Mathematical and Computational Methods for Stochastic/Deterministic Hybrid Systems

Lecture 1

Markos A. Katsoulakis

University of Massachusetts, Amherst

Some hybrid deterministic/stochastic systems

1. Microscopically active interface or boundary layer interacting with an adjacent "bulk" fluid phase.

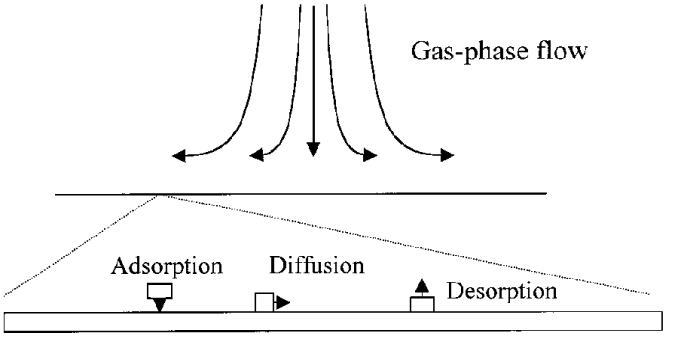
2. Rheology of polymers: *micro-macro* models.

Fluids equations at the macroscopic level coupled with kinetic or stochastic equations ruling the evolution of the fluid microstructure at the meso- or micro- scale, e.g. FENE-type models or coupled Monte Carlo with fluid dynamics.

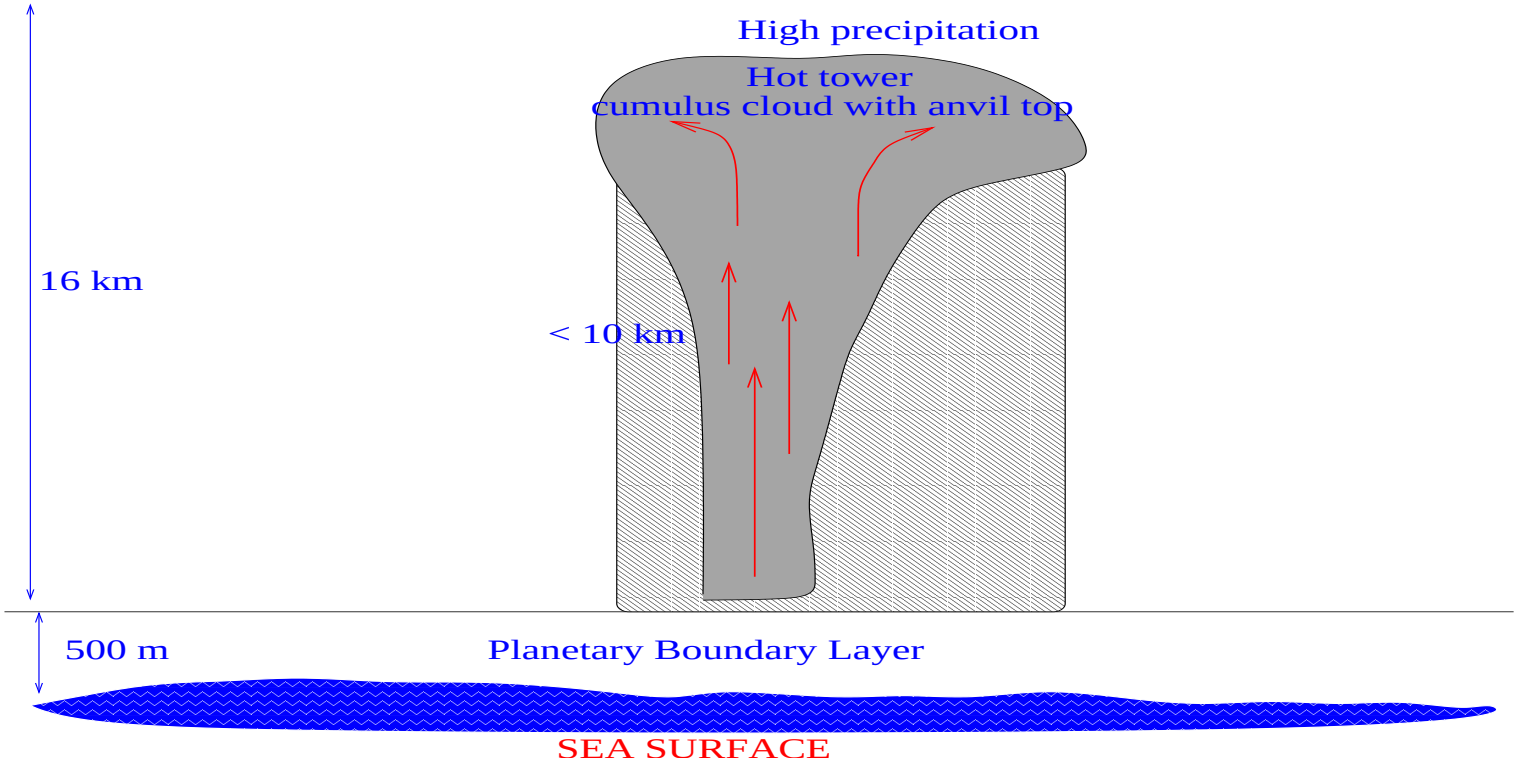
3. Stochastic Phase-Field models.

Solidification, dendritic growth in alloys.

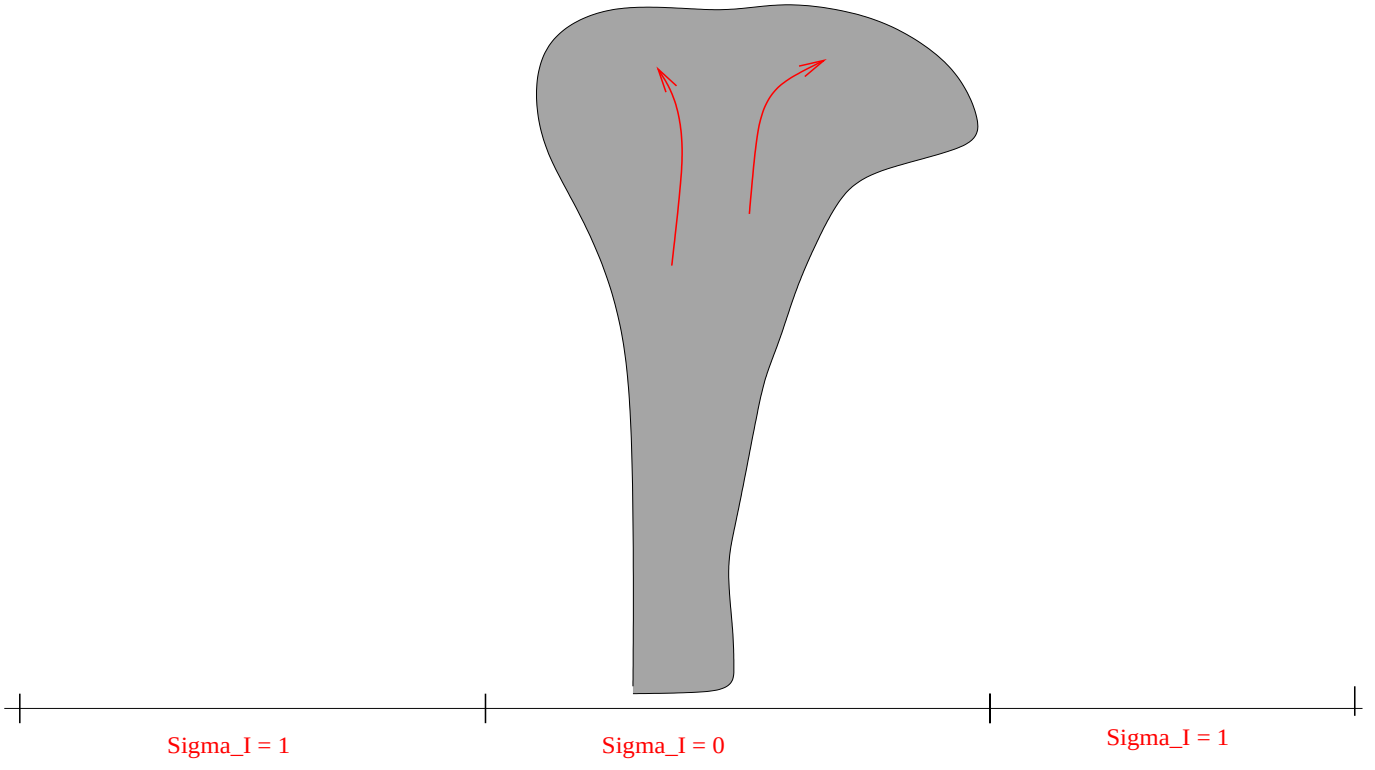
Surface processes: Catalysis, Chemical Vapor Deposition, epitaxial growth, etc.



Atmosphere/Ocean applications: Tropical convection.

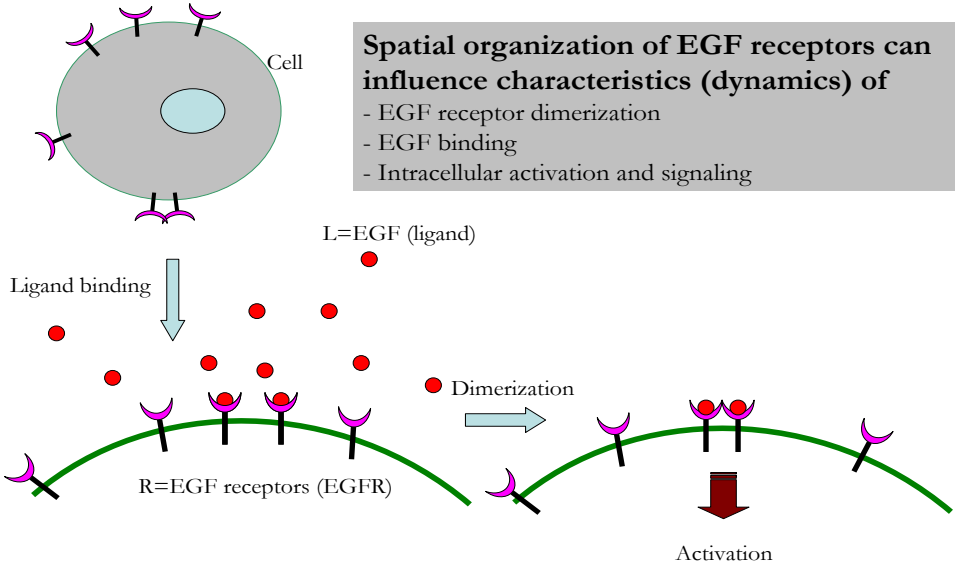


"Particles" and sub-grid scale effects: [Majda, Khouider, PNAS 2001]



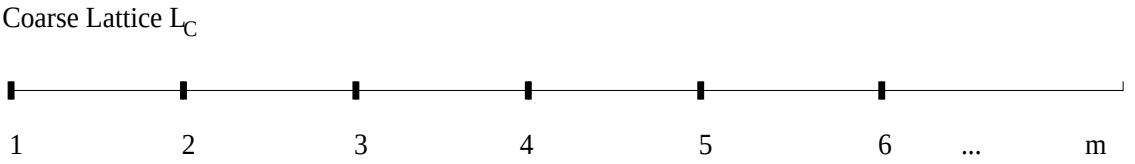
Cell Biology: Epidermal Growth Factor binding/dimerization

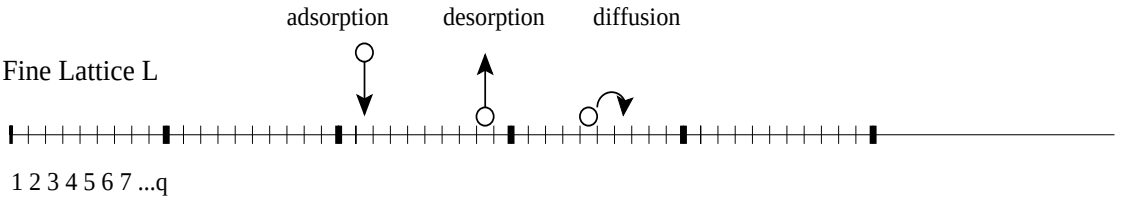
Early events of EGF signaling



Spatial organization of EGF receptors can influence characteristics (dynamics) of

- EGF receptor dimerization
- EGF binding
- Intracellular activation and signaling





$$\partial_t X = F[X, \sigma] \quad \text{(PDE/ODE system)}$$

$$\partial_t E f(\sigma) = E L_X f(\sigma) \quad \text{(stochastic model)}$$

X: Fluid/thermodynamic variables defined on top grid

L_X: generator of the subgrid stochastic process σ defined on the lower grid (subgrid)

Some challenges and questions:

- Disparity in scales and models; DNS require ensemble averages for a large system.
- Model reduction, however no clear scale separation: need hierarchical coarse-graining.
- Deterministic vs. stochastic closures; when is stochasticity important?
- Error control, stability of the hybrid algorithm; efficient allocation of computational resources: adaptivity, model and mesh refinement.

MODEL SYSTEM

$$\partial_t X = f(X, \bar{\sigma}) \quad \text{(ODE)}$$

$$\partial_t E f(\sigma) = E L_X f(\sigma) \quad \text{(stochastic lattice model)}$$

L_X: generator of a spatial stochastic process $\sigma_t(x)$.

$f(x) = f(x, \bar{\sigma})$: scalar bistable, saddle node, or spatially homogeneous complex Ginzburg-Landau equation (Hopf bifurcations), etc.

- $h = h(X)$: external field to the microscopic system.
- $\bar{\sigma} = \frac{1}{N} \sum_x \sigma_t(x)$: area fraction (spatial average).

Special case: well-mixed, catalytic reactors (CSTR)

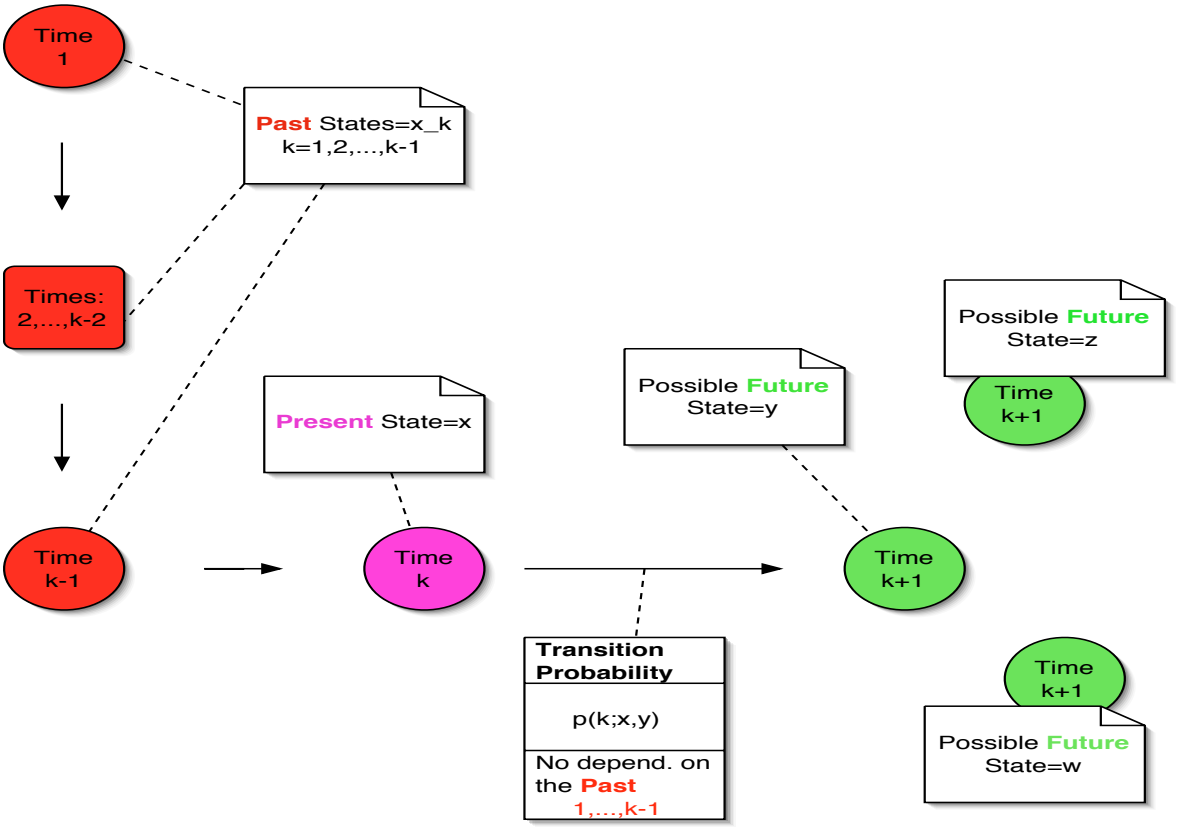
[joint work with A. Majda (Courant) and A. Sopasakis (UMass)]

I. Background material on Markov processes and Monte Carlo simulation

1. Discrete-time Markov Chains: "Stochastic, discrete-time dynamical systems"

Stochastic process $\{X_t : t = 1, 2, 3, ..., n, ...\}$, $t \sim$ equi-spaced time intervals; finite state space Σ:

$$X_t = x \in \Sigma = \{1, 2, ..., m\}$$


Time
1
Past States=x_k
k=1,2,...,k-1
Times:
2,...,k-2
Present State=x
Possible Future
State=y
Possible Future
State=z
Time
k+1
Time
k-1
Time
k
Time
k+1
Transition
Probability
p(k;x,y)
No depend. on
the Past
1,...,k-1
Time
k+1
Possible Future
State=w

Conditional Probablity: (dependent random variables)

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Markov property:

$$P(X_{t+1} = y \,|\, X_t = x, X_{t-1} = x_{t-1}, ..., X_1 = x_1) = P(X_{t+1} = y \,|\, X_t = x)$$

Transition probability matrix: $P = \{p(x,y)\}_{x,y \in \Sigma}$,

$$p(x,y) := P(X_{t+1} = y \,|\, X_t = x)$$

Note: $p(x,y) \geq 0$ and $\sum_{y \in \Sigma} p(x,y) = 1$

n-step transition probability matrix: $P^n = \{p^n(x,y)\}_{x,y \in \Sigma}$,

$$p^n(x,y) := P(X_{t+n} = y \,|\, X_t = x)$$

Evolution of the probability distribution:

$$
\begin{aligned}
p^2(x,y) = \quad & P(X_2 = y \,|\, X_0 = x) \\
= \quad & \sum_z P(X_2 = y\,, X_1 = z \,|\, X_0 = x) \\
= \quad & \sum_z P(X_2 = y \,|\, X_1 = z\,, X_0 = x) P(X_1 = z \,| X_0 = x) \\
= \quad & \sum_z P(X_2 = y \,|\, X_1 = z) P(X_1 = z \,| X_0 = x) \\
= \quad & \sum_z p(x,z) p(z,y)
\end{aligned}
$$

Chapman-Kolmogorov equation: $P^n = P \times P \times ... P$ (n-fold product) and $P^{n+m} = P^n P^m$, i.e.

$$
p^{n+m}(x,y) = \sum_{z \in \Sigma} p^n(x,z) p^m(z,y)
$$

Construction of sample paths:

assume $\Sigma = \{x_1, x_2, ..., x_m\}$; if $X_t = x$,

$$X_{t+1} = \begin{cases} x_1, & \text{if} \quad U \leq p(x, x_1) \\ x_2 & \text{if} \quad p(x, x_1) < U \leq p(x, x_1) + p(x, x_2) \\ \cdots & \cdots \\ x_m & \text{if} \quad p(x, x_1) + p(x, x_2) + ... + p(x, x_{m-1}) < U \leq 1 \end{cases}$$

where U is a uniformly distributed random variable in $(0, 1)$.

Example 1: Random walk on a (periodic) lattice

$\{\xi_i : i = 1, 2, ...\}$ i.i.d random variables with

$$P(\xi_i = \pm 1) = p^{\pm}, \quad p^+ + p^- = 1$$

Define

$$X_n = \sum_{i=1}^{n} \xi_i = X_{n-1} + \xi_n$$

Then X_n is a Markov chain, i.e. satisfies the Markov property (Exercise 1).

Transition matrix:

$$p(x, y) = \begin{cases} p^{\pm}, & \text{if} \quad y = x \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

Example 2: Discrete-in-time stochastic dynamical systems

$\{X_1, \xi_i : i = 1, 2, ...\}$ independent random variables

Define

$$X_{n+1} = f(X_n, \xi_n)$$

where $f = f(x, y)$ a given deterministic function. Then X_n is a Markov chain, i.e. satisfies the Markov property (Exercise 2).

- Analogous to deterministic discrete evolution

$$X_{n+1} = f(X_n)$$

Transition matrix:

$$p_n(x, y) = P(f_n(x, \xi_n) = y)$$

Some useful definitions

1. We say that $x, y \in \Sigma$ communicate if

$$p^n(x, y), p^m(y, x) > 0$$

for some m, n. If all states communicate the Markov chain is called irreducible.

2. Let $d(x)$ the greatest common divisor (called the period of the state x) of all $k \geq 1$ such that

$$P(X_{t+k} = x \,|\, X_t = x) > 0\,.$$

A Markov chain is called aperiodic if each state has period 1.

3. If $x \in \Sigma$ is revisited with probability 1 at some finite time is called recurrent; otherwise the state is called transient

Stationary distributions , long-time behavior and ergodicity

Let $\pi = \pi(x) \geq 0\,, \sum_{x\in\Sigma} \pi(x) = 1$ a probability distribution.

If the initial state X_0 is random and distributed according to π, then the distribution of X_n is

$$P(X_n = y \,|\, X_0 \sim \pi) = \sum_x \pi(x) p^n(x,y) = \pi P^n$$

- π is stationary (invariant) if $\pi P^n = \pi$ for all n; true if 1 is an eigenvalue of the matrix P with eigenvector π:

$$\pi P = \pi$$

- If X_t is aperiodic and irreducible then

$$\lim_{n\to\infty} p^n(x,y) = \pi(y)\,, \quad \text{for all } x \quad \text{(Ergodicity)}$$

- For simply irreducible chains we have:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} p^t(x,y) = \pi(y)\,, \qquad \text{for all } x \quad \text{(weak ergodicity)}$$

2. Continuous-time Markov Chains

Stochastic process $\{X_t : t \geq 0.\}$ taking finitely many values on the state space Σ: $X_t = x \in \Sigma$.

Markov property:

$$P(X_{t+s} = y \,|\, X_s = x, X_u = x_u\,, 0 \leq u < s) = P(X_{t+s} = y \,|\, X_s = x)$$

Transition probability matrix at time t: $P_t = \{p(t; x, y)\}_{x,y\in\Sigma}$,

$$p(t; x, y) := P(X_{t+s} = y \,|\, X_s = x)$$

Note: $p(t; x, y) \geq 0$ and $\sum_{y\in\Sigma} p(t; x, y) = 1$

Chapman-Kolmogorov equation: $P_{t+s} = P_t P_s$, i.e.

$$p(t + s; x, y) = \sum_{z\in\Sigma} p(t; x, z)p(s; z, y)\,. \qquad (C - K)$$

Continuous vs. Discrete time Markov Chains

Discretization: The C-K relation implies that $X_n := X_{nh}$ is a discrete-time Markov Chain with transition probability matrix $p(x,y) = p(h;x,y)$. Hence X_n is a discretization of the continuous time Markov Chain X_t.

Residence time τ_x: time spent by the process X_t at x; random waiting time between consecutive jumps.

Markov property implies:

$$P(\tau_x > t+s \,|\, \tau_x > s) = P(\tau_x > t)$$

i.e. τ_x is a "memoryless" distribution! Hence

$$P(\tau_x > t+s) = P(\tau_x > t+s\,, \tau_x > s) = P(\tau_x > t)P(\tau_x > s)$$

thus

$$P(\tau_x > t) = \exp\left(-\lambda(x)t\right), \quad \lambda(x) \geq 0$$

.

We can now define the transition matrix–corresponding to a "skeleton" Markov Chain

$$p(x,y) = P(X_{\tau_x} = y \,|\, X_0 = x)\,, y \neq x$$

we set $p(x,x) = 0$; note that we easily get that $\sum_y p(x,y) = 1$.

Define the **transition rates**

$$q(x,y) = \lambda(x) p(x,y)\,, x \neq y$$

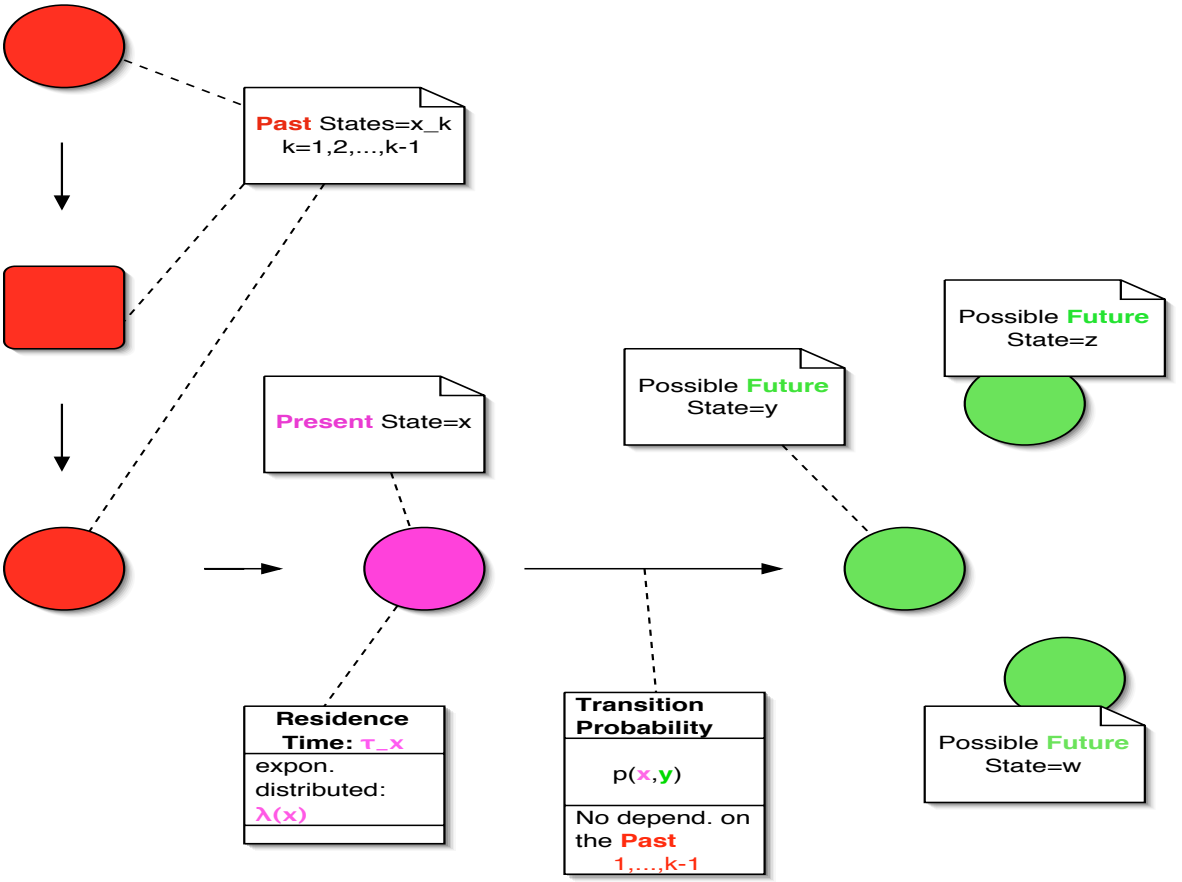
and

$$q(x,x) = -\sum_{y \neq x} q(x,y) = -\lambda(x)$$

denote the corresponding matrix $Q = (q(x,y)_{x,y \in \Sigma})$.

Building blocks of the continuous time chain:

$q(x,y)$ or equivalently $\lambda(x)\,, p(x,y)$ for all $x, y \in \Sigma$.


Past States=x_k
k=1,2,...,k-1
Present State=x
Possible Future
State=y
Possible Future
State=z
Residence
Time: τ_x
expon.
distributed:
λ(x)
Transition
Probability
p(x,y)
No depend. on
the Past
1,...,k-1
Possible Future
State=w

Construction of sample paths from the transition rates $q(x, y)$:

Step 1: Construct all residence times τ_x at each location x: exponentially distributed with rate $\lambda(x) = -q(x, x)$.

Step 2: Determine the next move, starting at $t = 0$ at the state $x \in \Sigma$ using the "skeleton" Markov Chain:

$$X_t = x\,, \quad \text{for} \quad 0 \le t < \tau_x$$

assume $\Sigma = \{1, 2, ..., m\}$;then

$$X_{\tau_x} = \begin{cases} 1, & \text{if} \quad U \le \frac{q(x,1)}{-q(x,x)} = p(x,1) \\ 2 & \text{if} \quad \frac{q(x,1)}{-q(x,x)} < U \le \frac{q(x,1)+q(x,2)}{-q(x,x)} \\ ... & ... \\ m & \text{if} \quad \frac{q(x,1)+q(x,2)+...+q(x,m-1)}{-q(x,x)} < U \le 1 \end{cases}$$

where U is a uniformly distributed random variable in $(0, 1)$.

Step 3: From Step 2 we obtain a new location $y \in \Sigma$; then

$$X_t = y\,, \quad \text{for} \quad \tau_x \le t < \tau_x + \tau_y$$

and proceed as before.

C-K and Master equations:

Forward C-K: $p(t+h;x,y) = \sum_z p(t;x,z)p(h;z,y)$

Can also show: $\frac{dp(t;x,y)}{dt}|_{t=0} = q(x,y)$

Thus,

$$\frac{dp(t;x,y)}{dt} = \sum_{z\in\Sigma} p(t;x,z)q(z,y)$$

or equivalently

$$\frac{dp(t;x,y)}{dt} = \sum_{\substack{z\in\Sigma \\ z\neq y}} p(t;x,z)q(z,y) - p(t;x,y)q(y,z)\,, \quad \text{(master equations)}$$

Generators:

Observables: $u(x,t) := \mathbf{E_x f(X_t)} = \sum_{y\in\Sigma} f(y)p(t;x,y).$

$$\begin{aligned}
\frac{d}{dt}E_x f(X_t) = & \sum_{y\in\Sigma} f(y)\frac{dp(t;x,y)}{dt} \\
= & \sum_{y\in\Sigma}\sum_{z\in\Sigma} p(t;x,z)q(z,y)f(y) \\
= & \sum_{z\in\Sigma} p(t;x,z) \sum_{y\in\Sigma, y\neq z} q(z,y)[f(y)-f(z)] \\
= & E_x Lf(X_t)\,, \quad \text{hence}
\end{aligned}$$

$$\mathbf{\frac{d}{dt}E_x f(X_t) = E_x Lf(X_t)}$$

where $\mathbf{L}$ the **generator** of X_t: $\mathbf{Lf(z) = \sum_{\substack{y\in\Sigma \\ y\neq z}} q(z,y)\Big(f(y) - f(z)\Big)}$

- The generator completely determines the Markov process X_t.

Example 1: Continuous-time random walk on a lattice

$\{\xi_i : i = 1, 2, ...\}$ i.i.d random variables with

$$P(\xi_i = \pm 1) = p^{\pm}, \quad p^+ + p^- = 1$$

$\{\tau_k : k = 1, 2, ...\}$ i.i.d exponentially distributed non-negative random variables (τ_k also independent of $\xi_i, i = 1, 2, ..., k$) with

$$P(\tau_k > t) = \exp(-\lambda t), \quad \lambda > 0$$

Consider the sequence of pairs

$$\left(X_n = \sum_{i=1}^{n} \xi_i, T_n = \sum_{k=1}^{n} \tau_k \right)$$

(X_n is a Markov Chain–see Example 1–also the pair is a Markov Chain) and define

$$X_t = X_n, \quad \text{if} \quad T_n \leq t < T_{n+1}$$

Then X_t is a continuous time Markov chain, i.e. satisfies the Markov property (Exercise 3).

Generator:

$$Lf(x) = \lambda p^+ (f(x+1) - f(x)) + \lambda p^- (f(x-1) - f(x))$$

Example 2: Birth-death processes (see next talk).

Markov process with random initial data, $\mu = \mu(x)\,, x \in \Sigma$:

$$\mu P_t(y) = \sum_x \mu(x) p(t; x, y) = P(X_t = y \,|\, X_0 \sim \mu)$$

Invariant (stationary) distribution: $\pi P_t = \pi$ for all t.

Since $\frac{dP}{dt} = PQ$, we have that $\pi \mathbf{Q} = \mathbf{0}$, hence

$$\sum_x \pi(x)\lambda(x)p(x,y) = \pi(y)\lambda(y)$$

A stronger condition is **detailed balance**:

$$\pi(x)\lambda(x)p(x,y) = \pi(\mathbf{x})\mathbf{q}(\mathbf{x},\mathbf{y}) = \pi(\mathbf{y})\mathbf{q}(\mathbf{y},\mathbf{x}) = \pi(y)\lambda(y)p(y,x)$$

Ergodicity: unique π such that $\lim_{tto\infty} \mu P_t = \pi$ for all μ.

1. Remarks on Monte Carlo simulation

1. Markov Chain Monte Carlo (MCMC)

Task: sample from a given a probability distribution $\pi = (\pi(x))_{x\in\Sigma}$.

Idea: construct a discrete-time Markov Chain with $P = (p(x,y))_{x,y\in\Sigma}$ having π as an stationary (invariant) distribution. Then generate π using ergodicity.

- Define $p(x,y) = r(x,y)a(x,y)$
- $R = (r(x,y))_{x,y\in\Sigma}$: transition probabilities for a *proposed* move y
- $a(x,y)$: *acceptance probability*
- $\frac{a(x,y)}{a(y,x)} = \frac{\pi(y)r(x,y)}{\pi(x)r(y,x)}$: detailed balance.

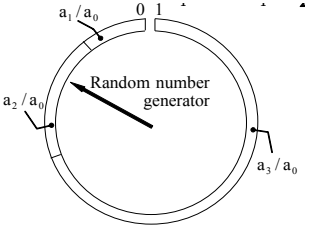
Often $a(x,y) << 1$ (e.g. low temperatures); many rejected moves hence convergence to π is slow. One way to resolve this....

2. Continuous Time Monte Carlo (CTMC)

Construct a continuous-time Markov Chain with invariant measure π;

- the random jump time is known (exponentially distributed, etc.) and defines the time step Δt of the simulator.
- no rejected moves
- one drawback: it may be costly to generate for a given x all $q(x, y) > 0$ for all y's accessible from x.
- CTMC is "real" dynamics.

Ref: Gillespie (chemical reactions); Bortz, Kalos, Lebowitz (Ising-type systems)



Pseudo-algorithm

Step 1: Construct all residence times τ_x at each location x: exponentially distributed with rate $\lambda(x) = -q(x,x)$.

Step 2: Determine the next move, starting at $t = 0$ at the state $x \in \Sigma = \{1, 2, ..., m\}$ using the "skeleton" Markov Chain:

$$X_t = x\,, \quad \text{for} \quad 0 \le t < \tau_x$$

$$X_{\tau_x} = \begin{cases} 1, & \text{if} \quad \mathbf{U} \le \frac{q(x,1)}{-q(x,x)} = p(x,1) \\ 2 & \text{if} \quad \frac{q(x,1)}{-q(x,x)} < \mathbf{U} \le \frac{q(x,1)+q(x,2)}{-q(x,x)} \\ \dots & \dots \\ m & \text{if} \quad \frac{q(x,1)+q(x,2)+\dots+q(x,m-1)}{-q(x,x)} < \mathbf{U} \le 1 \end{cases}$$

U: **uniformly distributed random variable** in $(0, 1)$

Step 3: From Step 2 we obtain a new location $y \in \Sigma$; then

$$X_t = y\,, \quad \text{for} \quad \tau_x \le t < \tau_x + \tau_y$$

and proceed as before.

2. Markov Chains with two time scales–Stochastic Averaging

Discrete-time Markov Chain with transition probability

$$\mathbf{P}^\epsilon = \mathbf{P} + \epsilon \mathbf{Q}\,, \quad \epsilon << 1$$

$P = (p(x,y))_{x,y\in\Sigma}$: transition probability matrix

$Q = (p(x,y))_{x,y\in\Sigma}$: a transition rate matrix

If P is irreducible (more general results exist!), the k-step transition matrix is expanded as:

$$(P^\epsilon)^k = \bar{\mathbf{P}} + \mathbf{\Psi_0(k)} + \epsilon \mathbf{\Phi_1(\epsilon k)} + \epsilon \Psi_1(k) + O(\epsilon^2)$$

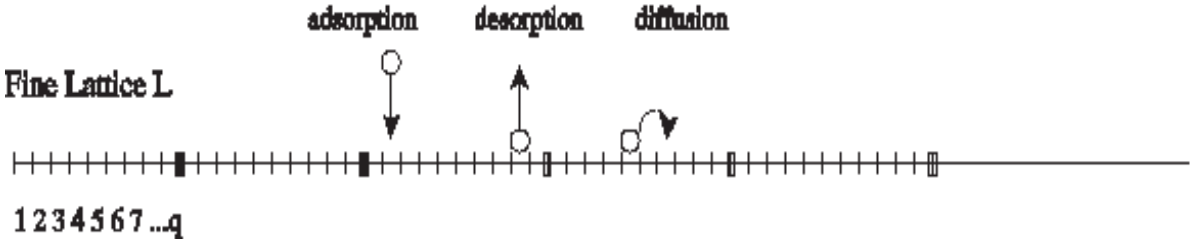
$\bar{\mathbf{P}} = \mathbf{I}\nu$: ν is the invariant distribution of P.

$\mathbf{\Phi_1(\epsilon k)}$: slow time-scale dynamics.

$\mathbf{\Psi_i(k)}$: initial layer terms

- Related results for contin.-time Markov Chains and SDE.
- **Ref:** In math, Khasminskii, Kurtz, Papanicolaou,... In EE, Phillips and Kokotovic,... In AOS, Majda, Timofeyev, Vanden-Eijnden,... Books: Yin and Zhang '04,...

3. Stochastic lattice dynamics–Ising Systems



- *Spin:* $\sigma(x) \in \{0, 1\}$ at the lattice site $x \in \mathbf{Z}^d$ (vacant vs. occupied sites).
- Potts, Heisenberg models.
- *Spin configuration:* $\sigma = \{\sigma(x) \mid x \in \Lambda \subset \mathbf{Z}^d\}$, $|\Lambda| = N$: total number of lattice sites.

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

- h: external field
- J: potential with interaction range L.

Canonical Gibbs measure:

at the inverse temperature $\beta = \frac{1}{kT}$,

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

[Probability of the configuration σ_0]

Partition function: $Z_{\Lambda,\beta} = \sum_{\sigma_0} \exp\Big\{ -\beta H_N(\sigma_0)\Big\} P_N(\sigma = \sigma_0)$

Prior distribution (no interactions, hight temp.):

$$P_N(\sigma = \sigma_0) = \Pi_{x\in\Lambda} P(\sigma(x) = \sigma_0(x))$$

where

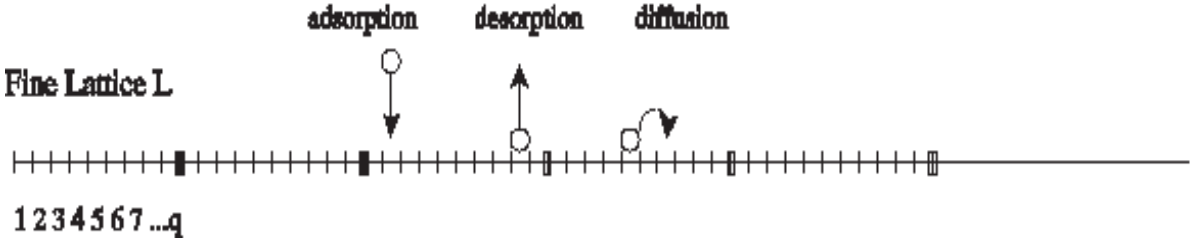
$$P(\sigma(x) = 1) = \frac{1}{2} \quad \text{and} \quad P(\sigma(x) = 0) = \frac{1}{2}.$$

i.e. the prior distribution is a product measure of Bernoulli distributions with parameter a.

Dynamics

A. Spin Flip – Adsorption/Desorption

B. Spin Exchange – Surface diffusion



Markov Chain modeling with state space

$$\Sigma = \text{set of all configurations } \sigma$$

A. Spin Flip–Adsorption/Desorption

Spin flips occur at each lattice site x in $[t, t+\Delta t]$ with probability

$$c(x,\sigma)\Delta t + O(\Delta t^2)$$

Generator: $L_N f(\sigma) = \sum_{x\in\Lambda} c(x,\sigma)[f(\sigma^x) - f(\sigma)]$.

Transition rate: $c(x,\sigma)$

Detailed balance law:

$$c(x,\sigma)\exp\Big(-\beta H(\sigma)\Big) = c(x,\sigma^x)\exp\Big(-\beta H(\sigma^x)\Big)$$

σ^x: configuration after a spin flip at x.

- Spin flip rate (Metropolis-type dynamics):

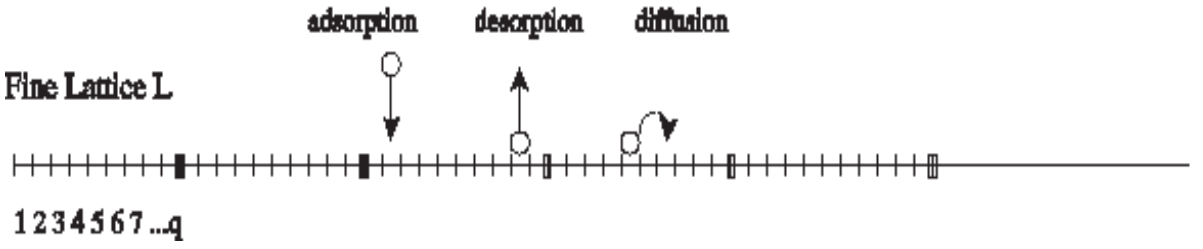
$$c(x,\sigma) = \Psi(-\beta\Delta_x H(\sigma)),$$

- $\Delta_x H(\sigma) = H(\sigma^x) - H(\sigma)$.
- $\beta > 0$: inverse temperature.
- Typical choices of Ψ's are:

  $\Psi(r) = (1+e^r)^{-1}$ (Glauber dynamics).

  $\Psi(r) = e^{-r^+}$ (Metropolis dynamics).

Arrhenius adsoprtion/desorption dynamics:



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

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

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

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

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

B. Spin Exchange Dynamics–Surface diffusion.

- **Dynamics:** Sequence of spin exchanges with nearest neighbors.

- Spin exchange rate (Metropolis-type dyn.):

$$c(x, y, \sigma) = \Psi(-\beta \Delta_{x,y} H(\sigma)),$$

- $\Delta_{x,y} H(\sigma) = H(\sigma^{(x,y)}) - H(\sigma)$.

  $\sigma^{(x,y)}$:config. after a spin exch. between $x\,, y$.

- Detailed balance.

- Typical choices of Ψ's are:

  $\Psi(r) = 2(1 + e^r)^{-1}$ (Kawasaki dynamics).

  $\Psi(r) = e^{-r^+}$ (Metropolis dynamics).

- Arrhenius dynamics.