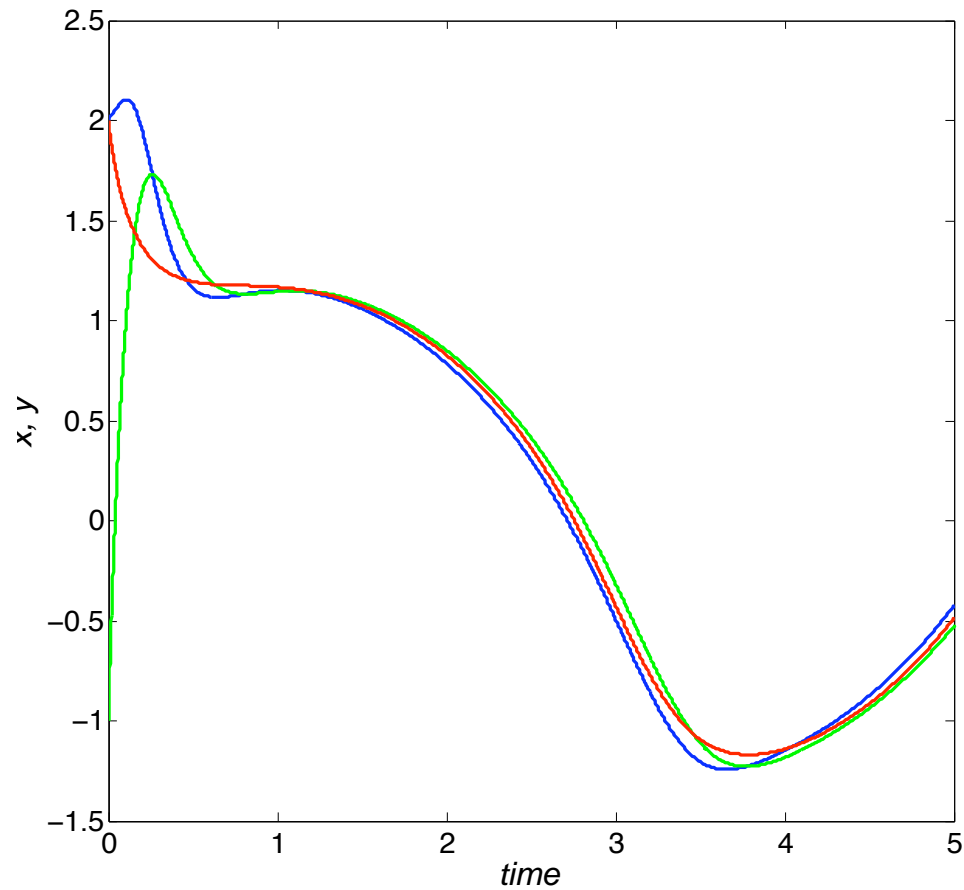# Multiscale Integrators for Systems with Disparate Time Scales
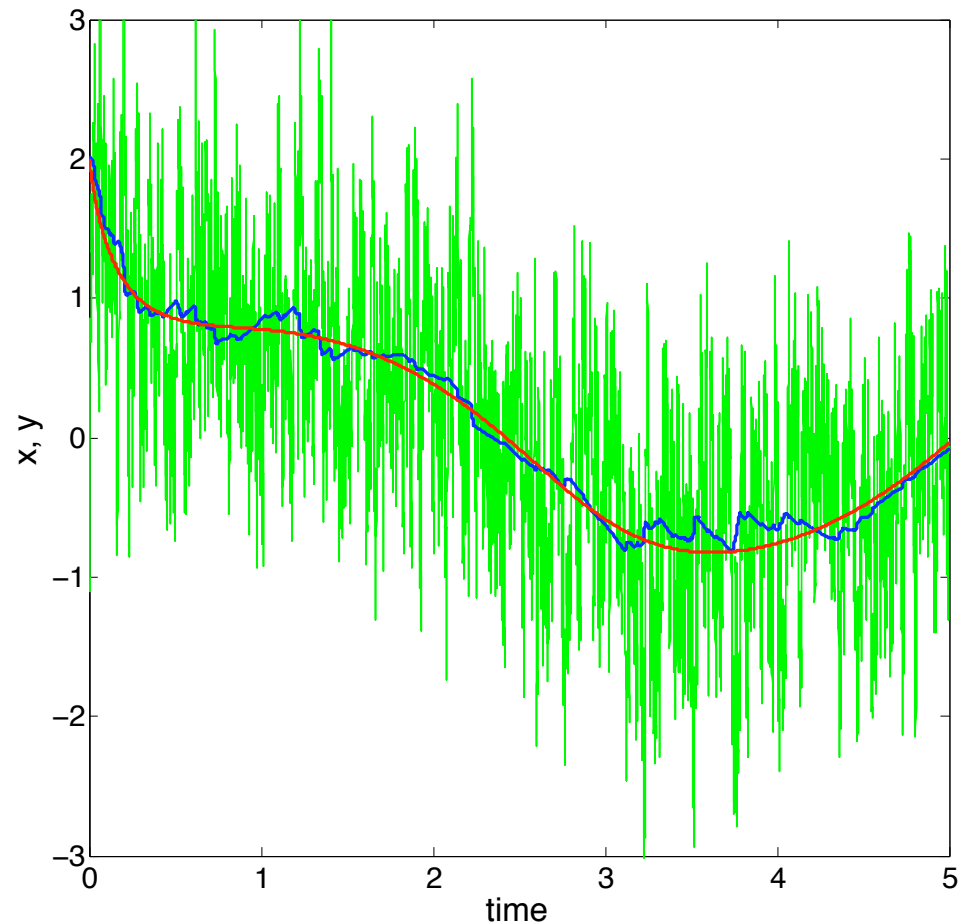
Eric Vanden-Eijnden

Courant Institute

The solution of

$$\begin{cases} \dot{X} = -Y^3 + \cos(t) + \sin(\sqrt{2}t) \\ \dot{Y} = -\varepsilon^{-1}(Y - X) \end{cases}$$

when $\varepsilon = 0.1$ and we took $X(0) = 2$, $Y(0) = -1$. X is shown in blue, and Y in green. Also shown in red is the solution of the limiting equation

$$\dot{X} = -X^3 + \cos(t) + \sin(\sqrt{2}t)$$

The solution of

$$
\begin{cases}
\dot{X} = -Y^3 + \cos(t) + \sin(\sqrt{2}t) \\
dY = -\varepsilon^{-1}(Y - X)dt + \varepsilon^{-1/2}dW
\end{cases}
$$

when $\varepsilon = 0.01$ and $X(0) = 2$, $Y(0) = -1$. X is shown in blue, and Y in green. Also shown in red is the solution of the limiting equation

$$
\dot{X} = -X^3 + X + \cos(t) + \sin(\sqrt{2}t)
$$

Notice how noisy Y is.

Outline

HMM-like multiscale integrators vs projective integration methods

What are they? Why are the different? When are they applicable?

Toward seamless multiscale integrators? Boosting method

Application to free energy calculations and phase-space exploration

Projective Integration Methods   (Gear & Kevrekidis 03; Eriksson, Johnson, & Logg 03)

Suitable for systems with widely separated time scales in which some slow variables exists which satisfy a closed ODE.

1. Make a few small time steps to estimate the rate of change of the slow variables via finite-difference;

2. Use this estimate to make a large extrapolation step;

3. Repeat.
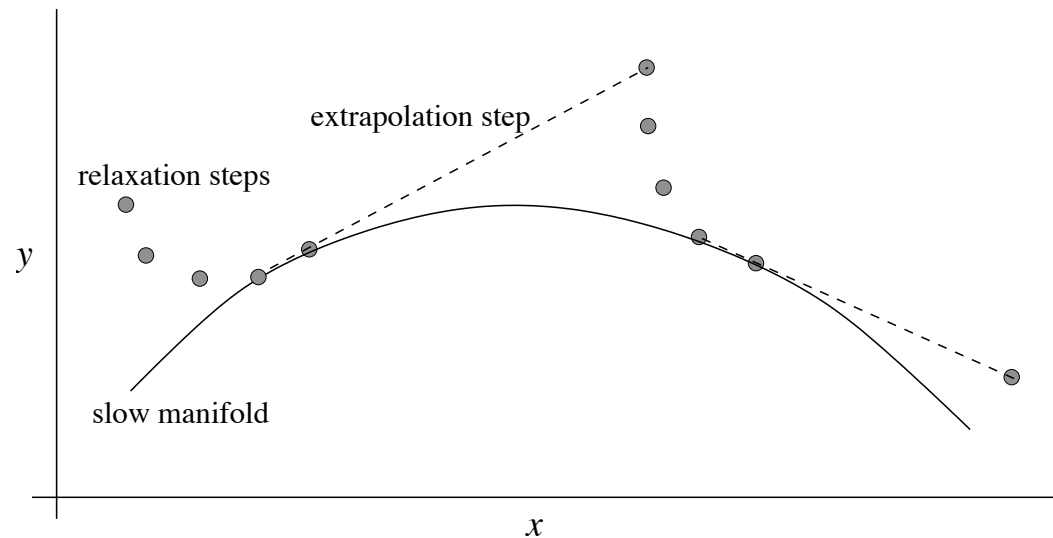
Key idea = extrapolation: *"The reader might think that these should be called 'extrapolation methods,' but that name has already been used [...]. Hence we call the proposed methods projective integration methods."* [from Gear & Kevrekidis *SIAM J. Sci. Comp.* **24**(4):109-110 (2003).]

Why do projective integration methods work?

Consider the ODE

$$\begin{cases} \dot{X} = f(X, Y) \\ \dot{Y} = -\varepsilon^{-1}(Y - \phi(X)) \end{cases} \tag{$\star$}$$

Typical stiff ODE with slow manifold:

If $\varepsilon \ll 1$, Y is very fast and it will adjust rapidly to the current value of X, i.e. after short $O(\varepsilon)$ transient we will have

$$Y = \phi(X) + O(\varepsilon) \text{ at all times.}$$

Then the equation for slow variables X reduces to

$$\dot{X} = f(X, \phi(X)) \tag{$\star\star$}$$

1. Simulate ($\star$) for a few micro-steps with time-step $\delta t = O(\varepsilon)$ (i.e small enough to resolve the fast motion) until the system reaches the slow manifold, i.e. $Y = \phi(X) + O(\varepsilon)$;

2. Extrapolate the last micro-step, i.e. make a macro-step with time-step $\Delta t = O(1)$, effectively simulating ($\star\star$);

3. Repeat.

Why is this better?

Basically, it only takes $O(\log \varepsilon^{-1})$ steps to reach an $O(1)$ time-scale.

Indeed $O(\log \varepsilon^{-1})$ is the number of micro-steps necessary to relax the system $O(\varepsilon)$ close to the slow manifold and be able to take an $O(1)$ extrapolation step.

In contrast, the naive explicit scheme

$$\begin{cases} X^{p+1} = X^p + \delta t \ f(X^p, Y^p), \\ Y^{p+1} = Y^p + \dfrac{\delta t}{\varepsilon} \ g(X^p, Y^p), \end{cases}$$

takes $O(\varepsilon^{-1})$ steps to reach an $O(1)$ time-scale.

In the context of stiff ODEs:
Projective integration methods = poor man's implicit scheme.

$$\begin{cases} X^{n+1} = X^n + \Delta t f(X^{n+1}, Y^{n+1}) \\ Y^{n+1} = Y^n - \dfrac{\Delta t}{\varepsilon}(Y^{n+1} - \phi(X^{n+1})) \end{cases}$$

When $\varepsilon \ll \Delta t \ll 1$:

$$\begin{cases} Y^{n+1} = \phi(X^n) + O(\Delta t) + O(\varepsilon), \\ X^{n+1} = X^n + \Delta t f(X^n, \phi(X^n)) + O(\Delta t^2) + O(\varepsilon) \end{cases}$$

Advantages: very simple to use and seamless. In particular projective integrators are applicable to

$$\begin{cases} \dot{X} = f(X, Y) \\ \dot{Y} = \dfrac{1}{\varepsilon} g(X, Y) \end{cases} \qquad \text{or} \qquad \dot{Z} = h(Z, \varepsilon)$$

provided that the solution is rapidly attracted to a slow manifold.

However: Other methods can be more efficient − Chebychev methods, implicit schemes, etc.

Projective integration methods can be extended to systems other than stiff ODEs.

In this case, they stop being seamless (one must know what the slow variables are) and they are based on the *same idea of extrapolation:*
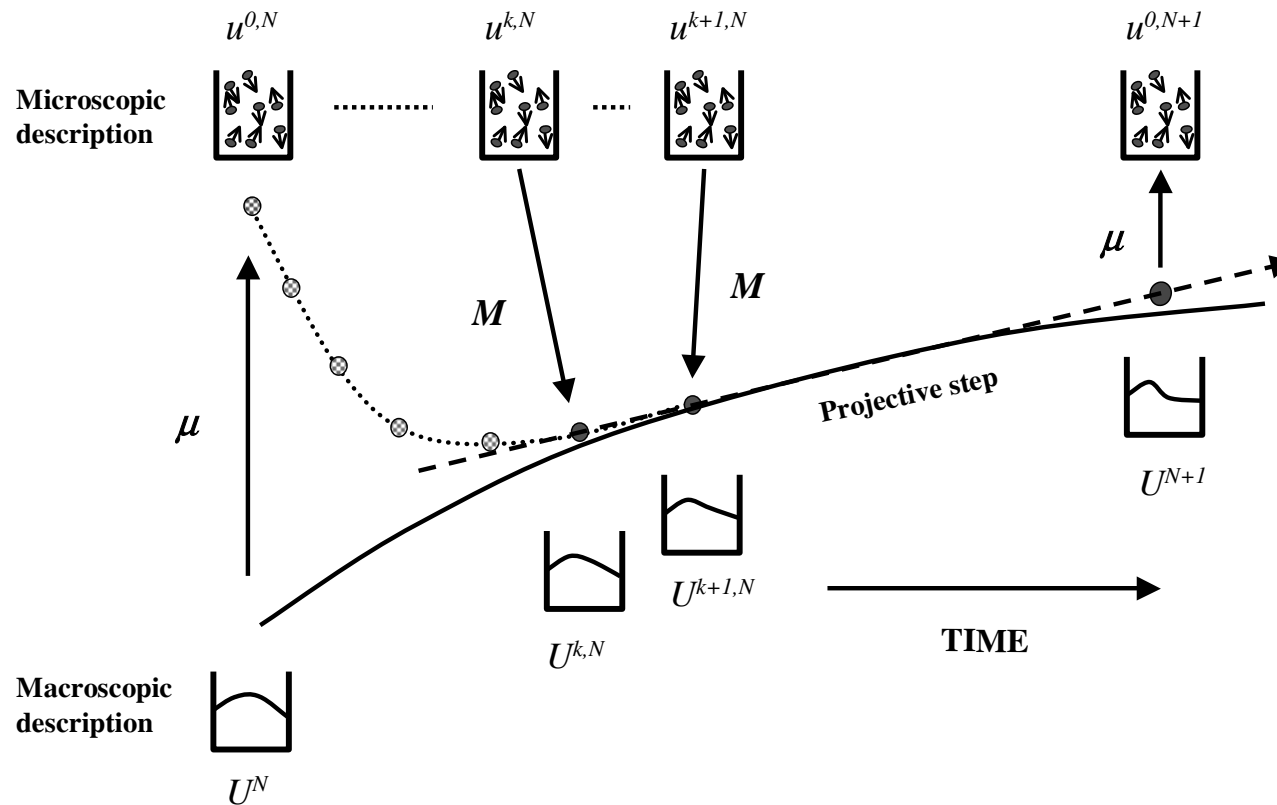
Projective integration methods assume the existence of a limiting equation for the slow variables but do not use it explicitly.
$\Rightarrow$ "Equation free approach".

HMM-like Integrators   (V.-E. 03; E & Engquist 03)

Unlike projective integration methods, HMM-like integrators explicitly use the limiting equation for the slow variables − antipodal to the equation free approach.

Key limit thm: Consider

$$\begin{cases} \dot{X} = f(X,Y), \\ dY = \varepsilon^{-1}g(X,Y)dt + \varepsilon^{-1/2}\sigma(X,Y)dW(t), \end{cases} \qquad (\star)$$

Assume that: (i) the evolution Y at every fixed $X = x$ is ergodic with respect to the probability distribution

$$d\mu_x(y)$$

and (ii)

$$F(x) = \int_{\mathbb{R}^m} f(x,y)d\mu_x(y) \quad \text{exists}$$

Then in the limit as $\varepsilon \to 0$ the evolution for X solution of $(\star)$ is governed by

$$\dot{X} = F(X)$$

In addition

$$F(x) = \int_{\mathbb{R}^m} f(x,y)d\mu_x(y) = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(x,Y_t^x)dt$$

where

$$dY^x = \varepsilon^{-1}g(x,Y^x)dt + \varepsilon^{-1/2}\sigma(x,Y^x)dW(t)$$

Basic HMM-like integrator

Use

$$X^{n+1} = X^n + \Delta t \ \tilde{F}^n, \qquad X_0 = X(t=0)$$

Here $\tilde{F}_n$ is an approximation of $F(X_n)$ obtained as

$$\tilde{F}^n = \frac{1}{M_T} \sum_{m=M}^{M+M_T} f(X^n, Y^{n,m})$$

where

$$Y^{n,m+1} = Y^{n,m} + \frac{\delta t}{\varepsilon} \ g(X^n, Y^{n,m}) + \sqrt{\frac{\delta t}{\varepsilon}} \ \sigma(X^n, Y^{n,m})\xi^m,$$

$$Y_{0,0} = Y(t=0), \ Y^{n+1,0} = Y^{n,M+M_T-1}$$

Why is this better?

Basically, because M and M_T are $O(1)$ in ε! In other words, one can reach an $O(1)$ time-scale with a $O(1)$ number of steps.

In contrast, the direct scheme

$$\begin{cases} X^{p+1} = X^p + \delta t \ f(X^p, Y^p), \\ Y^{p+1} = Y^p + \dfrac{\delta t}{\varepsilon} \ g(X^p, Y^p), \end{cases}$$

takes $O(\varepsilon^{-1})$ steps to reach an $O(1)$ time-scale.

Error estimate:

(E, Liu & V.-E. 03; E & Engquist 03)

Thm: *For any $T > 0$, there exists a constant $C > 0$ such that*

$$\mathbb{E}\left(\sup_{0 \leq n \leq T/\Delta t} |X(n\Delta t) - X^n|\right) \leq C\left(\sqrt{\varepsilon} + (\Delta t)^k + (\delta t/\varepsilon)^l + \sqrt{\frac{\varepsilon \Delta t}{M_T \delta t}}\right)$$

Summarizing:

Projective integration methods are *extrapolation methods*. In the context of stiff ODEs there are totally seamless; more generally, they require to know what the slow variables are but not what their limiting equation is.

Projective integration methods do not use explicitly the limiting equation for the slow variables.

HMM-like multiscale integrators use the limiting equation for the slow variables explicitly.

The key idea is integrate this equation by evaluating numerically and on-the-fly the coefficients in this equation when these are not available analytically in closed form.

Toward seamless multiscale integrators?

HMM-like integrators require to know explicitly what are the slow and fast variables. Can we do better? Can we extend the extrapolation strategy in a seamless way to systems more general than stiff ODEs?

Recall that

$$F(X^n) \approx \tilde{F}^n = \frac{1}{M} \sum_{m=M_T}^{M+M_T-1} f(X^n, Y^{n,m})$$

and observe: the HMM-like multiscale integrator works even if $M_T = 1$ (no time averaging) provided only that

$$\varepsilon \ll \frac{\varepsilon \Delta t}{M \delta t} \ll 1$$

The factor $\lambda = \Delta t / M \delta t$ also gives the efficiency boost the HMM-like multiscale integrators over a direct scheme.

Why is time-averaging unnecessary (since in this case the approximation on $\tilde{F}^n$ is very bad at each time step)?

HMM without time-averaging

$$\begin{cases} \dot{X} = f(X,Y) \\ dY = \dfrac{1}{\varepsilon}g(X,Y)dt + \dfrac{1}{\sqrt{\varepsilon}}\sigma(X,Y)dW(t) \end{cases} \qquad (\star)$$

Scheme:

1. Use:

$$Y^{n,m+1} = Y^{n,m} - \frac{\delta t}{\varepsilon}g(X^n, Y^{n,m}) + \sqrt{\frac{\delta t}{\varepsilon}}\sigma(X^n, Y^{n,m})\xi^{n,m}, \qquad Y^{n,0} = Y^{n-1,M}$$

2. Use

$$X^{n+1} = X^n + \Delta t\, f(X^n, Y^{n,M})$$

3. Repeat.

Denoting $\lambda = \Delta t/M\delta t$, this integrator approximates

$$\begin{cases} \dot{X} = f(X,Y) \\ dY = -\dfrac{1}{\lambda\varepsilon}g(X,Y)dt + \dfrac{1}{\sqrt{\lambda\varepsilon}}\sigma(X,Y)dW(t) \end{cases} \qquad (\star\star)$$

Hence, by the limit theorem, it approximates $(\star)$ and boost the efficiency by λ if

$$\varepsilon \ll \varepsilon\lambda \ll 1 \quad \text{or equivalently} \quad M\delta t \ll \Delta t \ll \varepsilon^{-1}M\delta t$$

Seamless multiscale algorithm − Boosting method

Consider

$$dZ = \frac{1}{\varepsilon}h_1(Z)dt + \frac{1}{\sqrt{\varepsilon}}h_2(Z)dW(t) + h_3(Z)dt + h_4(Z)d\bar{W}(t)$$

and assume that there exist slow variables $X = \phi(Z)$ which satisfies the closed SDE

$$dX = F(X)dt + G(X)dW(t) \qquad \text{as } \varepsilon \to 0$$

1. Use

$$Z^{n,m+1} = Z^{n,m} + \frac{\delta t}{\varepsilon}\, h_1(Z^{n,m}) + \sqrt{\frac{\delta t}{\varepsilon}}h_1(Z^{n,m})\xi^{n,m},$$

2. Use

$$Z^{n+1,0} = Z^{n,M} + \Delta t\, h_3(Z^{n,M}) + \sqrt{\Delta t}\, h_4(\tilde{Z}^{n,M})\eta^n$$

3. Repeat.

Seamless in that one does not need to know the slow variables X are. Works because it approximates

$$dZ = \frac{1}{\lambda\varepsilon}h_1(Z)dt + \frac{1}{\sqrt{\lambda\varepsilon}}h_2(Z)dW(t) + h_3(Z)dt + h_4(Z)d\bar{W}(t)$$

with $\lambda = \Delta t/M\delta t$.

Similar in spirit to Chorin's artificial compressibility method, the Car-Parrinello method in molecular dynamics, etc.

Application to free energy calculation and phase-space exploration

(joint work with L. Maragliano)

Definition: Consider a random variable $X \in \mathbb{R}^n$ with the *Boltzmann-Gibbs probability distribution*;

$$d\mu(x) = Z^{-1}e^{-\beta V(x)}dx \qquad Z = \int_{\mathbb{R}^n} e^{-\beta V(x)}dx$$

where $V(x)$ is the potential, $\beta > 0$ is the inverse temperature.

Let $\theta : \Omega \mapsto \mathbb{R}^N$ be a set of *collective variables* (aka vectorial reaction coordinates).

The *free energy* associated with θ is the function $G : \mathbb{R}^N \mapsto \mathbb{R}$ such that:

$e^{-\beta G(z)}dz$ *is the probability distribution of* $Z = \theta(X)$.

In formula:

$$G(z) = -\beta^{-1}\log Z^{-1}\int_{\mathbb{R}^n} e^{-\beta V(x)}\delta(\theta(x) - z)dx$$

$$= -\beta^{-1}\log Z^{-1}\int_{\Sigma(z)} e^{-\beta V(x)}J(x)d\sigma_{\Sigma(z)}(x)$$

where $\Sigma(z) = \{x : \theta(x) = z\}$, $J(x) = |\det M(x)|^{1/2}$, $M_{\alpha\beta}(x) = \nabla\theta_\alpha(x) \cdot \nabla\theta_\beta(x)$.

Consider a dynamical system modeled e.g. by the Langevin equation

$$\ddot{X}(t) = -\nabla V(X(t)) - \gamma \dot{X}(t) + \sqrt{2\gamma\beta^{-1}}\eta(t),$$

where γ is the friction tensor and $\eta(t) = \dot{W}(t)$ is a white-noise.

This system is ergodic with respect to the Boltzmann-Gibbs distribution;

Given any suitable $f : \Omega \mapsto \mathbb{R}$

$$\frac{1}{T}\int_0^T f(x(t))dt \to Z^{-1}\int_{\mathbb{R}^n} f(x)e^{-\beta V(x)}dx \quad \text{almost surely (a.s.) as } T \to \infty$$

In particular, letting $f = F \circ \theta$ for some $F : \mathbb{R}^N \mapsto \mathbb{R}$:

$$\frac{1}{T}\int_0^T F(\theta(X(t)))dt \to \int_{\mathbb{R}^N} F(z)e^{-\beta G(z)}dz \quad \text{a.s. as } T \to \infty$$

The free energy permits to organize the data in terms of some relevant observables θ.

Use extended system on state-space $(x, z) \in \mathbb{R}^n \times \mathbb{R}^N$ with potential

$$U(x, z) = V(x) + \tfrac{1}{2}k \sum_{\alpha=1}^{N}(\theta_\alpha(x) - z_\alpha)^2$$

$$\begin{cases} \alpha^2 \ddot{X}(t) = -\nabla V(X(t)) - k \sum_{\alpha=1}^{N}(\theta_\alpha(X(t)) - Z_\alpha(t))\nabla\theta_\alpha(X(t)) - \alpha\gamma\dot{X}(t) + \sqrt{2\alpha\gamma\beta^{-1}}\eta(t) \\ \ddot{Z}_\alpha(t) = k(\theta_\alpha(X(t)) - Z(t)) - \bar{\gamma}\dot{Z}_\alpha(t) + \sqrt{2\bar{\gamma}\bar{\beta}^{-1}}\bar{\eta}_\alpha(t) \end{cases}$$

Limit thm: As $\alpha \to 0$, the dynamics for $Z(t)$ is approximately

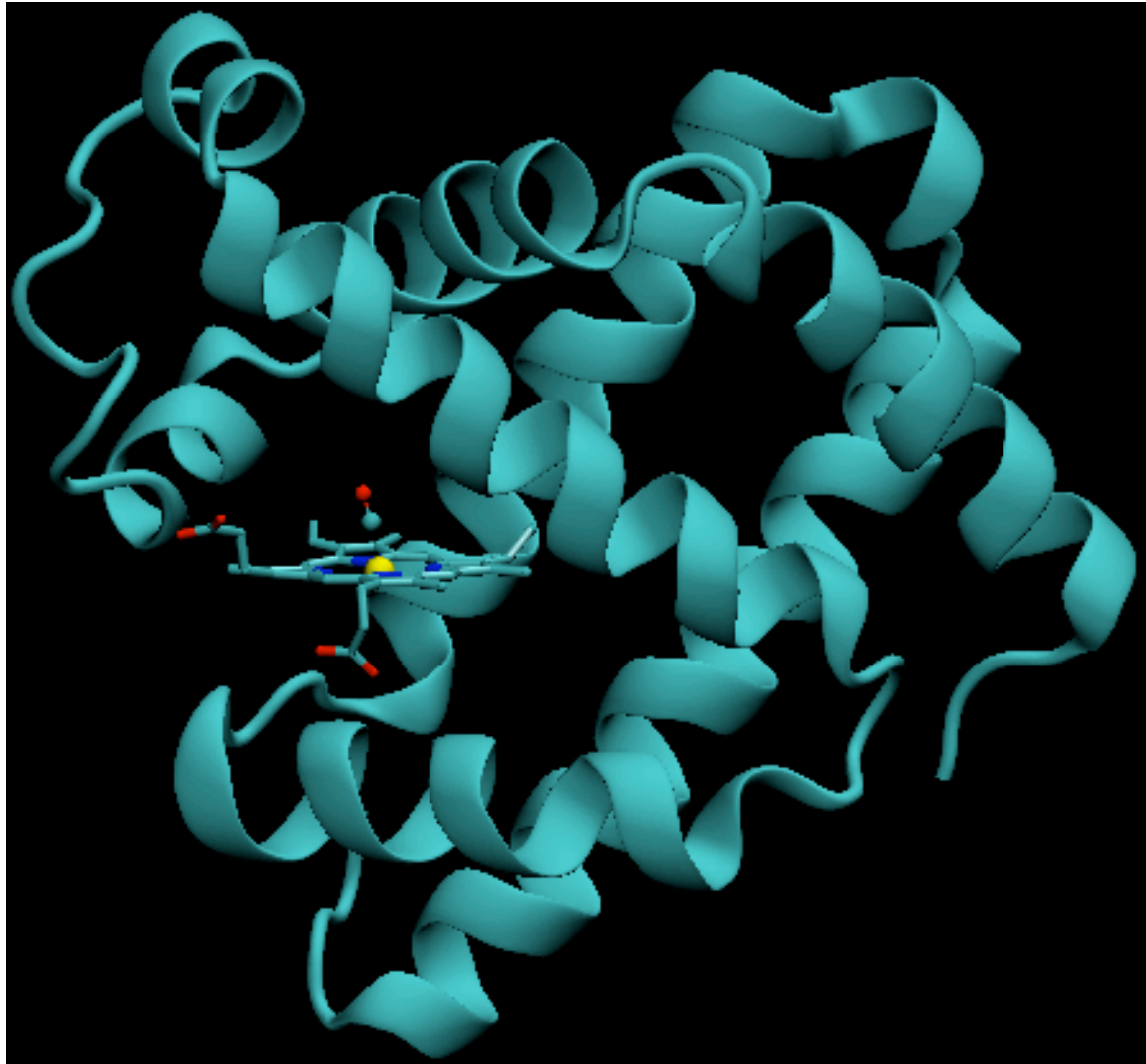$$\ddot{Z}(t) = -\nabla G_k(Z(t)) - \bar{\gamma}\dot{Z}(t) + \sqrt{2\bar{\gamma}\bar{\beta}^{-1}}\bar{\eta}(t)$$

where

$$G_k(z) = -\beta^{-1}\log Z_k^{-1} \int_{\mathbb{R}^n} \exp(-\beta V(x) - \frac{1}{2}\beta k \sum_\alpha (\theta_\alpha(x) - z_\alpha)^2)dx \to G(z) \qquad \text{as } k \to \infty$$

So: Simulate the system above with α small enough, k large enough and $\bar{\beta} < \beta$, and sample (i.e. bin) $Z(t)$ to retrieve $G(z)$ or (better) use thermodynamic integration along trajectory.

Remark: Use constraint instead of restraints to eliminate error in k (i.e. take $k \to \infty$ exactly).

Example: CO diffusion in Myoglobin

Crystal structure of Carbon-monoxide (CO) - bound Myoglobin, as deposited in the Protein Data Bank archive. The backbone chain is represented in "ribbons" style so to show the alpha-elices structure of MB. The heme is represented as sticks and the Iron atom in the middle as a yellow sphere. The ball-and-stick model is the CO molecule, bound to the Iron atom.

# Experiment using time-resolved X-ray crystallography
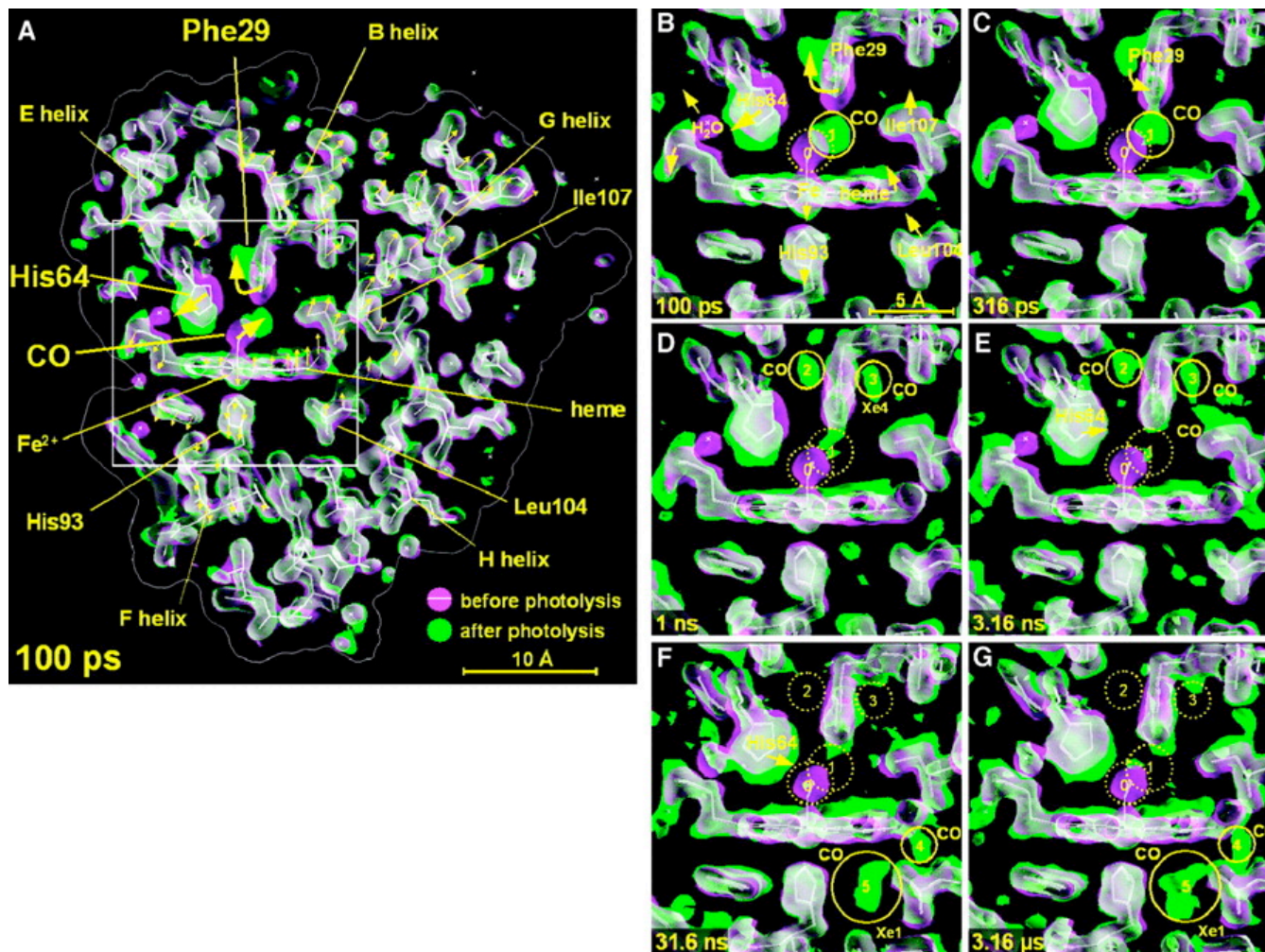
Time-resolved X-ray diffraction on photolyzed Mb.

Given the crystal with the bound CO, a laser flash to break the bond FE-CO. Then the diffraction patterns using a a time-delayed X-ray pulse on the crystal (picosecond Laue crystallography). Magenta is the pre-photolysis structure (CO-bound), and green are the structures at delayed times. When these are similar to pre-photolysis crystal structuresthey are colored white.

From F.Schotte *et al.* Science **300**: 1944–1947 (2003).

Free energy map of CO relative to binding site (Iron)
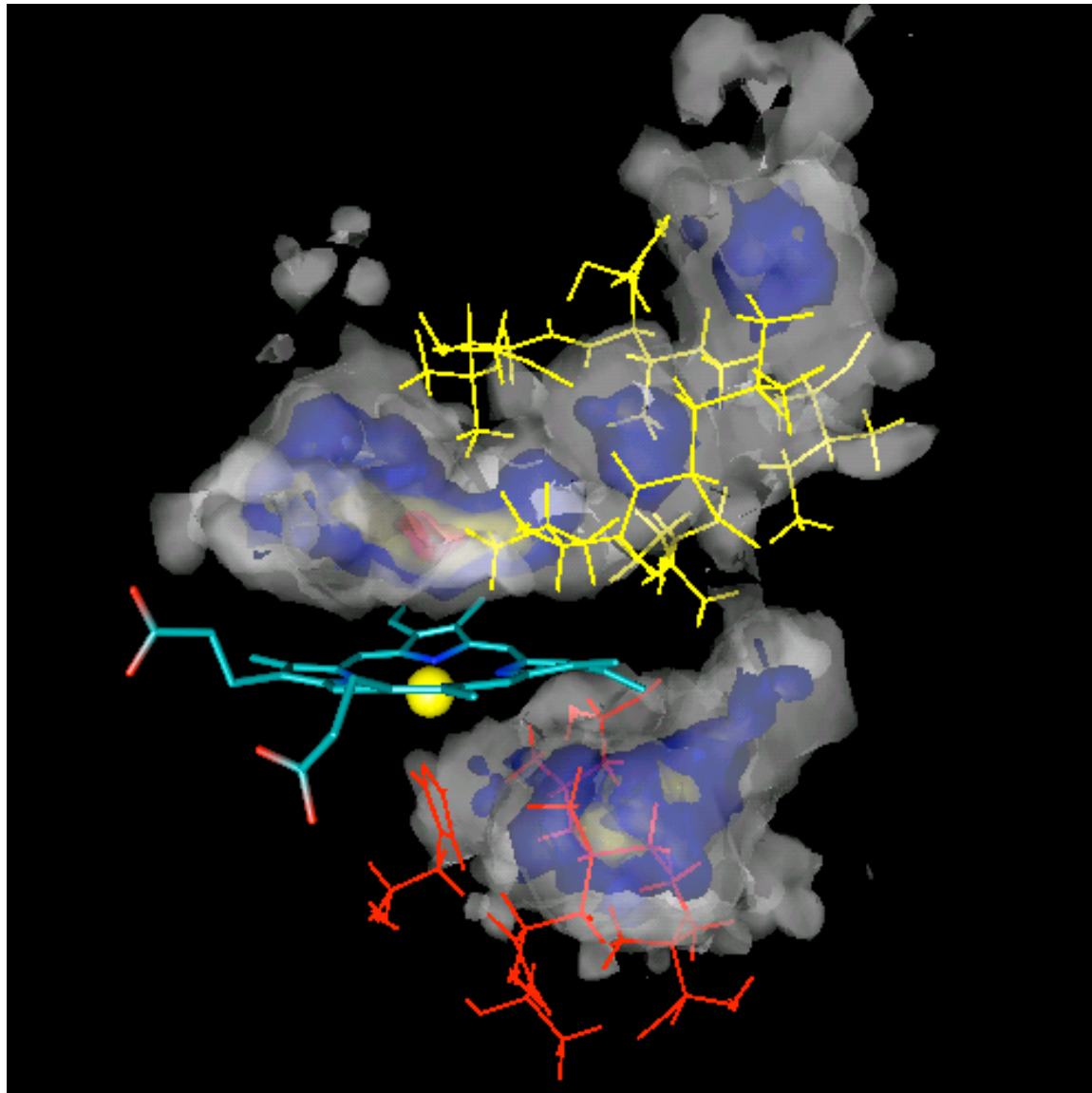
Sampling by one-sweep 5 different trajectories. In red are shown the residues that are used to define the Xe1 cavity (i.e. they surround the cavity), in yellow same for Xe4. The time-scale of the simulations is 100 picosecond; $k_B\bar{T} = 7k_BT$.

Explicit integrators for stiff ODEs:

C.W. Gear and I.G. Kevrekidis, "Projective methods for stiff differential equations: problems with gaps in their eigenvalue spectrum," *SIAM J. Sci. Comp.* **24**(4):109-110 (2003).

K. Eriksson, C. Johnson, and A. Logg, "Explicit Time-stepping for stiff ODEs," *SIAM J. Sci. Comp.* **25**(4):1142-1157 (2003).

V. I. Lebedev and S. I. Finogenov, "Explicit methods of second order for the solution of stiff systems of ordinary differential equations", *Zh. Vychisl. Mat. Mat Fiziki*, **16**:895–910 (1976).

A. Abdulle, "Fourth order Chebychev methods with recurrence relations", *SIAM J. Sci. Comput.* **23**:2041–2054 (2002).

HMM-like multiscale integrators for stochastic systems

E. Vanden-Eijnden, "Numerical techniques for multiscale dynamical systems with stochastic effects", *Comm. Math. Sci.*, **1**: 385–391 (2003).

W. E, D. Liu and E. Vanden-Eijnden, "Analysis of Multiscale Methods for Stochastic Differential Equations," *Comm. Pure App. Math.* **58**: 1544–1585 (2005).

E. Vanden-Eijnden, On HMM-like integrators and projective integration methods for systems with multiple time scales, *Comm. Math. Sci.* (2007).

HMM and Equation Free:

W. E, B. Engquist, The heterogeneous multi-scale methods, *Comm. Math. Sci.* **1** (2003), 87–133.

Kevrekidis, I. G.; Gear, C. W.; Hyman, J. M.; Panagiotis, G. K.; Runborg, O.; Theodoropoulos, C. Equation-Free, Coarse-Grained Multiscale Computation: Enabling Microscopic Simulators to Perform System-Level Analysis. *Comm. Math. Sci.* **1**: 87–133 (2003).

Applications to free energy calculations:

W. E and E. V.-E, *Metastability, conformation dynamics, and transition pathways in complex systems*, in: "Multiscale Modelling and Simulation" (S. Attinger and P. Koumoutsakos eds.) LNCSE **39**, Springer (2004).

G. Ciccotti, R. Kapral, and E. V.-E., "Blue Moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics," ChemPhysChem **6**, 1809–1814 (2005).

G. Ciccotti, T. Lelièvre, and E. V.-E. "Projection of diffusions on submanifolds: Application to mean force computation," to appear in: Comm. Pure App. Math.

E. V.-E. and G. Ciccotti, "Second-order integrators for Langevin equations with holonomic constraints," Chem. Phys. Lett. **429**, 310–316 (2006).

L. Maragliano and E. V.-E., "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations," Chem. Phys. Lett. **426**, 168–175 (2006).