]}}. \tag{1}$$

Error estimate of the type:

$$|\mathbb{E}_P[f] - \mathbb{E}_Q[f]| \leq C_f \Phi(\mathcal{R}(P \,||\, Q))$$

Theorem

$$|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| \leq \sqrt{\mathrm{Var}_P[f]} \sqrt{2\mathcal{R}(Q \,||\, P)} + \mathcal{O}(\mathcal{R}(Q \,||\, P)),$$

$\mathcal{O}(\mathcal{R}(P \,||\, Q))$ can be further quantified using the asymptotic expansions of $c^*(\rho)$.

$$\mathcal{R}(Q \,||\, P) = \int \log\left(\frac{dQ}{dP}\right) \, dQ$$

Sensitivity Analysis – Methods – Background

- Stochastic Sensitivity Analysis:
  - Observable-based:
    - Finite difference: (biased, problems with variance)

$$S_f(\theta,t) = \frac{\partial}{\partial\theta}\mathbb{E}_{P_t^\theta}[f] \approx \frac{1}{\epsilon}(\mathbb{E}_{P_t^{\theta+\epsilon}}[f] - \mathbb{E}_{P_t^\theta}[f])$$

$$S_f(\theta,t) = \frac{\partial}{\partial\theta}\mathbb{E}_{P_t^\theta}[f] \approx \frac{1}{2\epsilon}(\mathbb{E}_{P_t^{\theta+\epsilon}}[f] - \mathbb{E}_{P_t^{\theta-\epsilon}}[f])$$

    - pathwise methods (unbiased) [P. Glasserman (1991)]

$$\frac{\partial}{\partial\theta}\mathbb{E}[f_t(\theta)] = \mathbb{E}\Big[\frac{\partial}{\partial\theta}f_t(\theta)\Big]$$

    - Likelihood ratio method (unbiased) [P. Glynn, *Comm. ACM* (1990)]:

$$S(\theta,t) = \frac{\partial}{\partial\theta}\mathbb{E}_{P_t^\theta}[f] = \int f(x)\partial_\theta P_t^\theta(x)dx = \mathbb{E}_{P_t^\theta}[f\,\partial_\theta \log P_t^\theta]$$

Sensitivity analysis – Methods – Background

- Stochastic Sensitivity Analysis (Cont'd):
  - Density-based: Relative entropy, Fisher Information Matrix, Mutual Information.

    H. Liu, W. Chen, and A. Sudjianto, *J. Mech. Des.* (2006).
    N. Ludtke et al., *J. Royal Soc., Interface* (2008).
    A. J. Majda and B. Gershgorin, *Proc. Natl. Acad. Sci.* (2010).

  - The PDF is assumed known, e.g. a Gibbs equilibrium $\sim Ce^{-\beta H(\sigma)}$ or Gaussian fluctuations.

  - However, typically this is not the case in dynamics, non-equilibrium systems, non-gaussian fluctuations, etc.

Bounding sensitivity constant (robustness)

Sensitivity $S_\theta(f)$ constant of the observable f

$$|\mathbb{E}_{P^\theta}[f] - \mathbb{E}_{P^{\theta+\epsilon}}[f]| = S_f(\theta)\epsilon + o(\epsilon)$$

Computing directly the sensitivity of observables – difficult

Infinitesimal structure of $\mathcal{R}\left(P^\theta \,||\, P^{\theta+\epsilon}\right)$:

$$\mathcal{R}\left(P^\theta \,||\, P^{\theta+\epsilon}\right) = \frac{1}{2}\epsilon^T \mathbf{F}(P^\theta)\epsilon + \mathcal{O}(|\epsilon|^3)$$

Fisher Information Matrix (FIM):

$\mathbf{F}(P^\theta)_{ij} = \int \frac{\partial \log p_R^\theta}{\partial \theta_i} \frac{\partial \log p_R^\theta}{\partial \theta_j} p_R^\theta dR$

Bounding sensitivity constant (robustness)

Sensitivity $S_\theta(f)$ constant of the observable f

$$|\mathbb{E}_{P^\theta}[f] - \mathbb{E}_{P^{\theta+\epsilon}}[f]| = S_f(\theta)\epsilon + o(\epsilon)$$

Computing directly the sensitivity of observables – difficult

Infinitesimal structure of $\mathcal{R}\left(P^\theta \,||\, P^{\theta+\epsilon}\right)$:

$$\mathcal{R}\left(P^\theta \,||\, P^{\theta+\epsilon}\right) = \frac{1}{2}\epsilon^T \mathbf{F}(P^\theta)\epsilon + \mathcal{O}(|\epsilon|^3)$$

Fisher Information Matrix (FIM):
$\mathbf{F}(P^\theta)_{ij} = \int \frac{\partial \log p_R^\theta}{\partial \theta_i} \frac{\partial \log p_R^\theta}{\partial \theta_j} p_R^\theta dR$

Theorem (Stability/Sensitivity bound)

$$\frac{1}{\epsilon}|\mathbb{E}_{P^{\theta+\epsilon}}[f] - \mathbb{E}_{P^\theta}[f]| \leq \sqrt{\mathrm{Var}_{P^\theta}[f]}\sqrt{\mathbf{F}(P^\theta)}$$

Relative entropy on the path space

Markov chains on Σ:
$\{\sigma_n\}_{n\in\mathbb{Z}^+}$, $p^\theta(\sigma,\sigma')$, $\mu^\theta(\sigma)$
$\{\tilde\sigma_n\}_{n\in\mathbb{Z}^+}$, $\tilde p^\theta(\sigma,\sigma')$, $\tilde\mu^\theta(\sigma)$
Path measures:

$$Q^\theta(\sigma_0,\dots,\sigma_M)=\mu^\theta(\sigma_0)p^\theta(\sigma_0,\sigma_1)\dots p^\theta(\sigma_{M-1},\sigma_M)$$

Radon-Nikodym derivative

$$\frac{dQ^\theta}{d\tilde Q^\theta}(\{\sigma_n\})=\frac{\mu^\theta(\sigma_0)\prod_{i=0}^{M-1}p^\theta(\sigma_i,\sigma_{i+1})}{\tilde\mu^\theta(\sigma_0)\prod_{i=0}^{M-1}\tilde p^\theta(\sigma_i,\sigma_{i+1})}$$

Relative entropy

$$\mathcal{R}\left(Q^\theta\,||\,\tilde Q^\theta\right)=\int_\Sigma\dots\int_\Sigma\mu^\theta(\sigma_0)\prod_{i=0}^{M-1}p^\theta(\sigma_i,\sigma_{i+1})\log\frac{\mu^\theta(\sigma_0)\prod_{i=0}^{M-1}p^\theta(\sigma_i,\sigma_{i+1})}{\tilde\mu^\theta(\sigma_0)\prod_{i=0}^{M-1}\tilde p^\theta(\sigma_i,\sigma_{i+1})}\,d\sigma$$

Relative entropy decomposition

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})} \right) d\sigma_0 \dots d\sigma_M$$

Relative entropy decomposition

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})} \right) d\sigma_0 \dots d\sigma_M$$

$$\int_\Sigma p(\sigma, \sigma')\, d\sigma' = 1\,, \quad \int_\Sigma \mu(\sigma) p(\sigma, \sigma')\, d\sigma = \mu(\sigma')$$

Relative entropy decomposition

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})} \right) d\sigma_0 \dots d\sigma_M$$

$$\int_\Sigma p(\sigma, \sigma')\, d\sigma' = 1\,, \quad \int_\Sigma \mu(\sigma) p(\sigma, \sigma')\, d\sigma = \mu(\sigma')$$

$$\int_\Sigma \mu^\theta(\sigma_0) \log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)}\, d\sigma_0 + \sum_{i=0}^{M-1} \int_\Sigma \int_\Sigma \mu^\theta(\sigma_i) p^\theta(\sigma_i) \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}$$

$$= M \mathbb{E}^\theta_\mu \left[\int_\Sigma p^\theta(\sigma, \sigma') \log \frac{p^\theta(\sigma, \sigma')}{\tilde{p}^\theta(\sigma, \sigma')}\, d\sigma' \right] + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

Relative entropy decomposition

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})} \right) d\sigma_0 \dots d\sigma_M$$

$$\int_\Sigma p(\sigma, \sigma')\, d\sigma' = 1\,, \quad \int_\Sigma \mu(\sigma) p(\sigma, \sigma')\, d\sigma = \mu(\sigma')$$

$$\int_\Sigma \mu^\theta(\sigma_0) \log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)}\, d\sigma_0 + \sum_{i=0}^{M-1} \int_\Sigma \int_\Sigma \mu^\theta(\sigma_i) p^\theta(\sigma_i) \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}$$

$$= M \mathbb{E}^\theta_\mu \left[\int_\Sigma p^\theta(\sigma, \sigma') \log \frac{p^\theta(\sigma, \sigma')}{\tilde{p}^\theta(\sigma, \sigma')}\, d\sigma' \right] + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = M\mathcal{H}(Q^\theta \,||\, \tilde{Q}^\theta) + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

Relative entropy rate (RER)

Relative Entropy Rate (RER):

$$\mathcal{H}(p \,||\, q) = \lim_{T\to\infty} \frac{1}{T} \mathcal{R}\left(P_{[0,T]} \,||\, Q_{[0,T]}\right) .$$

- Markov chains

$$\mathcal{H}(p \,||\, q) = \int \mu(dx) \int p(x,x') \log \frac{p(x,x')}{q(x,x')}\, dx' = \mathbb{E}_\mu \Big[\int p(x,x') \log \frac{p(x,x')}{q(x,x')}\, d$$

  Note

$$\mathcal{H}(p \,||\, q) = \mathcal{R}\left(\mu \otimes p \,||\, \mu \otimes q\right)$$

- Continuous time Markov chains

$$\mathcal{H}(c \,||\, \tilde{c}) = \sum_{x\in\mathcal{X}} \sum_{x'\in\mathcal{X}} \mu(x) c(x,x') \log \frac{c(x,x')}{\tilde{c}(x,x')} - \sum_{x\in\mathcal{X}} \mu(x)(\lambda(x) - \tilde{\lambda}(x)) .$$

Properties:

- RER inherits properties of Relative Entropy:

$$\mathcal{H}(P \,||\, Q) = \mathcal{R}\left(\mu \otimes p \,||\, \mu \otimes q\right),$$

$$\mu \otimes p(A \times B) = \sum_{\sigma \in A} \mu(\sigma) \sum_{\sigma' \in B} p(\sigma, \sigma').$$

- Infers information regarding the path distribution: steady-state distribution + stationary dynamics.
- RER is an observable + statistical estimators $\Rightarrow$ computationally tractable using (fast, scalable, etc) molecular solvers.
- Not necessary to know the steady states μ explicitly: suitable for reaction networks, reaction-diffusion and other non-equilibrium systems.
- Applicable to the transient regime.

Bounds for the error

$$|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| \leq \sqrt{\mathrm{Var}_P[f]}\sqrt{2\mathcal{R}(Q\,||\,P)} + \mathcal{O}(\mathcal{R}(Q\,||\,P)),$$

and sensitivity

$$\frac{1}{\epsilon}|\mathbb{E}_{P^{\theta+\epsilon}}[f] - \mathbb{E}_{P^\theta}[f]| \leq \sqrt{\mathrm{Var}_{P^\theta}[f]}\sqrt{\mathbf{F}(P^\theta)}$$

Bounds for the error

$$|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| \leq \sqrt{\mathrm{Var}_P[f]}\sqrt{2\mathcal{R}(Q\,||\,P)} + \mathcal{O}(\mathcal{R}(Q\,||\,P)),$$

and sensitivity

$$\frac{1}{\epsilon}|\mathbb{E}_{P^{\theta+\epsilon}}[f] - \mathbb{E}_{P^\theta}[f]| \leq \sqrt{\mathrm{Var}_{P^\theta}[f]}\sqrt{\mathbf{F}(P^\theta)}$$

Similar bounds on the path space ?

Measurable functional $\mathcal{F}$ of the process $\{X_t\}_{t\geq 0}$

$$|\mathbb{E}_{Q_{[0,T]}}[\mathcal{F}] - \mathbb{E}_{P_{[0,T]}}[\mathcal{F}]| \leq \sqrt{\frac{1}{T}\mathrm{Var}_{P_{[0,T]}}[T\mathcal{F}]}\sqrt{\frac{2}{T}\mathcal{R}\left(Q_{[0,T]} \,||\, P_{[0,T]}\right)} + \mathcal{O}\Big(\frac{1}{T}\mathcal{R}\left(Q_{[0,T]} \,||\, P_{[0,T]}\right)\Big)$$

Recall for stationary process

$$\frac{1}{T}\mathcal{R}\left(Q_{[0,T]} \,||\, P_{[0,T]}\right) = \mathcal{H}(p \,||\, q) + \frac{1}{T}\mathcal{R}\left(\mu \,||\, \nu\right)$$

Particular class of observables: $\mathcal{F}(X) = \frac{1}{T}\sum_{k=0}^{T} f(X_k)$
$(\mathcal{F}(X) = \frac{1}{T}\int_0^T f(X_s)\,ds)$

$$\frac{1}{T}\mathrm{Var}_{P_{[0,T]}}[T\mathcal{F}] = \mathrm{Var}_{\mu}[f] + 2\sum_{k=1}^{T}(1 - \frac{k}{T})A_f(k) \equiv \tau_T(f)$$

$$A_f(t) \equiv \mathbb{E}_{P_{[0,T]}}[(X_0 - \mathbb{E}_{\mu}[X_0])(X_t - \mathbb{E}_{\mu}[X_t])]$$

Path-space Fisher Information Matrix (FIM)

Under a smoothness assumption on θ, (checkable, on the rates only!)

$$\mathcal{H}\left(Q^{\theta}_{0,M} \,||\, Q^{\theta+\epsilon}_{0,M}\right) = \frac{1}{2}\epsilon^{T}\mathbf{F}_{\mathcal{H}}\left(Q^{\theta}_{0,M}\right)\epsilon + O(|\epsilon|^{3})$$

where the Fisher Information Matrix is defined as

$$\mathbf{F}_{\mathcal{H}}\left(Q^{\theta}_{0,M}\right) = \mathbb{E}_{\mu^{\theta}}\left[\int_{E} p^{\theta}(\sigma,\sigma')\nabla_{\theta}\log p^{\theta}(\sigma,\sigma')\nabla_{\theta}\log p^{\theta}(\sigma,\sigma')^{T}\,d\,\sigma'\right]$$

- Spectral analysis of FIM gives the most/least sensitive directions.
- Derivative-free sensitivity analysis method.
- Characterizes robustness on parameter perturbations.
- Determines parameter identifiability, [e.g. Cramer-Rao Theorems].
- Optimal Experimental Design via path-wise FIM, e.g. D/A-optimality tests.

Sensitivity bounds – path-wise estimate under perturbation

$P \equiv P^\theta$ and $Q \equiv P^{\theta+\epsilon}$

$$\frac{1}{|\epsilon|}|\mathbb{E}_{Q_{[0,T]}}[\mathcal{F}]-\mathbb{E}_{P_{[0,T]}}[\mathcal{F}]| \leq \sqrt{\frac{1}{T}\mathrm{Var}_{P^\theta_{[0,T]}}[T\mathcal{F}]}\sqrt{e^T\left(\mathbf{F}_{\mathcal{H}}(p^\theta)+\frac{1}{T}\mathbf{F}(\mu^\theta)\right)e}+\mathcal{O}(\epsilon)\,,$$

Sensitivity index:

$$|S_{\mathcal{F}}(P_{[0,T]})| \leq \sqrt{\frac{1}{T}\mathrm{Var}_{P^\theta_{[0,T]}}[T\mathcal{F}]}\sqrt{e^T\left(\mathbf{F}_{\mathcal{H}}(p^\theta)+\frac{1}{T}\mathbf{F}(\mu^\theta)\right)e}$$

Ergodic-type observables: $\mathcal{F}(X) = \frac{1}{T}\sum_{t=1}^{T} f(X_t)$ as $T \to \infty$ perturbations $\epsilon = |\epsilon| e$ of the invariant measure μ^θ

$$\frac{1}{|\epsilon|}|\mathbb{E}_{\mu^{\theta+\epsilon}}[f] - \mathbb{E}_{\mu^\theta}[f]| \leq \sqrt{\tau(f)}\sqrt{e^T \mathbf{F}_{\mathcal{H}}(p^\theta) e} + \mathcal{O}(\epsilon)$$

or equivalently

$$|S_f(\mu^\theta)| \leq \sqrt{\tau(f)}\sqrt{e^T \mathbf{F}_{\mathcal{H}}(p^\theta) e}$$

Integrated Autocorrelation Time (IAT)

$$\tau(f) := \lim_{T\to\infty} \tau_T(f) = \mathrm{Var}_\mu^\theta[f] + 2\sum_{k=1}^{\infty} A_f(k)$$

Epidermal Growth Factor Receptor (EGFR)

Well-mixed reaction system

- The EGFR model describes signaling phenomena of (mammalian) cells.

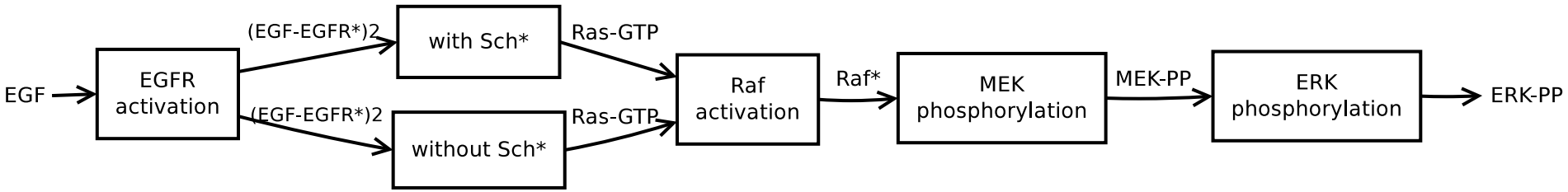


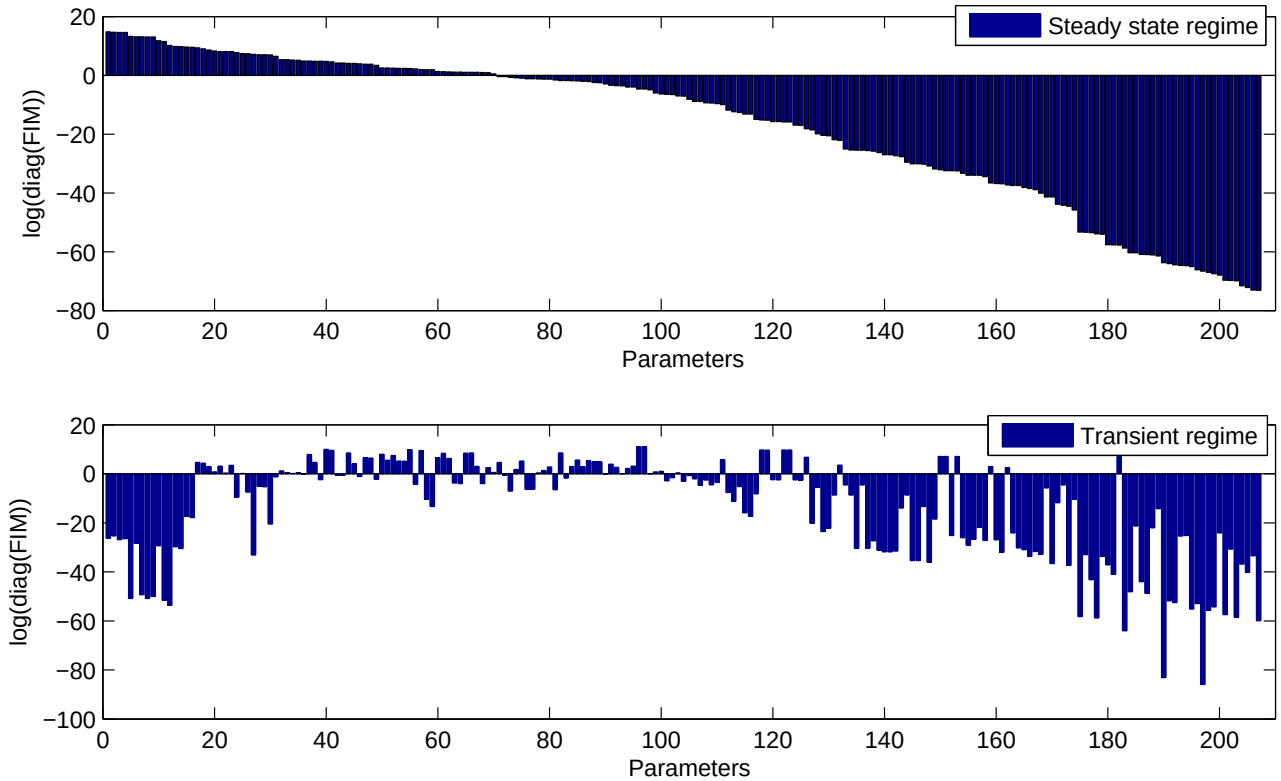
Figure: Building blocks of the EGFR reaction network.

- **94 species, 207 reactions, 207 parameters** (reaction constants).

Schoeberl B, C EJ, Gilles E, Muller G, *Nature Biotech.*, 2002.

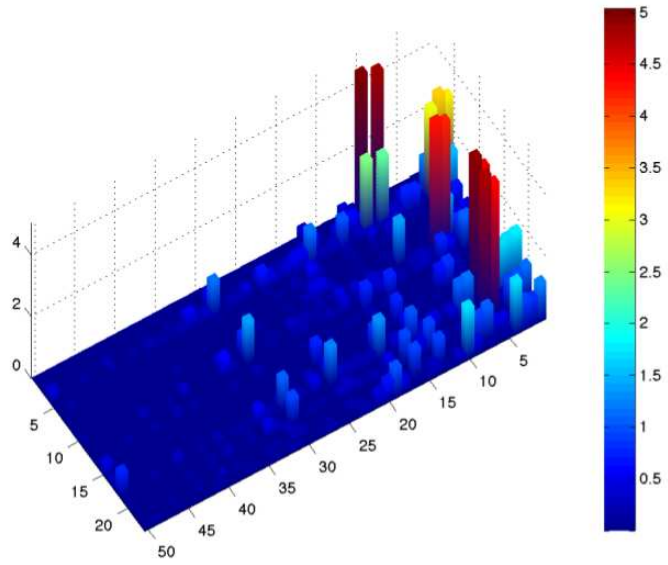
M. Katsoulakis, Y. Pantazis, D. Vlachos, *BMC Bioinformatics*, 2013

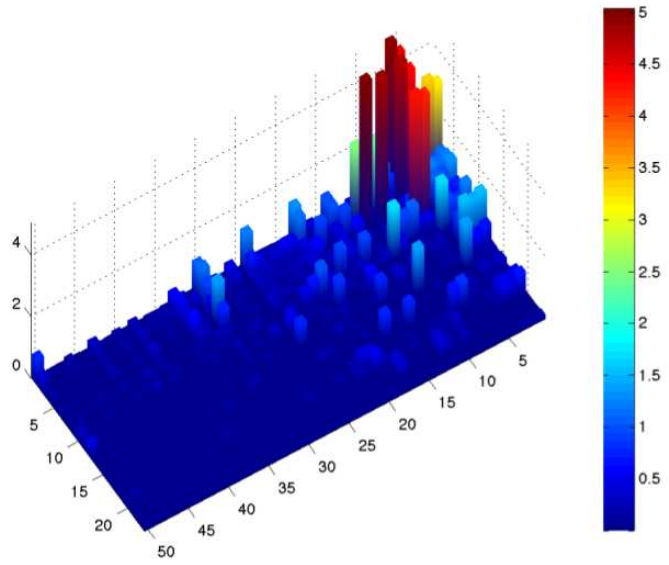
EGFR - FIM



Diagonal elements of the FIM computed at the steady state regime (upper plot) and at the transient regime (lower plot). Parameter sensitivities differ by orders of magnitude; most parameters insensitive.

EGFR – Goal oriented sensitivity





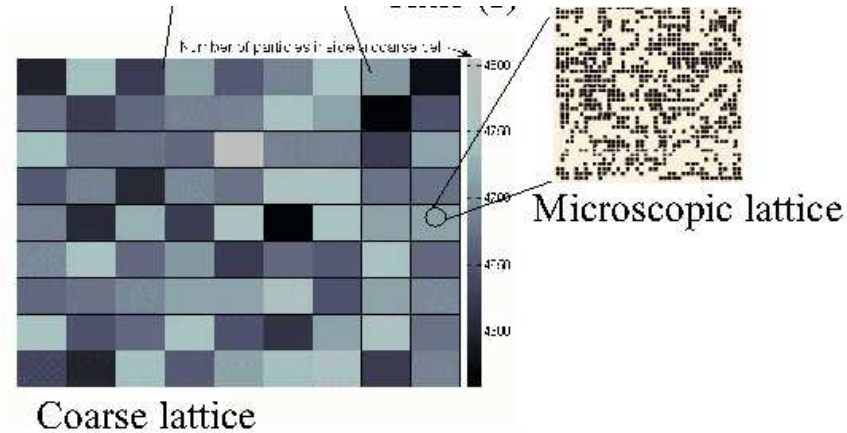
Ordering of parameters by their sensitivity bounds

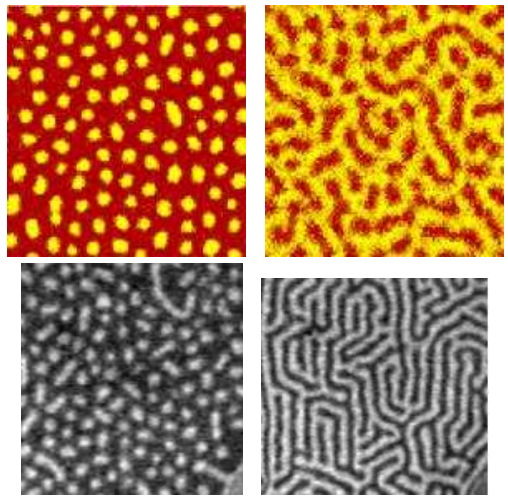
Lattice Models

Coarse-graining in spin systems

Examples: catalysis, epitaxial growth, micromagnetics, etc.

$$\text{block-spins } \{\sigma(x)\} \mapsto \text{block-spin } \{\eta(k)\} = \mathbf{T}\sigma = \sum_{x \in C_k} \sigma(x)$$





Patterning through self-assembly: CGMC simulations (top) vs experiment (bottom)

Intractable with conventional KMC due to μm scales

Sensitivity to entropic effects at finite temperature

statistical comparison

Lattice simulations with Ising and continuous spins

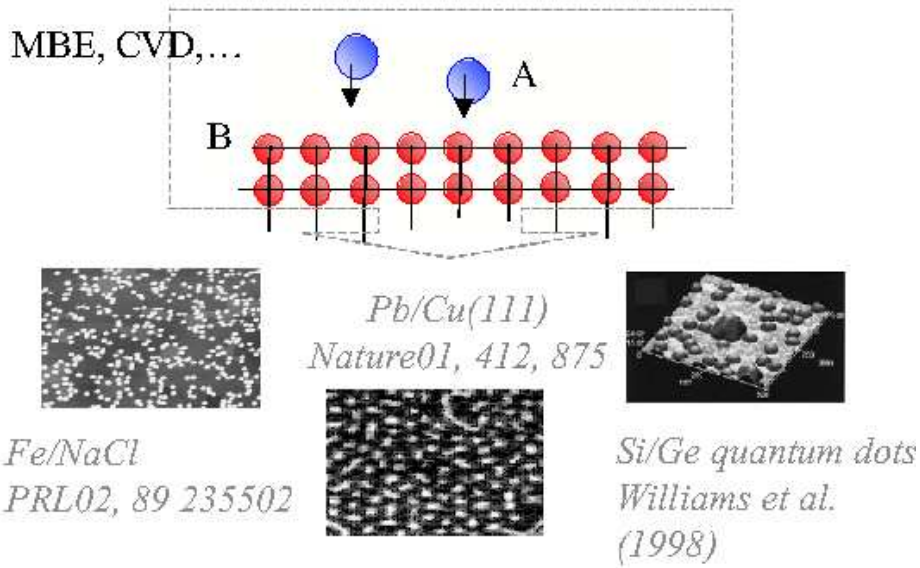
- emphasis on dynamics: Coarse-grained Monte Carlo (CGMC)
  Katsoulakis, Majda, Vlachos, Proc. Natl. Acad. Sci. (2003), Katsoulakis, PP, Sopasakis, SIAM Num. Anal. (2006); Are, Katsoulakis, PP, Rey-Bellet SIAM J.Sci.Comp. 2008; Sinno et al. J.Chem.Phys. 2008.

- equilibrium simulations and multi-resolution analysis:
  Ismail, Rutledge, Stephanopoulos, J. Chem. Phys. (2003)

- computational renormalization group – statistical/quantum field theory computations

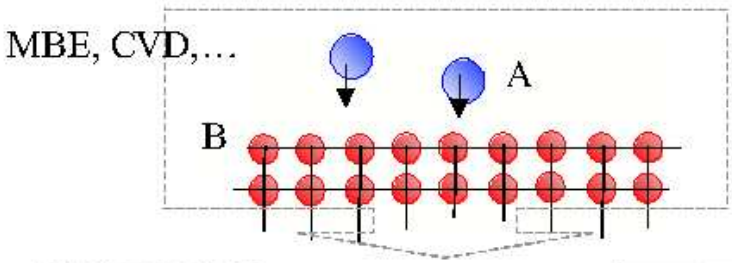
Surface processes on lattices

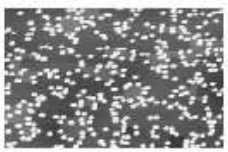
- Optical, magnetic, electronic devices, templating, catalysis
- non-uniform shape, size, spacing
- control the process to enable fabrication



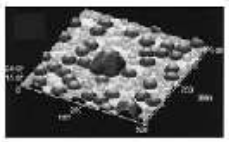
Surface processes on lattices

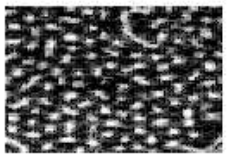
- Optical, magnetic, electronic devices, templating, catalysis
- non-uniform shape, size, spacing
- control the process to enable fabrication





Pb/Cu(111)
Nature01, 412, 875

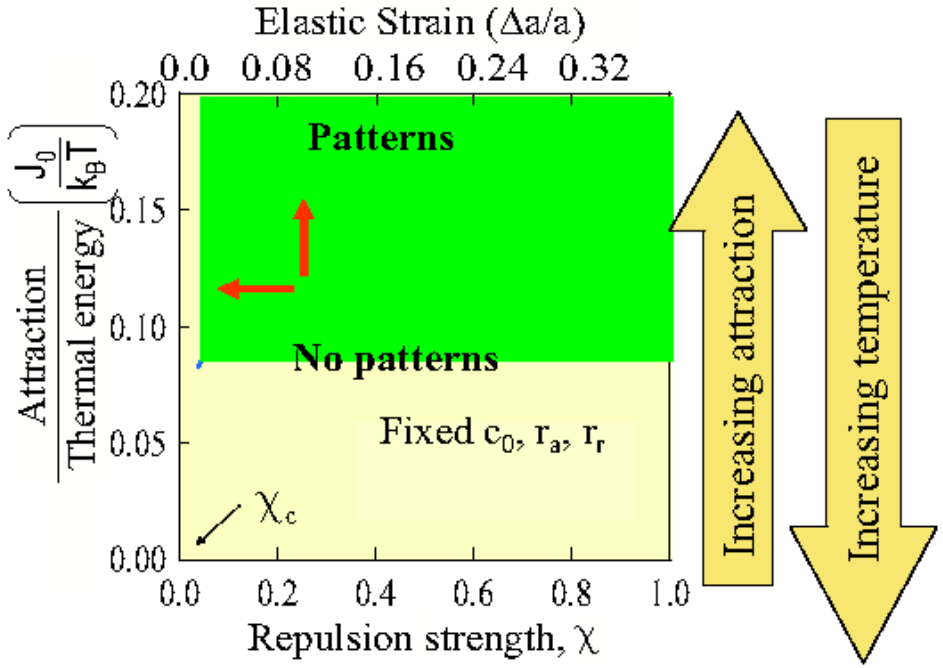




Fe/NaCl
PRL02, 89 235502

Si/Ge quantum dots
Williams et al.
(1998)

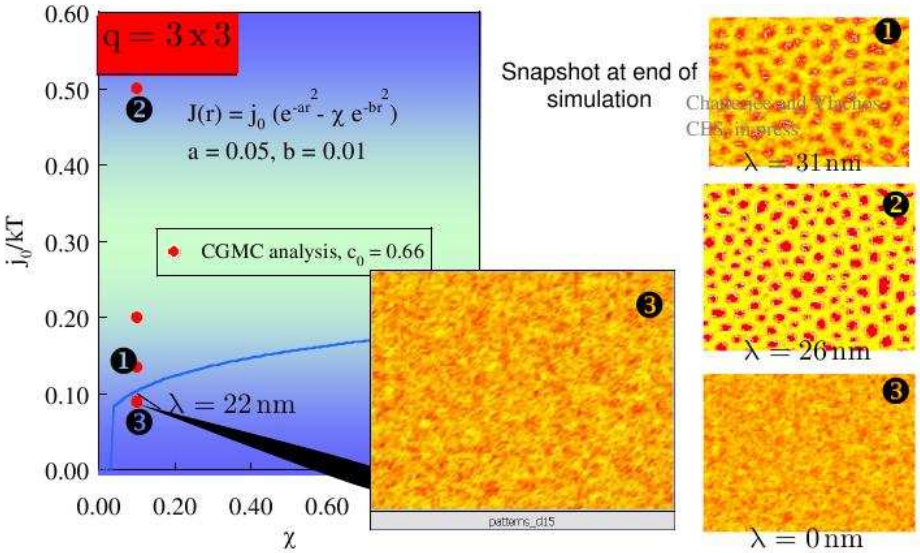
Control and Design



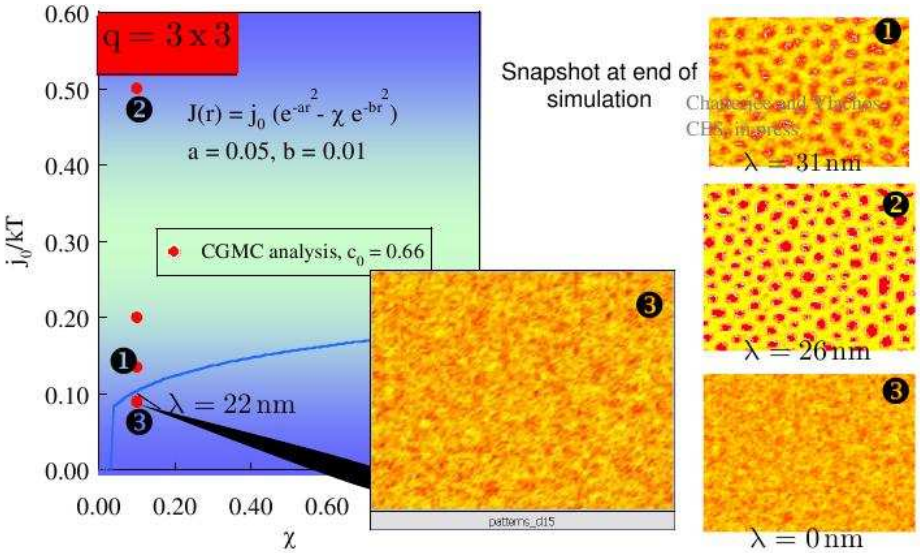
$\Lambda_N = \frac{1}{N}\mathbb{Z}^d \cap [0,1)^d$, $N \gg 1$
Configurations:
$\sigma \in \Sigma_N = \mathcal{S}^{\Lambda_N}$

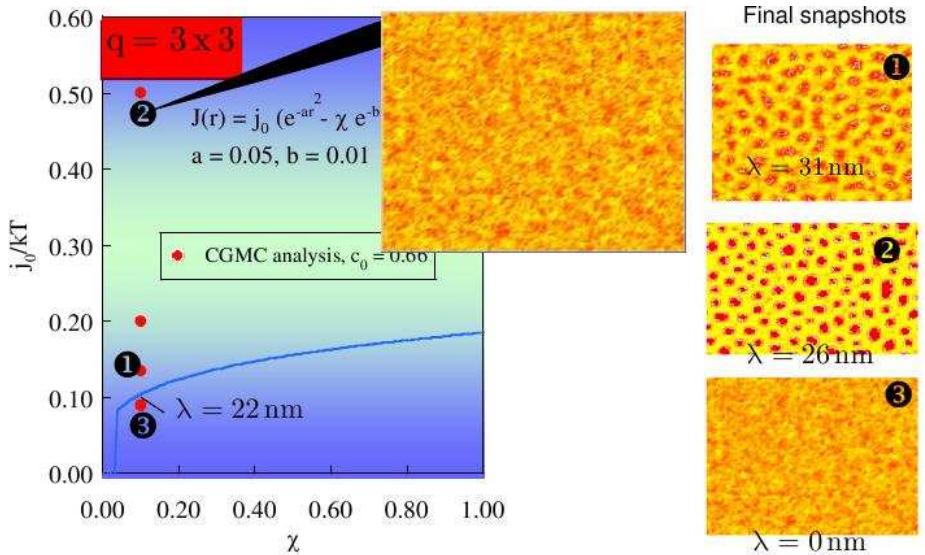
Kalligiannaki, Katsoulakis, PP, Vlachos, J. Comp. Phys. (2012); Kalligiannaki, Katsoulakis, PP, SIAM J. Sci. Comp. (2014)

Surface processes on lattices



Surface processes on lattices





Reaction kinetics in catalysis

CO oxidation reaction on Pt: kinetic Monte Carlo simulations

State space: $\sigma(x) \in \Sigma \equiv \{-1, 0, 1\}$
Ziff-Gulari-Barshad model:
$CO \rightarrow CO_{(ads)}$
$O_2 \rightarrow 2O_{(ads)}$
$CO_{(ads)} + O_{(ads)} \rightarrow CO_2$
diffusion of O

Events & Rates

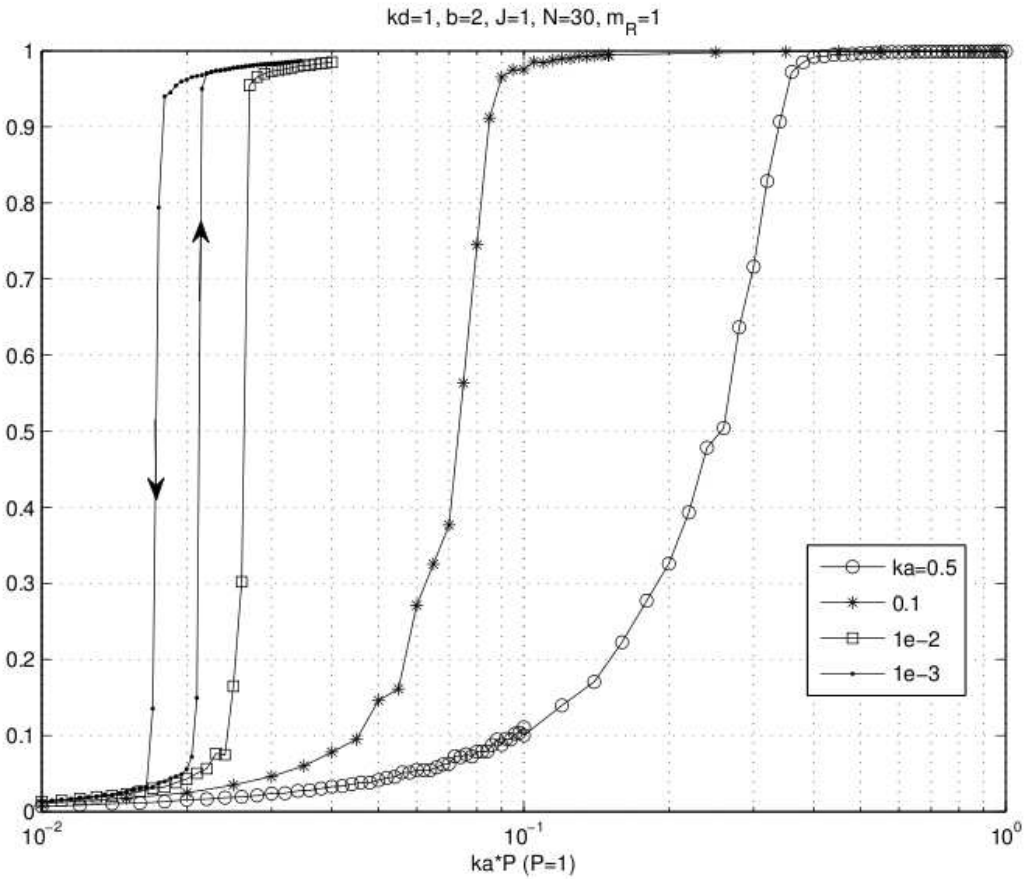
1. $\sigma(x) = 0$ (vacant site):
   (a) with a rate k_1 a CO particle adsorbs: $\sigma(x) = 0 \rightarrow \sigma(x) = 1$
   (b) if $\sigma(y) = 0$, $y = x^{nn}$ then two O_2 adsorb $(1 - k_1)$:
   $\sigma(x) = 0 \rightarrow \sigma(x) = -1$, $\sigma(x^{nn}) = 0 \rightarrow \sigma(x^{nn}) = -1$
2. $\sigma(x) = 1$ (CO molecule): if $\sigma(y) = -1$, $y = x^{nn}$ with the rate k_2: $CO+O_2$ and desorb:
   $\sigma(x) = 1 \rightarrow \sigma(x) = 0$, $\sigma(x^{nn}) = -1 \rightarrow \sigma(x^{nn}) = 0$
3. $\sigma(x) = -1$ (O_2 molecule): if $\sigma(y) = 1$, $y = x^{nn}$ with the rate k_2: $CO+O_2$ and desorb.
   $\sigma(x) = -1 \rightarrow \sigma(x) = 0$, $\sigma(x^{nn}) = 1 \rightarrow \sigma(x^{nn}) = 0$

Example: Reaction kinetics in catalysis

CO oxidation reaction on Pt: kinetic Monte Carlo simulations

State space: $\sigma(x) \in \Sigma \equiv \{-1, 0, 1\}$
Ziff-Gulari-Barshad model:
$CO \rightarrow CO_{(ads)}$
$O_2 \rightarrow 2O_{(ads)}$
$CO_{(ads)} + O_{(ads)} \rightarrow CO_2$
diffusion of O

Oxidation process – parallel KMC simulations


kd=1, b=2, J=1, N=30, m_R=1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10^-2
10^-1
10^0
ka*P (P=1)
ka=0.5
0.1
1e-2
1e-3

ZGB - Definition

- ZGB (Ziff-Gulari-Barshad) is a simplified spatio-temporal CO oxidation model without diffusion.
- Despite being an idealized model, the ZGB model incorporates basic mechanisms for the dynamics of adsorbate species during CO oxidation on catalytic surfaces.

Event	Reaction	Rate
1	$\emptyset \to CO$	$(1-\sigma(j)^2)k_1$
2	$\emptyset \to O_2$	$(1-\sigma(j)^2)(1-k_1)\frac{\#\text{vacant n.n.}}{\text{total n.n.}}$
3	$CO + O \to CO_2 + \text{des.}$	$\frac{1}{2}\sigma(j)(1+\sigma(j))k_2\frac{\#O\text{ n.n.}}{\text{total n.n.}}$
4	$O + CO \to CO_2 + \text{des.}$	$\frac{1}{2}\sigma(j)(\sigma(j)-1)k_2\frac{\#CO\text{ n.n.}}{\text{total n.n.}}$

Table: The rate of the kth event of the jth site given that the current configuration is σ is denoted by $c_k(j;\sigma)$ where n.n. stands for nearest neighbors.

ZGB - RER

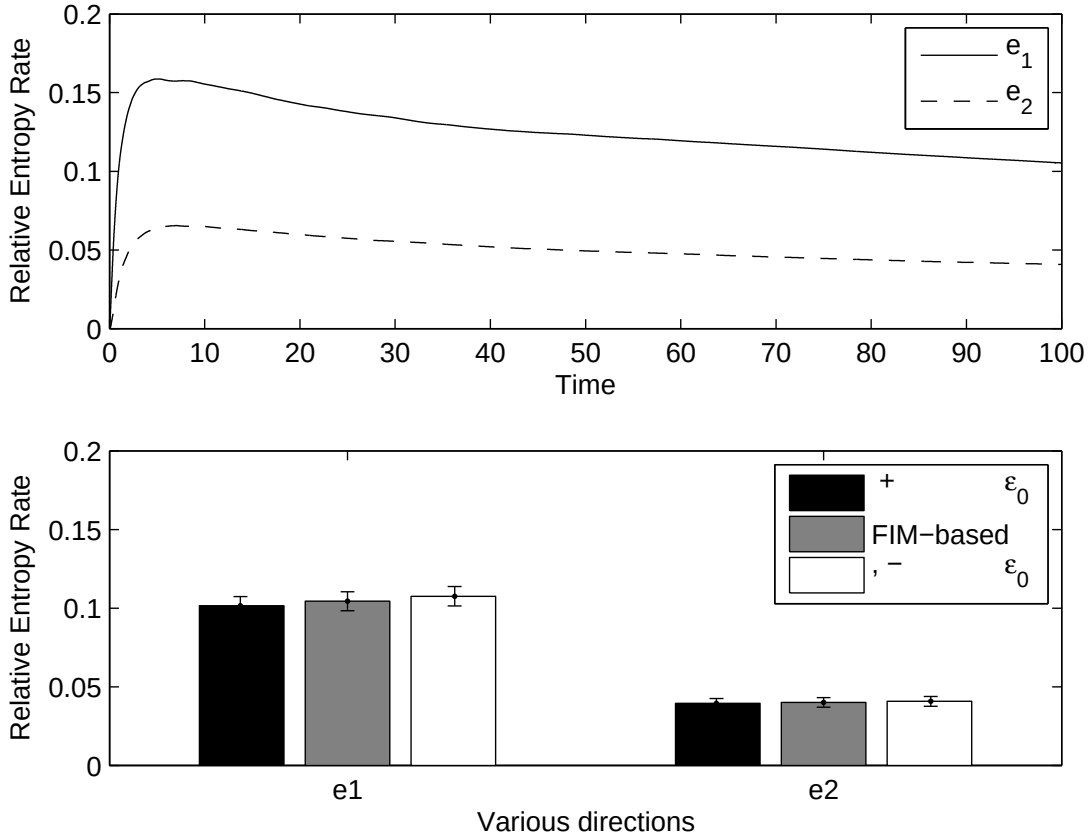


Figure: Upper plot: Relative entropy rate as a function of time for perturbations of both k_1 (solid line) and of k_2 (dashed line). An equilibration time until the process reach its metastable regime is evident. Lower plot: RER for various directions. The most sensitive parameter is k_1.

ZGB - Configurations

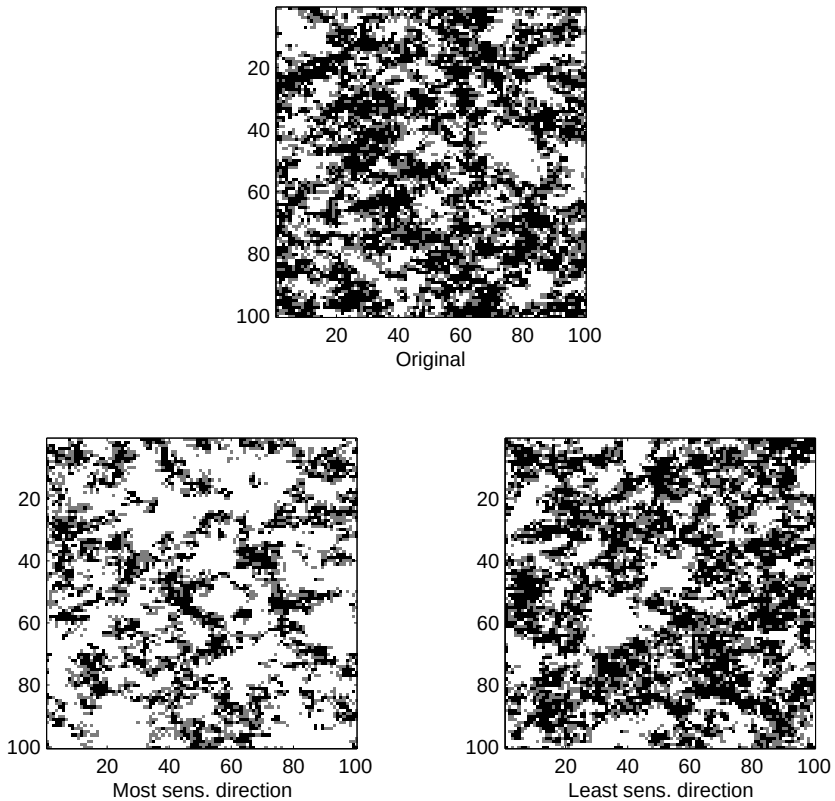
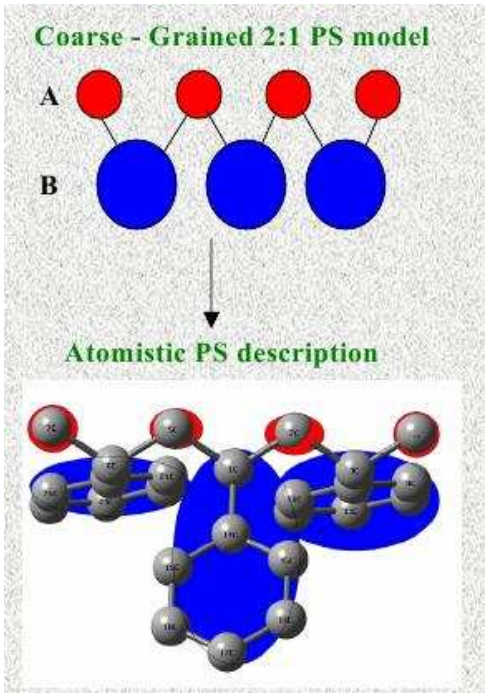


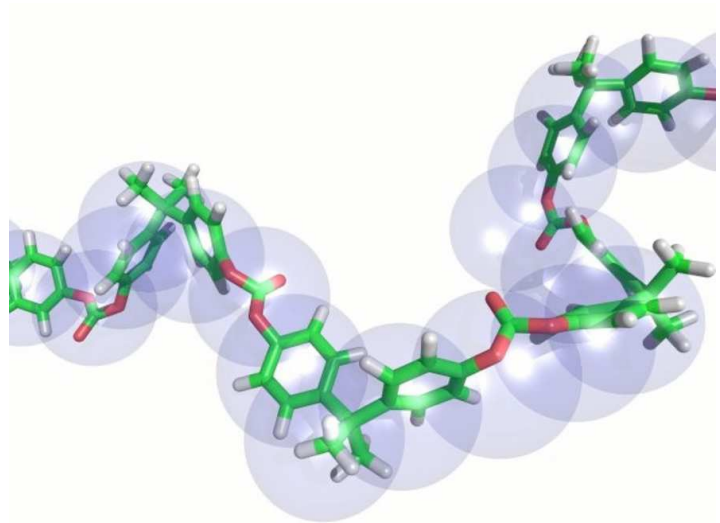
Figure: Typical configurations obtained by ϵ_0-perturbations of the most and least sensitive parameters. The comparison with the reference configuration reveals the differences between the most and least sensitive perturbation parameters.

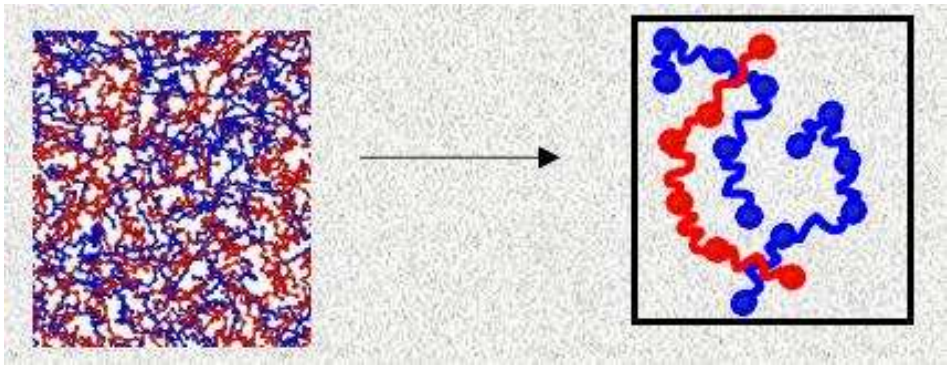
Off-lattice Models.

Coarse-graining in molecular simulations

$$\text{positions of atoms } \{X^{(k)}\} \mapsto \text{ positions of metaparticles } \{Q^{(p)}\} = \mathbf{T}X$$







Coarse-graining of polymers; DPD methods

Effective Hamiltonian $\bar{H}(P,Q)$ using simplifying assumptions

- *Parametric statistics approaches at equilibrium*
  Müller-Plathe, Chem. Phys. (2002), Kremer, Müller-Plathe, MRS Bull (2001), Shell (2008,2012), Zabaras (2012)
- *United Atom models* and McCoy-Curro scheme
  McCoy, Curro, Macromolecules (1998); Fukununaga, Takimoto, Doi, J. Chem. Phys. (2002)
- *Computational renormalization group*
  Brandt, Ron, JSP (2001); Bai, Brandt (2000)
- *Dissipative Particle Dynamics*
  Briels et al. J. Chem. Phys. (2001), Pivkin, Karniadakis J. Chem. Phys. (2002), Deserno et. al. Nature (2007).

Microscopic dynamics

$$\dot{q} = \nabla_p H(p,q)$$
$$\dot{p} = -\nabla_q H(p,q) - \gamma p + \sqrt{2/\beta}\,\dot{W}$$

CG map: $(P,Q) = \mathbf{T}(p,q)$

Effective equations of motion

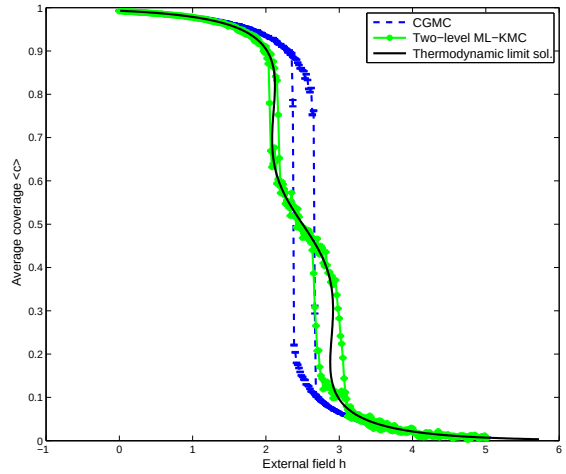
Coarse-grained Hamiltonian $\bar{H}(P,Q)$

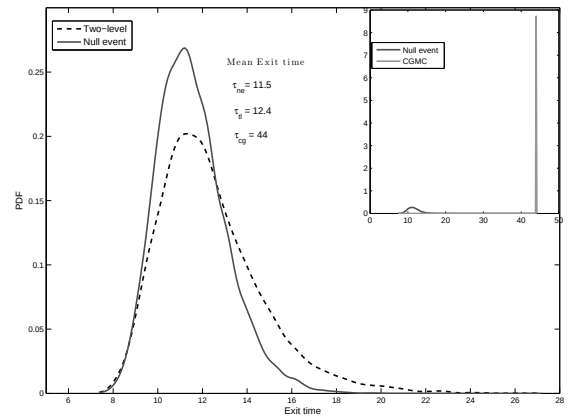
$$\dot{Q} = \nabla_P \bar{H}(P,Q)$$
$$\dot{P} = -\nabla_Q \bar{H}(P,Q) - \bar{\gamma} P + \bar{\sigma}\dot{W}$$

- $\bar{H}(P, Q)$ used in dynamics relevant only for long-time behaviour and approach to equilibrium
- ad hoc CG: wrong predictions of diffusion, crystallization, phase transitions

  Abrams, Kremer, J. Chem. Phys. (2001), Pivkin, Karniadakis J. Chem. Phys. (2002)
- without numerical analysis no indication of wrong phenomenon being deduced from simulation.
- adaptive change of CG difficult

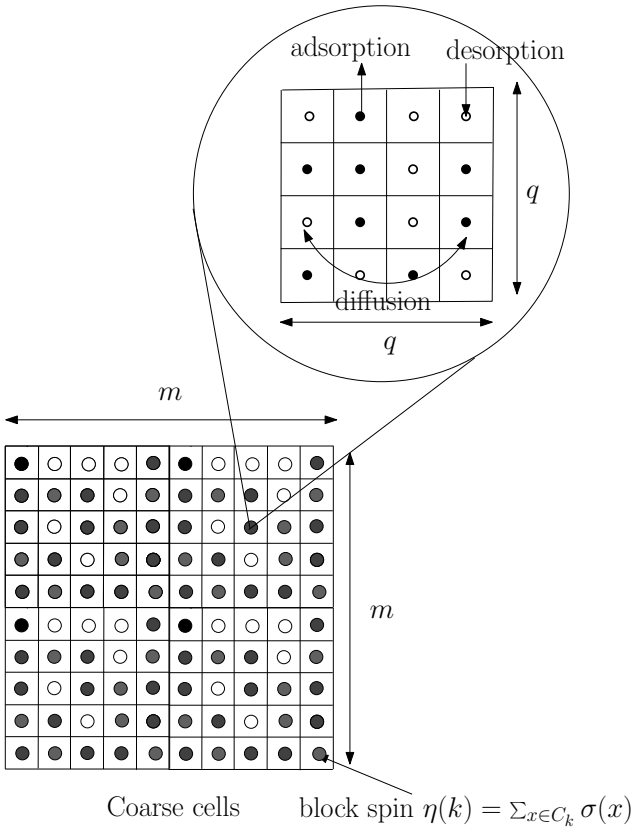
  Praprotnik,Matysiak,Kremer, Clementi, J. Phys. Cond. Matter (2007).





Coarse-graining in lattice systems

EXAMPLE: Block spins



Coarse map: $\mathbf{T} : \Sigma_N \to \bar{\Sigma}_M$

$$\{\eta(k)\} = \mathbf{T}\sigma = \sum_{x \in C_k} \sigma(x)$$

Microscopic process: $(\{\sigma_t\}_{t \geq 0}, \mathcal{L})$, $c(x, \sigma)$

Coarse process: $(\{\eta_t\}_{t \geq 0}, \bar{\mathcal{L}})$, $\bar{c}(k, \eta)$

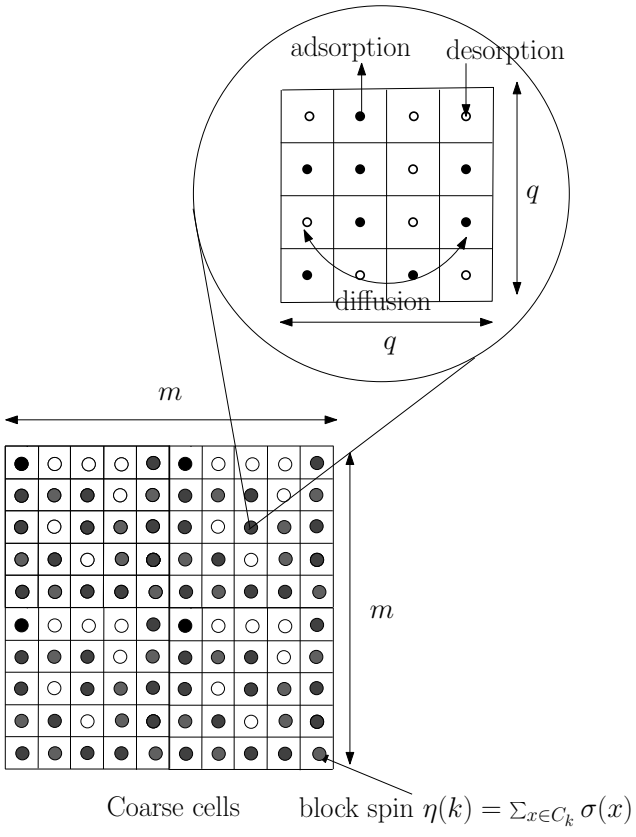
Reconstructed process: $(\{\tilde{\sigma}_t\}_{t \geq 0}, \tilde{\mathcal{L}})$, $\tilde{c}(x, \sigma)$

Error: at finite t and as $t \to \infty$

modeling error

Coarse-graining in lattice systems

EXAMPLE: Block spins



Coarse map: $\mathbf{T} : \Sigma_N \to \bar{\Sigma}_M$

$$\{\eta(k)\} = \mathbf{T}\sigma = \sum_{x \in C_k} \sigma(x)$$

Microscopic process: $(\{\sigma_t\}_{t \geq 0}, \mathcal{L})$, $c(x, \sigma)$

Coarse process: $(\{\eta_t\}_{t \geq 0}, \bar{\mathcal{L}})$, $\bar{c}(k, \eta)$

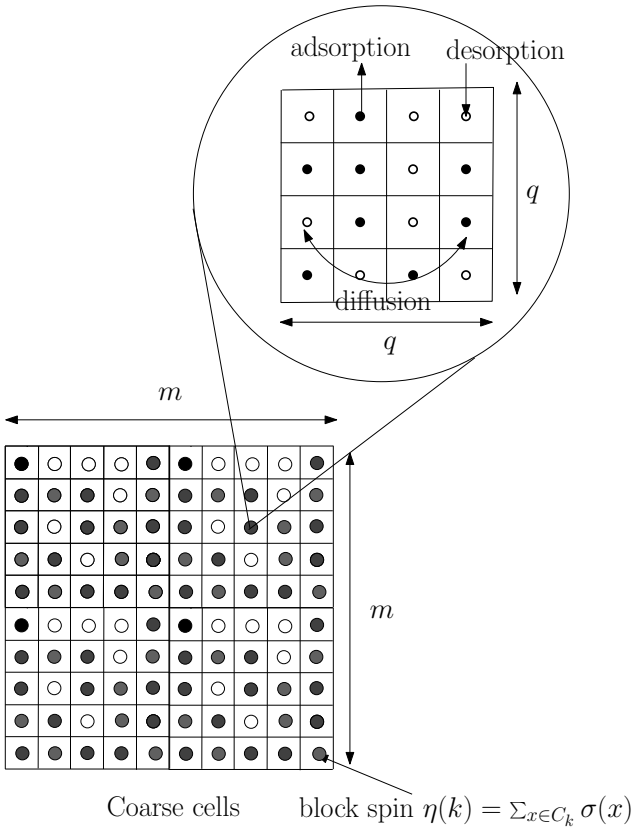
Reconstructed process: $(\{\tilde{\sigma}_t\}_{t \geq 0}, \tilde{\mathcal{L}})$, $\tilde{c}(x, \sigma)$

Error: at finite t and as $t \to \infty$

modeling error

Coarse-graining in lattice systems

EXAMPLE: Block spins



Coarse map: $\mathbf{T} : \Sigma_N \to \bar{\Sigma}_M$

$$\{\eta(k)\} = \mathbf{T}\sigma = \sum_{x \in C_k} \sigma(x)$$

Microscopic process: $(\{\sigma_t\}_{t \geq 0}, \mathcal{L})$, $c(x, \sigma)$

Coarse process: $(\{\eta_t\}_{t \geq 0}, \bar{\mathcal{L}})$, $\bar{c}(k, \eta)$

Reconstructed process: $(\{\tilde{\sigma}_t\}_{t \geq 0}, \tilde{\mathcal{L}})$, $\tilde{c}(x, \sigma)$

Error: at finite t and as $t \to \infty$

modeling error + numerical error + statistical error

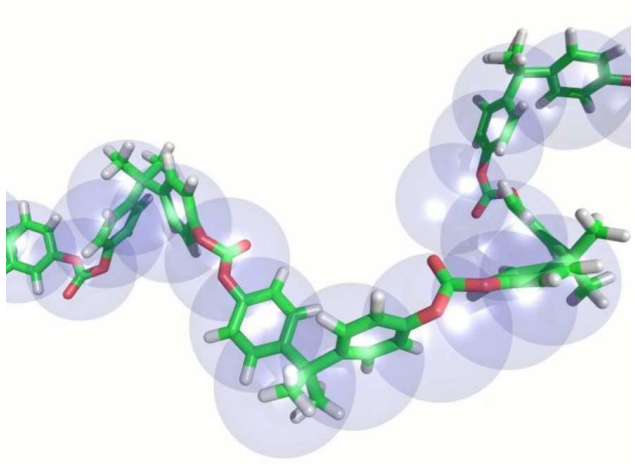
Coarse-Graining – Equilibrium

1. Coarse-graining of polymers; DPD methods

Briels, et. al. *J.Chem.Phys.* '01;
Doi et. al. *J.Chem.Phys.* '02;
Kremer et. al. *Macromolecules* '06;
Müller-Plathe *Chem.Phys.Chem* '02;
Laaksonen et. al. *Soft Matter* '03, etc;
Deserno et. al. *Nature* '07;
Espanol *J Chem. Phys. '07, '11*

Harmandaris *Macromolecules* Noid *J Chem. Phys.*

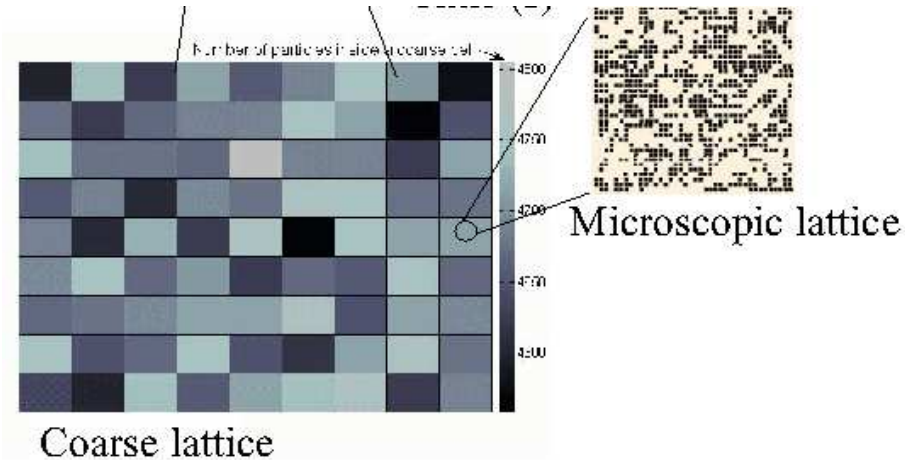
'13



2. Stochastic lattice dynamics/ KMC

Katsoulakis, Majda, Vlachos, *PNAS*'03;
Katsoulakis, P.P., Sopasakis, *SIAM Num. Anal.* '06;
Are, Katsoulakis, P.P., Rey-Bellet *SIAM J.Sci.Comp.* '08;

Sinno et al. *J.Chem.Phys.*'08, '13, *PRE '12*



Equilibrium

Invariant measure: $\mu \sim e^{-\beta H(\sigma)}$

Detailed balance for the coarse-grained process w.r.t. $\bar{\mu} \sim e^{-\beta \bar{H}(\eta)}$

Coarse-grained Hamiltonian $\bar{H}(\eta)$

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

Approximate $\bar{H}(\eta) \approx \bar{H}^{(0)}(\eta)$, i.e., $\bar{\mu}(d\eta) \approx \bar{\mu}^{(0)}(d\eta)$

Equilibrium

Invariant measure: $\mu \sim e^{-\beta H(\sigma)}$
Detailed balance for the coarse-grained process w.r.t. $\bar{\mu} \sim e^{-\beta \bar{H}(\eta)}$
Coarse-grained Hamiltonian $\bar{H}(\eta)$

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

Approximate $\bar{H}(\eta) \approx \bar{H}^{(0)}(\eta)$, i.e., $\bar{\mu}(d\eta) \approx \bar{\mu}^{(0)}(d\eta)$

Task: estimate & control the error in the relative entropy $\mathcal{R}\left(\bar{\mu} \,||\, \bar{\mu}^{(0)}\right)$
however $\bar{\mu}$ is unknown

"Lift" $\bar{\mu}^{(0)}$ to a new μ^{app} on the microscopic space

Reconstruction measures

Coarse-grained equilibrium measure:

$$\int f(\eta)\bar{\mu}(d\eta) = \int f(\mathbf{T}\sigma)\mu(d\sigma)$$

$\mu(d\sigma) = \frac{1}{Z} e^{-\beta \bar{H}(\sigma)} P(d\sigma)$ and $\bar{\mu}(d\eta) = \frac{1}{\bar{Z}} e^{-\beta \bar{H}(\eta)} \bar{P}(d\eta)$
Perfect reconstruction:

$$\mu(d\sigma) = e^{-\beta(H(\sigma)-\bar{H}(\eta))} P(d\sigma|\eta)\bar{\mu}(d\eta) \equiv \mu(d\sigma|\eta)\bar{\mu}(d\eta)$$

Approximate reconstruction:

$$\mu^{\text{app}}(d\sigma) = \nu(d\sigma|\eta)\bar{\mu}^{(0)}(d\eta)$$

$$\text{error} = \text{coarse-graining error} + \text{reconstruction error}$$

Katsoulakis, P.P., Rey-Bellet (2008), Trashorras, Tsagarogiannis (2010)

Example:

- $\mathcal{N}(\eta) = |\{\sigma | \mathbf{T}\sigma = \eta\}|$ and $\nu(d\sigma|\eta)$ is uniform
  Approximation at the fine level: $\mu^{\mathrm{app}}(\sigma) = \bar{\mu}^{(0)}(\mathbf{T}\sigma)\,\frac{1}{\mathcal{N}(\eta)}$

Example:

- $\mathcal{N}(\eta) = |\{\sigma | \mathbf{T}\sigma = \eta\}|$ and $\nu(d\sigma|\eta)$ is uniform
  Approximation at the fine level: $\mu^{\text{app}}(\sigma) = \bar{\mu}^{(0)}(\mathbf{T}\sigma)\,\frac{1}{\mathcal{N}(\eta)}$
- The relative entropy

$$\mathcal{R}\left(\mu \,||\, \mu^{app}\right) = \sum_{\sigma} \mu(\sigma) \log \frac{\mu(\sigma)}{\bar{\mu}^{(0)}(\mathbf{T}\sigma)} + \sum_{\sigma} \mu(\sigma) \log \mathcal{N}(\mathbf{T}\sigma)$$

Example:

- $\mathcal{N}(\eta) = |\{\sigma | \mathbf{T}\sigma = \eta\}|$ and $\nu(d\sigma|\eta)$ is uniform
  Approximation at the fine level: $\mu^{\text{app}}(\sigma) = \bar{\mu}^{(0)}(\mathbf{T}\sigma)\,\frac{1}{\mathcal{N}(\eta)}$
- The relative entropy

$$\mathcal{R}\left(\mu \,||\, \mu^{app}\right) = \sum_{\sigma} \mu(\sigma) \log \frac{\mu(\sigma)}{\bar{\mu}^{(0)}(\mathbf{T}\sigma)} + \sum_{\sigma} \mu(\sigma) \log \mathcal{N}(\mathbf{T}\sigma)$$

- Insert back $\mu(\sigma) = \frac{1}{Z} e^{-\beta H(\sigma)}$ etc.

$$\sum_{\sigma} \beta(\bar{H}^{(0)}(\mathbf{T}\sigma) - H(\sigma)) \frac{1}{Z} e^{-\beta H(\sigma)} + \log \frac{\bar{Z}^{(0)}}{Z} + \mathbb{E}_{\mu}[\log \mathcal{N}(\mathbf{T}\sigma)] =$$

$$\mathbb{E}_{\mu}[\beta(\bar{H}^{(0)} - H)] - \beta(\bar{A}^{(0)} - A) + \mathbb{E}_{\mu}[\log \mathcal{N}(\mathbf{T}\sigma)]$$

Helmholtz free energy $A \equiv U - TS = -\frac{1}{\beta} \log Z$

Inverse Monte Carlo

"Inverse thermodynamic problems"

- Build parametrized approximations $\bar{H}^{(0)}(\eta;\theta)$

Inverse Monte Carlo

"Inverse thermodynamic problems"

- Build parametrized approximations $\bar{H}^{(0)}(\eta;\theta)$
- Find optimal values of parameters θ^*, s.t., for selected ϕ_i

$$\min_{\theta} \sum_{i} \left| \mathbb{E}_{\mu}[\phi_i] - \mathbb{E}_{\bar{\mu}^{(0)}}[\phi_i] \right|^2$$

Review: F. Muller-Plathe Chem. Phys. Chem. (2002)

Inverse Monte Carlo

"Inverse thermodynamic problems"

- Build parametrized approximations $\bar{H}^{(0)}(\eta;\theta)$
- Find optimal values of parameters θ^*, s.t., for selected ϕ_i

$$\min_{\theta} \sum_{i} \left| \mathbb{E}_{\mu}[\phi_i] - \mathbb{E}_{\bar{\mu}^{(0)}}[\phi_i] \right|^2$$

Review: F. Muller-Plathe Chem. Phys. Chem. (2002)

- Parametrization depends on specific observable(s) ϕ_i.

Inverse Monte Carlo

"Inverse thermodynamic problems"

- Build parametrized approximations $\bar{H}^{(0)}(\eta;\theta)$
- Find optimal values of parameters θ^*, s.t., for selected ϕ_i

$$\min_{\theta} \sum_{i} \left| \mathbb{E}_{\mu}[\phi_i] - \mathbb{E}_{\bar{\mu}^{(0)}}[\phi_i] \right|^2$$

Review: F. Muller-Plathe Chem. Phys. Chem. (2002)

- Parametrization depends on specific observable(s) ϕ_i.
- Can we improve the "transferability" of the method ?

Inverse Monte Carlo

"Inverse thermodynamic problems"

- Build parametrized approximations $\bar{H}^{(0)}(\eta;\theta)$

- ϕ can be relative entropy

$$\mathbb{E}_\mu[\beta(\bar{H}^{(0)}(\theta) - H)] - \beta(\bar{A}^{(0)}(\theta) - A) + \mathbb{E}_\mu[\log \mathcal{N}(\mathbf{T}\sigma)]$$

optimality condition: $\nabla_\theta \mathcal{R} = \mathbb{E}_\mu[\nabla_\theta \bar{H}^{(0)}] - \mathbb{E}_{\bar{\mu}^{(0)}}[\nabla_\theta \bar{H}^{(0)}] = 0$

Parametrized CG – Approximations heuristics

$$\min_{\theta} \mathcal{R}\left(\mu \,||\, \mu^{\text{app}}(\theta)\right) \quad \text{or} \quad \min_{\theta} \mathcal{R}\left(\mu^{\text{app}}(\theta) \,||\, \mu\right)$$

Gibbs structure allows explicit calculations of $\mathcal{R}$

$$\mathcal{R}\left(\bar{\mu} \,||\, \bar{\mu}^{(0)}\right) \sim \mathbb{E}_{\mu}[\beta(\bar{H}^{(0)}(\theta) - H)] + \log \frac{\bar{Z}^{0}(\theta)}{Z}$$

Optimality condition: $\nabla_{\theta}\mathcal{R} = 0$

- Solution using typically gradient methods, Newton-Raphson, etc:
  M.S. Shell (2008, 2012), Noid (2012), Bilionis et al (2012), Zabaras et al (2013), a review Noid (2013)

- Is the parametric family $\bar{H}^{(0)}(\theta)$ rich enough?.

Cluster Expansion CG Hamiltonians

Are, Katsoulakis, P.P. , Rey-Bellet SIAM J. Sci. Comp. (2008); Katsoulakis, P.P. , Rey-Bellet, Tsagarogiannis Math. Comp. (2014)

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

Multi-body terms:

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

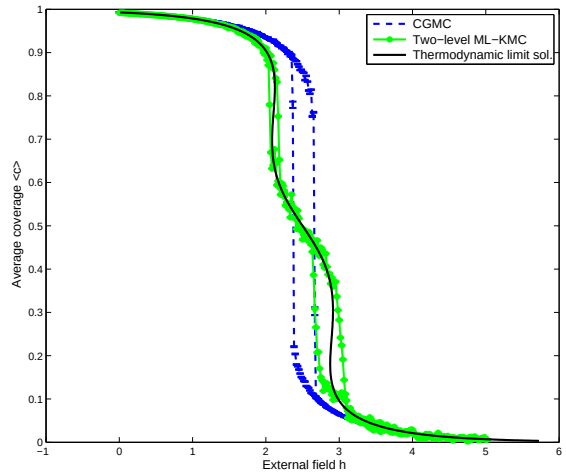
Cluster Expansion CG Hamiltonians

Are, Katsoulakis, P.P. , Rey-Bellet SIAM J. Sci. Comp. (2008); Katsoulakis, P.P. , Rey-Bellet, Tsagarogiannis Math. Comp. (2014)

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

Multi-body terms:

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



Typically omitted, but essential to capture phase transitions

Cluster Expansion CG Hamiltonians

Are, Katsoulakis, P.P. , Rey-Bellet SIAM Sci. Comp. (2008), Math. Comp. (2013)

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

Multi-body terms:

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

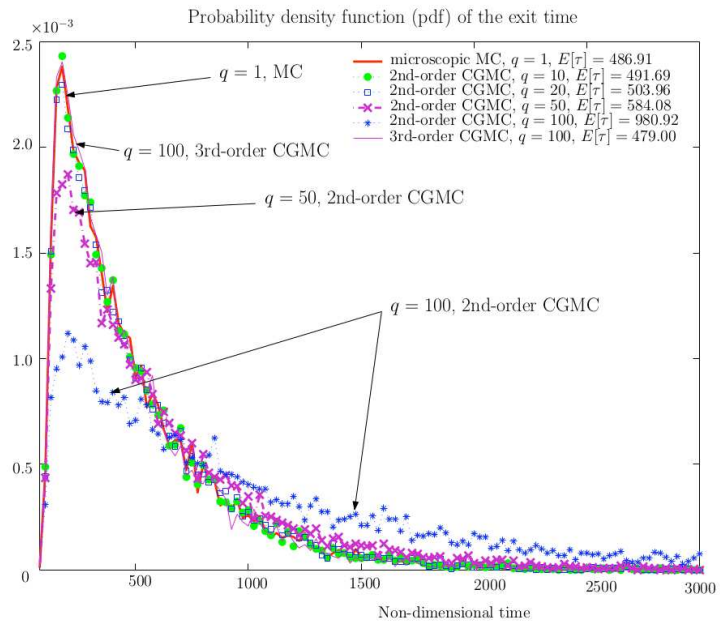
Cluster Expansion CG Hamiltonians

Are, Katsoulakis, P.P. , Rey-Bellet SIAM Sci. Comp. (2008), Math. Comp. (2013)

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

Multi-body terms:

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



Typically omitted, but essential to capture switching times etc.

Dynamics and Non-equilibrium steady states

- Continuous Time Markov Chain $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$
  $\sigma \in \Sigma \equiv \{0,1\}^{\Lambda_N}$, $\Lambda_N \subset \mathbb{Z}^d$

$$\mathbb{P}\left(\sigma_{t+\delta t} = \sigma' \,||\, \sigma_t = \sigma\right) = c(\sigma, \sigma')\delta t + o(\delta t)$$

Rates: $c(\sigma, \sigma')$

Dynamics and Non-equilibrium steady states

- Continuous Time Markov Chain $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$
  $\sigma \in \Sigma \equiv \{0,1\}^{\Lambda_N}$, $\Lambda_N \subset \mathbb{Z}^d$

$$\mathbb{P}\left(\sigma_{t+\delta t} = \sigma' \,||\, \sigma_t = \sigma\right) = c(\sigma, \sigma')\delta t + o(\delta t)$$

  Rates: $c(\sigma, \sigma')$

- Forward Kolmogorov Equation (aka Master Equation)

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

$$P(\sigma, 0; \zeta) = \delta(\sigma - \zeta)$$

Dynamics and Non-equilibrium steady states

- Continuous Time Markov Chain $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$
  $\sigma \in \Sigma \equiv \{0,1\}^{\Lambda_N}$, $\Lambda_N \subset \mathbb{Z}^d$

$$\mathbb{P}\left(\sigma_{t+\delta t} = \sigma' \,||\, \sigma_t = \sigma\right) = c(\sigma,\sigma')\delta t + o(\delta t)$$

  Rates: $c(\sigma,\sigma')$

- Forward Kolmogorov Equation (aka Master Equation)

$$\partial_t P(\sigma,t;\zeta) = \sum_{\sigma',\sigma'\neq\sigma} c(\sigma',\sigma)P(\sigma',t;\zeta) - \lambda(\sigma)P(\sigma,t;\zeta)\,,$$
$$P(\sigma,0;\zeta) = \delta(\sigma-\zeta)$$

- Simulation: Embedded Markov Chain $\{X_n\}_{n\geq 0} = \{\sigma_{n\delta t}\}$, $\sigma \to \sigma^{x,\omega}$

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

  Exponential clock: $\delta t \sim \mathrm{Exp}(\lambda(\sigma))$

Non-equilibrium steady states

Behavior as $t \to \infty$

$$\partial_t P(\sigma, t; \zeta) = \sum_{\sigma'} \left[c(\sigma', \sigma) P(\sigma', t; \zeta) - c(\sigma, \sigma') P(\sigma, t; \zeta) \right] ,$$

Stationary states: $\partial_t P = 0 \Longrightarrow \sum_{\sigma'} \mathbf{j}_s(\sigma', \sigma) = 0$
Current $\sigma' \to \sigma$: $\mathbf{j}_s(\sigma', \sigma) = c(\sigma', \sigma)\mu(\sigma') - c(\sigma, \sigma')\mu(\sigma)$

Reversible dynamics with the equilibrium $\mu(\sigma)$
Detailed Balance condition with respect to $\mu(\sigma)$ (e.g., $\mu \sim e^{-\beta H(\sigma)}$)

$$c(\sigma', \sigma)\mu(\sigma') = c(\sigma, \sigma')\mu(\sigma)$$

Irreversible dynamics $\Longrightarrow$ Non-equlibrium steady states

$$\sum_{\sigma'} \mathbf{j}_s(\sigma', \sigma) = \sum_{\sigma'} (c(\sigma', \sigma)\mu(\sigma') - c(\sigma, \sigma')\mu(\sigma)) = 0$$

irreversible rate loops, i.e., a non-zero current at stationary states.

Relative entropy on the path space

Markov chains on Σ:
$\{\sigma_n\}_{n\in\mathbb{Z}^+}$, $P^\theta(\sigma, d\sigma)$, $\mu^\theta(\sigma)$
$\{\tilde{\sigma}_n\}_{n\in\mathbb{Z}^+}$, $\tilde{P}^\theta(\sigma, d\sigma)$, $\tilde{\mu}^\theta(\sigma)$
Path measures:

$$Q^\theta(\sigma_0, \ldots, \sigma_M) = \mu^\theta(\sigma_0) p^\theta(\sigma_0, \sigma_1) \ldots p^\theta(\sigma_{M-1}, \sigma_M)$$

Radon-Nikodym derivative

$$\frac{dQ^\theta}{d\tilde{Q}^\theta}(\{\sigma_n\}) = \frac{\mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{\mu}^\theta(\sigma_0) \prod_{i=0}^{M-1} \tilde{p}^\theta(\sigma_i, \sigma_{i+1})}$$

Relative entropy

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \ldots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \log \frac{\mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{\mu}^\theta(\sigma_0) \prod_{i=0}^{M-1} \tilde{p}^\theta(\sigma_i, \sigma_{i+1})} d\sigma$$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}\right) d\sigma_0 \dots d\sigma_M$$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}\right) d\sigma_0 \dots d\sigma_M$$

$$\int_\Sigma p(\sigma, \sigma')\, d\sigma' = 1\,, \quad \int_\Sigma \mu(\sigma) p(\sigma, \sigma')\, d\sigma = \mu(\sigma')$$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}\right) d\sigma_0 \dots d\sigma_M$$

$$\int_\Sigma p(\sigma, \sigma')\, d\sigma' = 1\,, \quad \int_\Sigma \mu(\sigma) p(\sigma, \sigma')\, d\sigma = \mu(\sigma')$$

$$\int_\Sigma \mu^\theta(\sigma_0) \log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)}\, d\sigma_0 + \sum_{i=0}^{M-1} \int_\Sigma \int_\Sigma \mu^\theta(\sigma_i) p^\theta(\sigma_i) \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}$$

$$= M \mathbb{E}^\theta_\mu \left[\int_\Sigma p^\theta(\sigma, \sigma') \log \frac{p^\theta(\sigma, \sigma')}{\tilde{p}^\theta(\sigma, \sigma')}\, d\sigma'\right] + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = \int_\Sigma \dots \int_\Sigma \mu^\theta(\sigma_0) \prod_{i=0}^{M-1} p^\theta(\sigma_i, \sigma_{i+1}) \left(\log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)} + \sum_{i=0}^{i=M-1} \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}\right) d\sigma_0 \dots d\sigma_M$$

$$\int_\Sigma p(\sigma, \sigma')\, d\sigma' = 1\,, \quad \int_\Sigma \mu(\sigma) p(\sigma, \sigma')\, d\sigma = \mu(\sigma')$$

$$\int_\Sigma \mu^\theta(\sigma_0) \log \frac{\mu^\theta(\sigma_0)}{\tilde{\mu}^\theta(\sigma_0)}\, d\sigma_0 + \sum_{i=0}^{M-1} \int_\Sigma \int_\Sigma \mu^\theta(\sigma_i) p^\theta(\sigma_i) \log \frac{p^\theta(\sigma_i, \sigma_{i+1})}{\tilde{p}^\theta(\sigma_i, \sigma_{i+1})}$$

$$= M \mathbb{E}^\theta_\mu \left[\int_\Sigma p^\theta(\sigma, \sigma') \log \frac{p^\theta(\sigma, \sigma')}{\tilde{p}^\theta(\sigma, \sigma')}\, d\sigma'\right] + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = M\mathcal{H}(Q^\theta \,||\, \tilde{Q}^\theta) + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

Continuous time Markov chain

$\mathcal{D}_{[0,T]}$ (resp. $\tilde{\mathcal{D}}_{[0,T]}$) is the distribution of the process $\{\sigma_t\}_{t\in[0,T]}$ (resp. $\{\tilde{\sigma}_t\}_{t\in[0,T]}$) on the path space $\mathcal{Q}([0,T],\Sigma_N)$

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

The initial distribution is the stationary measure μ (resp. $\tilde{\mu}$).
Radon-Nikodym derivative:

$$\frac{d\mathcal{D}_{[0,T]}}{d\tilde{\mathcal{D}}_{[0,T]}} = \frac{\mu(\sigma_0)}{\tilde{\mu}(\sigma_0)} \exp\left\{-\int_0^T [\lambda(\sigma_s) - \tilde{\lambda}(\sigma_s)]\, ds + \int_0^T \log\frac{c(\sigma_{s-},\sigma_s)}{\tilde{c}(\sigma_{s-},\sigma_s)}\, dN_s\right\}$$

$N_s(\rho)$ – the number of jumps of the path σ_s up to time s.

(a) $N_t - \int_0^t \lambda(\rho_s)\, ds$ is a (zero mean) martingale

(b) exchanging $\int_0^T$ and $\mathbb{E}[\cdot]$

(c) stationarity

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

Hence:

$$\begin{aligned}
\mathcal{R}\left(\mathcal{D}_{[0,T]} \,||\, \tilde{\mathcal{D}}_{[0,T]}\right) &= \mathbb{E}_{\mathcal{D}}\left[\log \frac{d\mathcal{D}_{[0,T]}}{d\tilde{\mathcal{D}}_{[0,T]}}\right] = \mathbb{E}_{\mathcal{D}}\left[\log \frac{\mu}{\tilde{\mu}}\right] \\
&\quad + \mathbb{E}_{\mathcal{D}}\left[-\int_0^T [\lambda(\sigma_s) - \widetilde{\lambda}(\sigma_s)]\, ds + \int_0^T \lambda(\sigma_{s-}) \log \frac{c(\sigma_{s-}, \sigma_s)}{\widetilde{c}(\sigma_{s-}, \sigma_s)}\, ds\right] \\
&= T\mathbb{E}_{\mu}\left[\lambda(\sigma) - \tilde{\lambda}(\sigma) - \sum_{\sigma'} \lambda(\sigma) p(\sigma, \sigma') \log \frac{\lambda(\sigma) p(\sigma, \sigma')}{\tilde{\lambda}(\sigma)\tilde{p}(\sigma, \sigma')}\right] + \mathcal{R}\left(\mu \,||\, \tilde{\mu}\right) \\
&= T\mathcal{H}(\mathcal{D}_{[0,T]} | \tilde{\mathcal{D}}_{[0,T]}) + \mathcal{R}\left(\mu \,||\, \tilde{\mu}\right)
\end{aligned}$$

Relative Entropy Rate and Dynamics Parametrization

- Define parametrized CG transition probabilities $q^{\theta^*}(\sigma, \sigma')$:
  - Parametrized CG transition probabilities $\bar{p}^{\theta}(\eta, \eta')$
  - Reconstruction scheme: $\nu(\sigma'|\mathbf{T}\sigma')$, e.g. uniform: $\frac{1}{|\{\sigma:\mathbf{T}\sigma=\eta'\}|}$

Relative Entropy Rate and Dynamics Parametrization

- Define parametrized CG transition probabilities $q^{\theta^*}(\sigma, \sigma')$:
  - Parametrized CG transition probabilities $\bar{p}^{\theta}(\eta, \eta')$
  - Reconstruction scheme: $\nu(\sigma'|\mathbf{T}\sigma')$, e.g. uniform: $\frac{1}{|\{\sigma:\mathbf{T}\sigma=\eta'\}|}$
  - $q^{\theta}(\sigma, \sigma') = \nu(\sigma'|\mathbf{T}\sigma')\bar{p}^{\theta}(\mathbf{T}\sigma, \mathbf{T}\sigma')$,

Relative Entropy Rate and Dynamics Parametrization

- Define parametrized CG transition probabilities $q^{\theta^*}(\sigma, \sigma')$:
  - Parametrized CG transition probabilities $\bar{p}^{\theta}(\eta, \eta')$
  - Reconstruction scheme: $\nu(\sigma'|\mathbf{T}\sigma')$, e.g. uniform: $\frac{1}{|\{\sigma:\mathbf{T}\sigma=\eta'\}|}$
  - $q^{\theta}(\sigma, \sigma') = \nu(\sigma'|\mathbf{T}\sigma')\bar{p}^{\theta}(\mathbf{T}\sigma, \mathbf{T}\sigma')$,
- $\mathcal{R}\left(P \,||\, Q^{\theta}\right) =$ **Loss of Information (in time-series) due to CG**
- For long times $M >> 1$, RER is dominant:

$$\mathcal{R}\left(P \,||\, Q^{\theta}\right) = M\mathcal{H}(P \,||\, Q^{\theta}) + \mathcal{R}\left(\mu \,||\, \mu^{\theta}\right)$$

$$\mathcal{H}(P \,||\, Q^{\theta}) = \sum_{\sigma\in\Sigma} \mu(\sigma) \sum_{\sigma'\in\Sigma} p(\sigma, \sigma') \log \frac{p(\sigma, \sigma')}{q^{\theta}(\sigma, \sigma')}] \,.$$

Relative Entropy Rate and Dynamics Parametrization

- Define parametrized CG transition probabilities $q^{\theta^*}(\sigma, \sigma')$:
  - Parametrized CG transition probabilities $\bar{p}^{\theta}(\eta, \eta')$
  - Reconstruction scheme: $\nu(\sigma'|\mathbf{T}\sigma')$, e.g. uniform: $\frac{1}{|\{\sigma:\mathbf{T}\sigma=\eta'\}|}$
  - $q^{\theta}(\sigma, \sigma') = \nu(\sigma'|\mathbf{T}\sigma')\bar{p}^{\theta}(\mathbf{T}\sigma, \mathbf{T}\sigma')$,
- $\mathcal{R}\left(P \,||\, Q^{\theta}\right) =$ **Loss of Information (in time-series) due to CG**
- For long times $M >> 1$, RER is dominant:

$$\mathcal{R}\left(P \,||\, Q^{\theta}\right) = M\mathcal{H}(P \,||\, Q^{\theta}) + \mathcal{R}\left(\mu \,||\, \mu^{\theta}\right)$$

$$\mathcal{H}(P \,||\, Q^{\theta}) = \sum_{\sigma\in\Sigma} \mu(\sigma) \sum_{\sigma'\in\Sigma} p(\sigma, \sigma') \log \frac{p(\sigma, \sigma')}{q^{\theta}(\sigma, \sigma')}] \,.$$

- No need for explicit knowledge of NESS: suitable for reaction networks, driven systems, reaction-diffusion, etc.

Inverse Dynamic Monte Carlo

- Best-fit obtained by minimizing RER

$$\theta^* = \arg\min_\theta \mathcal{H}(P \,||\, Q^\theta)\,,$$

- Optimality condition $\nabla_\theta \mathcal{H}(P \,||\, Q^\theta) = 0$; minimization scheme:

$$\theta^{(n+1)} = \theta^{(n)} - \frac{\alpha}{n} G^{(n+1)}\,,$$

$\alpha > 0$ and $G^{(n+1)}$ being a suitable approximation of the gradient $\nabla_\theta \mathcal{H}(P \,||\, Q^\theta)$

Inverse Dynamic Monte Carlo

- Best-fit obtained by minimizing RER

$$\theta^* = \arg\min_\theta \mathcal{H}(P \,||\, Q^\theta)\,,$$

- Optimality condition $\nabla_\theta \mathcal{H}(P \,||\, Q^\theta) = 0$; minimization scheme:

$$\theta^{(n+1)} = \theta^{(n)} - \frac{\alpha}{n} G^{(n+1)}\,,$$

$\alpha > 0$ and $G^{(n+1)}$ being a suitable approximation of the gradient $\nabla_\theta \mathcal{H}(P \,||\, Q^\theta)$

- FIM revisited-Newton-Raphson:

$$G^n = \mathrm{Hess}(\mathcal{H}(P \,||\, Q^{\theta^n}))^{-1} \nabla_\theta \mathcal{H}(P \,||\, Q^{\theta^n})\,.$$

$$\mathbf{F}_{\mathcal{H}}\left(Q^\theta\right) = \mathrm{Hess}(\mathcal{H}(P \,||\, Q^\theta)) = -\mathbb{E}_\mu \left[\sum_{\sigma'} p(\sigma,\sigma) \nabla_\theta^2 \log q^\theta(\sigma,\sigma')\right]\,.$$

Data-based parametrization of CG dynamics

- Unbiased estimator for RER,

$$\hat{\mathcal{H}}_N(P|\,Q^\theta) := \frac{1}{N}\sum_{i=1}^{N} \log \frac{p(\sigma_i,\sigma_{i+1})}{q^\theta(\sigma_i,\sigma_{i+1})}\,,$$

- Minimization of RER:

$$\min_\theta \hat{\mathcal{H}}_N(P|\,Q^\theta) = \max_\theta \frac{1}{N}\sum_{i=1}^{N} \log q^\theta(\sigma_i,\sigma_{i+1}) - \frac{1}{N}\sum_{i=1}^{N} \log p(\sigma_i,\sigma_{i+1})\,,$$

- *Coarse-grained path space Log-Likelihood* maximization

$$\max_\theta L(\theta;\{\sigma_i\}_{i=0}^N) := \max_\theta \frac{1}{N}\sum_{i=1}^{N} \log \bar{p}^\theta(\mathbf{T}\sigma_i,\mathbf{T}\sigma_{i+1})\,.$$

- No need for microscopic reconstruction: $q^\theta(\sigma,\sigma') = \nu(\sigma'|\mathbf{T}\sigma')\bar{p}^\theta(\mathbf{T}\sigma,\mathbf{T}\sigma')$

Fisher Information and Parameter Identifiability

Since RER is a relative entropy, $\mathcal{H}(P \,||\, Q) = \mathcal{R}(\mu \otimes p \,||\, \mu \otimes q)$:

- Asymptotic Gaussianity of the Maximum Likelihood Estimator:

$$\hat{\theta}_N \to \theta^* \text{ a.s. and } N^{1/2}(\hat{\theta}_N - \theta^*) \rightharpoonup N(0, \mathbf{F}_{\mathcal{H}}{}^{-1}(Q^{\theta^*})),$$

- Variance determined by the path-space FIM $\mathbf{F}_{\mathcal{H}}\left(Q^{\theta^*}\right)$, or asymptotically by $\mathbf{F}_{\mathcal{H}}\left(Q^{\hat{\theta}_N}\right)$.
- Estimating the FIM $\mathbf{F}_{\mathcal{H}}\left(Q^{\hat{\theta}_N}\right)$ provides rigorous error bars on computed optimal parameter values θ^*.

Katsoulakis, PP, *J.Chem. Phys.* (2013)

Relative Entropy Rate (RER) $\mathcal{H}$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = M\mathcal{H}(Q^\theta \,||\, \tilde{Q}^\theta) + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

$$\mathcal{H}(Q^\theta \,||\, \tilde{Q}^\theta) = \mathbb{E}_\mu^\theta \left[\int_\Sigma p^\theta(\sigma,\sigma') \log \frac{p^\theta(\sigma,\sigma')}{\tilde{p}^\theta(\sigma,\sigma')}\, d\sigma'\right]$$

Relative Entropy Rate (RER) $\mathcal{H}$

$$\mathcal{R}\left(Q^\theta \,||\, \tilde{Q}^\theta\right) = M\mathcal{H}(Q^\theta \,||\, \tilde{Q}^\theta) + \mathcal{R}\left(\mu^\theta \,||\, \tilde{\mu}^\theta\right)$$

$$\mathcal{H}(Q^\theta \,||\, \tilde{Q}^\theta) = \mathbb{E}^\theta_\mu \left[\int_\Sigma p^\theta(\sigma,\sigma') \log \frac{p^\theta(\sigma,\sigma')}{\tilde{p}^\theta(\sigma,\sigma')}\, d\sigma'\right]$$

- RER is an observable $\Rightarrow$ tractable and statistical estimators are available.
- Contains information not only for the invariant measure but also for the dynamics.
- No need for explicit knowledge of NESS (stationary measure): suitable for reaction networks, driven and/or reaction-diffusion systems, etc.

Examples: **Statistical estimators**

$$\mathcal{H}(\mathcal{D}_{[0,T]}|\tilde{\mathcal{D}}^{\theta}_{[0,T]}) = \mathbb{E}_{\mu}\left[\sum_{\sigma'} c(\sigma,\sigma')\log\frac{c(\sigma,\sigma')}{\tilde{c}(\sigma,\sigma';\theta)} - (\lambda(\sigma)-\tilde{\lambda}(\sigma;\theta))\right]$$

Estimator I:

$$\widehat{\mathcal{H}}^{(n)}_1 = \frac{1}{T}\sum_{k=0}^{n-1}\delta\tau_i\left[\sum_{\sigma'} c(\sigma_k,\sigma')\log\frac{c(\sigma_k,\sigma')}{\widetilde{c}(\sigma_k,\sigma')} - (\lambda(\sigma_k)-\widetilde{\lambda}(\sigma_k))\right]$$

Estimator II:

$$\widehat{\mathcal{H}}^{(n)}_2 = \frac{1}{n}\sum_{k=0}^{n-1}\log\frac{c(\sigma_k,\sigma_{k+1})}{\widetilde{c}(\sigma_k,\sigma_{k+1})} - \frac{1}{T}\sum_{k=0}^{n-1}\delta\tau_k(\lambda(\sigma_k)-\widetilde{\lambda}(\sigma_k))$$

Pantazis, Katsoulakis J. Chem. Phys. (2013)

Multi-scale Diffusions and Stochastic Averaging

- Coarse-graining for diffusion processes on $\mathbb{R}^n \times \mathbb{R}^m$

$$dX_t = a(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t\,, \quad X_0 = x\,.$$

Multi-scale Diffusions and Stochastic Averaging

- Coarse-graining for diffusion processes on $\mathbb{R}^n \times \mathbb{R}^m$

$$dX_t = a(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t\,, \quad X_0 = x\,.$$

- Approximating Markov Chain:

$$X^{n+1} = X^n + a(X^n)h + \sqrt{2\beta^{-1}}Z\sqrt{h}\,,$$

Multi-scale Diffusions and Stochastic Averaging

- Coarse-graining for diffusion processes on $\mathbb{R}^n \times \mathbb{R}^m$

$$dX_t = a(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t\,, \quad X_0 = x\,.$$

- Approximating Markov Chain:

$$X^{n+1} = X^n + a(X^n)h + \sqrt{2\beta^{-1}}Z\sqrt{h}\,,$$

- CG (reduced) dynamics: $\bar{x} \equiv \mathbf{P}x \in \mathbb{R}^n$, $\tilde{x} \equiv \mathbf{P}^\perp x \in \mathbb{R}^m$

$$\mathbf{P}X^{n+1} = \mathbf{P}X^n + \mathbf{P}a(X^n)h + \sqrt{2\beta^{-1}}\mathbf{P}Z\sqrt{h}\,,$$

Multi-scale Diffusions and Stochastic Averaging

- Coarse-graining for diffusion processes on $\mathbb{R}^n \times \mathbb{R}^m$

$$dX_t = a(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t\,, \quad X_0 = x\,.$$

- Approximating Markov Chain:

$$X^{n+1} = X^n + a(X^n)h + \sqrt{2\beta^{-1}}Z\sqrt{h}\,,$$

- CG (reduced) dynamics: $\bar{x} \equiv \mathbf{P}x \in \mathbb{R}^n$, $\tilde{x} \equiv \mathbf{P}^\perp x \in \mathbb{R}^m$

$$\mathbf{P}X^{n+1} = \mathbf{P}X^n + \mathbf{P}a(X^n)h + \sqrt{2\beta^{-1}}\mathbf{P}Z\sqrt{h}\,,$$

- Markovian approximation:

$$\bar{X}^{n+1} = \bar{X}^n + \bar{a}(\bar{X}^n;\theta)h + \sqrt{2\beta^{-1}}\bar{Z}\sqrt{h}\,,$$

Relative entropy rate: Approximating Markov chain

$$p_h(x,x')dx' \sim e^{-\frac{\beta}{h}|x'-x-ha(x)|^2}dx' , \quad \bar{p}_h(\bar{x},\bar{x}';\theta) \sim e^{-\frac{\beta}{h}|\bar{x}'-\bar{x}-\bar{a}(\bar{x};\theta)|^2}d\bar{x}'$$

$$q_h(x,x';\theta) = \bar{p}_h(\mathbf{P}x,\mathbf{P}x';\theta)\nu(x'|\mathbf{P}x')$$

$$\mathcal{H}(P \,||\, P^\theta) = \int\int \mu(x)p_h(x,x')\log\frac{p_h(x,x')}{q_h(x,x';\theta)}dx'\,dx\,.$$

$$\min_\theta \mathcal{H}(P \,||\, P^\theta) \iff \min_\theta \int |\bar{a}(\mathbf{P}x;\theta) - \mathbf{P}a(x)|^2\mu(x)\,dx$$

"Force-matching"

Example: Two-scale systems - Stochastic Averaging

$$
\begin{aligned}
dX_t^\epsilon &= a(X^\epsilon, Y^\epsilon)dt + dW_t^1 \\
dY_t^\epsilon &= \epsilon^{-1} b(X^\epsilon, Y^\epsilon)dt + \epsilon^{-1/2} dW_t^2 \,,
\end{aligned}
$$

Theory: asymptotics $\epsilon \to 0$ – averaging principle (Khasminskii, etc)

$$
d\bar{X}_t = \bar{a}(\bar{X}_t) + dW_t \,, \quad \bar{a}(x) = \lim_{\epsilon \to 0} \int a(x,y)\, \mu_x^\epsilon(dy)
$$

Minimization of RER: $\mu^\epsilon(dx\, dy) = \bar{\mu}^\epsilon(dx)\mu(dy|x)$ and for $\epsilon \ll 1$
$\mu^\epsilon(dx\, dy) \approx \bar{\mu}(dx)\mu_x(dy)$

Example: Two-scale systems - Stochastic Averaging

$$dX_t^\epsilon = a(X^\epsilon, Y^\epsilon)dt + dW_t^1$$
$$dY_t^\epsilon = \epsilon^{-1} b(X^\epsilon, Y^\epsilon)dt + \epsilon^{-1/2} dW_t^2 \,,$$

Theory: asymptotics $\epsilon \to 0$ – averaging principle (Khasminskii, etc)

$$d\bar{X}_t = \bar{a}(\bar{X}_t) + dW_t \,, \quad \bar{a}(x) = \lim_{\epsilon \to 0} \int a(x,y)\, \mu_x^\epsilon(dy)$$

Minimization of RER: $\mu^\epsilon(dx\, dy) = \bar{\mu}^\epsilon(dx)\mu(dy|x)$ and for $\epsilon \ll 1$
$\mu^\epsilon(dx\, dy) \approx \bar{\mu}(dx)\mu_x(dy)$

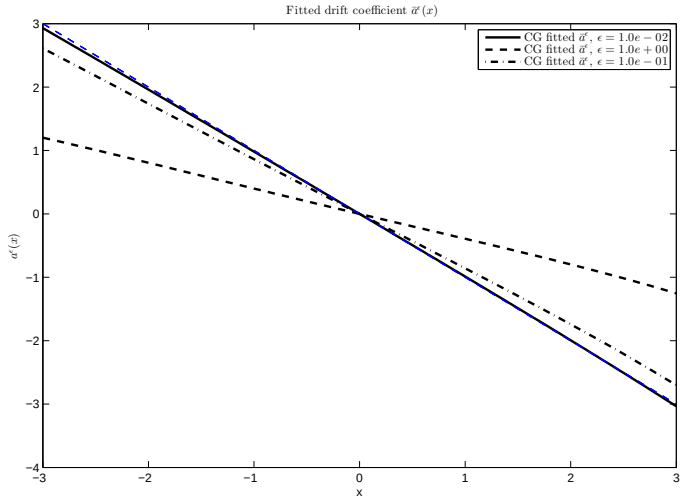
$$\min_{\bar{a}} \int |a(x,y) - \bar{a}(x)|^2 \mu_x(dy)\bar{\mu}(x)\, dx \,,$$

Unique minimizer as $\epsilon \to 0$

$$\bar{a}(x) = \int a(x,y)\mu_x(dy) \,.$$

Example:

$$a(x,y) = -y - y^3$$
$$b(x,y) = x - y - y^3$$
$$\mu_x(dy) \sim e^{-\frac{1}{2}(y-x)^2 - \frac{1}{4}y^4}$$



$\epsilon \to 0$

$$\bar{a}(x) = -x$$

Example: $a(x,y) = -y - y^3$, $b(x,y) = x - y - y^3$

$\mu_x(dy) \sim e^{-\frac{1}{2}(y-x)^2 - \frac{1}{4}y^4}$

$\epsilon \to 0$ $\bar{a}(x) = -x$

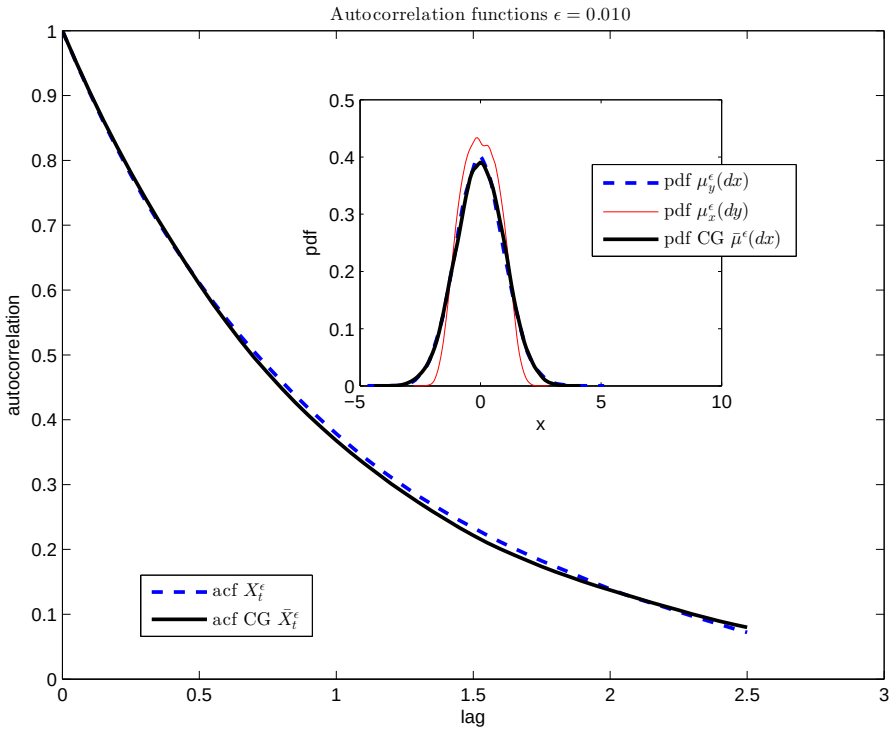


Figure: Autocorrelation function of the CG stationary process $\bar{X}_t^\epsilon$

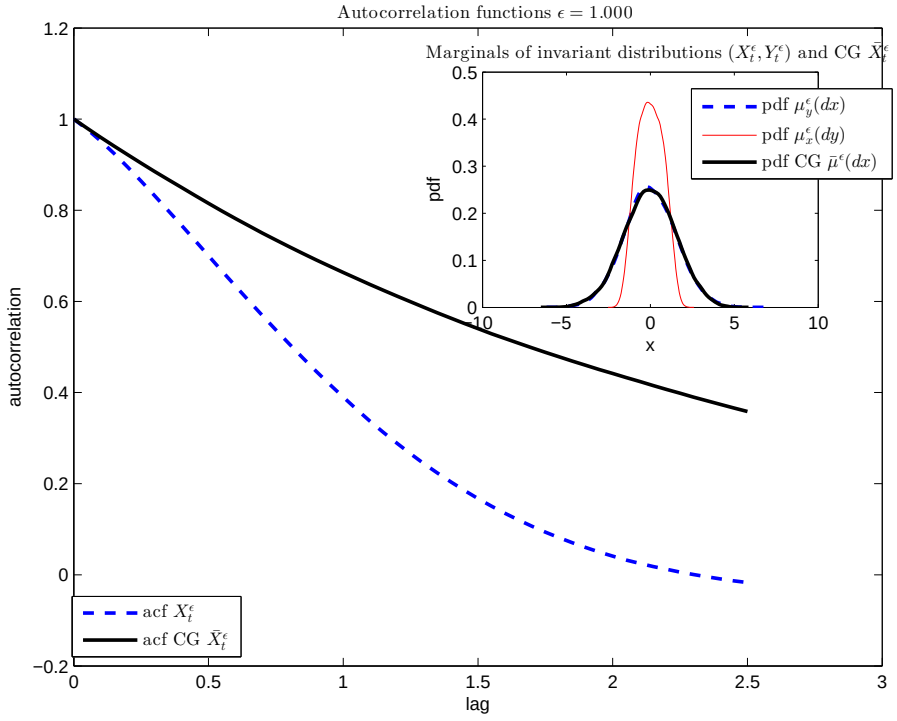


Figure: Autocorrelation function of the CG stationary process $\bar{X}_t^\epsilon$

Driven Arrhenius diffusion

- Rates: Exchange dynamics with the migration rate to n.n. site $|x-y|=1$

$$c(x,y,\sigma) = d\, e^{-\beta(U(x,\sigma))}[\sigma(x)(1-\sigma(x+1)) + \sigma(x)(1-\sigma(x-1))]$$

- Energy barrier: $U(x,\sigma) = \sum_{z\neq x} J(x-z)\sigma(z) - h$
  $J(z) = J_0$, for $|z| \leq L$ and $J = 0$ otherwise.
- Coarse-grained potential:

$$\bar{U}(k,\eta) = \sum_l \bar{J}(k,l)\eta(k) + \bar{J}(0,0)(\eta(k)-1) - \bar{h}$$

- Coarse-grained rates: assume *local equilibrium*, $\sigma(x) \approx q^{-1}\eta(k)$

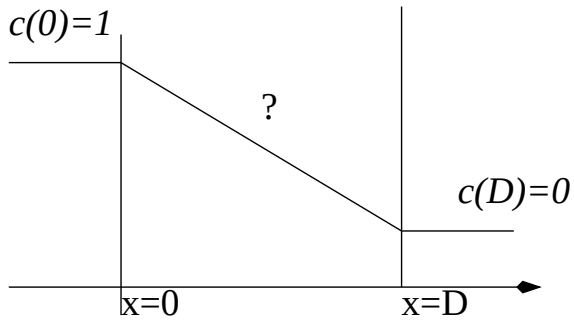
$$\bar{c}(k,l,\eta) = \frac{1}{q}\eta(k)(q-\eta(l))\, d\, e^{-\beta\bar{U}(k,\eta)}$$

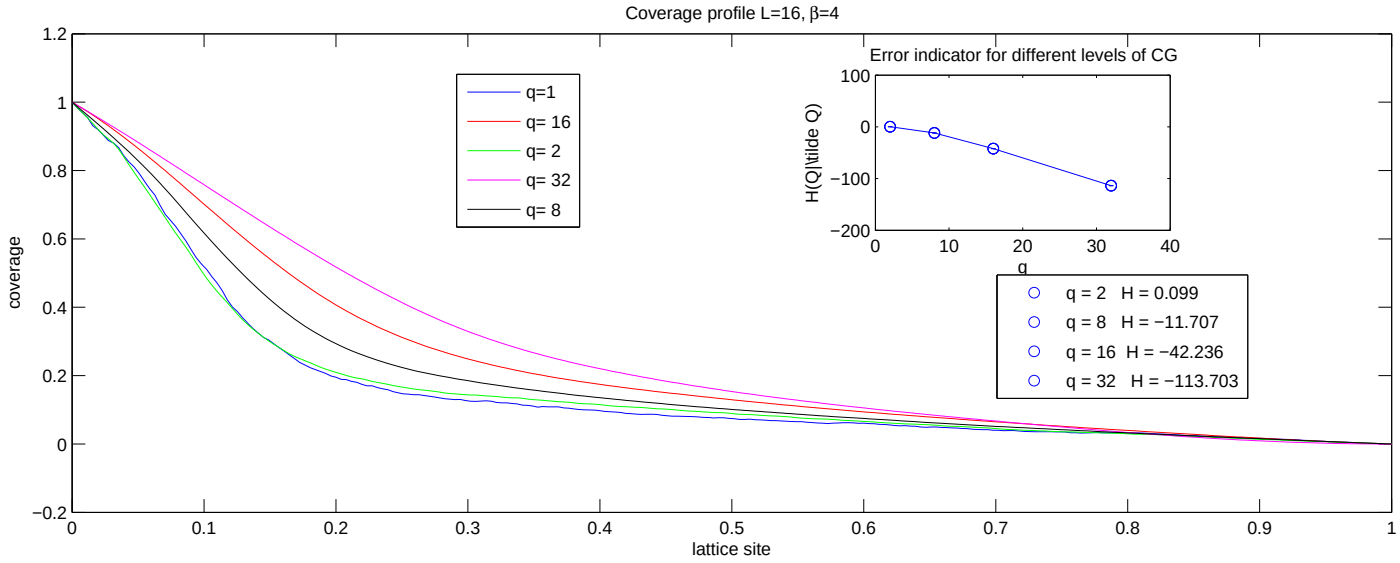
The generator $\bar{\mathcal{L}}$:

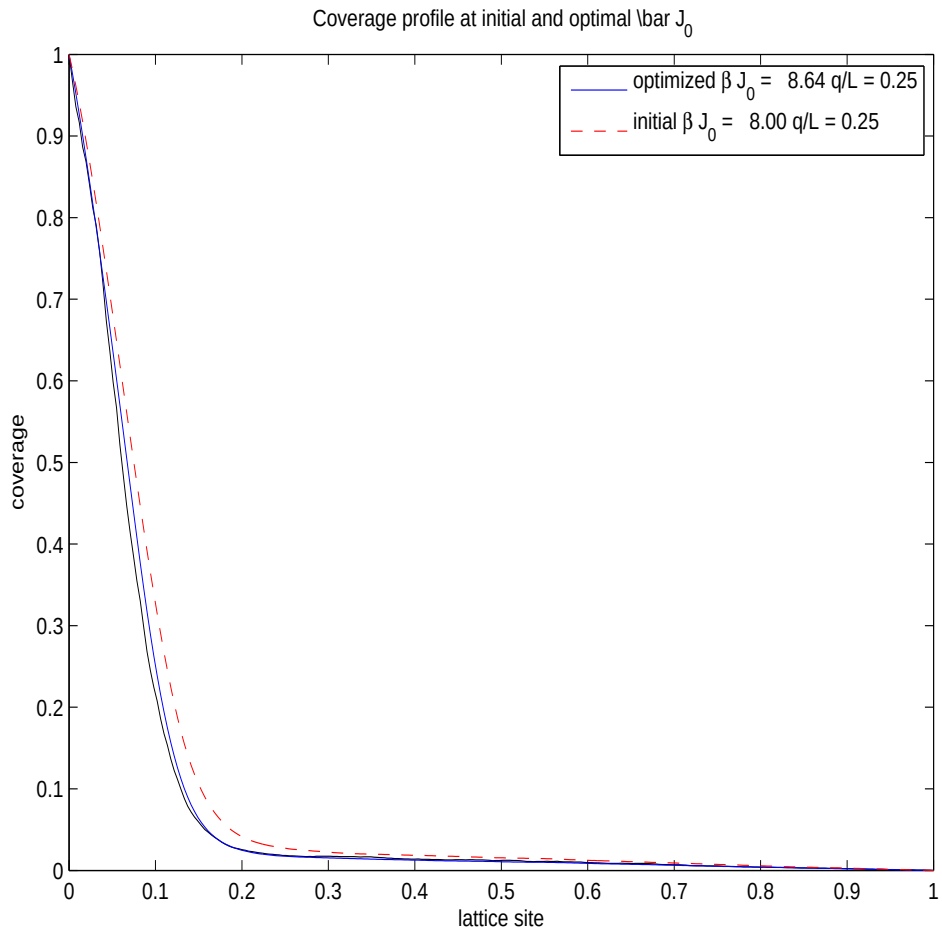
$$\bar{\mathcal{L}}g(\eta) = \sum_{k,l} \bar{c}(k,l,\eta)[g(\eta+\delta_l-\delta_k) - g(\eta)]$$

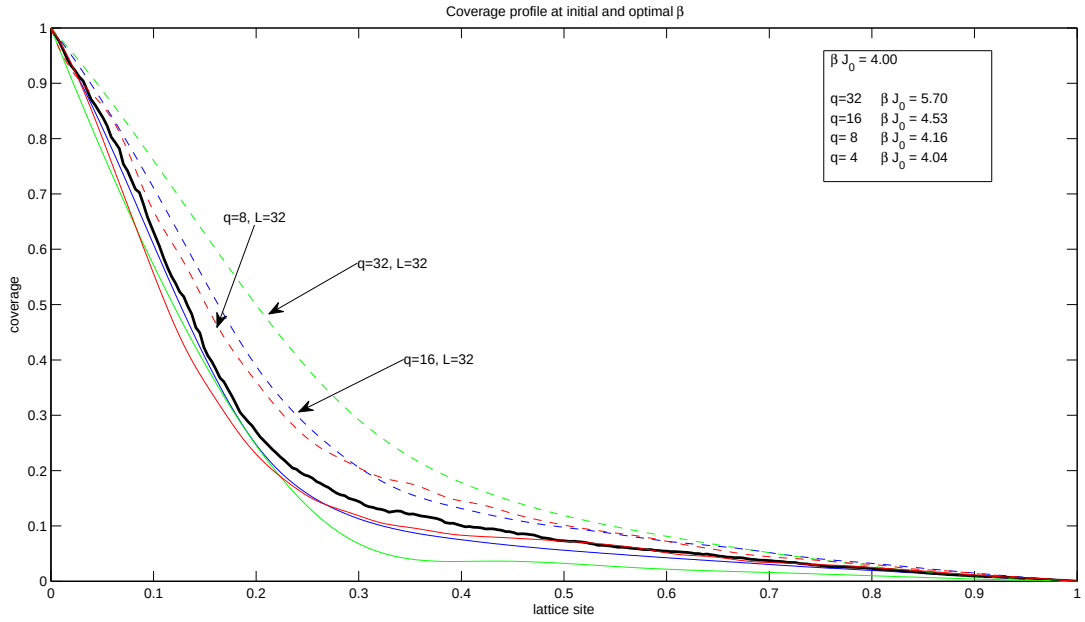
Non-equilibrium stationary states

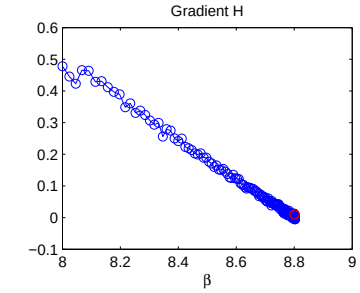
Bounded domain with a gradient in concentrations

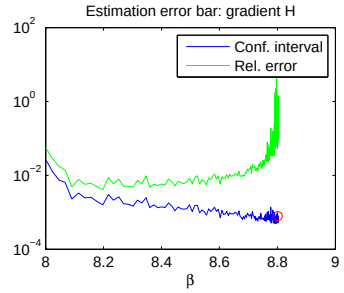


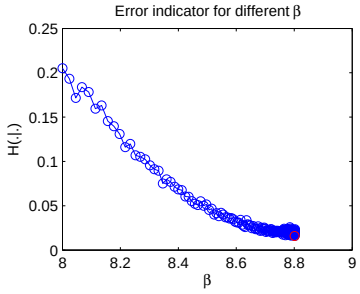

Coverage profile L=16, β=4
coverage
lattice site
q=1
q= 16
q= 2
q= 32
q= 8
Error indicator for different levels of CG
H(Q|ltilde Q)
q
q = 2 H = 0.099
q = 8 H = −11.707
q = 16 H = −42.236
q = 32 H = −113.703

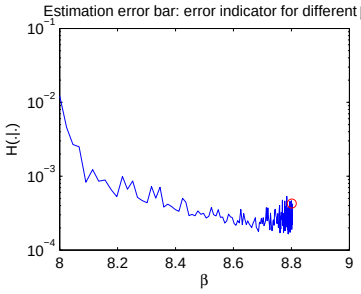

Coverage profile at initial and optimal \bar J0
optimized β J0 = 8.64 q/L = 0.25
initial β J0 = 8.00 q/L = 0.25
coverage
lattice site
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

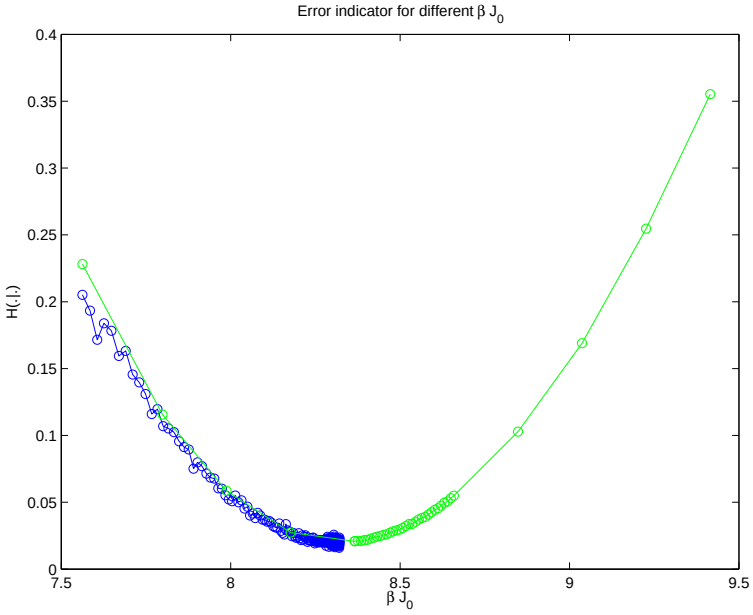

Coverage profile at initial and optimal β
β J0 = 4.00
q=32 β J0 = 5.70
q=16 β J0 = 4.53
q= 8 β J0 = 4.16
q= 4 β J0 = 4.04
q=8, L=32
q=32, L=32
q=16, L=32
coverage
lattice site
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Gradient H
β


Estimation error bar: gradient H
Conf. interval
Rel. error
β


Error indicator for different β
H(·,·)
β


Estimation error bar: error indicator for different β
H(·,·)
β


Error indicator for different β J0
H(·,·)
β J0

Computational example

- **Process:** deposition on a lattice, long-range interactions L, block-spin CG q
  – continuous time Markov jump process
  – Arrhenius dynamics
- **Simulator:** kinetic Monte Carlo
- **Tests:** coarse observable = total coverage c_t
  - phase diagram wrt external field h
  - the average time to phase transition from *low* to *high* coverage
  - adaptive CG for phase diagrams

Are, Katsoulakis, PP, Rey-Bellet SIAM J. Sci. Comp., (2008)

Base CG Hamiltonian $\bar{H}^{(0)}$:
pair-interactions of $\eta(k)$ and $\eta(l)$ with the potential $\bar{J}(k-l)$
$\Longrightarrow$ compressed interaction kernel J using the Haar basis

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

$$\bar{J}(k,k) = J(0,0) = \frac{1}{q(q-1)} \sum_{x \in C_k} \sum_{y \in C_k, y \neq x} J(x-y)\,.$$

$$\bar{H}^{(1)} = \bar{H}^{(1,1)} + \bar{H}^{(1,2)}$$

(i) $\bar{H}^{(1,1)}$ – correction to 2-body interactions

(ii) $\bar{H}^{(1,2)}$ – 3-body interactions

$$
\begin{aligned}
-\bar{H}^{(1,1)}(\eta) &= \frac{\beta}{8}\sum_{k} 4j^2_{kk}(-E_4(k)+E_2(k)) + 2j^1_{kk}(E_4(k)-2E_2(k)+1) + \\
&+ \frac{\beta}{2}\sum_{k<l} j^2_{kl}(E_2(k)-2E_2(k)E_2(l)+E_2(l)) + \\
&+ \frac{\beta}{2}\sum_{k<l} j^1_{kl}(1+E_2(k)E_2(l)-E_2(k)-E_2(l)) + \\
&+ \frac{\beta}{2}\sum_{k,l\neq k} j^2_{kkl}(-E_3(k)E_1(l)+2E_1(k)E_1(l)-E_3(l)E_1(k)) \\
\bar{H}^{(1,2)}(\eta) &= \beta\sum_{k_1}\sum_{k_2>k_1}\sum_{k_3>k_2}[j^2_{k_1k_2k_3}(-E_1(k_1)E_2(k_2)E_1(k_3)+E_1(k_1)E_1(k_3)) + \\
&+ j^2_{k_2k_3k_1}(\ldots k_1,\, k_2,\, k_3 \text{ permut.}\ldots) \\
&\quad + j^2_{k_3k_1k_2}(\ldots k_1,\, k_2,\, k_3 \text{ permut.}\ldots)] \\
E_r(k) &\equiv E_r(\eta(k)) = (2\eta(k)/q-1)^r + o_q(1)
\end{aligned}
$$

"Moments" of interaction potential J:

$$
\begin{aligned}
j^1_{kl} &= \sum_{x\in C_k}\sum_{y\in C_l}(J(x-y)-\bar{J}(k,l))^2\,,\\
j^2_{kl} &= \sum_{x\in C_k}\sum_{y,y'\in C_l}(J(x-y)-\bar{J}(k,l))(J(x-y')-\bar{J}(k,l))\\
j^2_{k_1k_2k_3} &= \sum_{x\in C_{k_1}}\sum_{y\in C_{k_2}}\sum_{z\in C_{k_3}}(J(x-y)-\bar{J}(k_1,k_2))(J(y-z)-\bar{J}(k_2,k_3))
\end{aligned}
$$

Another view: Multiresolution analysis

$\bar{J}$ = projection of J on scaling functions of Haar system

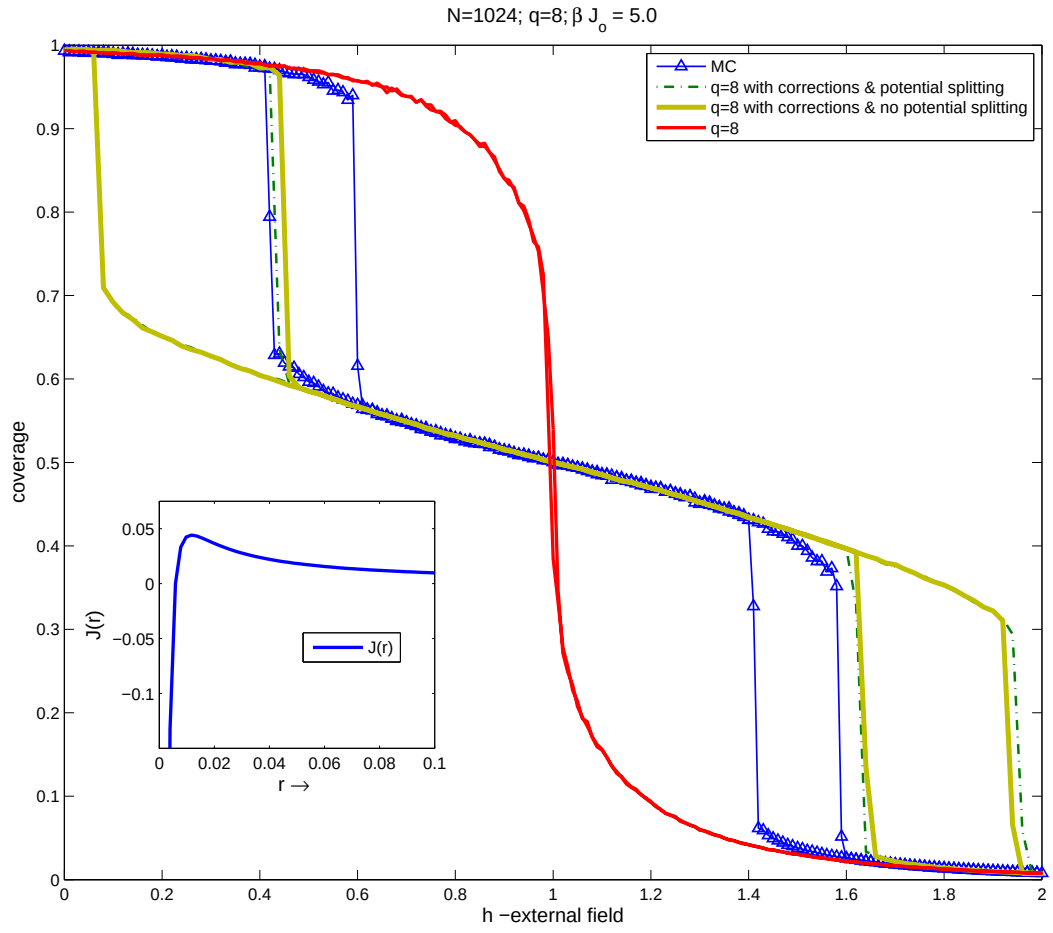
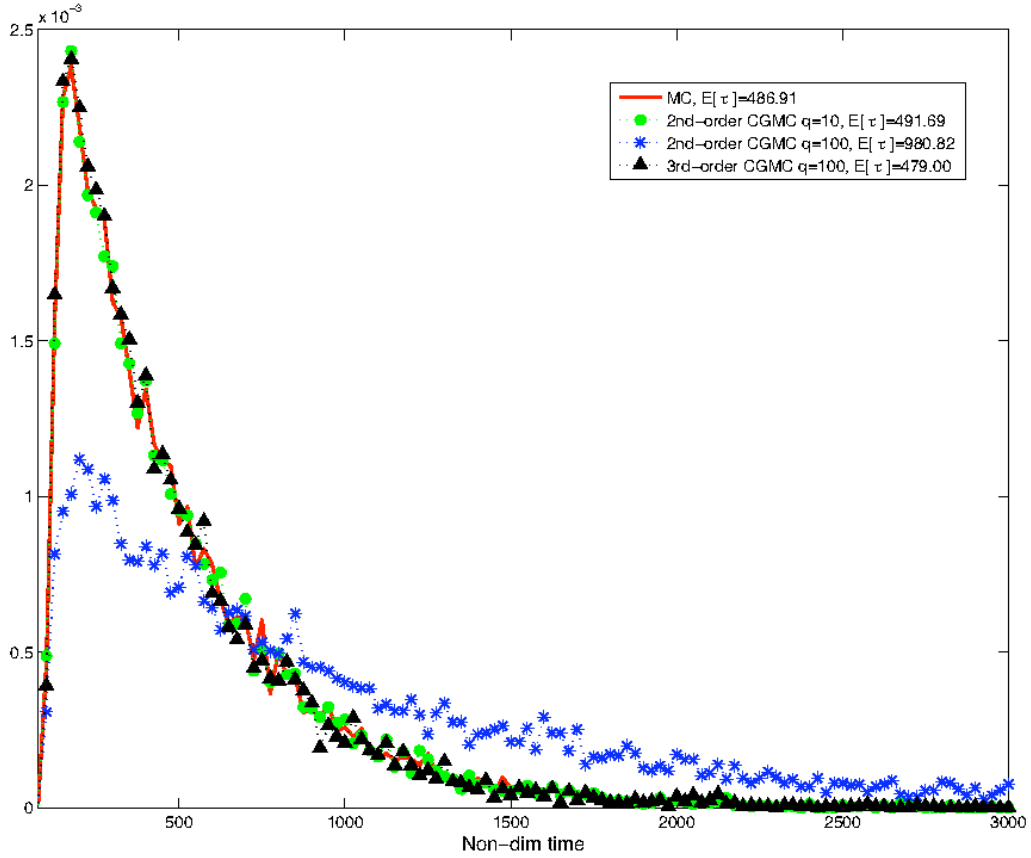
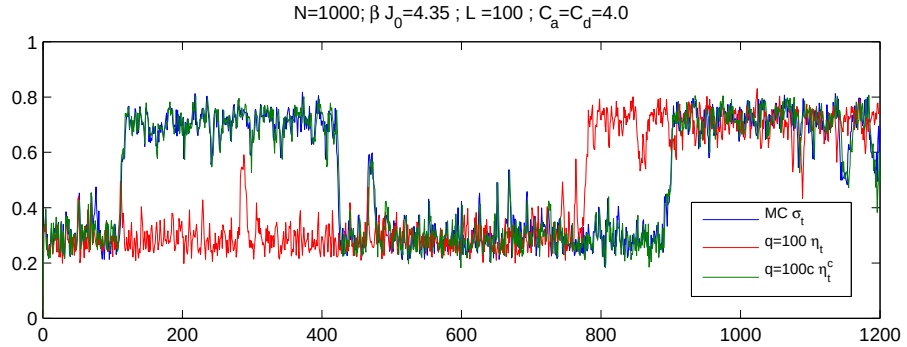


Figure: Comparison of hysteresis using the potential of the microscopic process (MC), the coarse-grained process $q = 8$, the coarse-grained process $q = 8$ with corrections.

Rare events – exit times





Multi-body interactions

Table: Approximation of $\bar{\tau}_T$, $||\rho_\tau^q - \rho_\tau||_{L^1}$

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

CGMC without corrections

L	q	$\bar{\tau}_T$	$\|\|\rho_\tau^q - \rho_\tau\|\|_{L^1}$	Rel. Err.
100	1	486.9	0	0
100	50	584.1	0.0074	20.17%
100	100	980.9	0.0246	101.82%

CGMC with corrections

L	q	$\bar{\tau}_T$	$\|\|\rho_\tau^q - \rho_\tau\|\|_{L^1}$	Rel. Err.
100	50	480.8	0.0025	1.08%
100	100	479.0	0.0028	1.45%

$N = 1000, \beta J_0 = 8.0, h = 0.38375$

CGMC without corrections

L	q	$\bar{\tau}_T$	$\|\rho_\tau^q - \rho_\tau\|_{L^1}$	Rel. Err.
100	1	367.9	0	0
100	50	569.4	0.0131	54.78%
100	100	1482.23	0.0416	302.9%

CGMC with corrections

L	q	$\bar{\tau}_T$	$\|\rho_\tau^q - \rho_\tau\|_{L^1}$	Rel. Err.
100	50	335.9	0.0042	8.68%
100	100	290.6	0.0072	21.00%

Efficiency ?

Table: CPU cost comparisons of different CG algorithms

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

Process	CPU (secs)
q=1 (no coarse-graining)	322192
q=8	5232
q=8c (no splitting)	69473
q=8c (splitting)	6900

Table: Computational complexity of evaluating the Hamiltonian

	Count	Speed-up
Microscopic $q = 1$: $H_N(\sigma)$	$\mathcal{O}(NL^d)$	1
Scheme 2nd order: $\bar{H}_M^{(0)}$	$\mathcal{O}(ML^d/q^d)$	$\mathcal{O}(q^{2d})$
Scheme 3rd order: $\bar{H}_M^{(0)} + \bar{H}_M^{(1)}$	$\mathcal{O}(ML^{2d}/q^{2d})$	$\mathcal{O}(q^{3d}/L^d)$

Lattice Models: Dynamics

- Continuous Time Markov Chain $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$
  $\sigma \in \Sigma \equiv \{0,1\}^{\Lambda_N}$, $\Lambda_N \subset \mathbb{Z}^d$

$$\mathbb{P}\left(\sigma_{t+\delta t} = \sigma' \,||\, \sigma_t = \sigma\right) = c(\sigma, \sigma')\delta t + o(\delta t)$$

Rates: $c(\sigma, \sigma') \equiv c(x, \omega; \sigma)$

Lattice Models: Dynamics

- Continuous Time Markov Chain $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$
  $\sigma \in \Sigma \equiv \{0,1\}^{\Lambda_N}$, $\Lambda_N \subset \mathbb{Z}^d$

$$\mathbb{P}\left(\sigma_{t+\delta t} = \sigma' \,||\, \sigma_t = \sigma\right) = c(\sigma, \sigma')\delta t + o(\delta t)$$

  Rates: $c(\sigma, \sigma') \equiv c(x, \omega; \sigma)$

- Forward Kolmogorov Equation (aka Master Equation)

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

$$P(\sigma, 0; \zeta) = \delta(\sigma - \zeta)$$

Lattice Models: Dynamics

- Continuous Time Markov Chain $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$
  $\sigma \in \Sigma \equiv \{0,1\}^{\Lambda_N}$, $\Lambda_N \subset \mathbb{Z}^d$

$$\mathbb{P}\left(\sigma_{t+\delta t} = \sigma' \,||\, \sigma_t = \sigma\right) = c(\sigma, \sigma')\delta t + o(\delta t)$$

  Rates: $c(\sigma, \sigma') \equiv c(x, \omega; \sigma)$

- Forward Kolmogorov Equation (aka Master Equation)

$$\partial_t P(\sigma, t; \zeta) = \sum_{\sigma', \sigma' \neq \sigma} c(\sigma', \sigma) P(\sigma', t; \zeta) - \lambda(\sigma) P(\sigma, t; \zeta)\,,$$
$$P(\sigma, 0; \zeta) = \delta(\sigma - \zeta)$$

- Simulation: Embedded Markov Chain $\{X_n\}_{n\geq 0} = \{\sigma_{n\delta t}\}$, $\sigma \to \sigma^{x,\omega}$

$$p(\sigma, \sigma^{x,\omega}) = \frac{c(x, \omega; \sigma)}{\lambda(\sigma)}\,, \quad \lambda(\sigma) = \sum_x \sum_\omega c(x, \omega; \sigma)$$

  Exponential clock: $\delta t \sim \mathrm{Exp}(\lambda(\sigma))$

Generator of the process

- Evolution of observables (Backward Kolomogorov Equation):

$$u(\zeta,t) = \mathbb{E}_\zeta[f(\sigma_t)] \equiv \sum_\sigma f(\sigma) P(\sigma,t;\zeta)$$

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

Generator of the process

- Evolution of observables (Backward Kolomogorov Equation):

$$u(\zeta,t) = \mathbb{E}_\zeta[f(\sigma_t)] \equiv \sum_\sigma f(\sigma) P(\sigma,t;\zeta)$$

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

- Generator:

$$\mathcal{L} f(\sigma) = \sum_{\sigma'} c(\sigma,\sigma')[f(\sigma') - f(\sigma)] = \sum_x \sum_\omega c(x,\omega;\sigma)[f(\sigma^{x,\omega}) - f(\sigma)]$$

- Markov semigroup $\mathbf{P}_t = e^{t\mathcal{L}}$

$$\delta_x f(\sigma) = f(\sigma^{x,\omega}) - f(\sigma)$$

$$C_\infty(\Sigma) = \{f \in C_b(\Sigma) \,||\, \sum_x \|\delta_x(f)\|_\infty < \infty\}$$

Events

- *Adsorption/desorption:*

$$\sigma^x = \begin{cases} 1 - \sigma(x) & \text{if } z = x \\ \sigma(z) & \text{if } z \neq x \end{cases}$$

- *Diffusion (spin exchange, Kawasaki dynamics):*

$$\sigma^{x,y}(z) = \begin{cases} \sigma(y) & \text{if } z = x \\ \sigma(x) & \text{if } z = y \\ \sigma(z) & \text{if } x \neq y \end{cases}$$

- *Multicomponent reactions.*
$\sigma(x) \in \{0, 1, \ldots, K\}$

$$\sigma^{(x,k)}(z) = \begin{cases} \sigma(z) & \text{if } z \neq x, y, \\ k & \text{if } z = x. \end{cases}$$

- *Reactions involving particles with internal degrees of freedom.*

$$\sigma^{(x,y,k,l)}(z) = \begin{cases} \sigma(z) & \text{if } z \neq x, y, \\ k & \text{if } z = x, \\ l & \text{if } z = y, \end{cases}$$

Rate functions I

EXAMPLE: Arrhenius dynamics for adsorbtion/desorption

- Transition rate to the gas phase: $c(x,\sigma) = c_2\sigma(x)e^{-\beta U(x,\sigma)}$
- Energy barrier: $U(x,\sigma) = \sum_{z\neq x} J(x-z)\sigma(z) - h$
- Transition rates:

$$c(\sigma,\sigma^x) = c_1(1-\sigma(x)) + c_2\sigma(x)e^{-\beta U(x,\sigma)}$$

- Reversible w.r.t. Gibbs measure $\mu \sim e^{-\beta H(\sigma)}$
  Detailed balance

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

CG and acceleration of KMC

Stochastic Simulation Algorithm (SSA)

Gillespie, JCP (1976), chemical reactions in well-mixed systems.

Step 1: Update. (a) Calculate: $c(y,\sigma)$, $\forall y \in \Lambda_N$
(b) Calculate:

$$\lambda_x(\sigma) = \sum_{y<x} c(y,\sigma)\,, \qquad \lambda(\sigma) = \sum_{y\in\Lambda_N} c(y,\sigma)$$

Step 2: Search. $u_1 \sim \mathrm{U}([0,1))$ and search for $x \in \Lambda_N$ such that

$$\lambda_{x-1}(\sigma) < \lambda(\sigma)u_1 \leq \lambda_x(\sigma)$$

Step 3: Time. $t \leftarrow t + \delta t$, $\delta t \sim \mathrm{Exp}(\lambda(\sigma))$
$\delta t = -\log(u_2)/\lambda(\sigma)$, $u_2 \sim \mathrm{U}([0,1))$
$\sigma_{t+\delta t} = \sigma^x$

Kinetic Monte Carlo Implementation

n-fold Algorithm (aka BKL)

Bortz, Kalos, Lebowitz, JCP (1975), Ising spin lattice systems.

Step 1: Update. (a) Calculate $c(y,\sigma)$, $\forall y \in \Lambda_N$

Step 2: Search. Group sites $x \in \Lambda_N$ in classes D_i, $i = 1, \ldots, n$, define

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

Generate $u \sim \mathrm{U}([0,1))$ and search for $i = 1, \ldots, n$ s.t.

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

then choose $x \in D_i$ uniformly.

Step 3: Time $t \leftarrow t + \delta t$, $\delta t \sim \mathrm{Exp}(Q_n(\sigma))$

$\sigma_{t+\delta t} = \sigma^x$

Kinetic Monte Carlo Implementation

Uniformization – Null-event Algorithm

Choose $\lambda(\sigma) \leq \lambda^*$ and $\{Y_n\}$ such that

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

Bounds: $U^* = \min_{x,\sigma} U(x,\sigma)$, $\quad \lambda^{*,\mathrm{loc}} = d_0 \max\{1, e^{-\beta U^*}\}$

Step 1: Search/Update. Select: $x \in \Lambda_N$ uniformly
Calculate: $c(x,\sigma)$

Step 2: Time/Accept/Reject. $t \leftarrow t + \delta t$, $\delta t \sim \mathrm{Exp}(\lambda^{*,\mathrm{loc}})$
Generate $u \in \mathrm{U}([0,1))$
If $c(x,\sigma) \geq \lambda^{*,\mathrm{loc}} u$ then $\sigma_{t+\delta t} = \sigma^x$
If $c(x,\sigma) < \lambda^{*,\mathrm{loc}} u$ then $\sigma_{t+\delta t} = \sigma$

Spatial two-level kinetic Monte Carlo

- continuous time Markov jump process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$

$$\mathcal{L}f(\sigma) = \sum_{\sigma'} c(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

Embedded Markov Chain $\{X_n = \sigma_{n\delta t}\}_{n\geq 0}$

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

Spatial two-level kinetic Monte Carlo

- continuous time Markov jump process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$

$$\mathcal{L}f(\sigma) = \sum_{\sigma'} c(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

Embedded Markov Chain $\{X_n = \sigma_{n\delta t}\}_{n\geq 0}$

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

- construct an *approximating* process $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$

$$\widetilde{\mathcal{L}}f(\sigma) = \sum_{\sigma'} \widetilde{c}(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

Spatial two-level kinetic Monte Carlo

- continuous time Markov jump process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$

$$\mathcal{L}f(\sigma) = \sum_{\sigma'} c(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

Embedded Markov Chain $\{X_n = \sigma_{n\delta t}\}_{n\geq 0}$

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

- construct an *approximating* process $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$

$$\widetilde{\mathcal{L}}f(\sigma) = \sum_{\sigma'} \widetilde{c}(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

Spatial two-level kinetic Monte Carlo

- continuous time Markov jump process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$

$$\mathcal{L}f(\sigma) = \sum_{\sigma'} c(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

Embedded Markov Chain $\{X_n = \sigma_{n\delta t}\}_{n\geq 0}$

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

- construct an *approximating* process $(\{\tilde{\sigma}_t\}_{t\geq 0}, \tilde{\mathcal{L}})$

$$\tilde{\mathcal{L}}f(\sigma) = \sum_{\sigma'} \tilde{c}(\sigma, \sigma')(f(\sigma') - f(\sigma))$$

- use coupling with the coarse process $(\{\eta_t\}_{t\geq 0}, \bar{\mathcal{L}})$ $\mathbf{T} : \Sigma \to \bar{\Sigma}$, $\mathbf{T}\sigma = \eta$. The coarse generator $\bar{\mathcal{L}}$ with the rates $\bar{c}(\eta', \eta)$

$$\bar{c}(\eta, \eta')c_{\mathrm{rf}}(\sigma'|\eta', \sigma) = \tilde{c}(\sigma, \sigma')(= c(\sigma, \sigma'))$$

Approximating embedded Markov chain $\{\tilde{X}_n\}_{n\geq 0}$:
Coarse level: $\eta \to \eta' \in \bar{\Sigma}$

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

Microscopic level: $\sigma' \in \Sigma$, s.t. $\mathbf{T}\sigma' = \eta'$
accept with the probability

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

or reject with the probability

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

Time step: $t \leftarrow t + \delta t$, $\delta t \sim \mathrm{Exp}(\tilde{\lambda}^*(\sigma))$

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

Lemma

For any $\sigma \in \Sigma$ we have $\widetilde{\lambda}(\sigma) \leq \widetilde{\lambda}^*(\sigma) \equiv \bar{\lambda}(\eta)\lambda_{\mathrm{rf}}(\sigma)$.
If we assume exact sampling then $\lambda(\sigma) \leq \widetilde{\lambda}^*(\sigma)$.

Rejection rate:

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

$$= 1 - \sum_{\sigma' \in \Sigma} \frac{\widetilde{c}(\sigma, \sigma')}{\bar{\lambda}(\eta)\lambda_{\mathrm{rf}}(\sigma)} = 1 - \frac{\widetilde{\lambda}(\sigma)}{\bar{\lambda}(\eta)\lambda_{\mathrm{rf}}(\sigma)} .$$

Note: Lumpable process with respect $\eta = \mathbf{T}\sigma$

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

then $c_{\mathrm{rf}}(\sigma' | \eta', \sigma) = 1/|\{\sigma' : \mathbf{T}\sigma' = \eta'\}|$ for all $\sigma' \in \{\sigma' : \mathbf{T}\sigma' = \eta'\}$ such that $\sum_{\{\sigma' : \mathbf{T}\sigma' = \eta'\}} c_{\mathrm{rf}}(\sigma' | \eta', \sigma) = 1$, $\lambda_{\mathrm{rf}}(\sigma) = 1$, $\widetilde{\lambda}(\sigma) = \bar{\lambda}(\eta)$ thus

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

Lemma

Let the coarse rates define an approximately lumpable process, that is

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

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

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

Approximate two-level coarse-grained dynamics

Example: Arrhenius spin-flip dynamics

$$U(x,\sigma) = U^{(s)}(x,\sigma) + U^{(l)}(x,\sigma)$$

Coarse rates:

$$\bar{c}_a(k,\eta) = c_1\left(q - \eta(k)\right), \quad \bar{c}_d(k,\eta) = c_2\eta(k)e^{-\beta\,\bar{U}^{(l)}(k,\eta)},$$

$$\bar{U}^{(l)}(k,\eta) = \sum_{\substack{l\in\Lambda_M^c \\ l\neq k}} \bar{J}(k,l)\eta(l) + \bar{J}(k,k)(\eta(k)-1) - \tfrac{1}{2}\bar{h}(k).$$

Reconstruction rates:

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

$$\begin{aligned} U^{(s)}(x,\sigma) &= \sum_{y\neq x, y\in\Lambda_N} K(x-y)\sigma(y) - \frac{1}{2}h(x), \\ U^{(l)}(x,\sigma) &= \sum_{y\neq x, y\in\Lambda_N} J(x-y)\sigma(y) - \frac{1}{2}h(x). \end{aligned}$$

Markov process $(\{\widetilde{\sigma}_t\}_{t\geq 0}, \widetilde{\mathcal{L}})$:

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

$$\widetilde{U}(x,\sigma) = U^{(s)}(x,\sigma) + \bar{U}^{(l)}$$

Detailed balance: $\widetilde{c}(x,\sigma)$ satisfy the detailed balance condition with

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

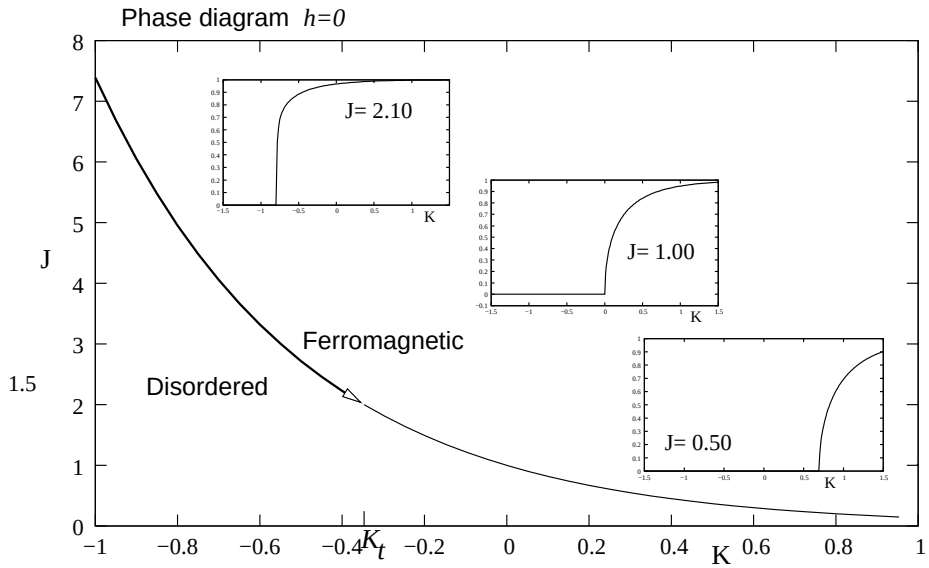
and $\widetilde{Z}_N$ is the normalization constant corresponding to the Hamiltonian

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

Benchmark: Long and short-range interactions

$$\beta H(\sigma) = -\frac{K}{2}\sum_{x}\sum_{|x-y|=1}\sigma(x)\sigma(y) - \frac{J}{2N}\sum_{x}\sum_{x\neq y}\sigma(x)\sigma(y) - h\sum_{x}\sigma(x)$$

Exactly solvable in $d = 1, 2$ with the explicitly given total coverage $c(K, J, h; \beta)$



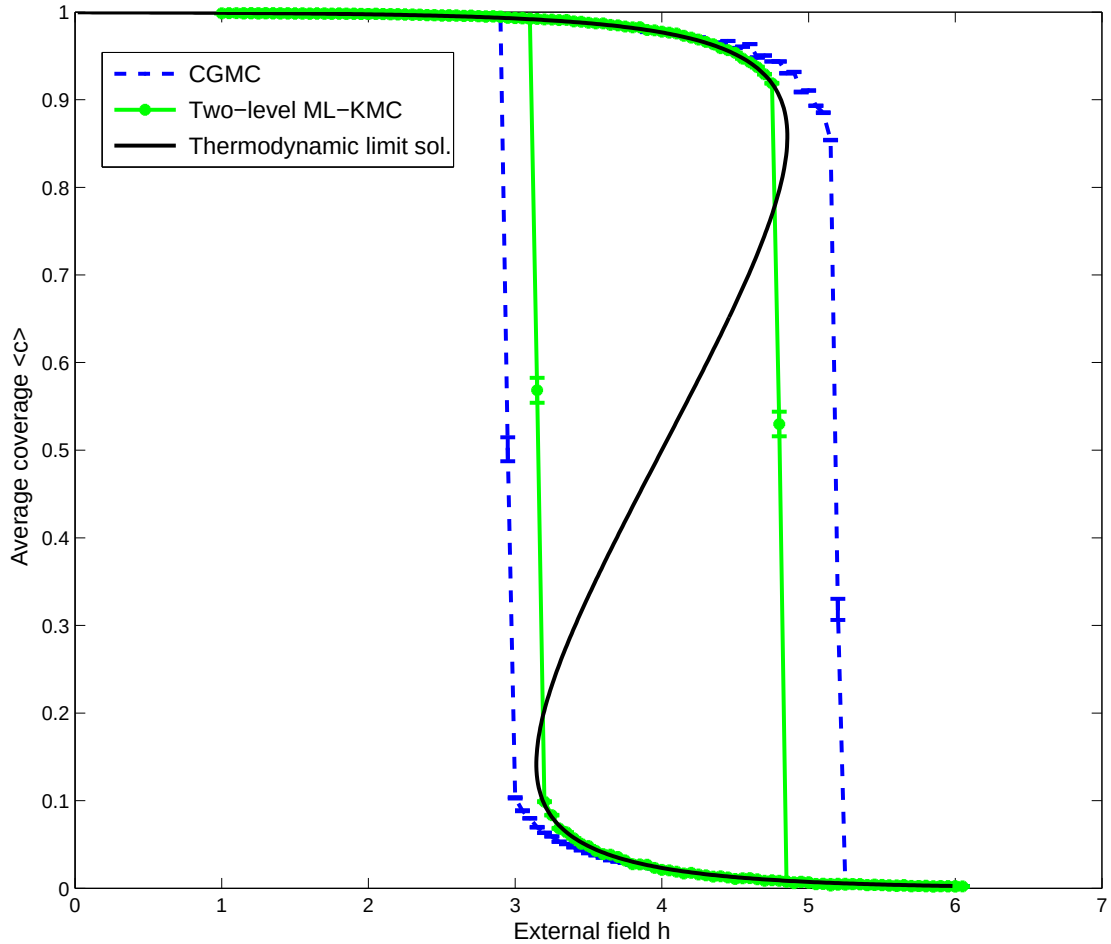


Figure: $K = 3,\ J = 5,\ L = N,\ N = 1024.$

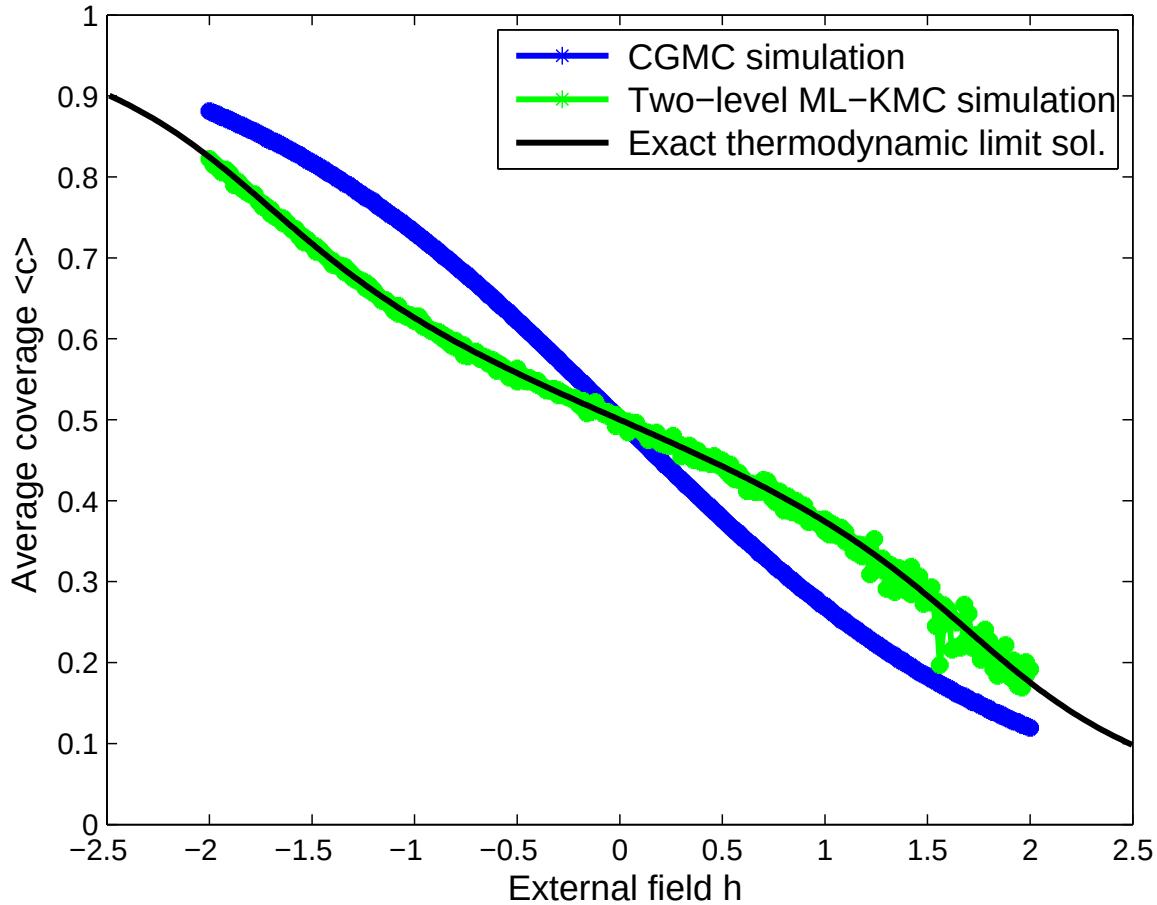


Figure: $K = -5,\ J = 5,\ L = 20,\ N = 1024,\ q = N.$

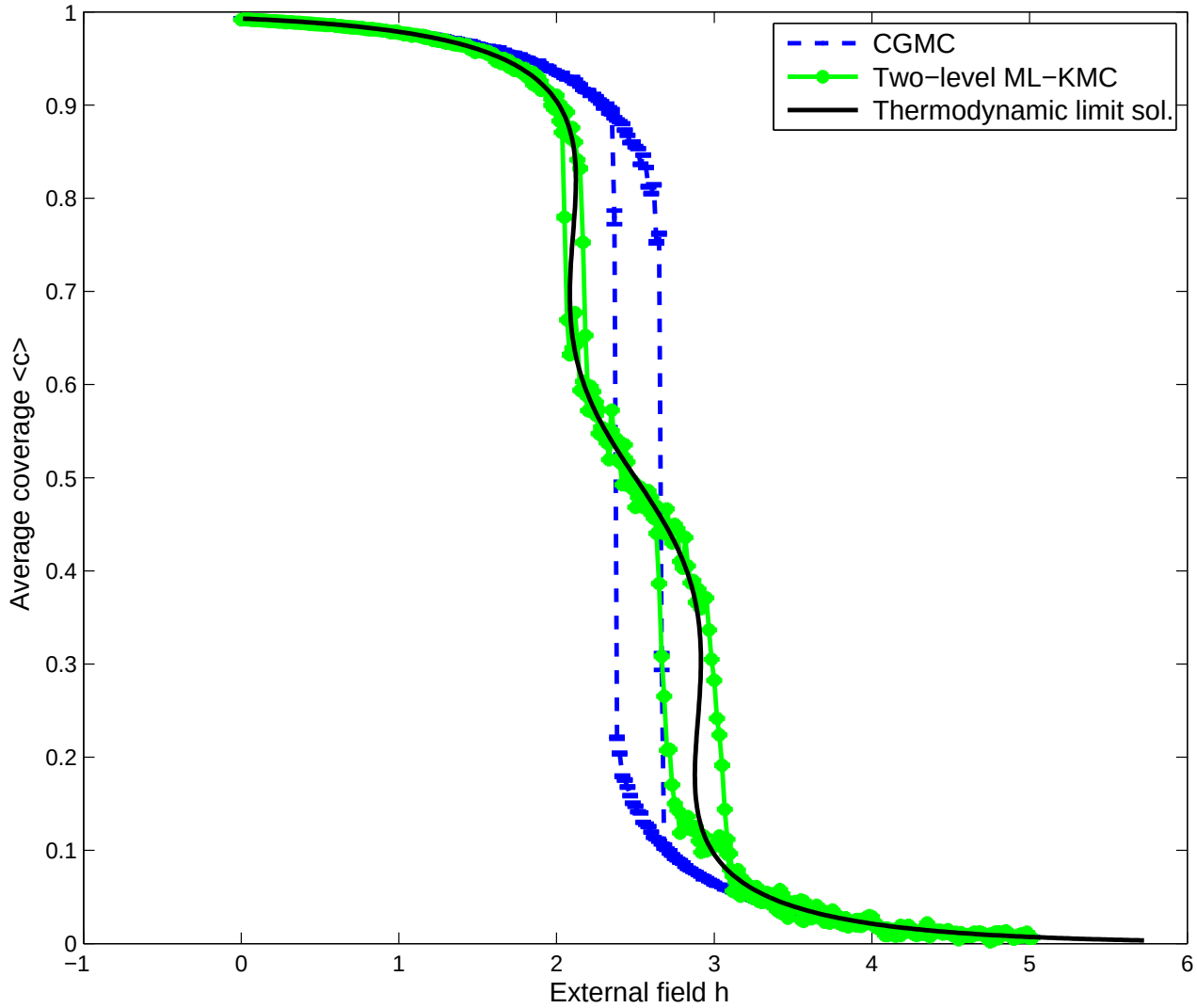


Figure: $K = -5,\ J = 10,\ L = 100,\ N = 1024,\ q = N.$

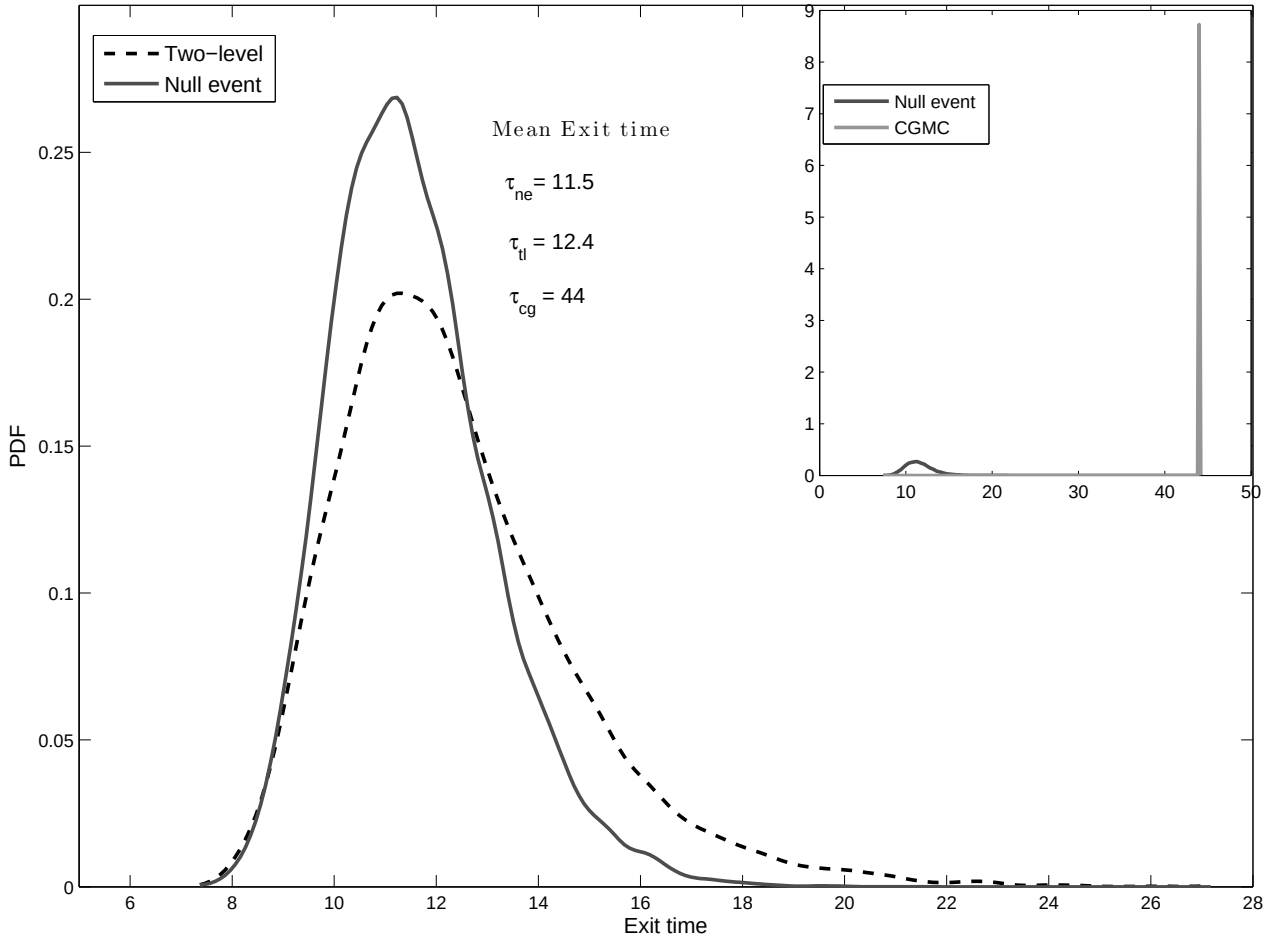


Figure: Comparing the probability density function of the exit time. $K = 3$, $J = 5$, $h = 3.1$, $L = 100$, $N = 1024$, $q = N$.

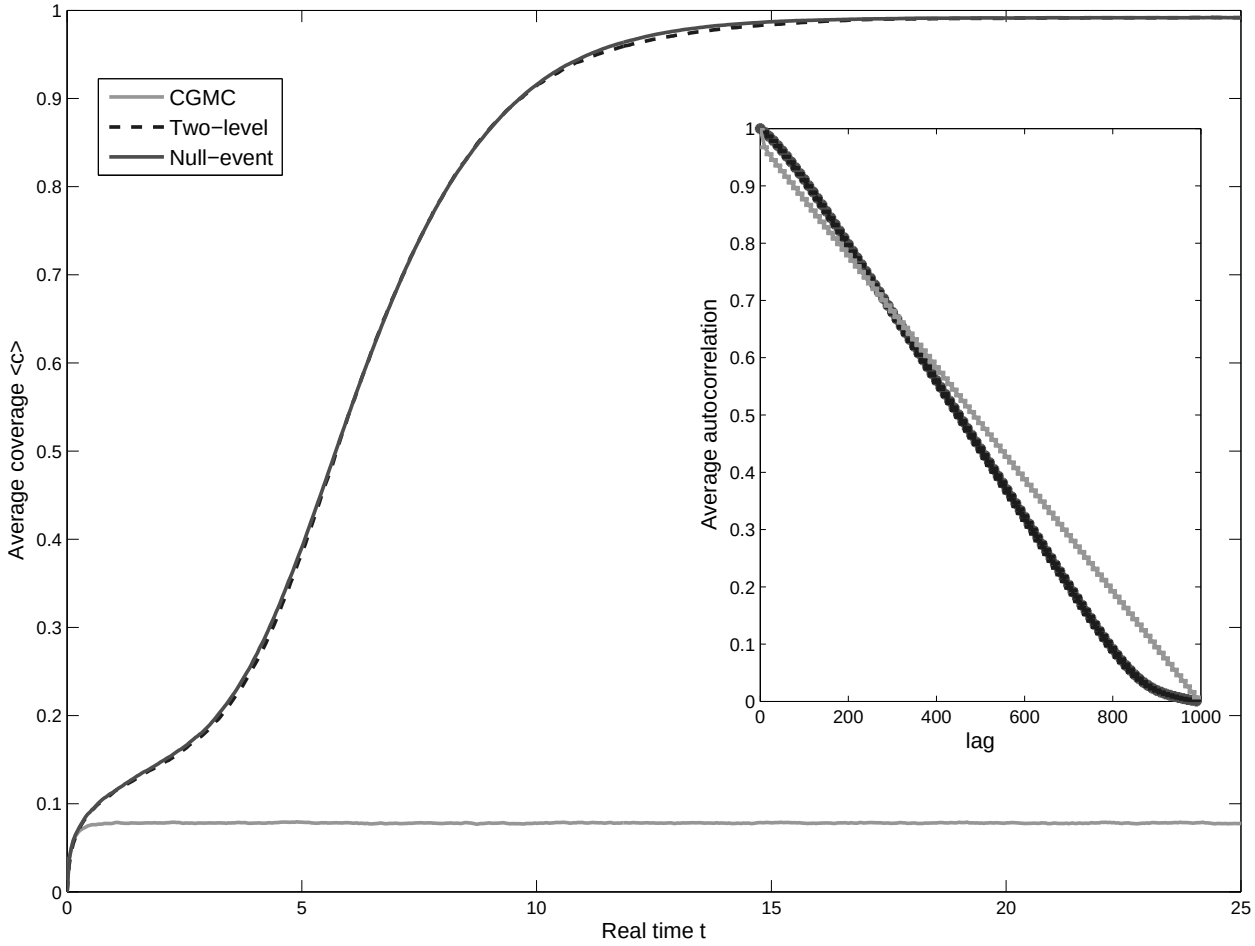


Figure: Average coverage trajectory. $K = 3$, $J = 5$, $h = 3.1$, $L = 100$, $N = 1024$, $q = N$.

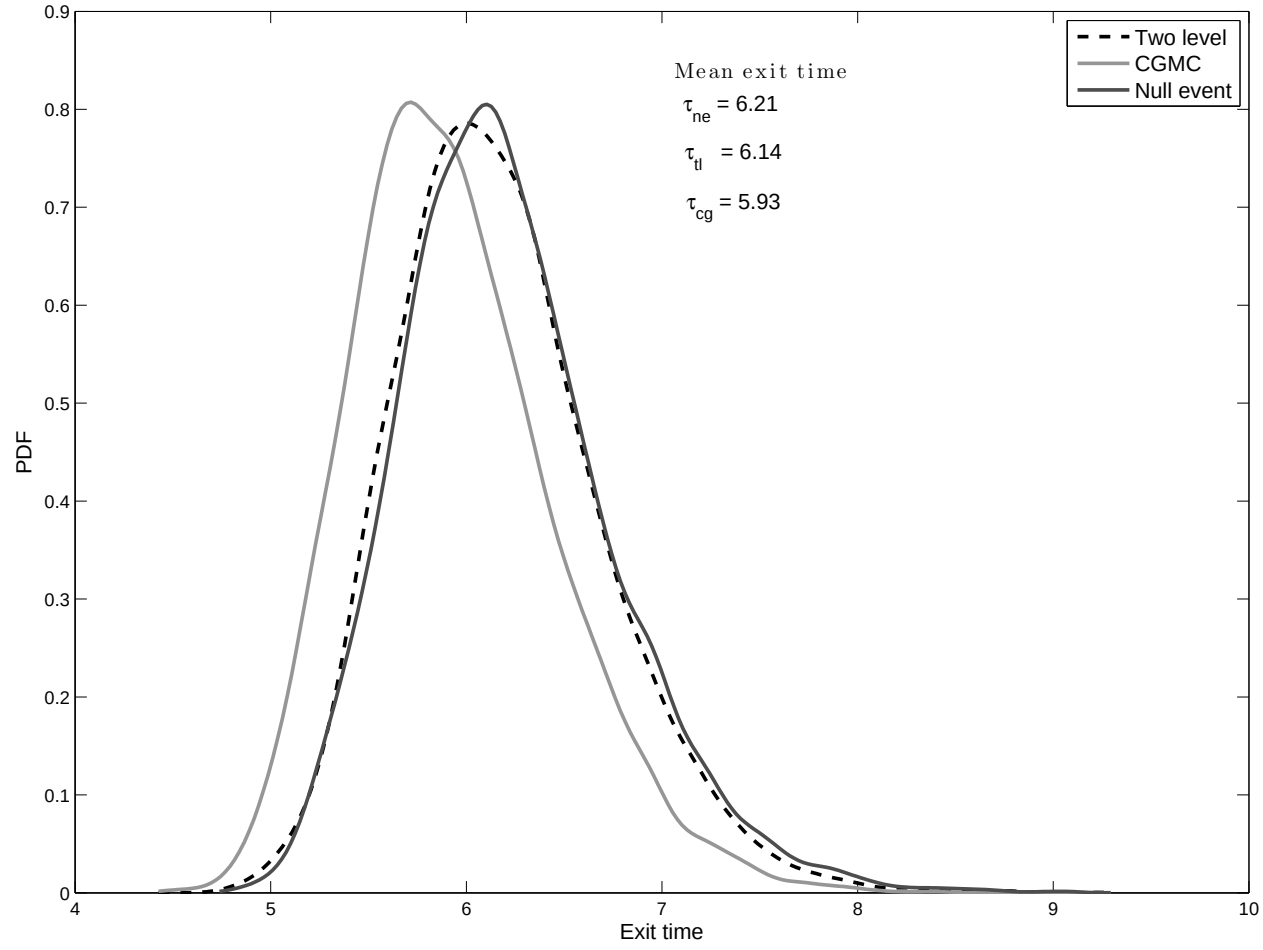


Figure: Comparing the probability density function of the exit time. $K = 3$, $J = 5$, $h = 2.1$, $L = 100$, $N = 1024$, $q = N$.

Table: Approximation of the mean exit time τ. $J = 5$, $q = N$, $N = 1024$ fixed.

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

Table: CPU time (seconds): The evolution final time $T = 20$, $K = 1$, $J = 5$, $h = 2.5$, $L = N$, and $q = N$

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

Path Space Relative Entropy

- Infers information about the path distribution: it contains information not only for the invariant measure but also for the dynamics.

Path Space Relative Entropy

- Infers information about the path distribution: it contains information not only for the invariant measure but also for the dynamics.

- No need for explicit knowledge of invariant measure. Thus, it is suitable for reaction networks and non-equilibrium steady state systems.

Path Space Relative Entropy

- Infers information about the path distribution: it contains information not only for the invariant measure but also for the dynamics.
- No need for explicit knowledge of invariant measure. Thus, it is suitable for reaction networks and non-equilibrium steady state systems.
- Relative entropy rate $\mathcal{H}$ is an observable $\Rightarrow$ tractable and statistical estimators can provide easily and efficiently its value using KMC solvers.

Path Space Relative Entropy

- Infers information about the path distribution: it contains information not only for the invariant measure but also for the dynamics.
- No need for explicit knowledge of invariant measure. Thus, it is suitable for reaction networks and non-equilibrium steady state systems.
- Relative entropy rate $\mathcal{H}$ is an observable $\Rightarrow$ tractable and statistical estimators can provide easily and efficiently its value using KMC solvers.
- Minimizing the error in $\mathcal{H}$ gives optimal parametrization similar to max-likelihood parameter estimation.

Path Space Relative Entropy

- Infers information about the path distribution: it contains information not only for the invariant measure but also for the dynamics.
- No need for explicit knowledge of invariant measure. Thus, it is suitable for reaction networks and non-equilibrium steady state systems.
- Relative entropy rate $\mathcal{H}$ is an observable $\Rightarrow$ tractable and statistical estimators can provide easily and efficiently its value using KMC solvers.
- Minimizing the error in $\mathcal{H}$ gives optimal parametrization similar to max-likelihood parameter estimation.
- Fisher information matrix allows for parameter identifiability in parameterization of dynamics [analogue to Cramer-Rao Theorems]

Conclusions

- Path-Space Information Theory Methods for Hi-Dim. stochastic systems [in state & parameter space]
  - Sensitivity Analysis, Robustness, Parameter Identifiability
  - "Best-fit" Coarse-grained Dynamics to fine-scale data

Conclusions

- Path-Space Information Theory Methods for Hi-Dim. stochastic systems [in state & parameter space]
  - Sensitivity Analysis, Robustness, Parameter Identifiability
  - "Best-fit" Coarse-grained Dynamics to fine-scale data
- Mathematical/Computational tools
  - Relative Entropy Rate, Path-Space Fisher Information Matrix
  - Efficient statistical estimators for RER
  - Sensitivity bounds for observables

Conclusions

- Path-Space Information Theory Methods for Hi-Dim. stochastic systems [in state & parameter space]
  - Sensitivity Analysis, Robustness, Parameter Identifiability
  - "Best-fit" Coarse-grained Dynamics to fine-scale data
- Mathematical/Computational tools
  - Relative Entropy Rate, Path-Space Fisher Information Matrix
  - Efficient statistical estimators for RER
  - Sensitivity bounds for observables
- Further Research
  - Observables and risk-sensitive bounds
  - Synergies with other SA methods: Goal-oriented stochastic coupling methods
  - Global Sensitivity Analysis and Bayesian perspective (prior knowledge on parameters)
  - SA for complex stochastic dynamics (non-gaussian behavior, intermittency, memory, etc.)

References

- **Sensitivity Analysis in Path Space (information-theoretic, goal-oriented)**

  Sensitivity bounds and error estimates for stochastic models, P. Dupuis, M.A. Katsoulakis, Y. Pantazis, P. Plechac, preprint 2014

  A Relative Entropy Rate Method for Path Space Sensitivity Analysis of Stationary Complex Stochastic Dynamics, Y. Pantazis, M. Katsoulakis, *J. Chem. Phys.* (2013).

  Parametric Sensitivity Analysis for Biochemical Reaction Networks based on Pathwise Information Theory, M. Katsoulakis, D. Vlachos, Y. Pantazis, *BMC Bioinformatics*, (2013).

  Spatial multi-level interacting particle simulations and information theory-based error quantification, E. Kalligiannaki, M. K., P. Plechac, *SIAM Sci.Comp.*, to appear (2014).

  Goal-oriented sensitivity analysis for lattice kinetic Monte Carlo simulations, M. Katsoulakis and G. Arampatzis, *J. Chem. Phys.*, to appear, (2014).

- **Coarse-graining (information-theoretic, path-space, error analysis)**

  Information-theoretic tools for parametrized coarse-graining of non-equilibrium extended systems, M. Katsoulakis, P. Plechac, *J. Chem. Phys.* (2013).

  Coarse-graining schemes for stochastic lattice systems with short and long-range interactions, M. Katsoulakis, P. Plechac, L. Rey-Bellet and D. Tsagkarogiannis, *Math. Comp.*, to appear (2014).

  Spatial two-level interacting particle simulations and information theory-based error quantification, E. Kalligiannaki, M. K. , P. Plechac *SIAM J. Sci. Comp.*, to appear (2014).

- **Enabling tools, parallel KMC and numerical analysis**

  Hierarchical fractional-step approximations and parallel kinetic Monte Carlo algorithms, G. Arampatzis, M. Katsoulakis, P. Plechac, Michela Taufer and Lifan Xu *Journal of Computational Physics*, (2012).

  Parallelization, processor communication and error analysis in lattice kinetic Monte Carlo, G. Arampatzis, M. Katsoulakis and P. Plechac, *SIAM Numerical Analysis*, to appear (2014).