Stochastic multiscale modeling of spatially distributed biological systems

Jeremy Edwards[1], Markos Katsoulakis[2], Dion Vlachos[3], James Faeder[4]

[1]Molecular Genetics and Microbiology, Cancer Research and Treatment Center, University of New Mexico

[2]Dept. of Mathematics and Statistics, Univ. of Massachusetts

[3]Dept. of Chemical Engineering, University of Delaware

[4]Theoretical Biology and Biophys. Group, Los Alamos Nat. Lab

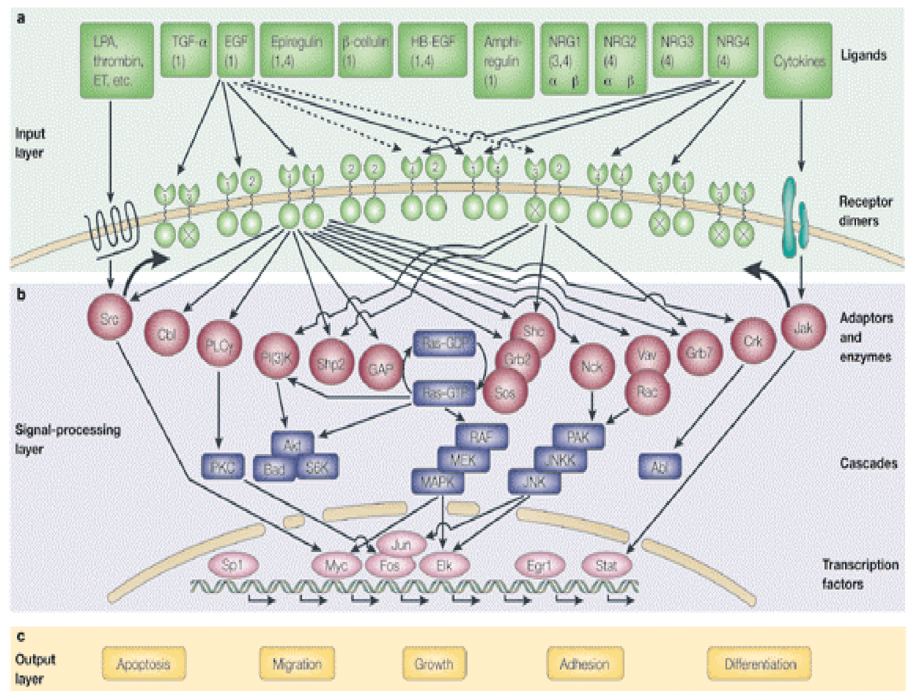
Supported by DOE (DE-FG02-05ER25702)

Outline

- Examples of multiple scales from cellular biology
  - Molecular scale and stochastic effects are crucial
- Multiscale mathematical and computational methods
  - Concepts
  - Mathematical tools
  - Numerical examples for model systems
- Comparison to real experimental data
- On-going and future work

EGF Receptor Signaling Network

- Important biological role
  - Trigger a rich network of signaling pathways and regulate functions such as proliferation, differentiation and migration.
  - Dysregulations in the signaling pathway lead to a variety of cancers: endometrial, breast, lung, prostate, colon, ovary, bladder, head and neck.
- Important as drug targets
  - EGFR is a target for anti-cancer drugs: ease of manipulation in the extracellular domain



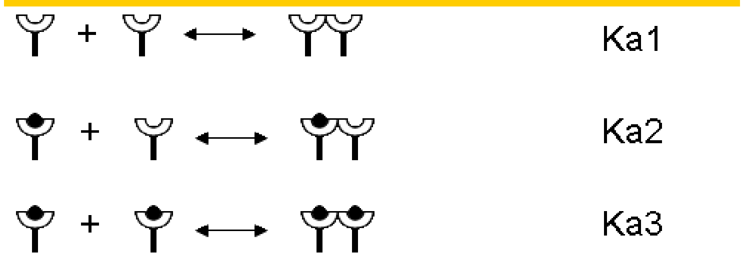
Highly simplified signaling map of some of the proteins activated by the 4-member family of EGF receptors

Figure taken from Yarden and Sliwkowski, Nature Reviews Molecular Cell Biology, 2001

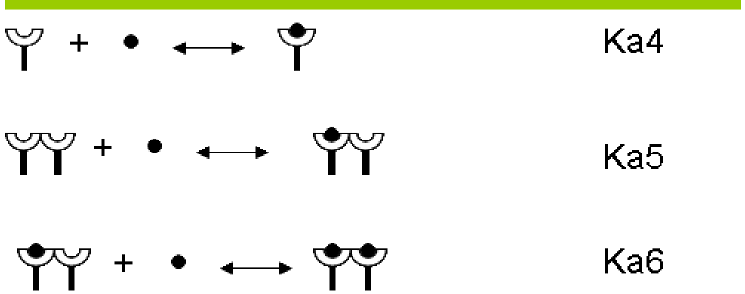
EGFR pathway

- Adsorption and desorption of ligand
- Surface diffusion
- Surface dimerization

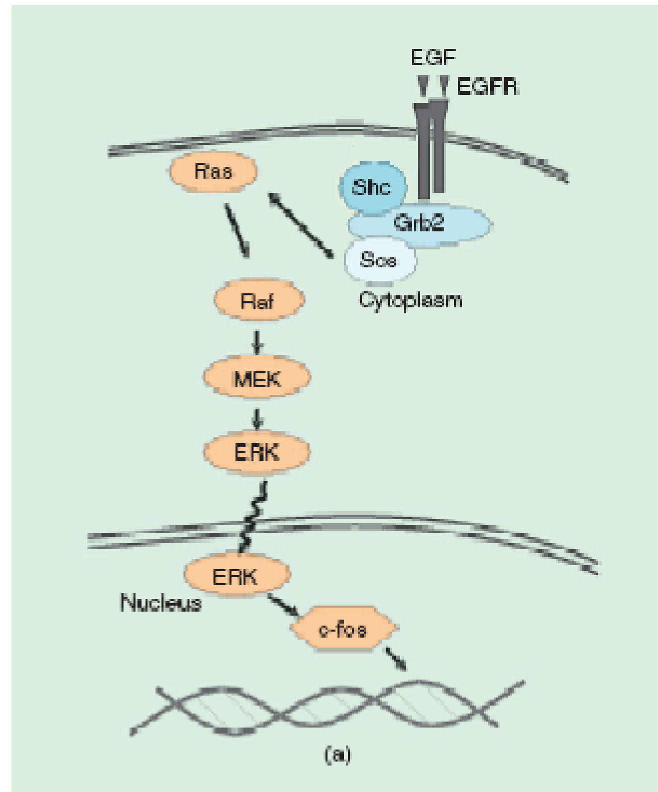
Dimerization Reactions



Ligand Binding Reactions

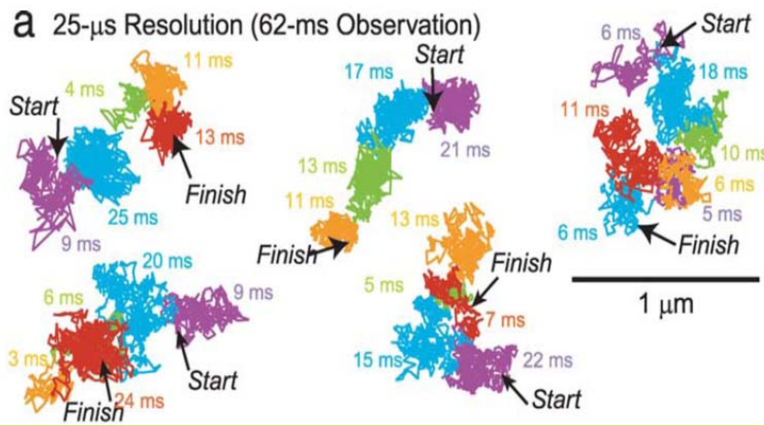


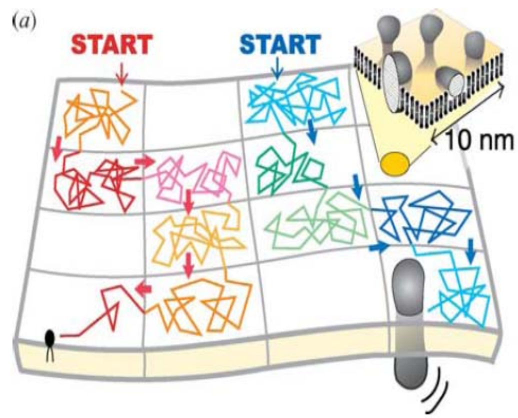
- Phosphorylation, internalization, etc.



J. Saez-Rodriguez, et al. (2004)

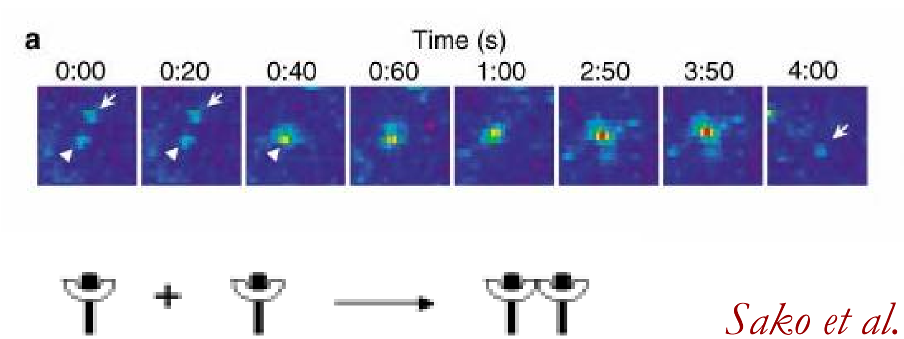
New Imaging Data at the Receptor Level





Single particle tracking (SPT) experiments track receptor movement at micro-second scale are indicating compartments in the plasma membrane

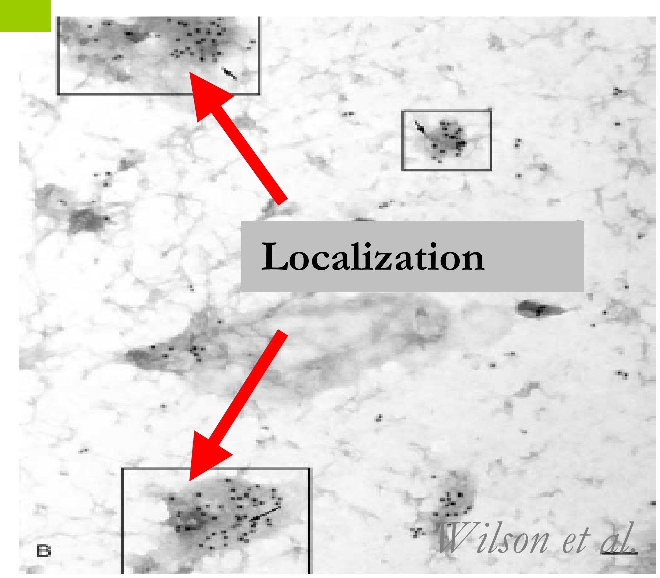
Kusumi et al.



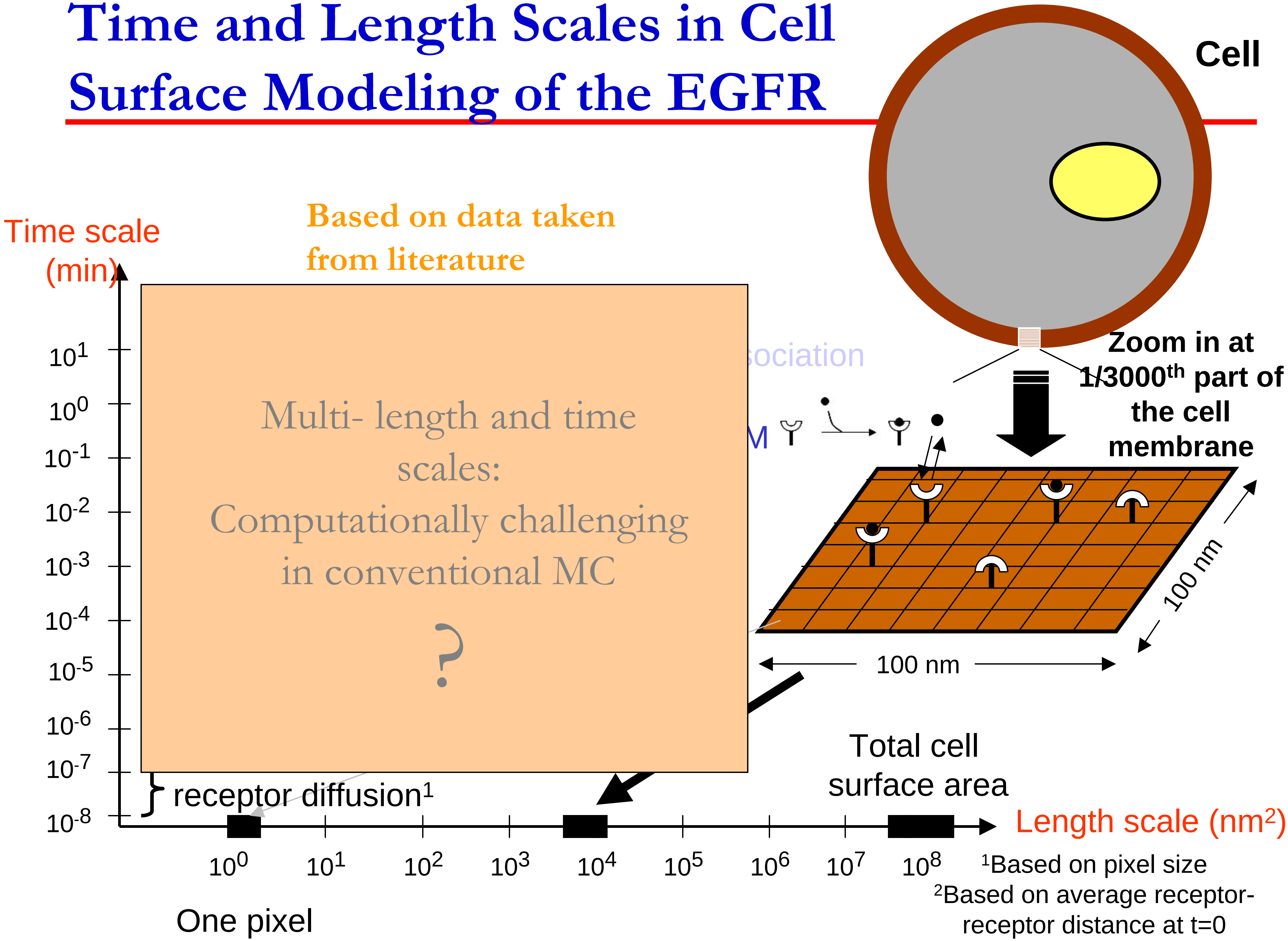
Sako et al.

SPT experiments tracking dimerization reaction events at the molecular level

Heterogeneity in receptor distribution: EM image suggesting localization of receptors



Time and Length Scales in Cell Surface Modeling of the EGFR



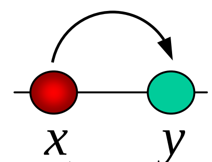
[1]Based on pixel size

[2]Based on average receptor-receptor distance at t=0

A hierarchy of coarse grained processes and KMC simulations

Microscopic Stochastic Process

Hamiltonian: $$H(\sigma) = -\frac{1}{2}\sum_{x\in M}\sum_{y\neq x} J(x-y)\sigma(x)\sigma(y) + \sum_{x\in M} h\sigma(x)$$

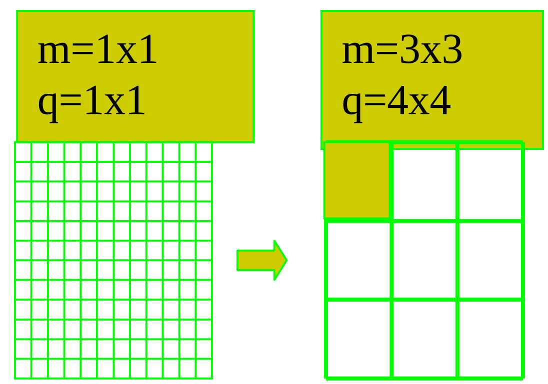


Microscopic processes/transition probabilities, e.g., surface diffusion $$\nu(x\rightarrow y,\sigma) = \nu_o\sigma(x)(1-\sigma(y))\exp[-\beta(U_o + U(x))]$$

Mesoscopic Stochastic Process

$$\eta_k = \sum_{y\in D_k}\sigma(y)$$

q=number of points in a coarse cell



$$\bar{U}_i = \sum_{k\in M_c}\bar{J}_{ik}\eta_k + \bar{J}_{00}(\eta_i - 1) - h$$

$$\bar{H}(\eta) = -\sum_{k\in M_c}\bar{U}_k\eta_k = -\frac{1}{2}\sum_{i\in M_c}\sum_{\substack{k\in M_c,\\ k\neq i}}\bar{J}_{ik}\eta_k\eta_i - \frac{\bar{J}_{00}}{2}\sum_{i\in M_c}\eta_i(\eta_i - 1) + \sum_{k\in M_c} h\eta_k$$

$$\bar{\nu}(k\rightarrow i,\eta) = \nu_o\eta_k(q-\eta_i)\exp[-\beta(U_o + \bar{U}_k)]$$

Katsoulakis et al., *PNAS* **100,** 782 (2003)

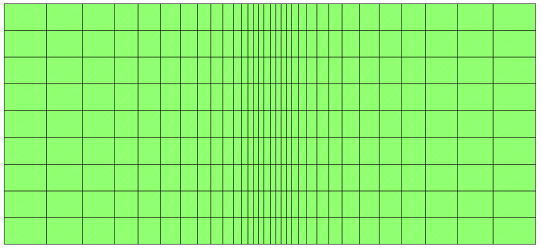
J. Comp. Phys. **186**, 250 (2003)

Features of coarse grained processes

- We never pass to any continuum limit
  - **Full hierarchy** from microscopic MC to global MF
- **Stochastic closure (local equil./LMF approximation)**
- **Detailed balance** is used as a design tool in deriving transition probabilities from the microscopics (self consistent fluctuations)
- Haar wavelet basis is used in estimating the coarse grained interactions
- Large Deviation Principles demonstrate proper **rare event** description
- **The method is exact when the potential length is infinite**
- Implementation is straightforward/**CPU savings: $q^{2\text{-}4}$**

Spatial acceleration methods

- Spatial adaptivity[1]
  - Error estimates guide mesh refinement



Non-uniform mesh

[1] Chatterjee et al., *JCP* **121**, 11420 (2004); *PRE* **71**, 0267021 (2005)

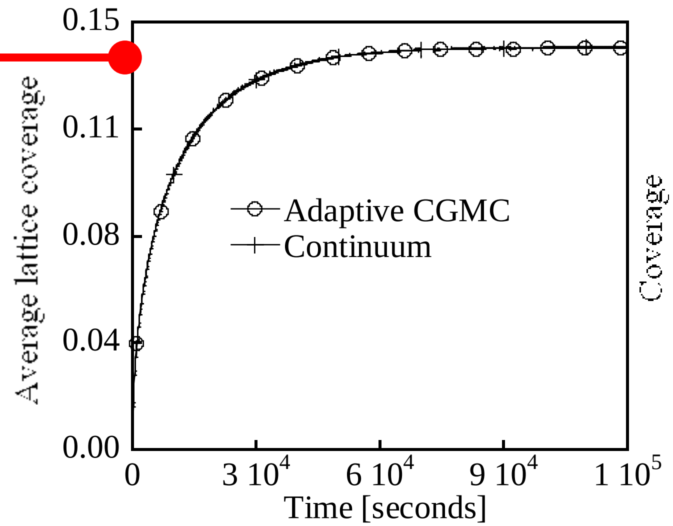
Example of diffusion-reaction in a 1 mm domain (pellet) using ACGMC

- Coarse-graining reduces computational intensity of ACGMC

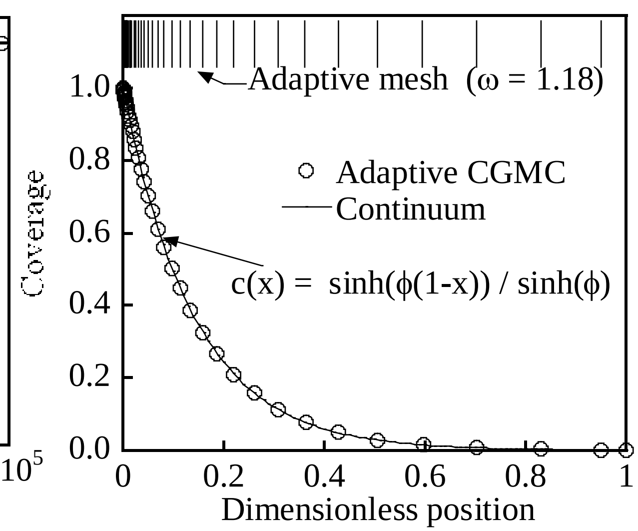
CPU time:

ACGMC: 6 h
KMC: 10^{10} y (est.)

Transient



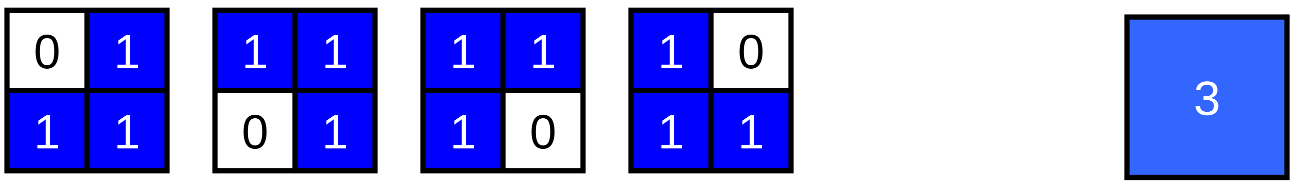
Steady state



- A posteriori error estimates are well-established in FEM. How does one optimally adapt the mesh in MC?

Information theory in ACGMC

- Approach is completely different from finite elements
- Microscopic Gibbs measure $\mu_{micro} = Z_{micro}^{-1} e^{-H_{micro}/kT} P_{micro}(\underline{\sigma})$
- Coarse-graining results in loss of information (degrees of freedom)



MICROSCOPIC (4 states)

COARSE-GRAINED (1 state)

- Adaptive CG Gibbs measure $\mu_{CG} = Z_{CG}^{-1} e^{-H_{CG}/kT} P_{CG}(\underline{\eta})$
- Relative information entropy, R
  - distance between the two distributions

$$R = \left\langle \log \frac{\mu_{CG}}{\mu_{micro}} \right\rangle$$

- Express in terms of difference of Hamiltonians

$$R = \langle \Delta H \rangle + \left\langle \log \frac{P_{CG}(\underline{\eta})}{\sum_{cell} e^{-\Delta H/kT} P_{micro}(\underline{\sigma})} \right\rangle \qquad \Delta H = H_{micro} - H_{CG}$$

Information theory in ACGMC

- Microscopic solution is unknown! Estimate upper bounds

$$R = \langle \Delta H \rangle + \left\langle \log \frac{P_{CG}(\underline{\eta})}{\sum_{cell} e^{-\Delta H / kT} P_{micro}(\underline{\sigma})} \right\rangle \rightarrow R \le c \left\langle \Lambda(\underline{\eta}) \right\rangle$$

$$\Lambda(\underline{\eta}) = \text{Upper bound}(\Delta H)$$

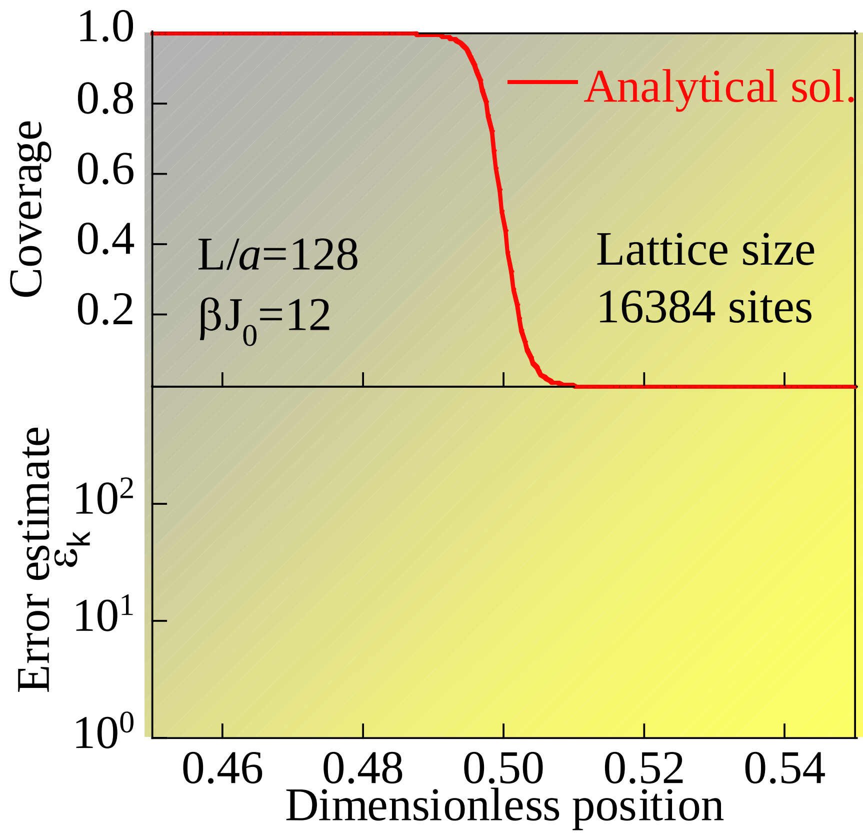
$$\left\langle \Lambda(\underline{\eta}) \right\rangle = 4 \sum_{cells,k} \left\{ \begin{array}{l} \frac{j_{kk}}{q_k(q_k - 1)} \left\langle \eta_k (q_k - \eta_k) \left[\eta_k(\eta_k - 1) + (q_k - \eta_k)(q_k - \eta_k + 1) \right] \right\rangle + \\ \sum_{int\,eracting\,cells,l} \frac{j_{kl}}{q_k q_l} \left\langle q_k{}^2 \eta_l (q_l - \eta_l) - 2 \eta_k \eta_l (q_k - \eta_k)(q_l - \eta_l) \right\rangle \end{array} \right\}$$

- Overall: $\left\langle \Lambda(\underline{\eta}) \right\rangle = \sum_{cells,k} \left\langle \varepsilon_k \right\rangle$

A posteriori **error estimate based on the computed solution only![1]**

Chatterjee et al. *Int. J. Multiscale Comp. Eng.* **3**, 59 (2005)

Use error estimates for mesh adaptation



Standing wave in 1D

- Adsorption, desorption
- Strong long-ranged attractive potential
- Continuum (analytical) solution

$$\theta(x) = \frac{1}{2}\left[\left(2\theta_+ - 1\right)\tanh\left(\beta J_0\left(2\theta_+ - 1\right)x\right) + 1\right]$$

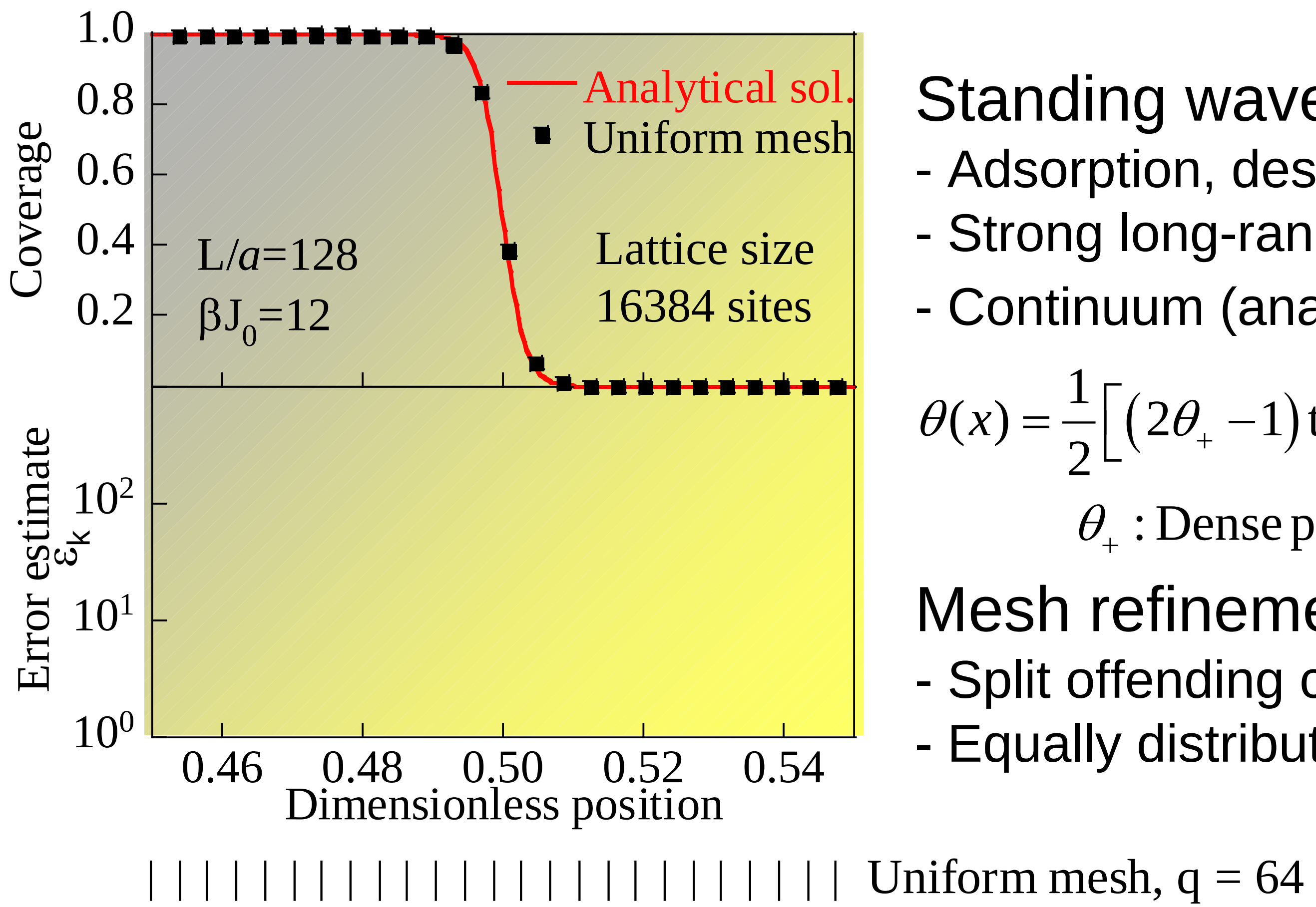
θ_+ : Dense phase coverage

Mesh refinement

- Split offending coarse cells into two
- Equally distribute order parameters

Chatterjee et al., *Phys. Rev. E* **71**, 0267021 (2005)

Use error estimates for mesh adaptation



Standing wave in 1D

- Adsorption, desorption
- Strong long-ranged attractive potential
- Continuum (analytical) solution

$$\theta(x) = \frac{1}{2}\left[\left(2\theta_+ - 1\right)\tanh\left(\beta J_0\left(2\theta_+ - 1\right)x\right) + 1\right]$$

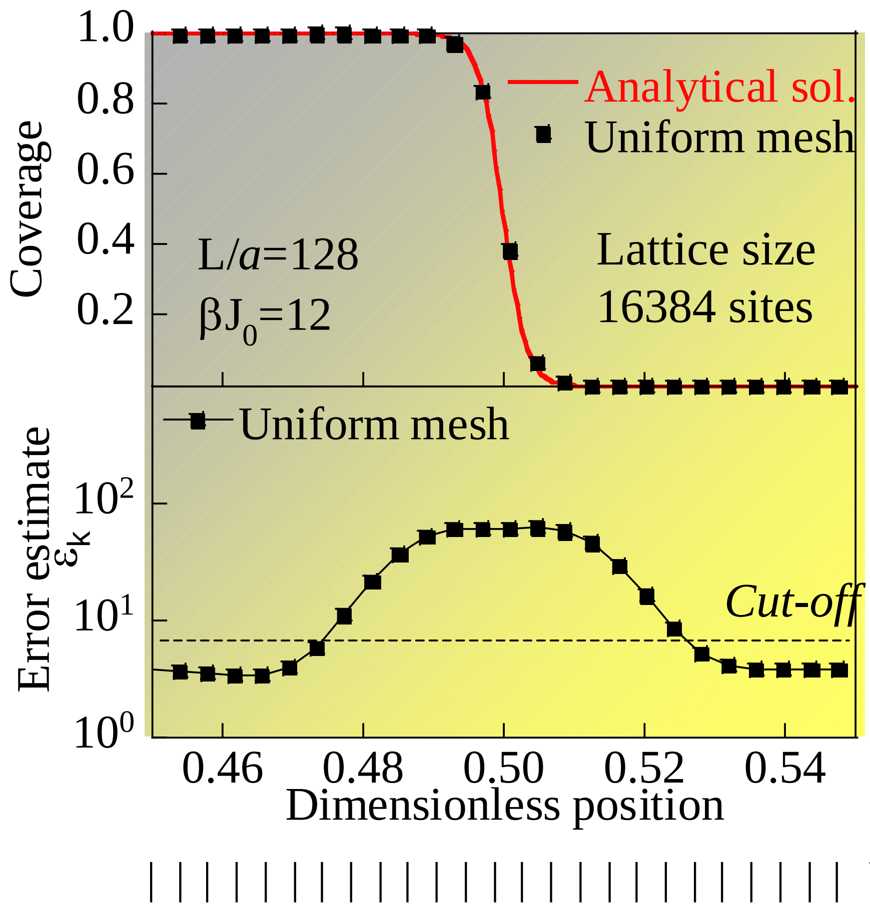
θ_+ : Dense phase concentration

Mesh refinement

- Split offending coarse cells into two
- Equally distribute order parameters

Chatterjee et al., *Phys. Rev. E* **71**, 0267021 (2005)

Use error estimates for mesh adaptation



Uniform mesh, q = 64

Standing wave in 1D

- Adsorption, desorption
- Strong long-ranged attractive potential
- Continuum (analytical) solution

$$\theta(x) = \frac{1}{2}\left[\left(2\theta_+ - 1\right)\tanh\left(\beta J_0\left(2\theta_+ - 1\right)x\right) + 1\right]$$

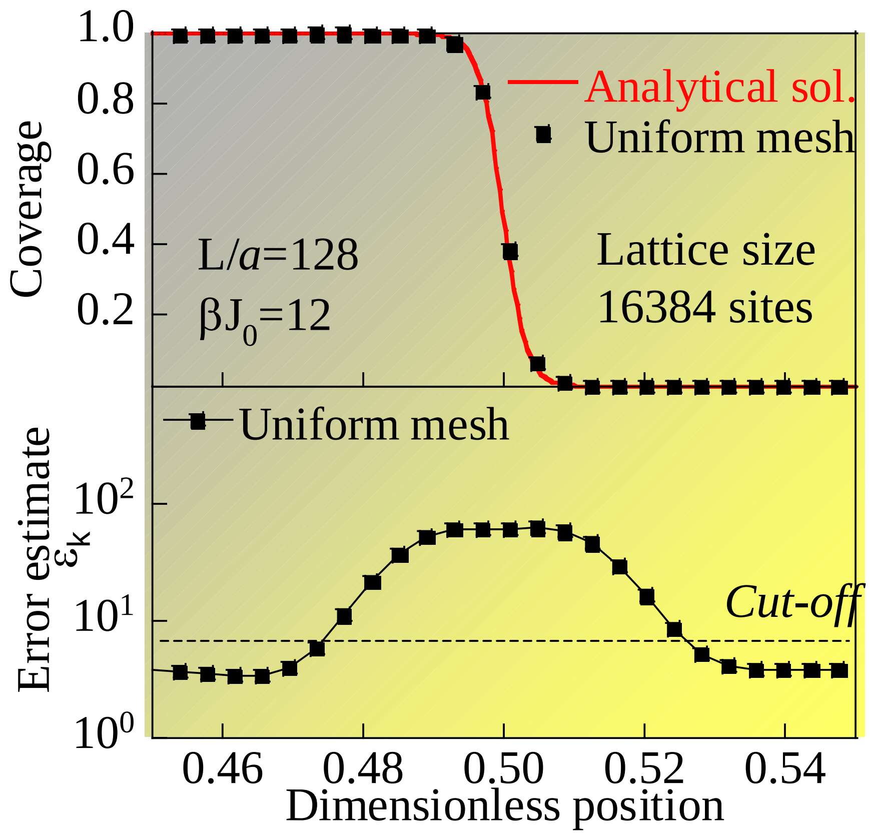
θ_+ : Dense phase concentration

Mesh refinement

- Split offending coarse cells into two
- Equally distribute order parameters

Chatterjee et al., *Phys. Rev. E* **71**, 0267021 (2005)

Use error estimates for mesh adaptation



Standing wave in 1D

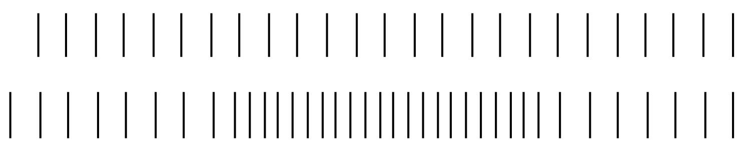
- Adsorption, desorption
- Strong long-ranged attractive potential
- Continuum (analytical) solution

$$\theta(x)=\frac{1}{2}\Big[\big(2\theta_+-1\big)\tanh\Big(\beta J_0\big(2\theta_+-1\big)x\Big)+1\Big]$$

θ_+ : Dense phase concentration

Mesh refinement

- Split offending coarse cells into two
- Equally distribute order parameters

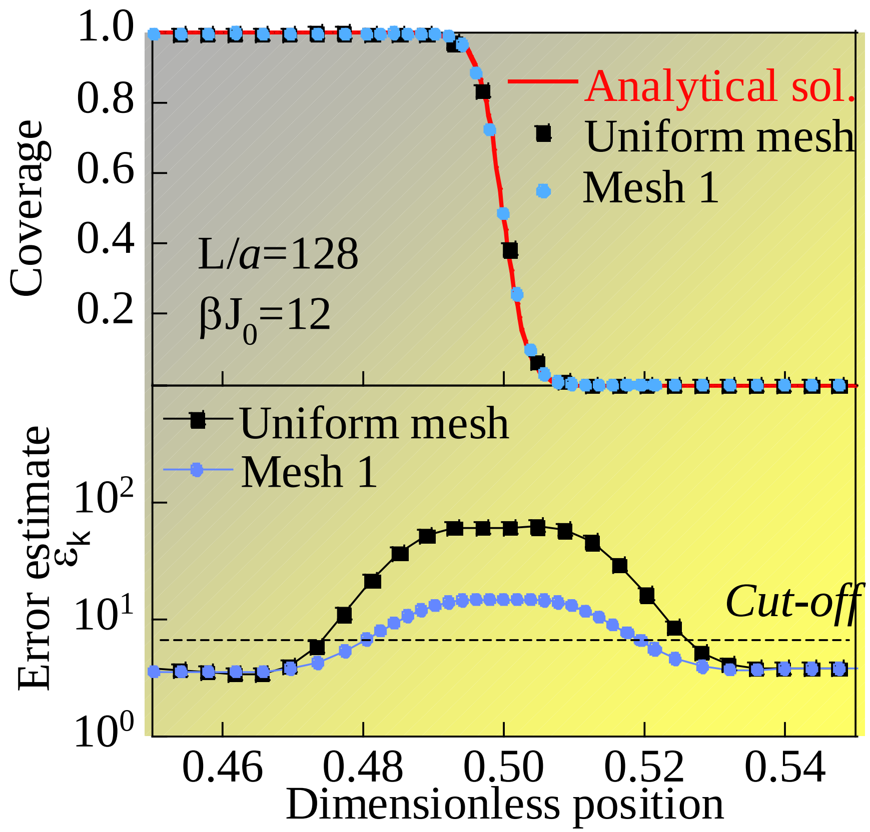


Uniform mesh, q = 64

Mesh 1

Chatterjee et al., *Phys. Rev. E* **71**, 0267021 (2005)

Use error estimates for mesh adaptation



Standing wave in 1D

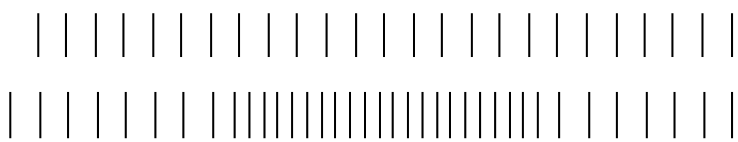
- Adsorption, desorption
- Strong long-ranged attractive potential
- Continuum (analytical) solution

$$\theta(x) = \frac{1}{2}\left[\left(2\theta_+ - 1\right)\tanh\left(\beta J_0\left(2\theta_+ - 1\right)x\right) + 1\right]$$

θ_+ : Dense phase concentration

Mesh refinement

- Split coarse cells into two
- Equally distribute order parameters

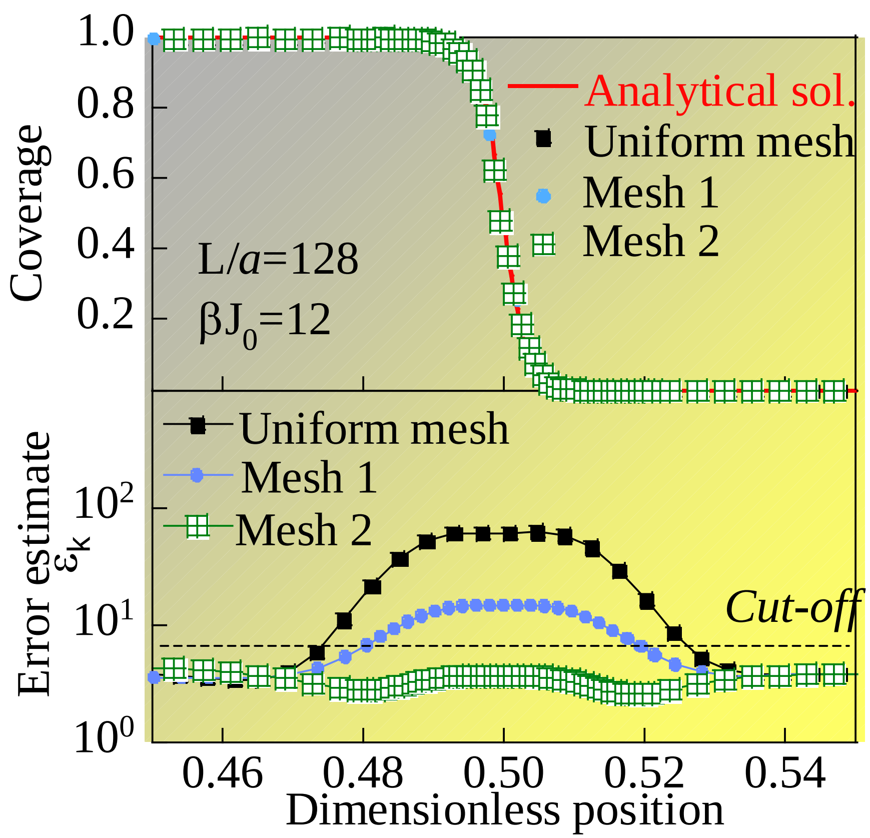


Uniform mesh, $q = 64$

Mesh 1

Chatterjee et al., *Phys. Rev. E* **71**, 0267021 (2005)

Use error estimates for mesh adaptation



Standing wave in 1D

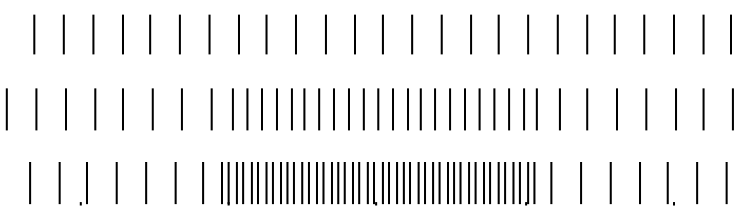
- Adsorption, desorption
- Strong long-ranged attractive potential
- Continuum (analytical) solution

$$\theta(x) = \frac{1}{2}\left[\left(2\theta_+ - 1\right)\tanh\left(\beta J_0\left(2\theta_+ - 1\right)x\right) + 1\right]$$

θ_+ : Dense phase concentration

Mesh refinement

- Split coarse cells into two
- Equally distribute order parameters



Uniform mesh, q = 64

Mesh 1

Mesh 2

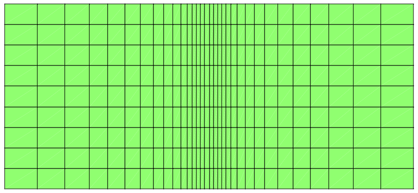
ACGMC: 30 minutes

KMC: 90 days (est.)

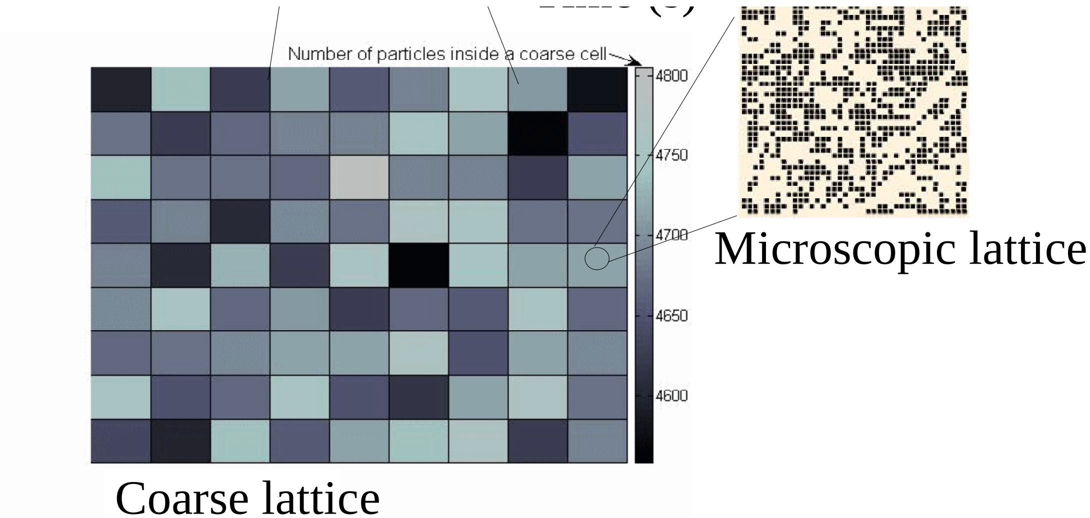
Chatterjee et al., *Phys. Rev. E* **71**, 0267021 (2005)

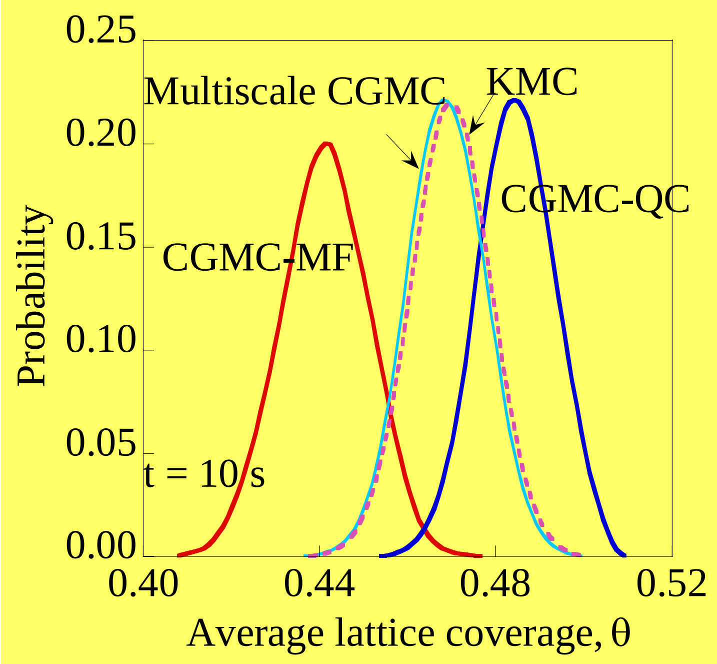
Spatial acceleration methods

- Spatial adaptivity[1]
  - Error estimates guide mesh refinement
- Multiscale MC methods for high accuracy
  - Higher order closures[2]
  - Multigrid/Relaxation criteria[2]
  - Rigorous cluster expansions[3]



Non-uniform mesh





[1] Chatterjee et al., *JCP* **121**, 11420 (2004); *PRE* **71**, 0267021 (2005)

[2] Chatterjee and Vlachos, JCP **124**, 0641101 (2006)

[3] Katsoulakis, Plechac, et al., ESAIM, Math Model. Num. Anal., to appear

Higher-order coarse-grained MC: Rigorous cluster expansions

- Corrections around $\bar{H}^o$ from prior work of Katsoulakis and Vlachos: Renormalization Group Map

$$H^c(\eta) = \bar{H}^o(\eta) - \frac{1}{\beta}\log E[e^{-\beta(H_N - \bar{H}^o)} |_\eta]$$

- Heuristics (expansion of exp and log)

$$= \quad E[\Delta H |_\eta] + E[(\Delta H)^2 |_\eta] - E[\Delta H |_\eta]^2 + O((\Delta H)^3)$$

  – Formal calculations inadequate since: $\Delta H = H_N - \bar{H}^o = N \cdot O(\varepsilon)$

  – Rigorous analysis: Cluster expansion around $\bar{H}^o$

[3] Katsoulakis, Plechac, Rey-Bellet, Tsagkarogiannis, ESAIM, Math Model. Num. Anal., to appear

Higher-order coarse-grained MC: Rigorous cluster expansions – Cont.

- Corrections to the Hamiltonian – **many body** terms

$$H^c(\eta) = \bar{H}^o(\eta) + \bar{H}^1(\eta) + \ldots$$

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

$$E_r(k) = E_r(\eta(k)) = (2\eta(k)/q - 1)^r + O_q(1)$$

- Moments of the potential

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

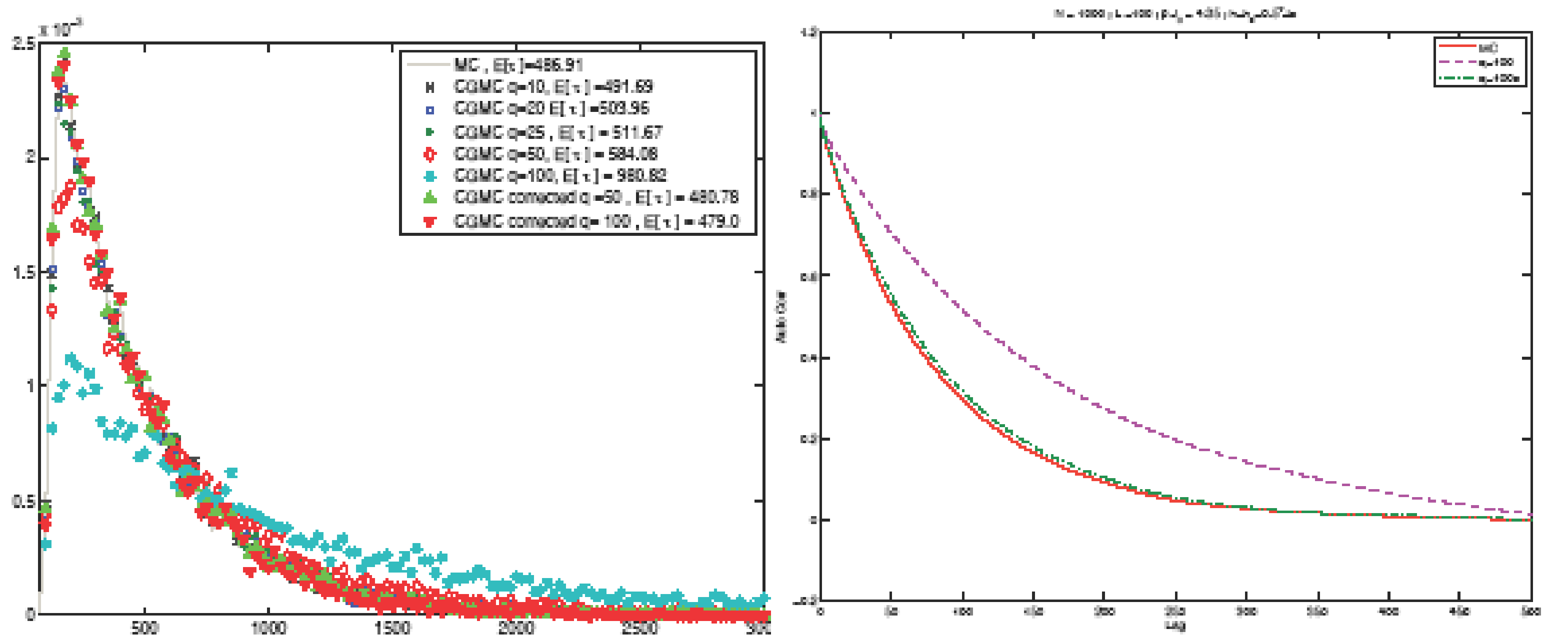
- **Clusters expansions give**
  - Sharp a posteriori error estimates/Adaptivity
  - Higher-order CGMC schemes

A numerical example

- Phase transitions between two states: PDF of switching times

Demonstration: Metastability for **CG Arrhenius dynamics**

Switching Time PDFs/Autocorr.: *with and w/o corrections*



joint work with Sasanka Are (UMass)

Error quantification in CG schemes

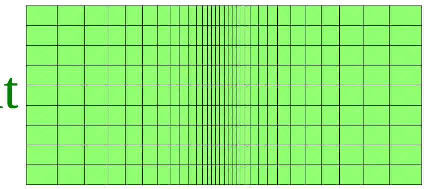
- **Theorem**: ***A priori*** error estimate
  - Define a small parameter: $\varepsilon \equiv C\beta \frac{q}{L} \| J' \|_1$
  - Then, specific relative entropy: $R(\mu_{M,q,\beta} | \mu_{N,\beta} \circ T^{-1}) = O(\varepsilon^{\alpha+2})$

    α=order of truncation in cluster expansion
  - $T\sigma$=projection on coarse variables = $\sum_{y \in D_k} \sigma(y)$
- Error estimates for observables – **'quantity of interest'**[4]
  - Difficulty in dynamics: $T\sigma$ is not a Markov process
- **Reverse CG map** – Microscopic reconstruction possible[4]
  - Reverse the CG via the conditional prior
  - Numerical error estimate for the reconstruction

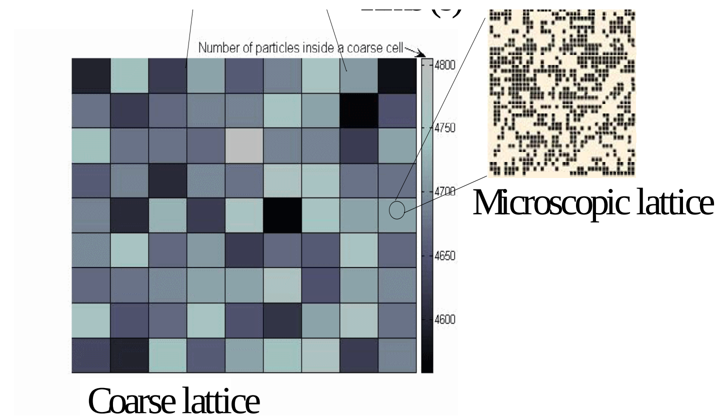
[4] Katsoulakis, Plechac, Sopasakis, SIAM Num. Anal. (2006)

Spatial acceleration methods

- Spatial adaptivity[1]
  - A posteriori error estimates guide mesh refinement and generation of phase diagrams
- Multiscale MC methods for high accuracy
  - Higher order closures[2]
  - Multigrid/Relaxation criteria[2]
  - Rigorous cluster expansions/Theory[3,4]
- Multicomponent interacting systems[5]
- Time scale acceleration



Non-uniform mesh



[1] Chatterjee et al., *JCP* **121**, 11420 (2004); *PRE* **71**, 0267021 (2005); Katsoulakis, Plechac, et al., J. Non Newtonian Fluid Mech. (2007)

[2] Chatterjee and Vlachos, JCP **124**, 0641101 (2006)

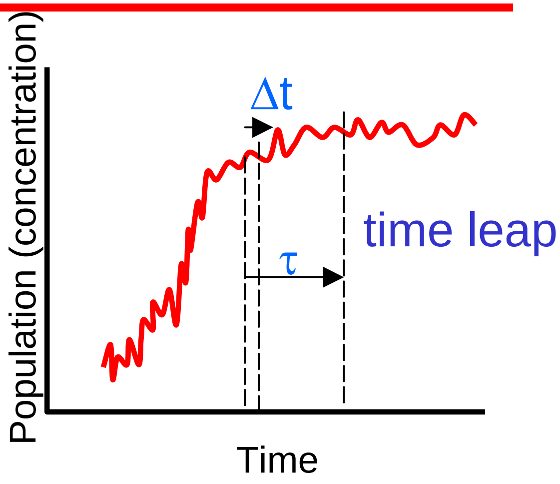
[3] Katsoulakis, Plechac, et al., *ESAIM, Math Model. Num. Anal.*, to appear

[4] Katsoulakis, Plechac, Sopasakis, SIAM Num. Anal. (2006)

[5] Chatterjee and Vlachos, JCP, accepted

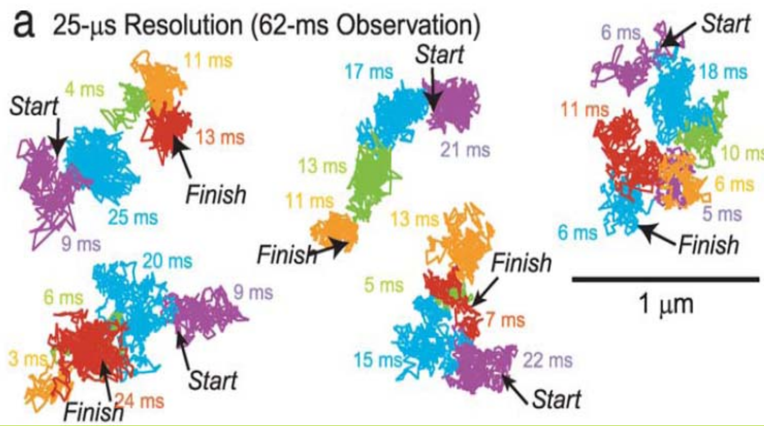
Time scales acceleration methods

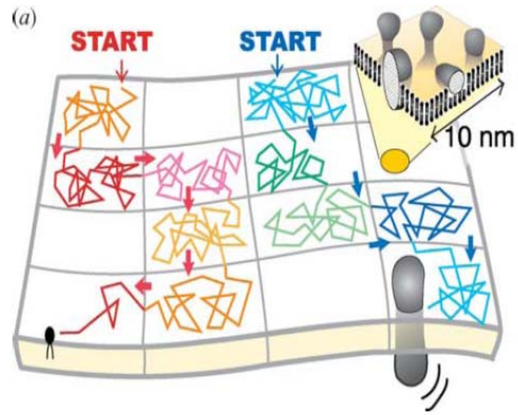
- Binomial τ-leap method: fire multiple events[6,7]
  - Time step increment based on stability
- Stochastic low dimensional manifold (SLDM): overcome stiffness
  - Computational singular perturbation (CSP) assisted partitioning[8]
  - Statistical criteria for convergence to the LDM[8,9]
- Hybrid multiscale Monte Carlo method[9]
  - Deal with simultaneous separation of time scales and populations
- *Incorporate these algorithms in spatial MC[10]*



[6] Chatterjee et al., *J. Chem. Phys.* **122**, 024112 (2005)
[7] Chatterjee et al., *Bioinformatics* **21**(9), 2136 (2005)
[8] Samant and Vlachos, *J. Chem. Phys.* **123**, 144114 (2005)
[9] Samant et al. BCM Bioinformatics, accepted
[10] Chatterjee and Vlachos, *J. Comp. Phys.* **211**, 596 (2006)

New Imaging Data at the Receptor Level

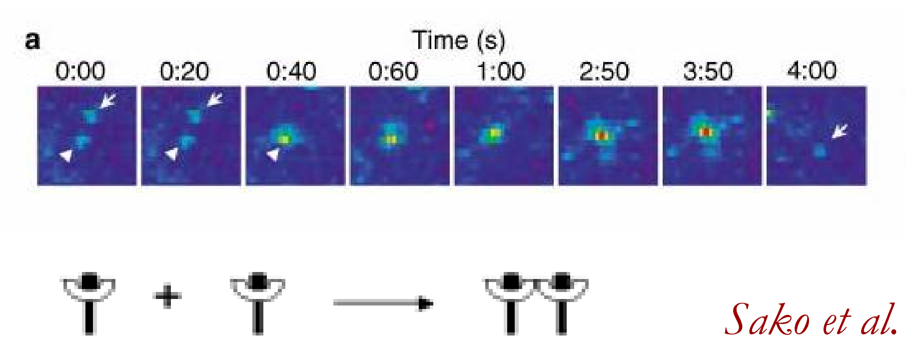




Single particle tracking (SPT) experiments track receptor movement at micro-second scale are indicating compartments in the plasma membrane

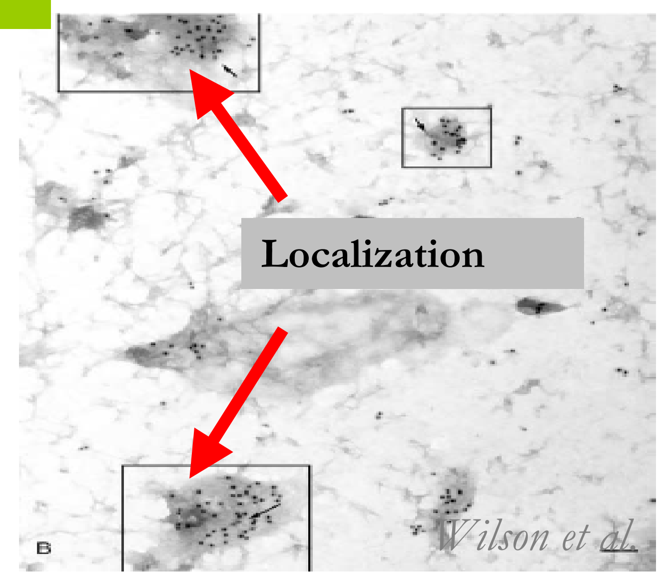
Kusumi et al.

Heterogeneity in receptor distribution: EM image suggesting localization of receptors



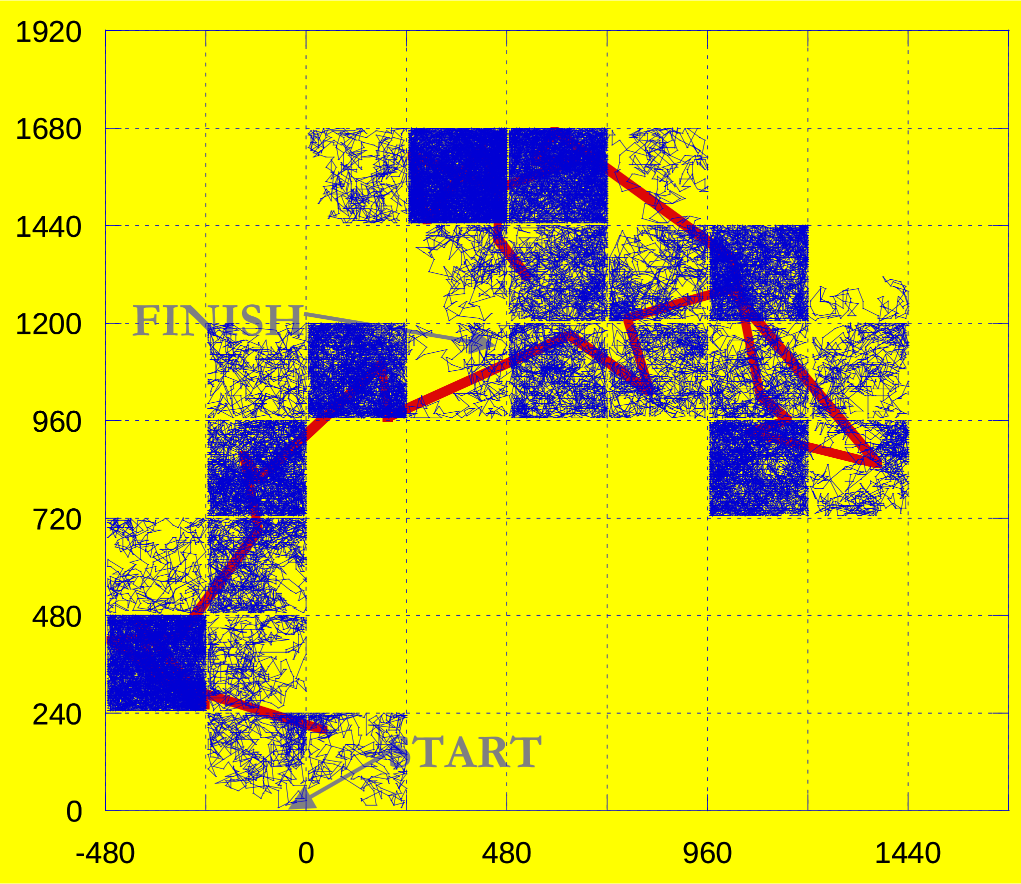
Sako et al.

SPT experiments tracking dimerization reaction events at the molecular level



Single Particle Simulation Reveals Hop Diffusion

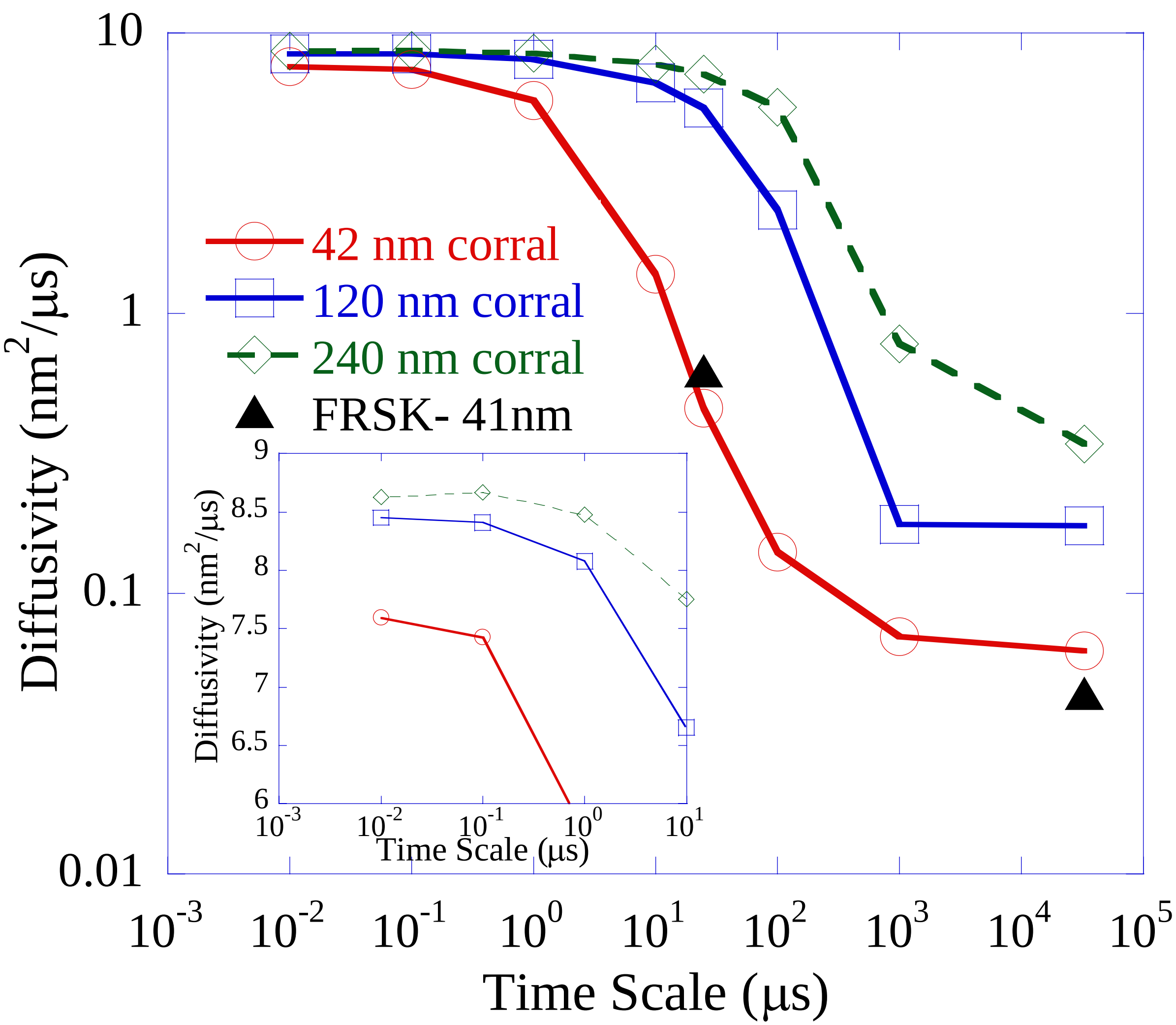
- At a low frame rate (33 ms), simple Brownian Motion.
- At a high frame rate (25 µs), corrals trapping receptors.



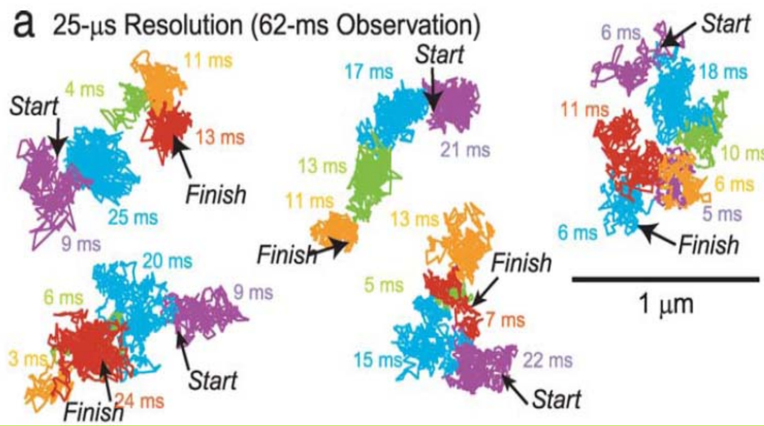
Position in Nanometers

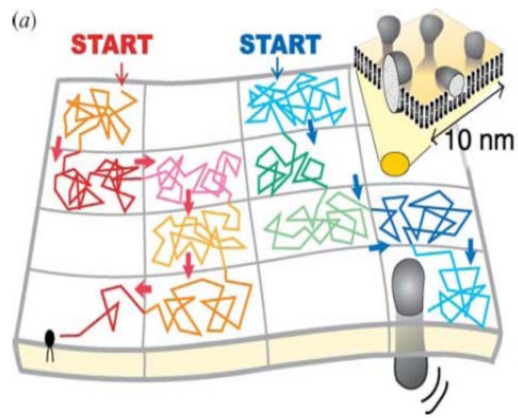
Effect of Sampling Time

- D depends on time resolution of data collection
- Good agreement with data
- Kusumi's data give macro D and a transition D
- At short times, receptors are unaffected by observation time but affected slightly by corral size
  - D can be lower than the intrinsic value!



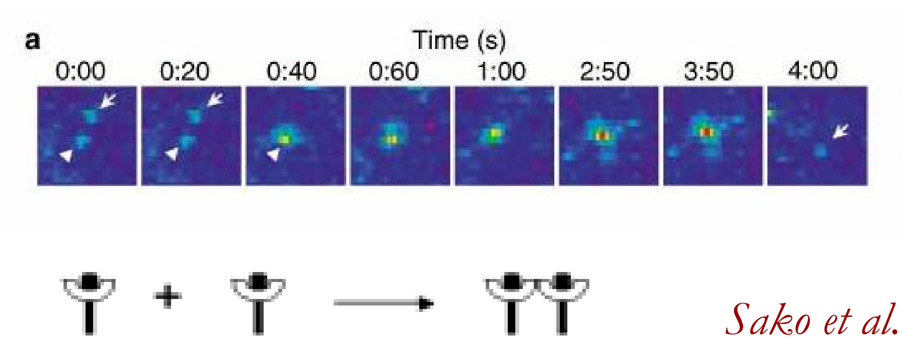
Imaging Data at the Receptor Level





Single particle tracking (SPT) experiments track receptor movement at micro-second scale are indicating compartments in the plasma membrane

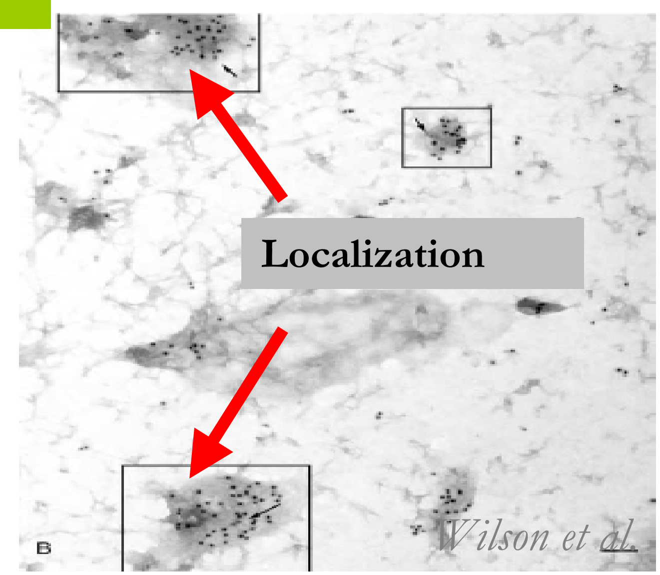
Kusumi et al.



Sako et al.

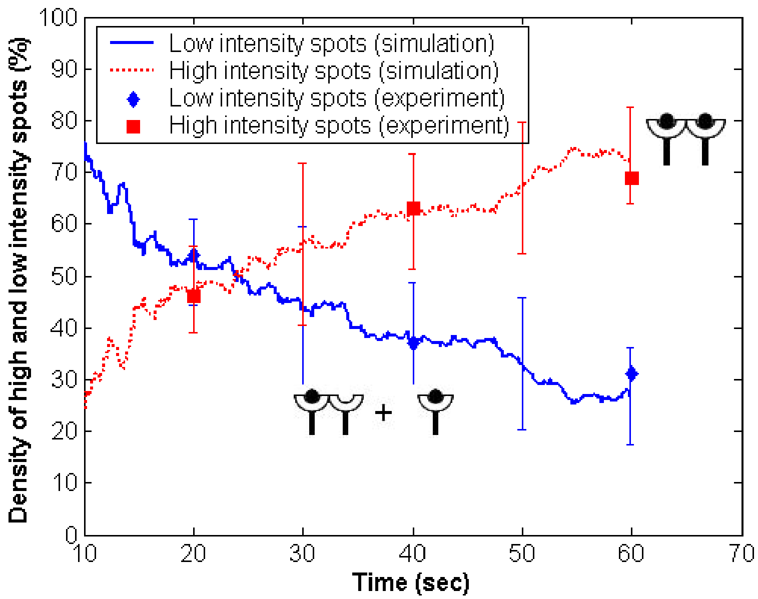
SPT experiments tracking dimerization reaction events at the molecular level

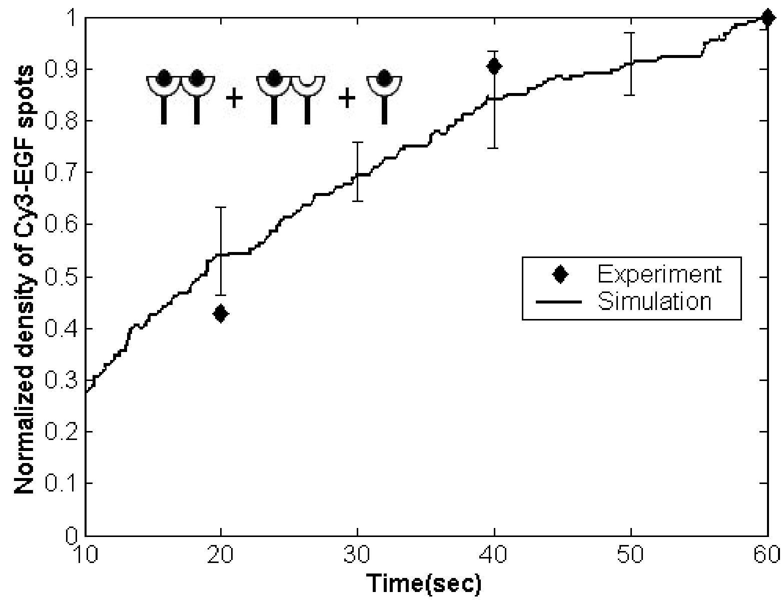
Heterogeneity in receptor distribution: EM image suggesting localization of receptors



Wilson et al.

Comparison of MC with Single Particle Tracking Experiment Data

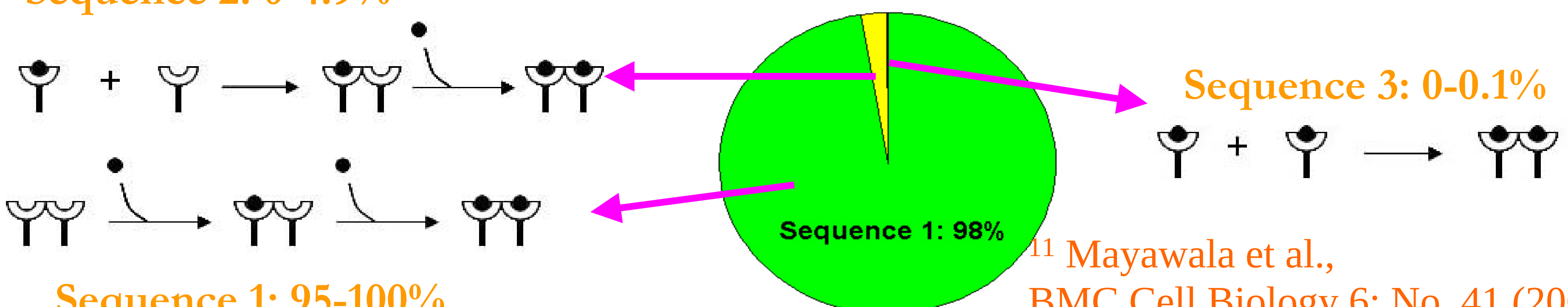




Points: Experimental data

Curves and error bars: Simulation results

Sequence 2: 0-4.9%



Sequence 1: 95-100%

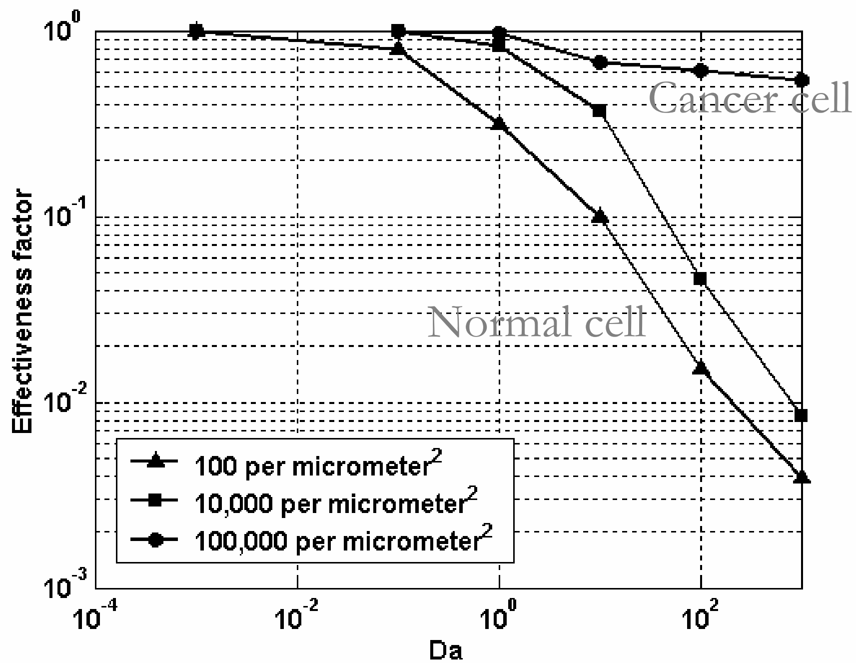
[11] Mayawala et al., BMC Cell Biology 6: No. 41 (2005)

Effect of Localization on EGFR Dimerization Rate

- EGFR Diffusivity: **10^{-9}-10^{-11}** cm^2/s.

  Range is likely due to presence of membrane microdomains (lipid rafts).
- Localization is unlikely to cause a significant increase in EGFR dimerization rates in normal cells.
- in cancer cells, localization leads to 1-2 orders of magnitude increase in rate!

[12] Mayawala et al., *Biophys. Chem.* **121**, 194–208 (2006)

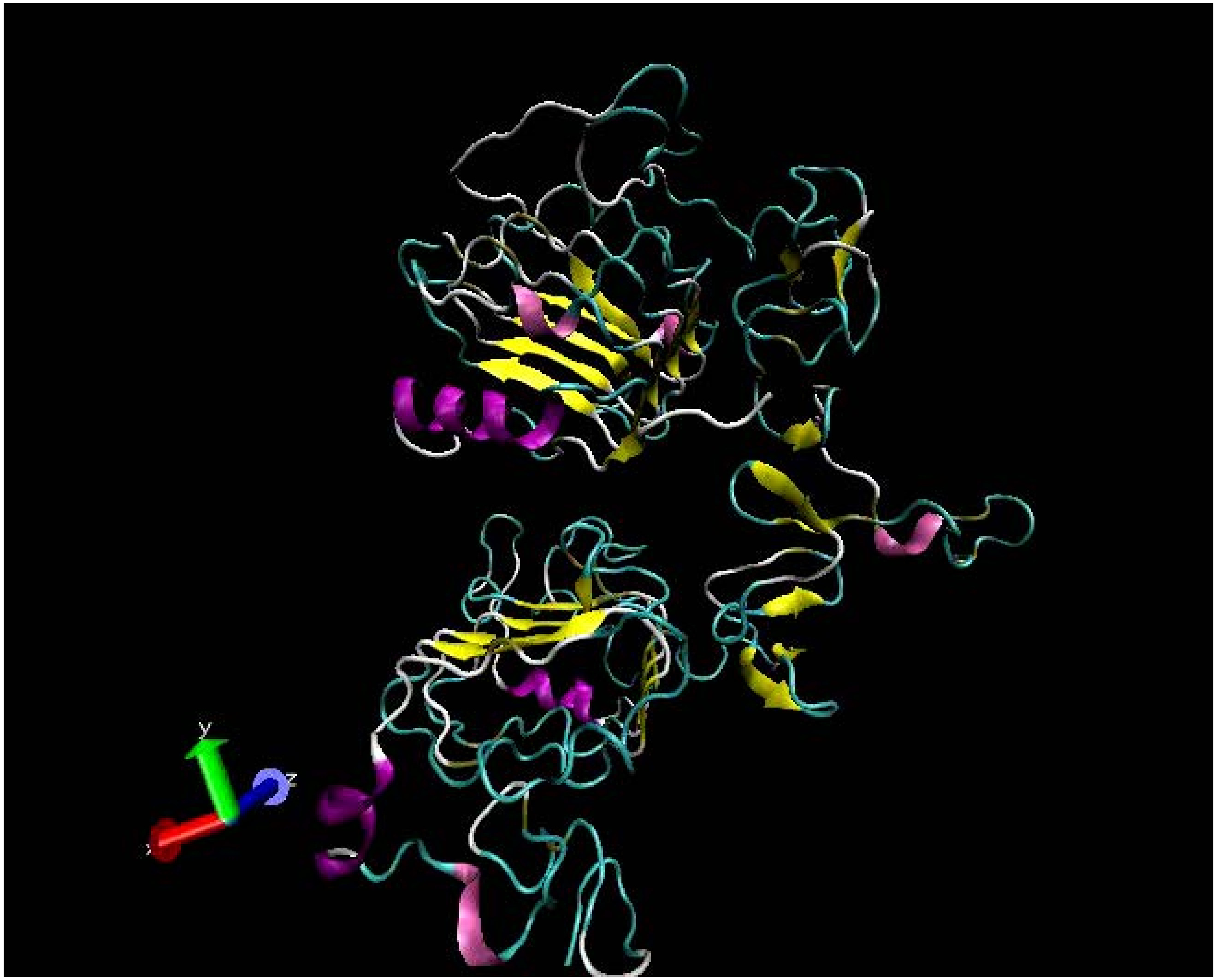


Y-axis Effectiveness factor: Measure of reduction in dimerization rate due to diffusion limitation.

X-axis Damköhler number, Da: Ratio of time scales of diffusion and reaction.

100 EGFR/μm^2 :	Avg. density in normal cells
10,000 EGFR/μm^2 :	Avg. density in cancer cells
or	Localization in normal cells
100,000 EGFR/μm^2 :	Localization in cancer cells

Molecular dynamics of extra-cellular part of EGFR



Summary of accomplishments

- Develop a multiscale stochastic framework for spatiotemporal dynamics
  - Adaptivity
  - *A posteriori* error estimates
  - Cluster expansion-based corrections
  - Multigriding
  - Temporal acceleration
- Modeled real biological systems
  - Good agreement with experimental data
  - Plasma membrane heterogeneity and localization can lead to substantial increase in dimerization rates in cancer cells
- The framework applies to many other areas
  - Growth of nanomaterials
  - Separations (e.g., hydrogen production)
  - Distributed and portable energy production (e.g., alternative fuels, biofuels)

Acknowledgements

- Collaborators
  - Petr Plechac (UTN/ORNL)
  - Luc Rey-Bellet (UMass)
  - Andy Majda (Courant)
  - Dimitri Tsagkarogiannis (Max Planck)
  - Jose Trashorras (Paris IX)
  - Anders Szepessy (KTH/Sweden)
- Students and postdocs (UD, UMass, UNM)
  - Kapil Mayawala
  - Anne-Marie Niehaus
  - Abhijit Chatterjee
  - Alex Sopasakis
  - Altaf Karim
  - Michelle Costa
  - Sasanka Are