Heat bath models for nonlinear partial differential equations

Controlling statistical properties in extended dynamical systems

Ben Leimkuhler
University of Edinburgh

Funding: NAIS (EPSRC) & NWO

MIGSEE 2015

This Talk

Averaging against a Gibbs invariant state

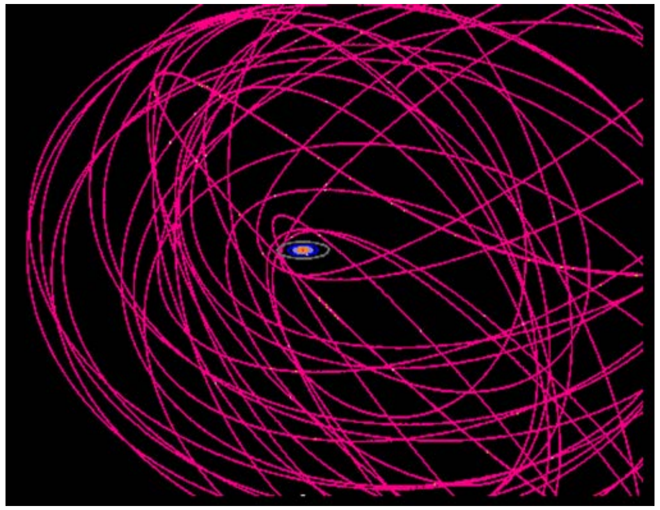
Degenerate dynamics,
Thermostats

Ergodicity

Examples

P. Simon de Laplace, *A Philosophical Essay on Probabilities* (from a lecture in 1795)

"…given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it, nothing would be uncertain…"



D. Ruelle, ***Chance and Chaos:***

Most systems are fundamentally chaotic and the chaotic nature of the dynamical system is related to a statistical interpretation of the associated physical system.

The Problem

Given a model system, typically deterministic, that does not do what we want it to, say

$$\dot{x} = f(x)$$

we wish to

- modify to preserve a **prescribed invariant distribution** (normally using stochastics)
- control **perturbation of dynamics** (or averaged dynamics)
- Design accurate **numerical methods**

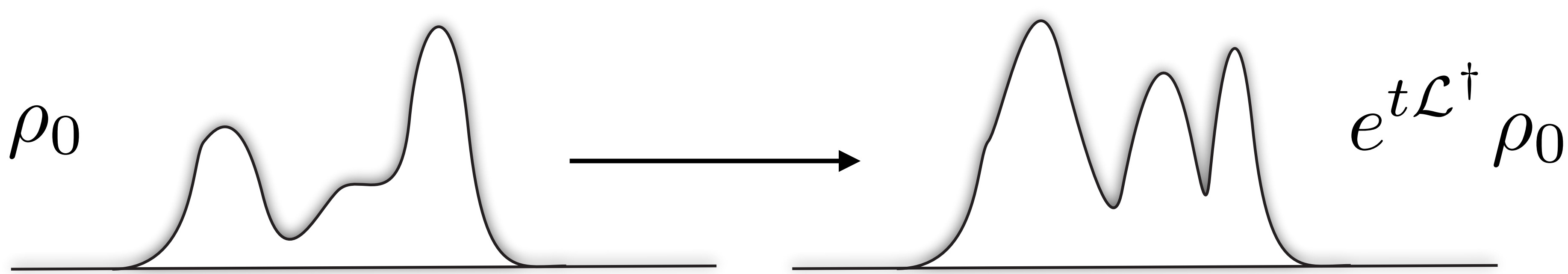
Continuity Equation

dynamics

$$\dot{x} = f(x)$$

continuity equation

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho f) =: \mathcal{L}^\dagger \rho$$

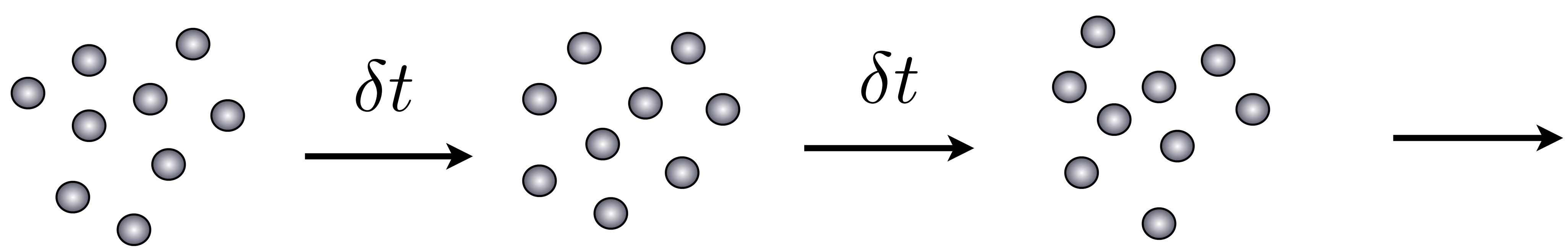


Molecular Dynamics

dynamics of nuclei of a collection of atoms treated classically

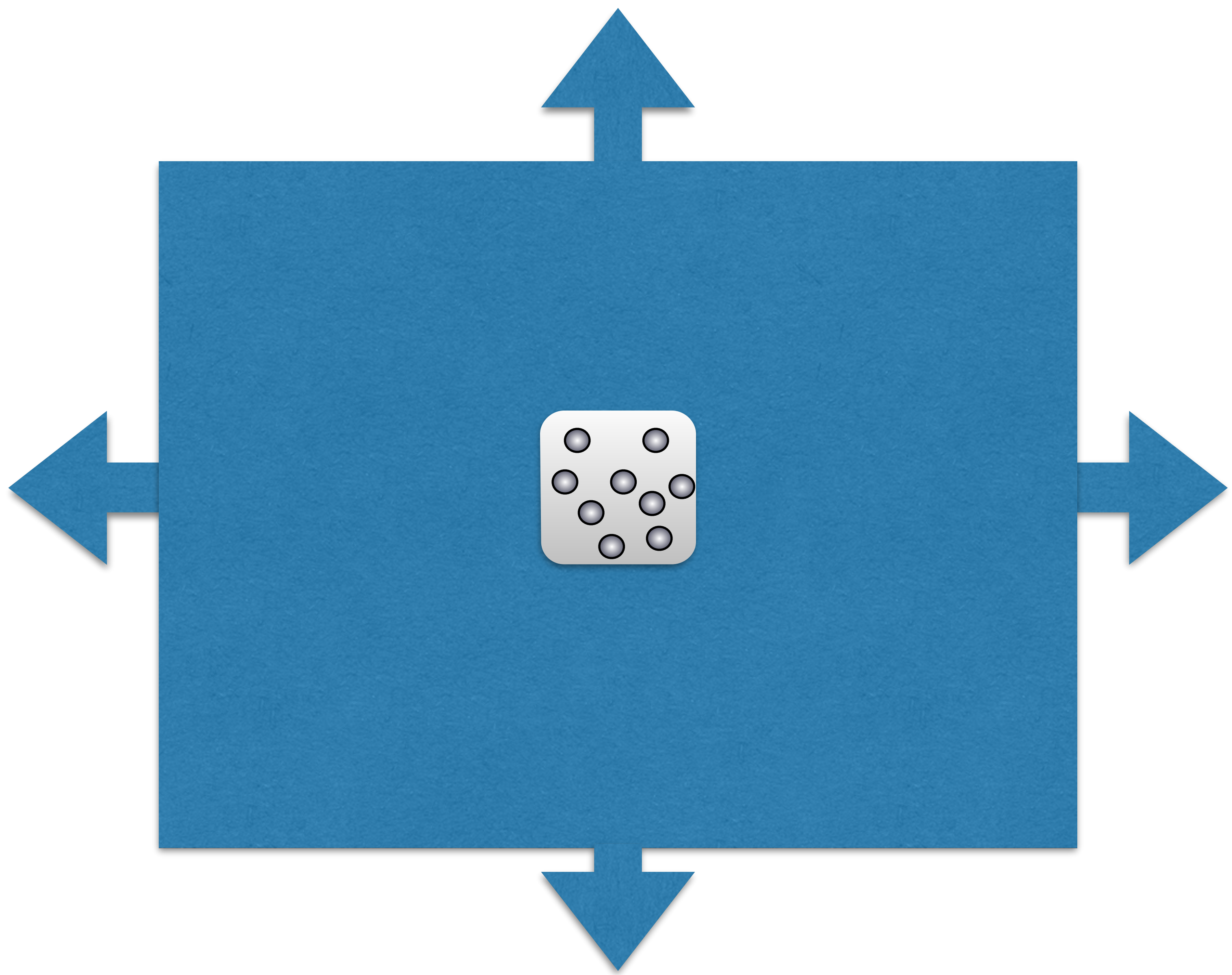
Normally, one solves something near to Newton's equations of motion using a numerical method.

$$M\ddot{q} = F$$



Canonical Ensemble

Describes the situation of a small number of atoms in contact with a much larger reservoir at fixed energy



J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, Scribner, 1902.

The distribution represented by

$$\eta = \log P = \frac{\psi - \epsilon}{\Theta}, \tag{90}$$

Energy

or

$$P = e^{\frac{\psi - \epsilon}{\Theta}}, \tag{91}$$

where Θ and ψ are constants, and Θ positive, seems to represent the most simple case conceivable, since it has the property that when the system consists of parts with separate energies, the laws of the distribution in phase of the separate parts are of the same nature,— a property which enormously simplifies

Note: rigorous derivations are very difficult!


The quick brown fox jumps over the lazy
dog and feels as if he were in the seventh
heaven of typography together with Her-
mann Zapf, the most famous artist of the

Ensemble Averages

Hamiltonian (energy function)

$$H(q,p) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + U(q_1, q_2, \ldots, q_N)$$

Average of an Observable f

Gibbs
(canonical) density ρ_β

$$\langle f(q,p)\rangle_\beta = Z^{-1} \int_D f(q,p) e^{-\beta H(q,p)} \mathrm{d}q\mathrm{d}p$$

(or some other weighted average over states)

$$\langle f(q,p)\rangle_\beta = Z^{-1}\int_D f(q,p)e^{-\beta H(q,p)}\mathrm{d}q\mathrm{d}p$$

Dynamics based methods

Generate lots of positions (or positions and momenta) which are consistent with the target distribution

Average the values of the observable with respect to this collection.

$$\langle f(q,p)\rangle_\beta \approx \frac{1}{\nu}\sum_{n=1}^{\nu} f(q^n,p^n)$$

Ergodicity - convergence as $\nu \to \infty$

How to get molecular dynamics to sample (ergodically) from the canonical (or other) ensemble?

$$\rho_\beta \propto \exp(-H/k_B T), \qquad \beta = (k_B T)^{-1}$$

Deterministic Hamiltonian dynamics does preserve this ensemble, but it is not ergodic.

Idea of a '**thermostat**': find a modified system whose invariant distribution is uniquely ρ_β

Normally we need to use **stochastic noise** to achieve this...Stochastic Differential Equations (SDEs)

Ex: Brownian dynamics

$$\mathrm{d}X = -\nabla U(X)\mathrm{d}t + \sqrt{2}\mathrm{dW}$$

invariant measure: $\rho_{\mathrm{eq}} = e^{-U}$

under certain (mild) conditions on U
unique steady state of the Fokker-Planck equation:

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^*_{\mathrm{BD}}\rho$$

$$\mathcal{L}^*_{\mathrm{BD}}\rho = -\nabla \cdot [\rho \nabla U] + \Delta \rho$$

J. Perrin, ***Atoms***, van Nostrand, 1916, esp. Chapters 3 and 4

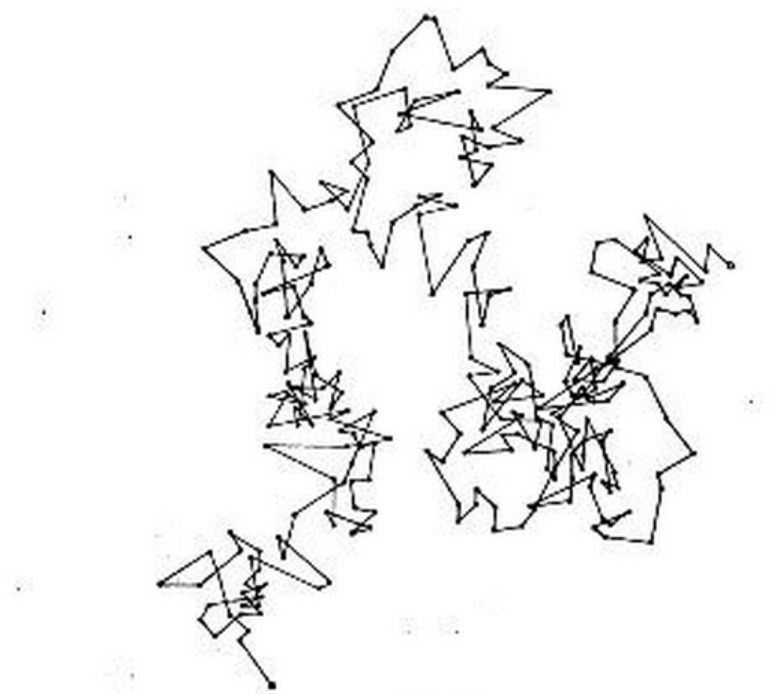
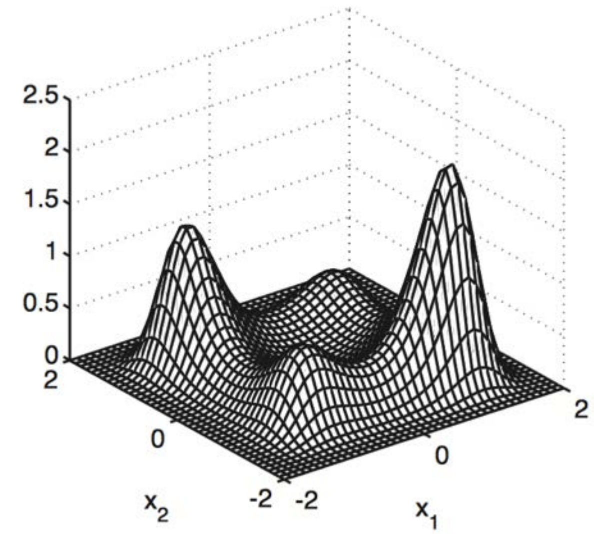


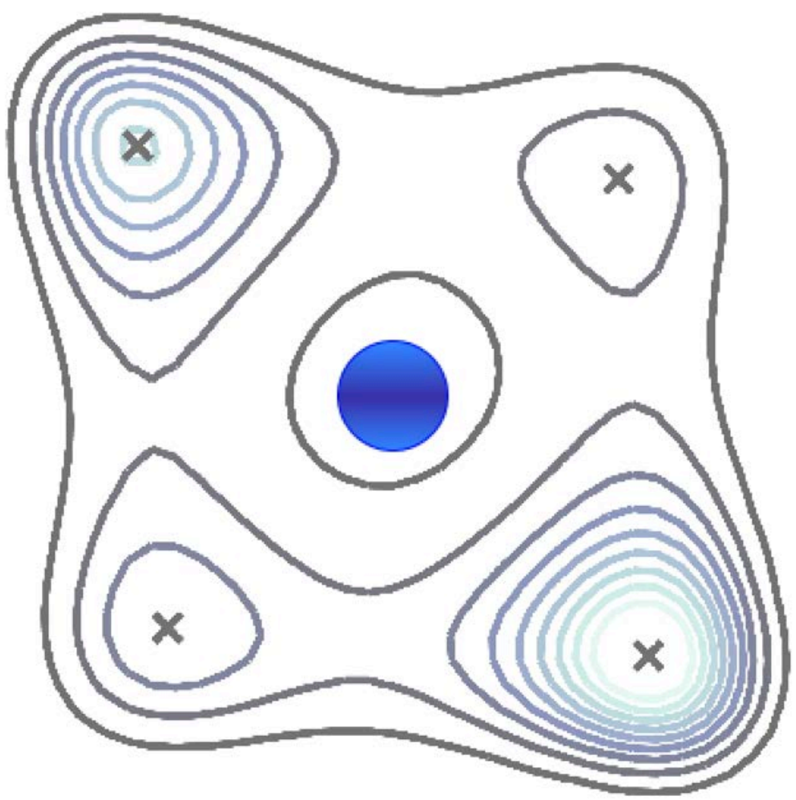
FIG. 9.

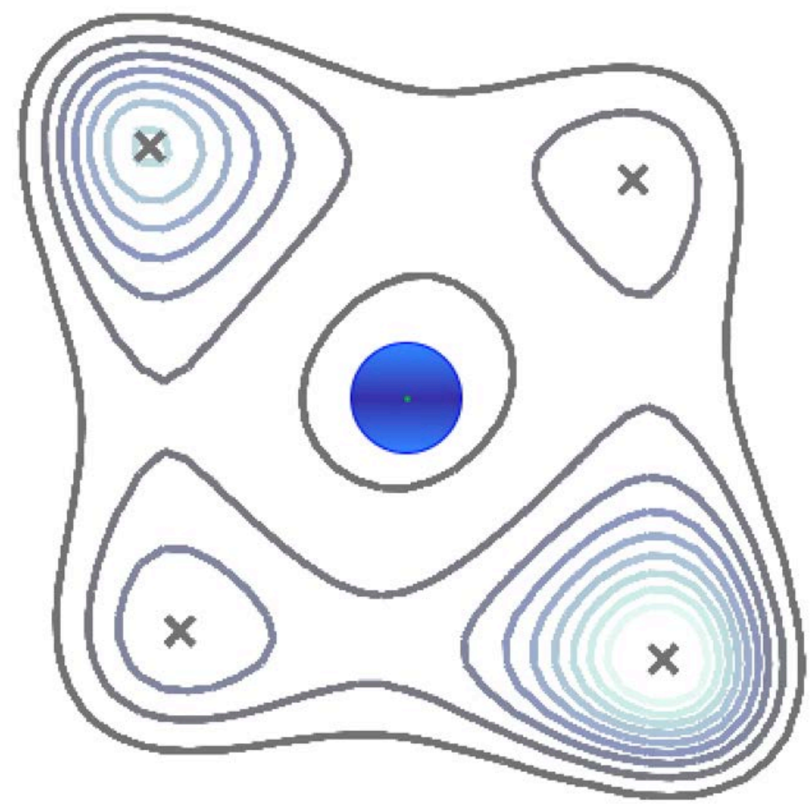
Brownian dynamics is a non-inertial stochastic model…the particles don't carry momentum.

In practice it is very useful (and more realistic) for the atoms to have momentum. This allows the paths to model dynamical paths. In many cases this enhances the rate of exploration of phase space.

Ex: Rare Event Sampling (java by Gilles Vilmart, Geneva)







Brownian Dynamics

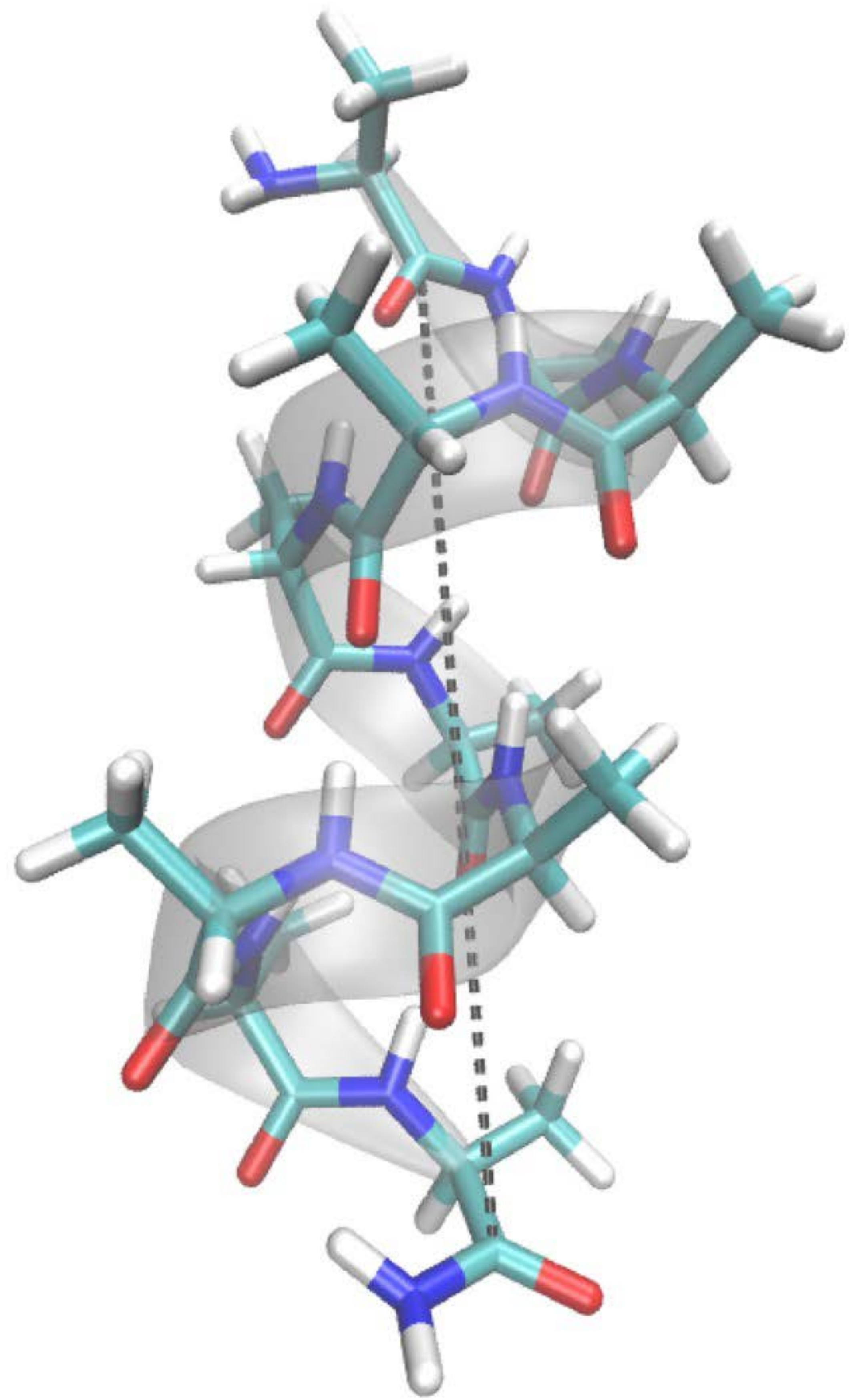
Langevin (Inertial) Dynamics

“decalanine”

10 amino acid chain

Easily knots to intermediate
states which are
hard to escape

“metastable” landscape



Ex: Langevin Dynamics

$$\mathrm{d}q = M^{-1}p\,\mathrm{d}t$$

$$\mathrm{d}p = [-\nabla U(q) - \gamma p]\mathrm{d}t + \sqrt{2\gamma k_B T}M^{1/2}\mathrm{d}W$$

γ = friction parameter

Fokker-Planck Operator:

$$\mathcal{L}^*_{\mathrm{LD}}\eta = -(M^{-1}p)\cdot\nabla_q\eta + \nabla U\cdot\nabla_p\eta + \gamma\nabla_p\cdot(p\eta) + \gamma k_B T\Delta\eta$$

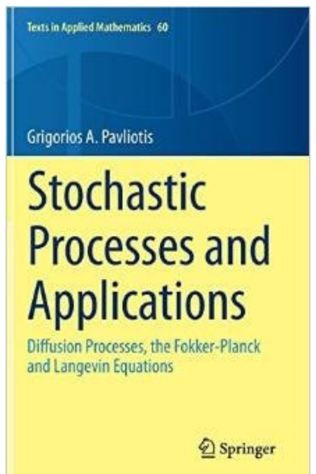
mass weighted partial Laplacian

Preserves Gibbs distribution:

$$\mathcal{L}^*_{\mathrm{LD}}\rho_\beta = 0$$

Reading List

G. Pavliotis, ***Stochastic Processes and Applications,*** Springer, 2014

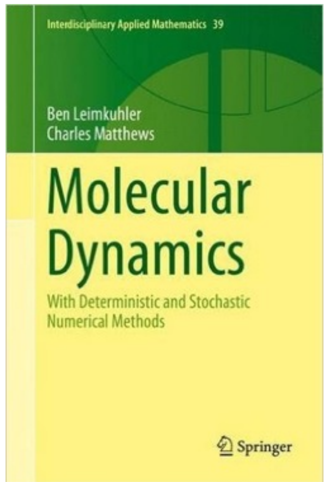


Readable introduction to the modern treatment of SDEs, especially Brownian and Langevin dynamics and Fokker-Planck equations

---

B. Leimkuhler, C. Matthews, ***Molecular Dynamics*****,** Springer, 2015

A mathematical introduction to molecular dynamics,
including both deterministic and stochastic models, and numerical methods



Ergodic Properties of SDEs

From L.&Matthews, drawing on work of Stuart, Mattingley…

Definition 6.1 (Hörmander's Condition). Let $\mathcal{U} \subset \mathbb{R}^m$ be open. The C^∞ vector fields $\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_r : \mathcal{U} \to \mathbb{R}^m$ are said to satisfy Hörmander's condition at $\boldsymbol{z} \in \mathcal{U}$ if the vector space spanned by the collection of iterated Lie brackets

$$\boldsymbol{b}_0, \boldsymbol{b}_1 \ldots, \boldsymbol{b}_r, [\boldsymbol{b}_0, \boldsymbol{b}_1], [\boldsymbol{b}_0, \boldsymbol{b}_2], \ldots, [\boldsymbol{b}_{r-1}, \boldsymbol{b}_r], [\boldsymbol{b}_0, [\boldsymbol{b}_0, \boldsymbol{b}_1]] \ldots$$

(evaluated at $\boldsymbol{z}$) is $\mathbb{R}^m$.

Theorem 6.2 (Smoothness of invariant measures). *Let an SDE be defined on $\mathbb{R}^m$:*

$$\mathrm{d}\mathbf{X} = \boldsymbol{b}_0(\mathbf{X})\mathrm{d}t + \sum_{i=1}^{r} \boldsymbol{b}_i \mathrm{dW}_i,$$

where $\boldsymbol{b}_0$ is C^∞ and $\boldsymbol{b}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ satisfy Hörmander's condition for all $\mathbf{X} \in \mathbb{R}^m$. Suppose there is a smooth invariant measure μ_* with associated density ρ_* such that $\mu_*(O) > 0$ for any open set $O \subset \mathbb{R}^m$. Suppose further that there is a smooth function ϕ : such that (a) $\phi > 0$, (b) $\phi(\boldsymbol{z}) \to +\infty$ as $\|\boldsymbol{z}\| \to \infty$, and (c) the estimate

$$\mathcal{L}\phi \le d - \alpha\phi$$

holds for some constants $\alpha, d > 0$, where $\mathcal{L}$ is the generator associated to the SDE. (ϕ is referred to as a Lyapunov function.) Then ρ_* is the unique distributional solution of

$$\mathcal{L}^\dagger \rho = 0.$$

Langevin dynamics

[Stuart, Mattingley, Higham '02]

$$H = p^2/2 + U(x)$$
$$f(x) = -U'(x)$$

$$\mathrm{d}x = p\mathrm{d}t$$

$$\mathrm{d}p = f(x)\mathrm{d}t - p\mathrm{d}t + \sqrt{2}\mathrm{dW}$$

✔ HC:

$$\boldsymbol{b}_0 = (p, f(x) - p); \qquad \boldsymbol{b}_1 = (0, 1)$$

$$[\boldsymbol{b}_0, \boldsymbol{b}_1] = -\begin{bmatrix} 0 & 1 \\ f'(x) & -1 \end{bmatrix} \boldsymbol{b}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

✔ invariant measure:

$$\rho_* = e^{-H}$$

positive measure on open sets

✔ Lyapunov function

Therefore, Langevin dynamics is ergodic

Properties of $\mathcal{L}_{\mathrm{LD}}$

- Discrete Spectrum, Spectral Gap
- Hypocoercive (but degenerate in the limit of small friction)
- Ergodic

$$\lim_{t\to\infty} \langle f, \rho(\cdot, t)\rangle = \langle f, \rho_\beta\rangle$$

- Exponential convergence in an appropriate norm

$$\|e^{t\mathcal{L}}\|_\bullet \leq K e^{-\lambda_\gamma t} \quad \lambda_\gamma > 0$$

Invariant measures and discretization

Langevin dynamics Fokker-Planck operator:

$$\mathcal{L}^*_{\mathrm{LD}}\eta = -(M^{-1}p)\cdot\nabla_q\eta + \nabla U\cdot\nabla_p\eta + \gamma\nabla_p\cdot(p\eta) + \gamma k_B T\Delta\eta$$

Fokker-Planck equation: $$\frac{\partial\rho}{\partial t} = \mathcal{L}^*\rho \quad \mathcal{L}^*\rho_{\mathrm{eq}} = 0$$

Numerical discretization:

$$\hat{\mathcal{L}}^*_{\delta t} \neq \mathcal{L}^* \qquad \hat{\mathcal{L}}^*_{\delta t}\hat{\rho}_{\delta t} = 0$$

$$\hat{\rho}_{\delta t} \neq \rho_{\mathrm{eq}}$$

highly degenerate diffusions

L., Phys Rev E, 2010

$$\mathrm{d}X = f(X)\mathrm{d}t + g(X,\Xi)\mathrm{d}t$$

$$\mathrm{d}\Xi = g(X,\Xi)\mathrm{d}t \boxed{-\gamma\Xi\mathrm{d}t + \sqrt{2\gamma}\mathrm{dW}}$$

OU

- design to preserve extended Gibbs distribution

$$\tilde{\rho} = \rho_*(X)e^{-\Xi^2/2}$$

- 'weak' coupling to stochastic perturbation

Nose-Hoover

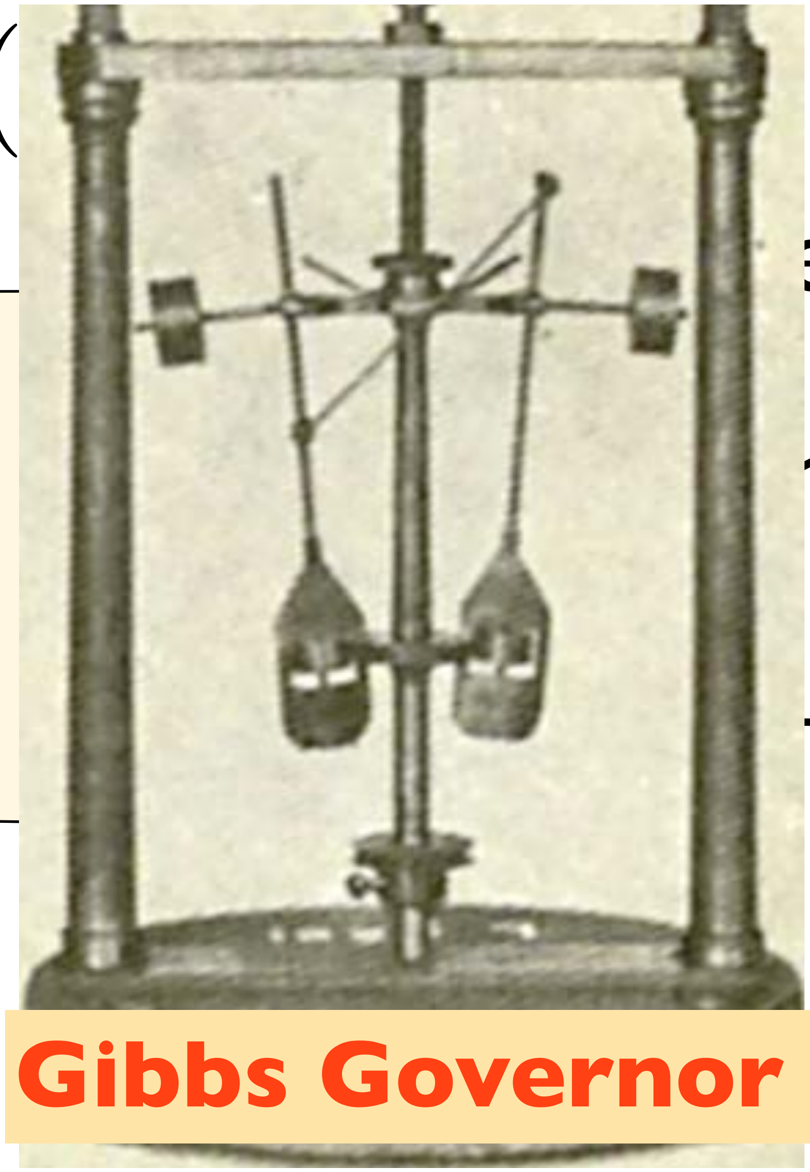
a "Virtual Gibbs Governor"

$$H = \frac{p^2}{2} + U($$

$$\dot{q} =$$

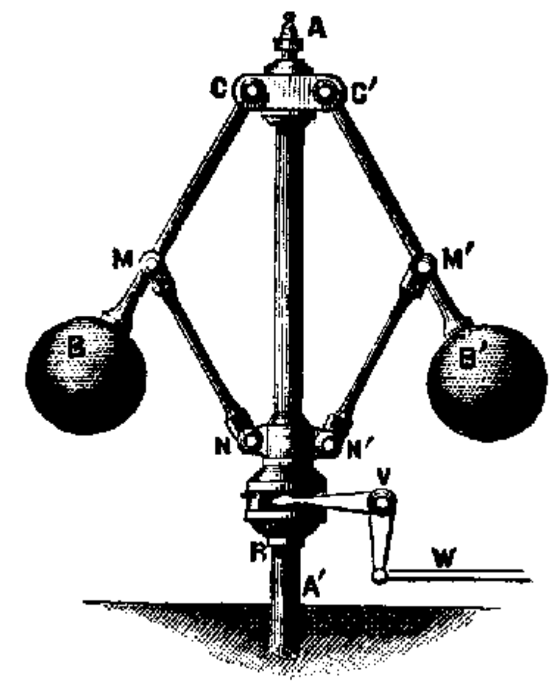
$$\dot{p} = \quad - \xi p$$

$$\mu\dot{\xi} =$$



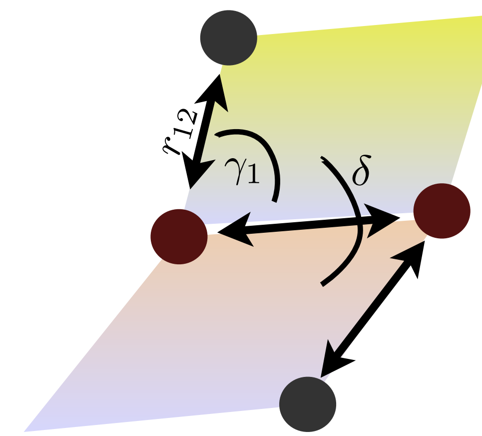
ewtonian dynamics

'governor



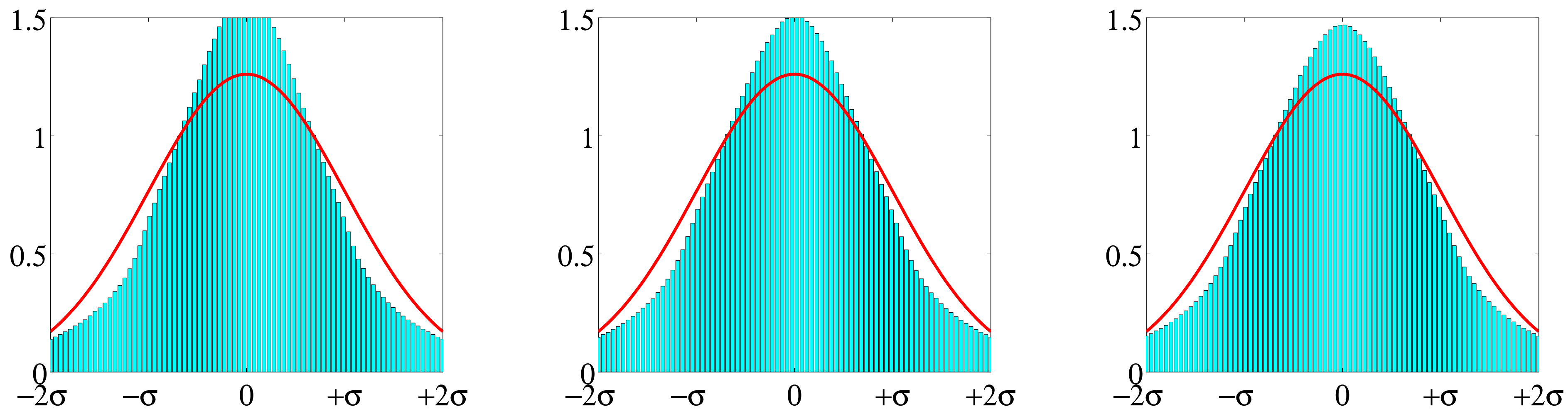
preserves $\tilde{\varrho} = \exp(-\beta H - \alpha\xi^2/2)$

Ex: NH for Butane



Error due to discretization
+ lack of ergodicity

lack of ergodicity only



decreasing stepsize

Legoll, Luskin and Moeckel 2006, 2009:
Nonergodicity of NH in large mu limit

Nosé-Hoover-Langevin

L., Noorizadeh and Theil, J Stat Phys 2009

NH + OU

$$\mathrm{d}q = p\mathrm{d}t$$

$$\mathrm{d}p = -U'(q)\mathrm{d}t - \xi p\mathrm{d}t$$

$$\mathrm{d}\xi = (p^2 - \theta)\mathrm{d}t - \gamma\xi\mathrm{d}t + \sqrt{2\gamma}\mathrm{dW}$$

Clearly preserves $\rho = e^{-H/\theta} \times e^{-\xi^2}$

but... **is it ergodic?**

$$\mathrm{d}\boldsymbol{z} = \boldsymbol{F}(\boldsymbol{z})\mathrm{d}t + \xi \boldsymbol{G}(\boldsymbol{z})\mathrm{d}t,$$

$$\mathrm{d}\xi = g(\boldsymbol{z})\mathrm{d}t - \gamma\xi\mathrm{d}t + \sigma \mathrm{dW},$$

Prop: Let the given system preserve

$$e^{-\beta H} \times e^{-\xi^2/2}$$

Suppose the system is defined on $\mathcal{M} \times R$
where $\mathcal{M}$ is a smooth compact submanifold
Further suppose that the Lie algebra spanned by
$\boldsymbol{F}, \boldsymbol{G}$ spans $\mathrm{T}\mathcal{M}$ at every point of $\mathcal{M}$

Then the given system is ergodic on $\mathcal{M}$

NHL on a harmonic system

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{p} \\ -\boldsymbol{A}\boldsymbol{q} \end{bmatrix}, \qquad \boldsymbol{G} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{p} \end{bmatrix}$$

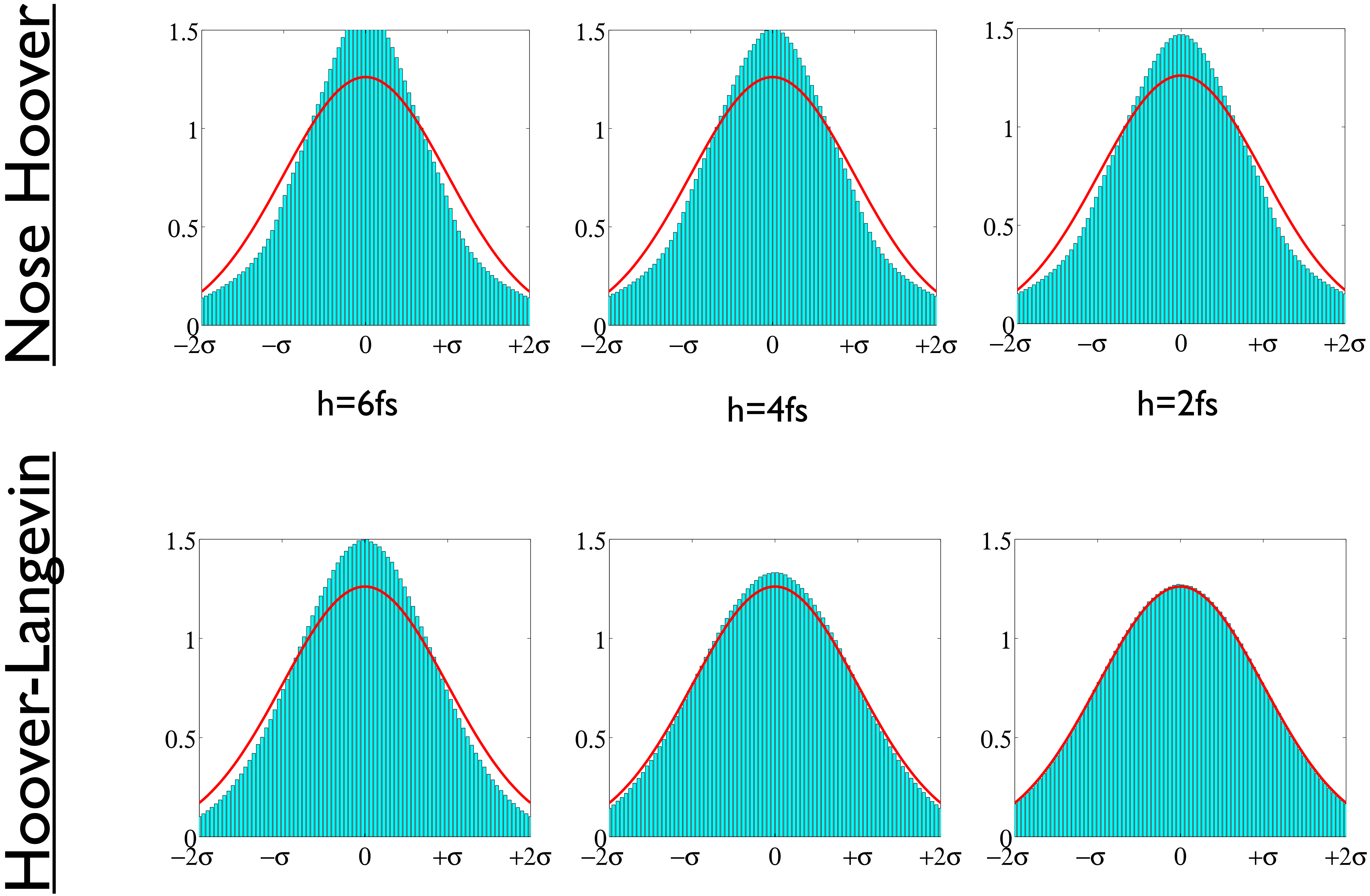
$S\{\boldsymbol{F}, \boldsymbol{G}\} =$ Lie algebra generated by $\boldsymbol{F}, \boldsymbol{G}$

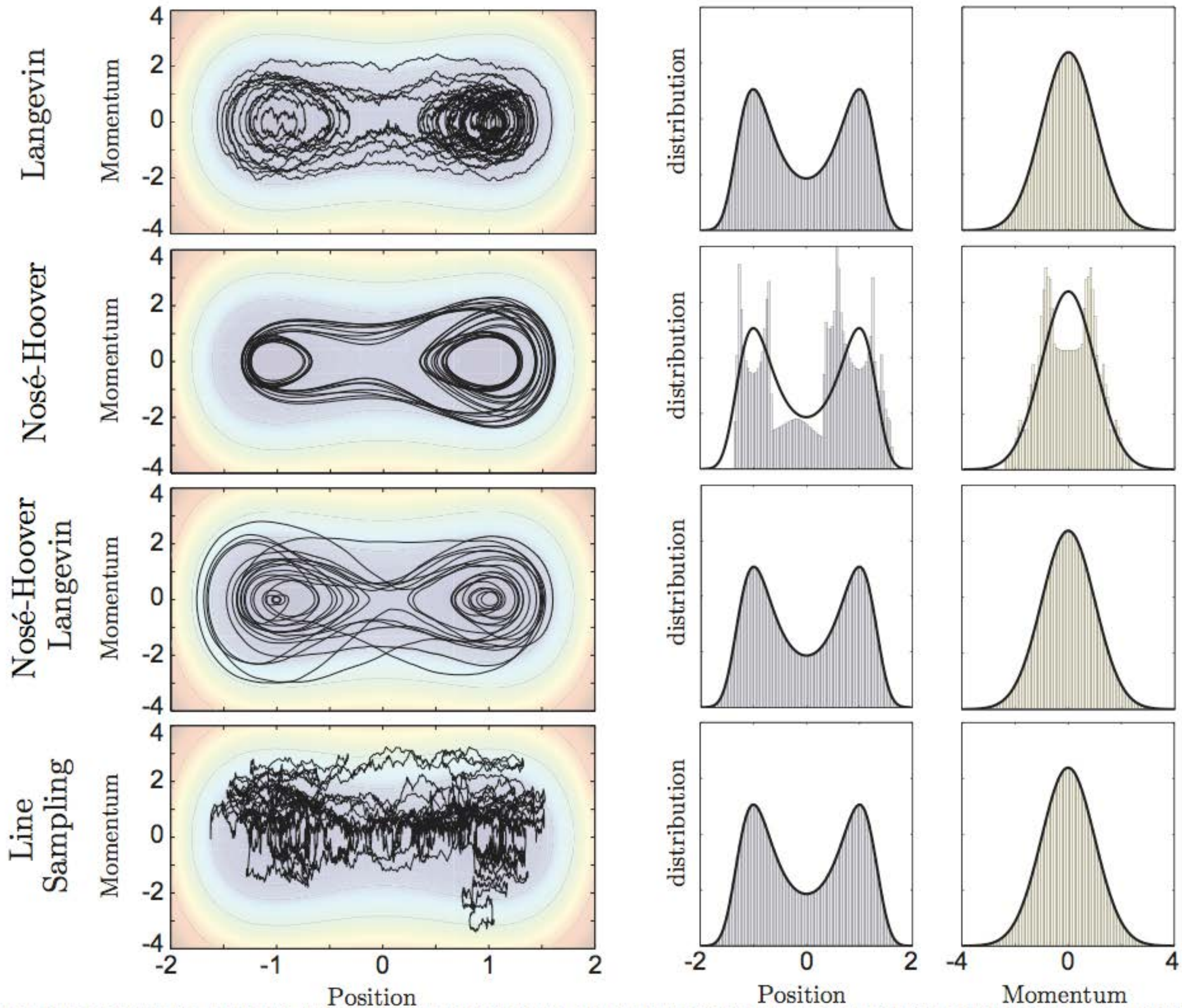
Prop:

$$\boldsymbol{C}_k = \begin{bmatrix} \boldsymbol{A}^{k-1}\boldsymbol{p} \\ \boldsymbol{A}^{k}\boldsymbol{q} \end{bmatrix} \in S\{\boldsymbol{F}, \boldsymbol{G}\}$$

$$\boldsymbol{D}_k = \begin{bmatrix} \boldsymbol{A}^{k}\boldsymbol{q} \\ -\boldsymbol{A}^{k}\boldsymbol{p} \end{bmatrix} \in S\{\boldsymbol{F}, \boldsymbol{G}\}$$

Butane Momentum Distributions



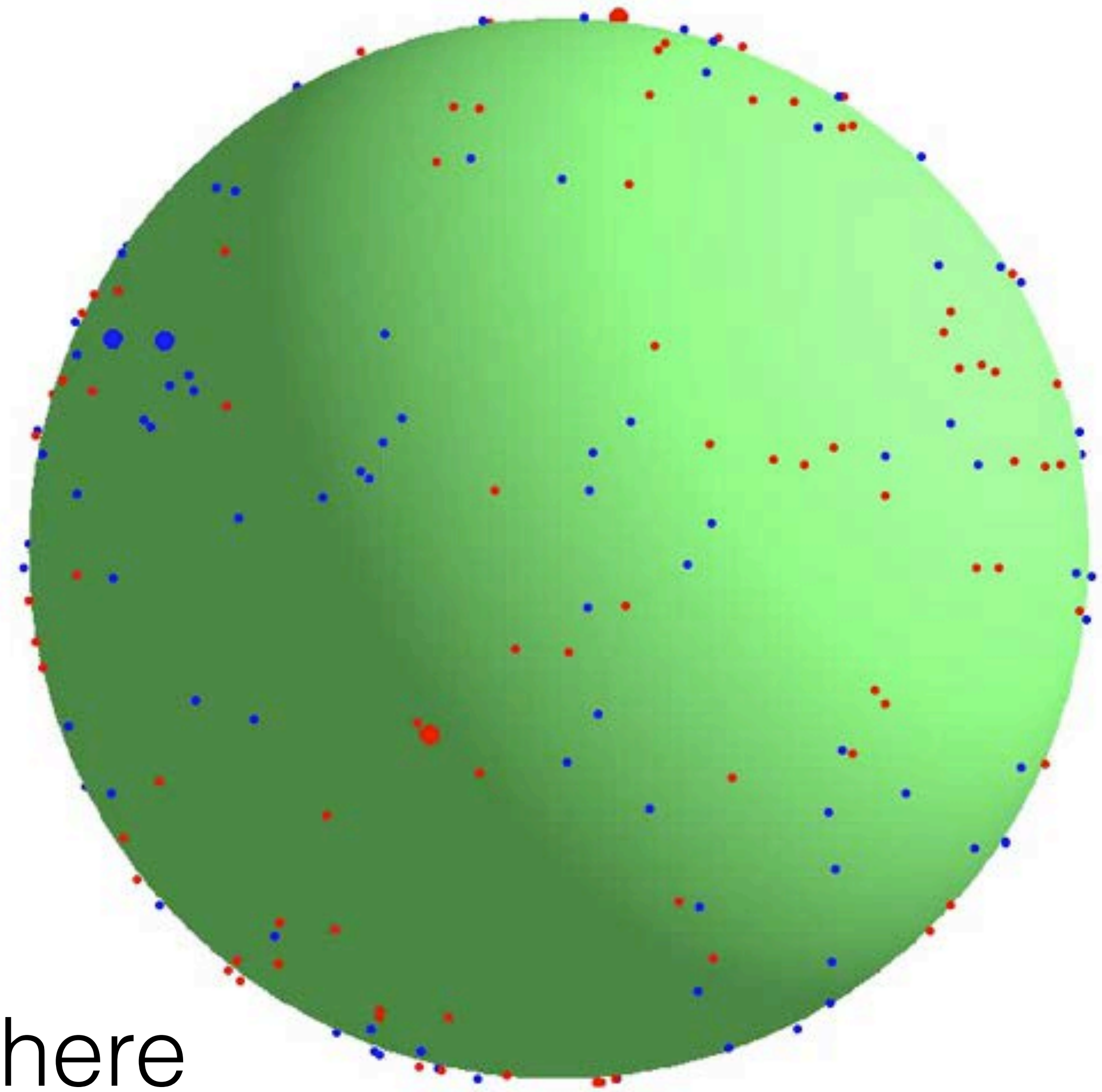


Langevin
Nosé-Hoover
Nosé-Hoover Langevin
Line Sampling
Momentum
distribution
Position
Momentum

Applications in Fluid Dynamics

Dubinkina, Frank and L. SIAM MMS 2011

Frank, L. & Myerscough 2014

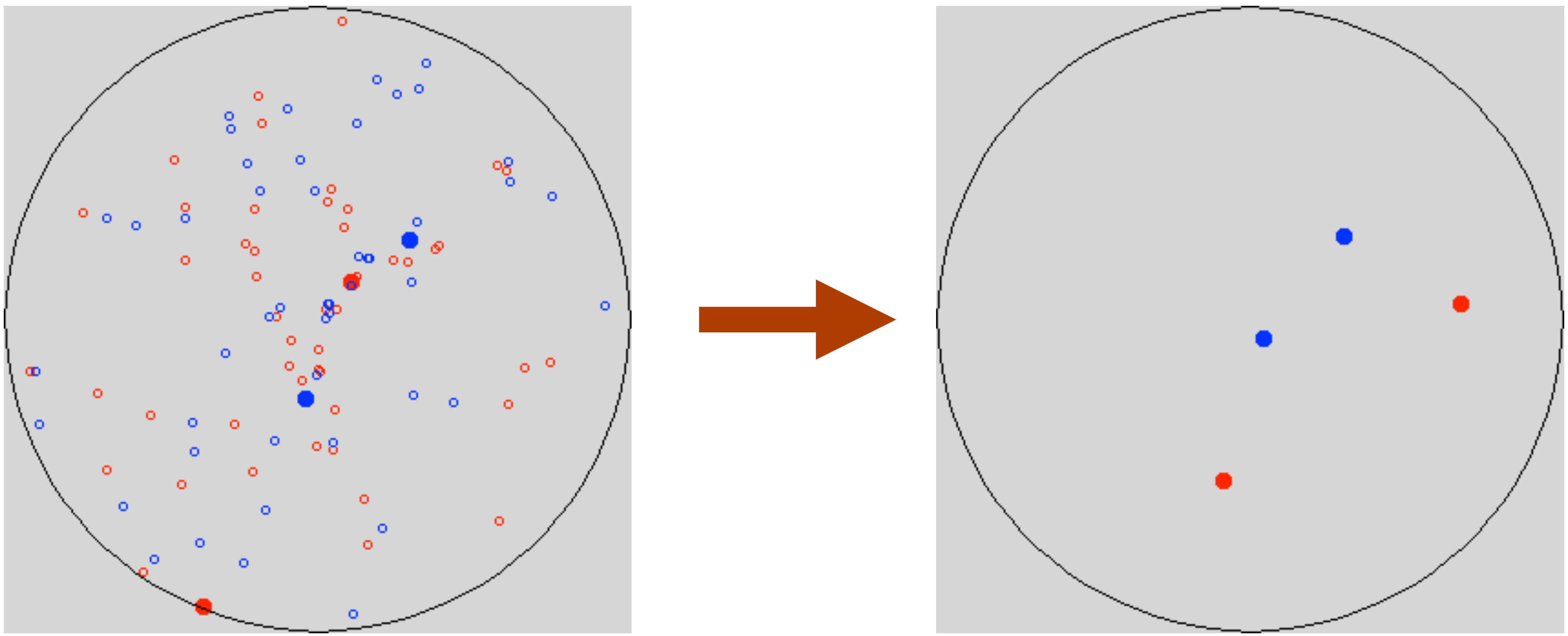


Vortex discretization
of QG PV formulation
of fluid dynamics on rotating sphere

Design controls to match prescribed (possibly evolving) distributional constraints

Reduced Model for Point Vortices

Restricted to Disk



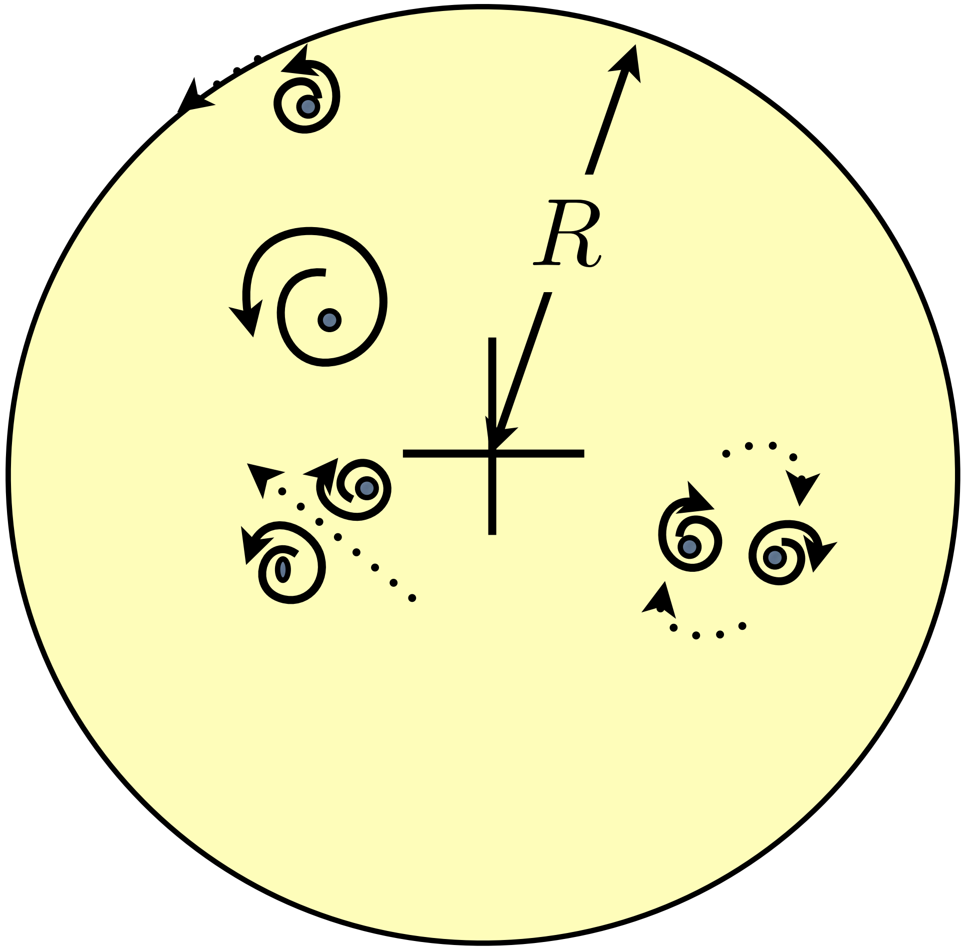
- Singular point vorticity field
- Heterogeneous: +/- orientation, strong/weak
- Reduce to only strong vortices, preserve their statistics
- Weak vortices → reservoir (canonical ensemble)

Reduced Model for Point Vortices

A point vortex model for N vortices in a cylinder

$$H = -\frac{1}{4\pi}\sum_{i<j}\Gamma_i\Gamma_j \ln(|r_i - r_j|^2)$$

$$+ \textit{ boundary terms}$$



$$\Gamma_i \dot{r}_i = J\nabla_{r_i} H$$

$$q(x,t) = \sum_i \Gamma_i \delta(x - r_i(t))$$

$$\Delta\psi(x,t) = q(x,t)$$

$$u(x,t) = \nabla^{\perp}\psi(x,t)$$

H unbounded under collisions

Choosing a Density

Maximize relative entropy

$$S[\rho] = -\int \rho \ln \rho/\pi$$

subject to constraints:

$$\mathscr{C}_1, \mathscr{C}_2, \ldots$$

E.g. prescribed bath energy variance,
defined averages of certain functions.

$$\rho = \pi \times \lambda_0 e^{\lambda_1 C_1 + \lambda_2 C_2 + \ldots + \lambda_m C_m}$$

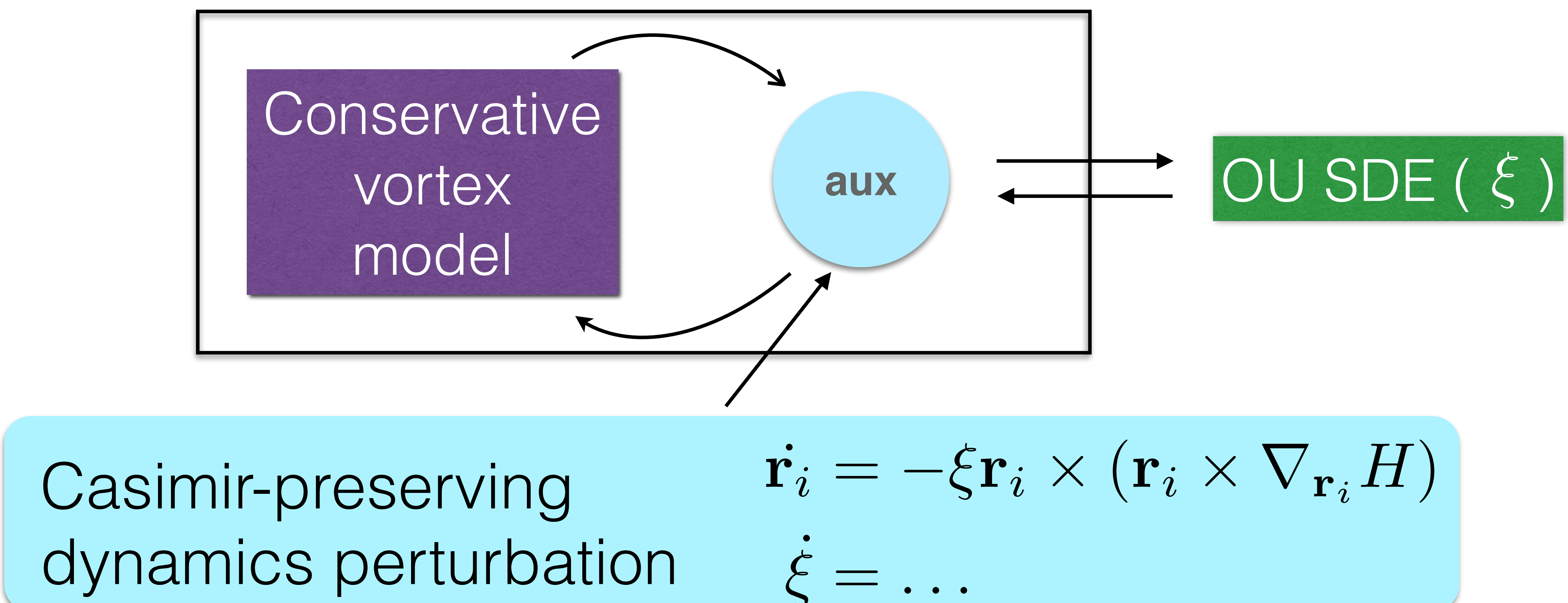
The coefficients may or may not be fixed/known…

Reduced Model for Point Vortices

Simulations by Bühler (2002): 4 strong, 96 weak vortices, sign indefinite. We established correspondence between Buhler's simulation and a model with 4 vortices+ a finite bath reservoir

$$\rho(X) \propto \exp(-\beta H(X) - \gamma H(X)^2)$$

Coupling method:



Numerical Study

Numerical comparisons with Bühler's simulation

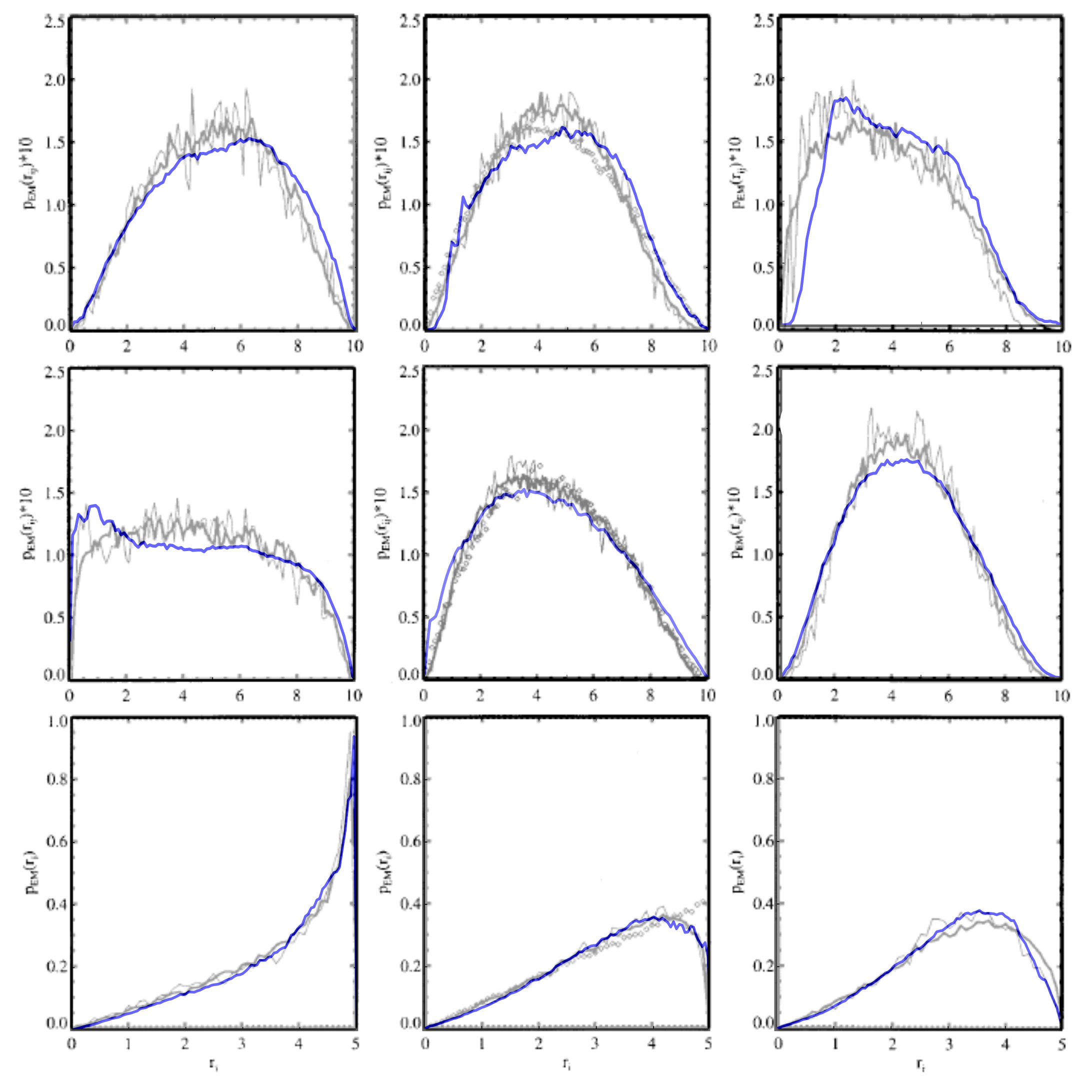
β	E	Coll.
$\beta>0$	E<0	+/-
$\beta<0$	E>0	++ --

$\beta > 0$ $\quad$ $\beta \approx 0$ $\quad$ $\beta < 0$

$|x_i - x_j|_{\pm\pm}$

$|x_i - x_j|_{\pm\mp}$

$|x_i|$



Burgers-KdV Equation

[Bajars, Frank and L., Nonlinearity, 2013]

<u>Rationale</u>

Discretized PDE models, e.g. Euler fluid equations, have a multiscale structure

Energy flows from low to high modes
"turbulent cascade"

Under discretization, the cascade is destabilized leading either to an artificial increase in energy at fine scales, or, if dissipation is added, an artificial decrease

Burgers/KdV model

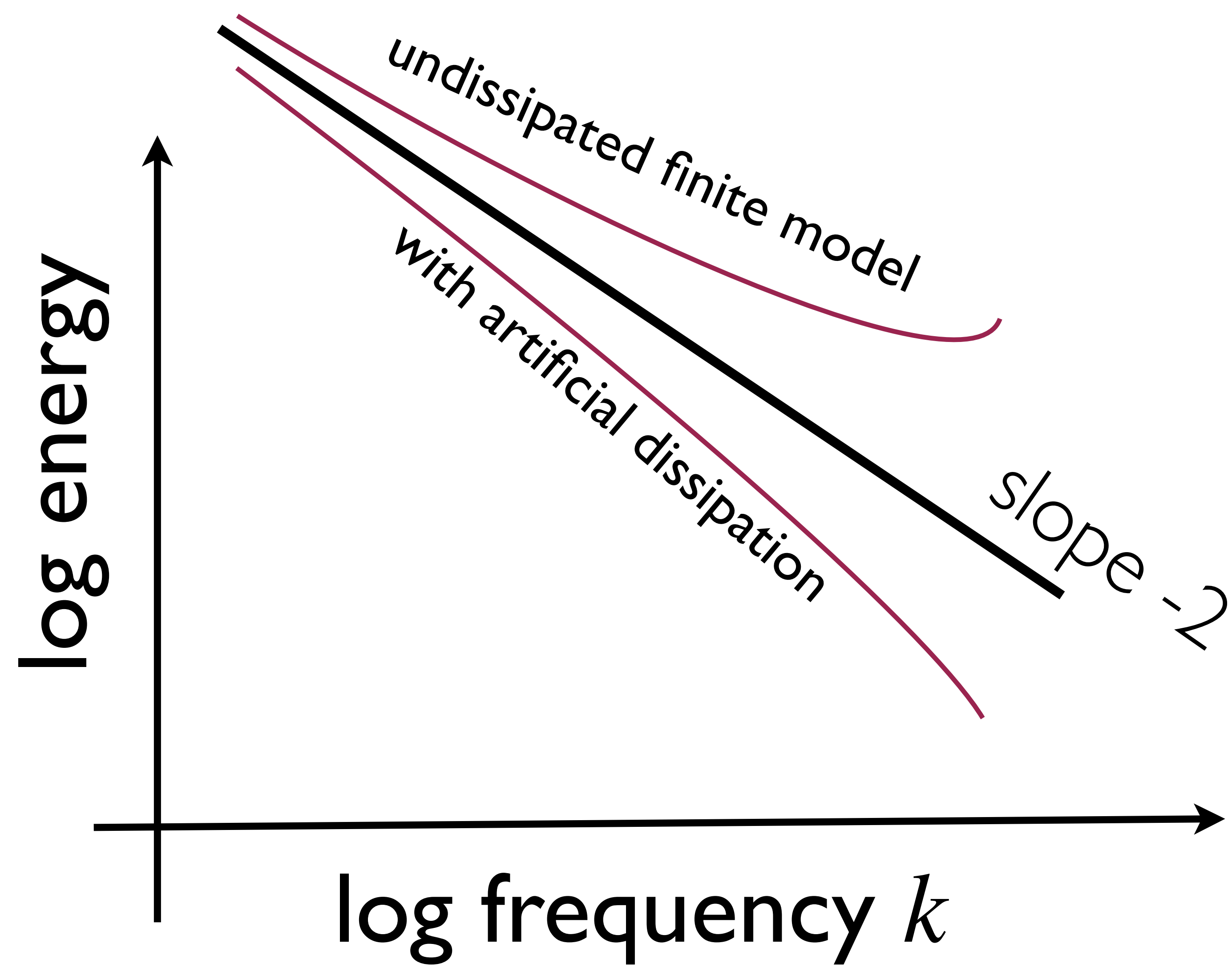
$$u_t + uu_x + \mu u_{xxx} = 0$$

Hamiltonian system $\quad u_t = -D_x \delta\mathcal{H}/\delta u$

energy $\quad \mathcal{H} = \int \frac{1}{6}u^3 - \frac{\mu}{2}u_x^2$

Energy Cascade in a Forced Burgers Eqn.

$$E_k \sim k^{-2}$$



Ergodicity

E. & Mattingley 2002

for truncated, incompressible Navier-Stokes,
stochastic forcing of two low wave number modes & viscous damping in the high modes ⇒ ergodicity

We have the OPPOSITE situation. We allow the system to be naturally driven at low modes and apply a thermostat only in the high wave numbers.

Moreover, all our calculations have to happen on the hypersphere which makes calculation of commutators extremely difficult.

Truncated, discrete model

$$\frac{\mathrm{d}\hat{u}_n}{\mathrm{d}t} = -\frac{in}{2}\left(\sum_{|n-m|\le N} \hat{u}_{n-m}\hat{u}_m\right) + in^3\mu\hat{u}_n$$

$$H = \frac{\pi}{3} \sum_{\substack{\ell+m+n=0 \\ |\ell|,|m|,|n|\le N}} \hat{u}_\ell \hat{u}_m \hat{u}_n - \mu\pi \sum_{|\ell|\le N} \ell^2 \hat{u}_\ell \hat{u}_\ell^*$$

$$\hat{u}_n = a_n + ib_n \qquad \hat{u}_\ell^* = \hat{u}_{-\ell}$$

Two other first integrals

total momentum M, total kinetic energy E

Mixed Probability Distribution

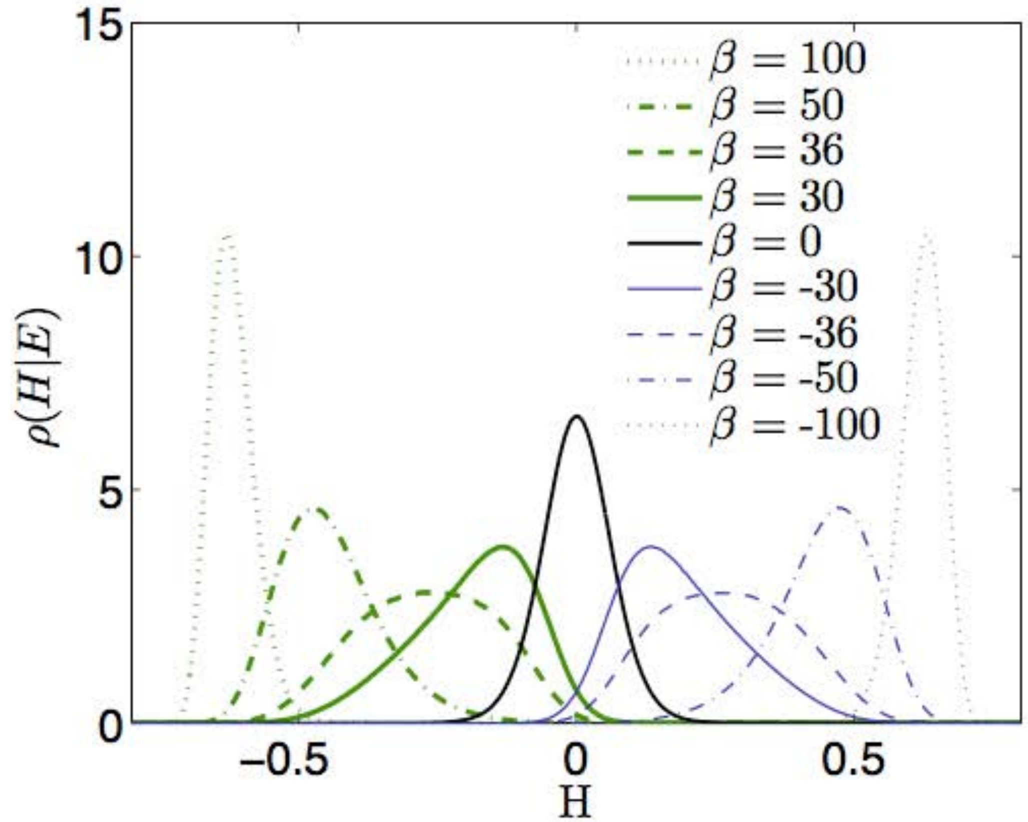
Proposed 'mixed' distribution:

$$\rho = \exp(-\beta H)\delta(E - E_0)\delta(M)$$

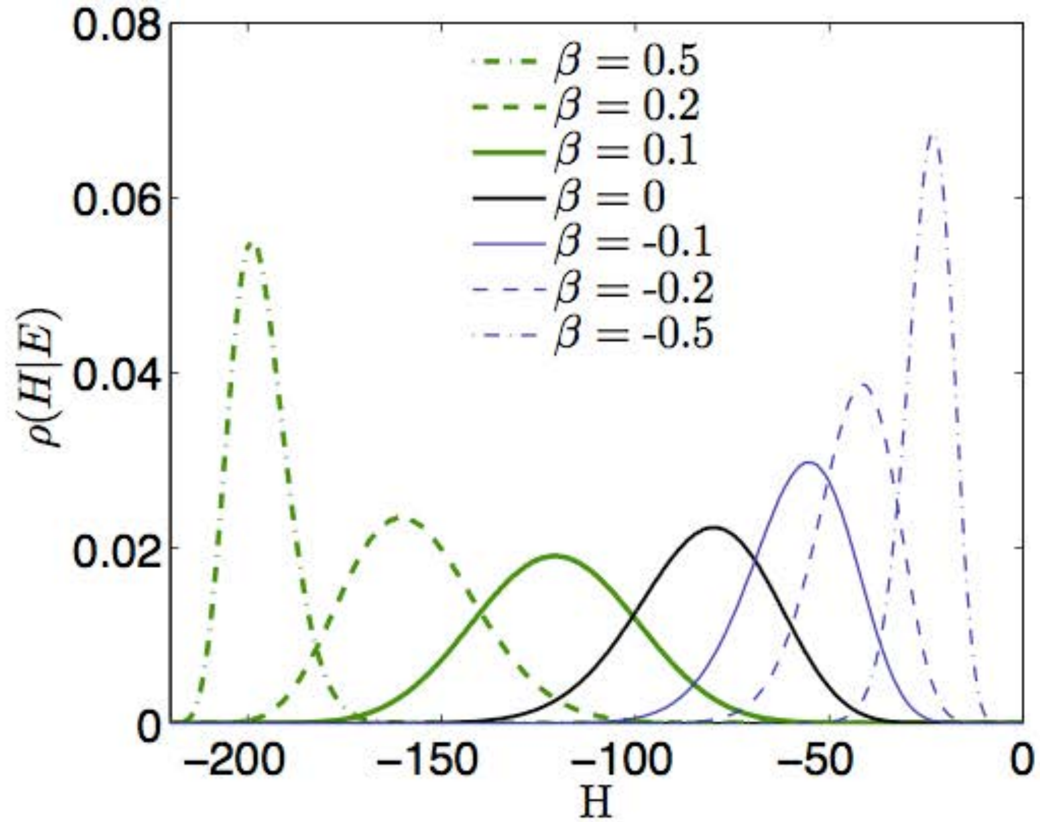
The incorporation of constraints complicates the use of Langevin dynamics, and, in particular, its numerical discretization.

However, we can easily design degenerate thermostats to cope with the constraint manifold.

pdfs (in H)



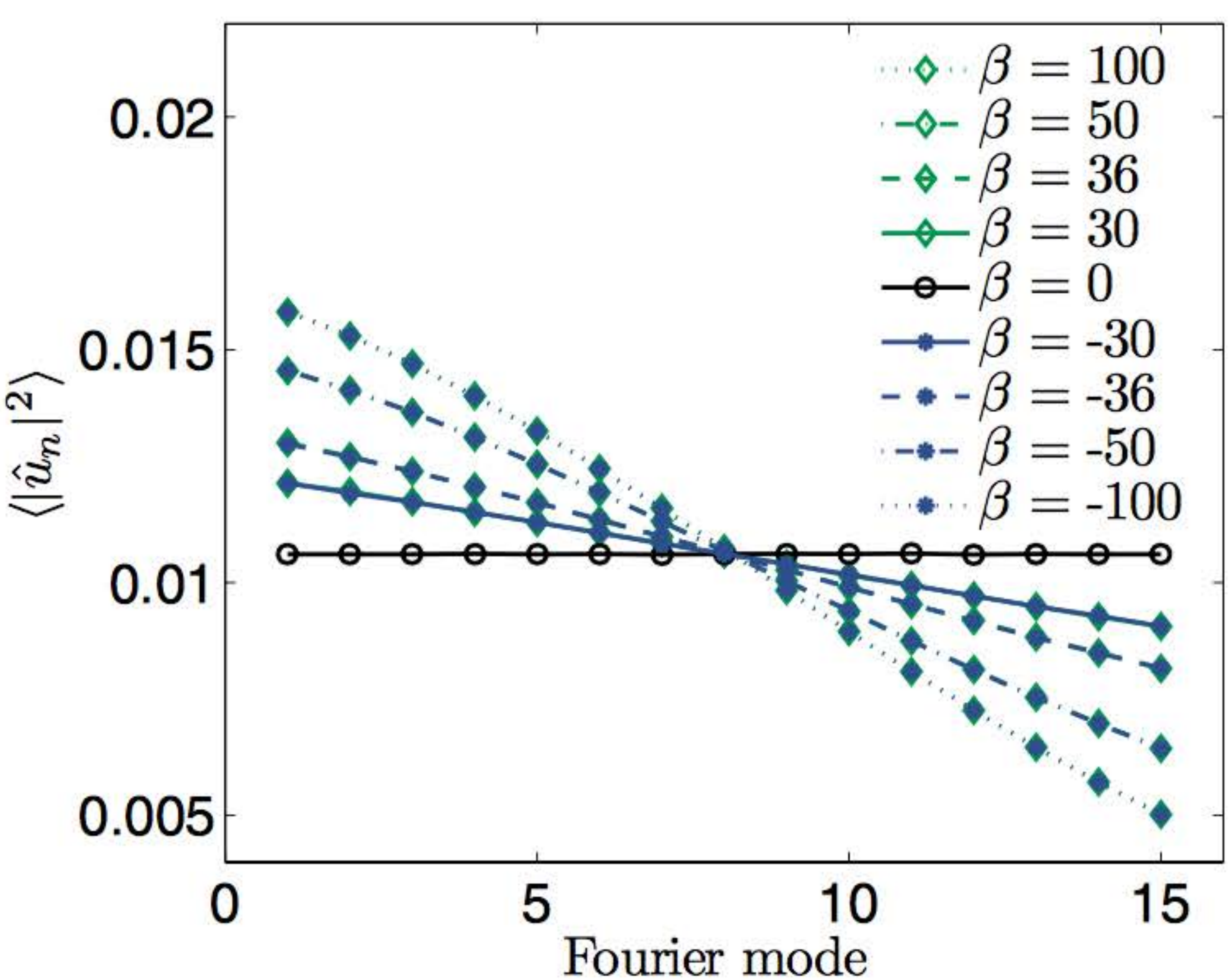
Burgers



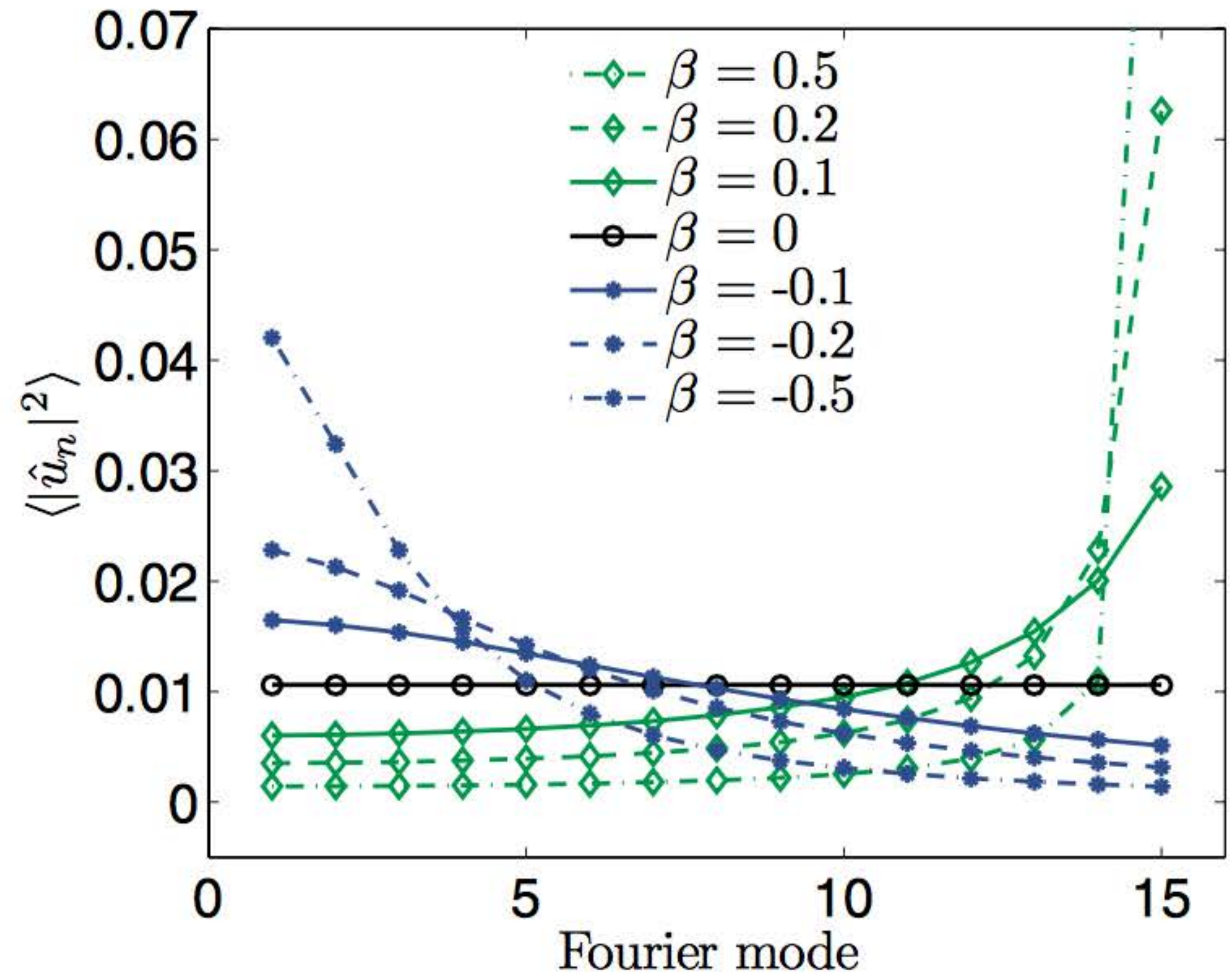
KdV

Burgers equation: N = 15, E_0 =1, β = -30

kinetic energy spectrum



Burgers



KdV

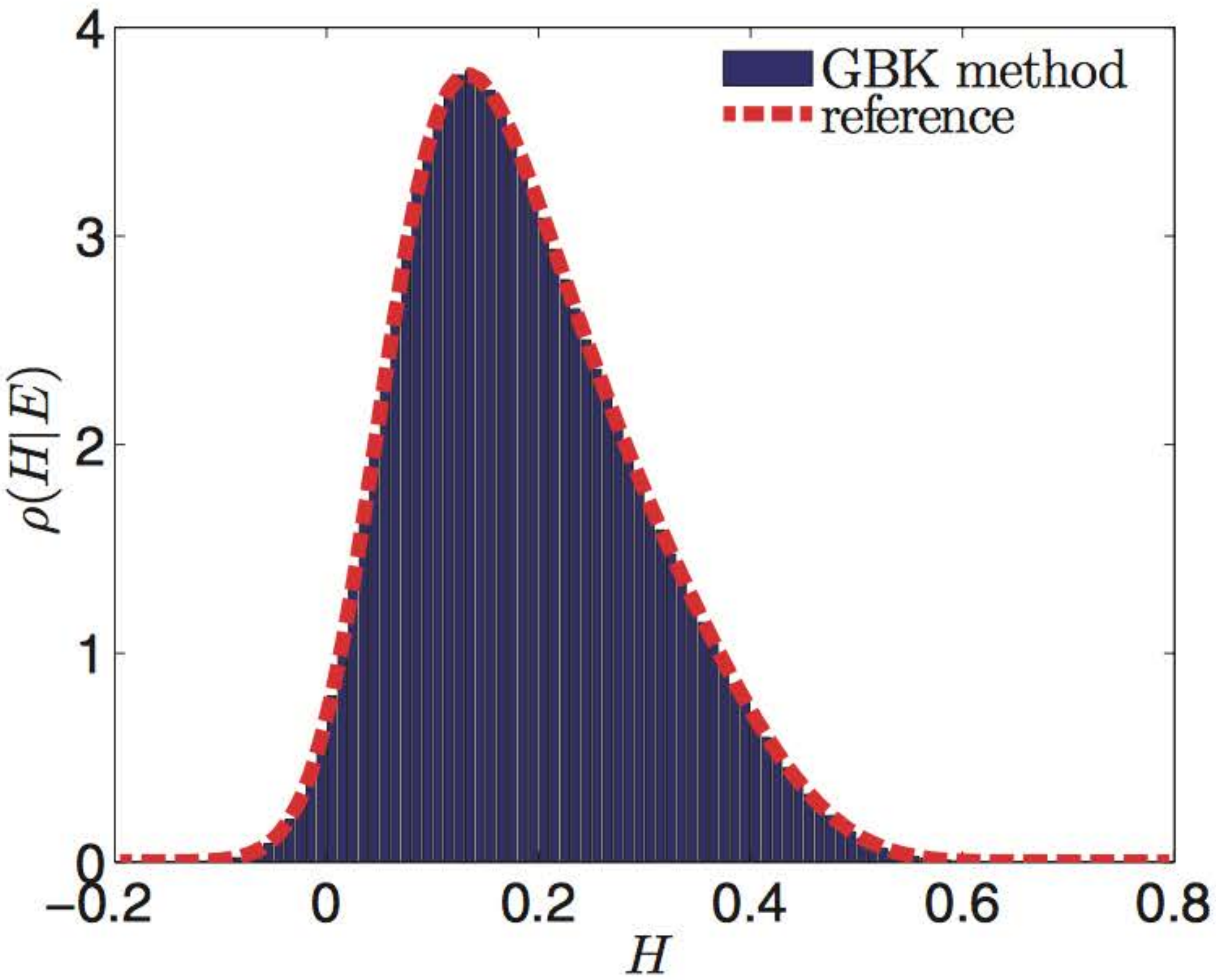
$$
\begin{aligned}
\mathrm{d}X &= f(X)\mathrm{d}t + \sum_{k=1}^{d_T} \xi_k g_k(X)\mathrm{d}t, \\
\mathrm{d}\xi_k &= h_k(X)\mathrm{d}t - \gamma\xi_k \mathrm{d}t + \sigma \mathrm{d}w_k, \quad k = 1, \ldots, d_T
\end{aligned}
$$

One example of the coupling functions (projecting u^2 to a truncated Fourier representation)

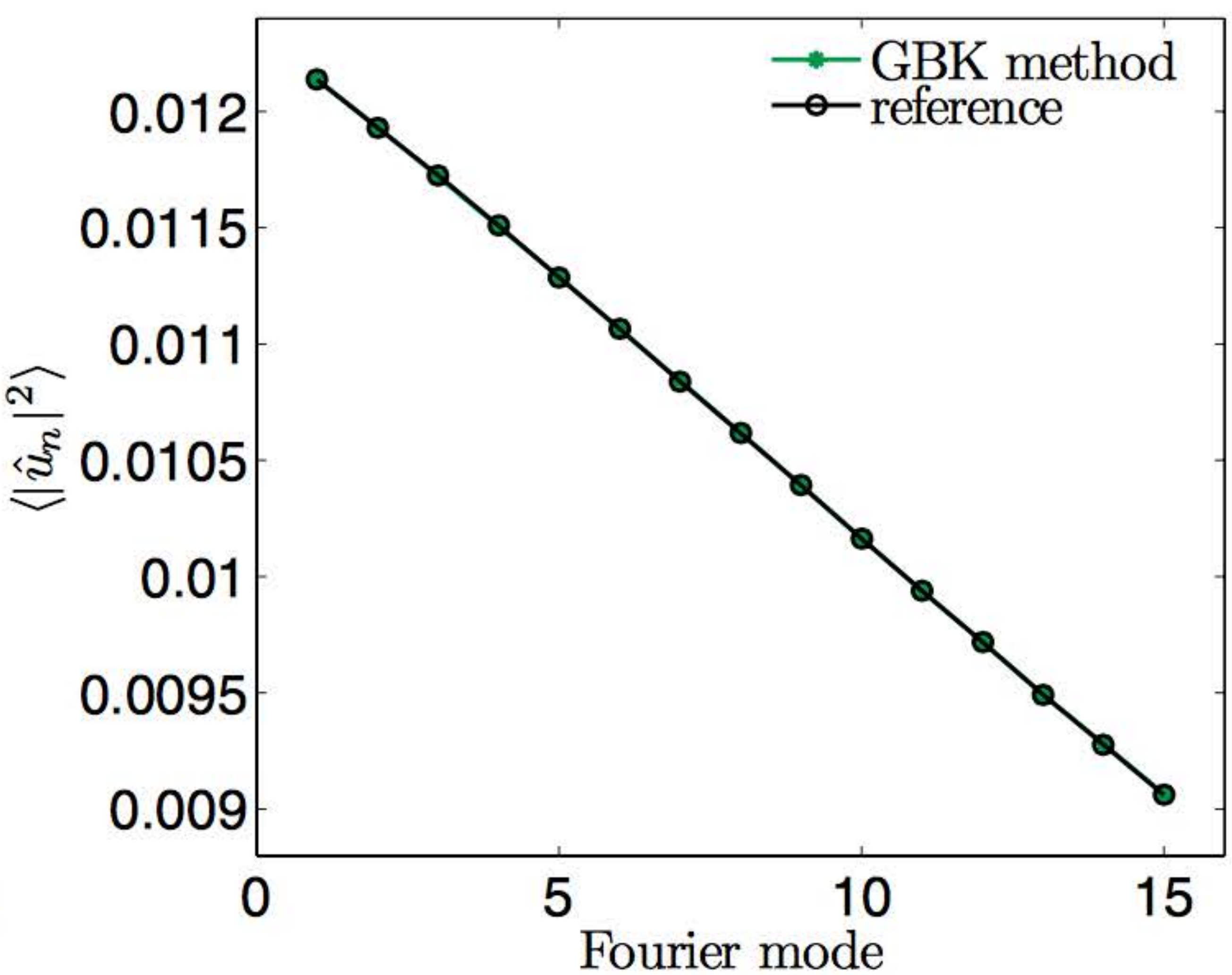
$$
g(u_N) = u_N^2 - \frac{1}{2\pi}\int_0^{2\pi} u_N^2 \, \mathrm{d}x - \frac{1}{2E} u_N \int_0^{2\pi} u_N^3 \, \mathrm{d}x
$$

$$
h(u_N) = -\frac{1}{\alpha}\left(\frac{N}{E}\int_0^{2\pi} u_N^3 \, \mathrm{d}x + \beta \int_0^{2\pi} \frac{\delta H}{\delta u_N} g_N(u_N) \, \mathrm{d}x\right)
$$

energy dist

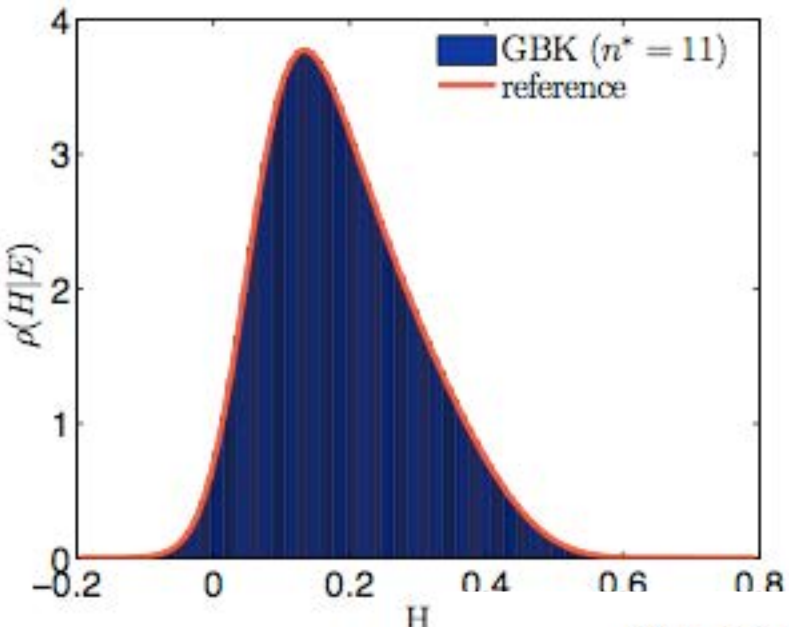


KE spectrum

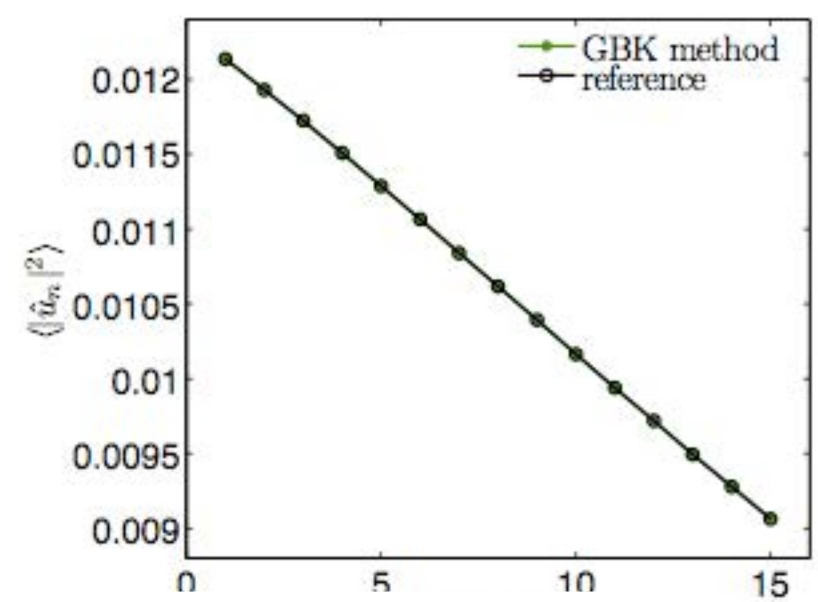


GBK(n=m)*: results using a thermostat applied only to modes *m…N*

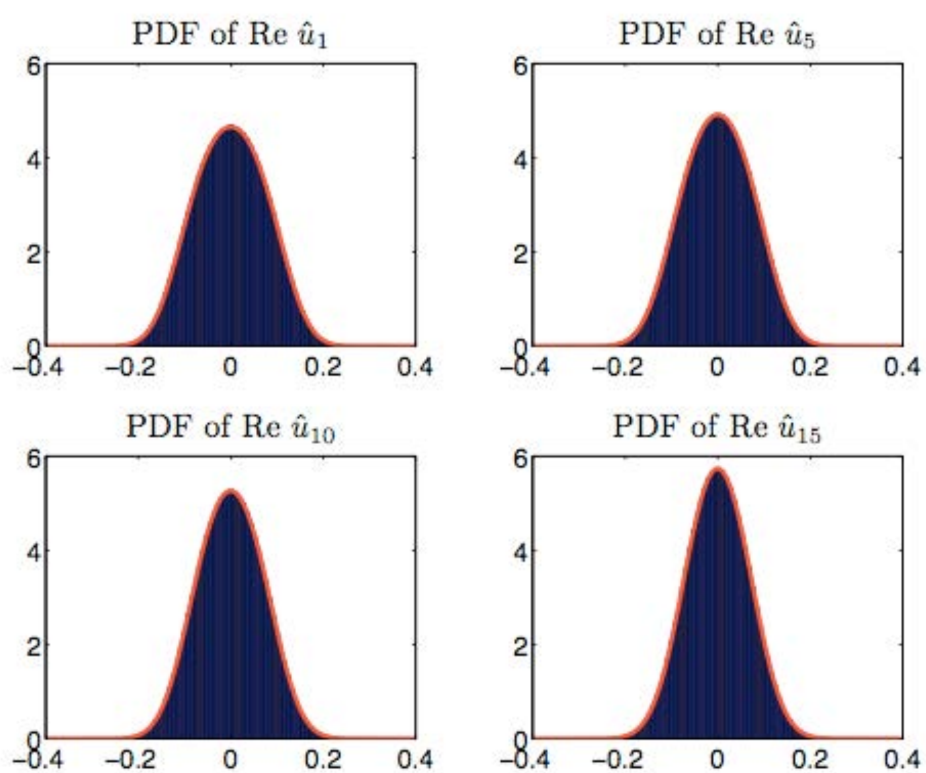
H Dist



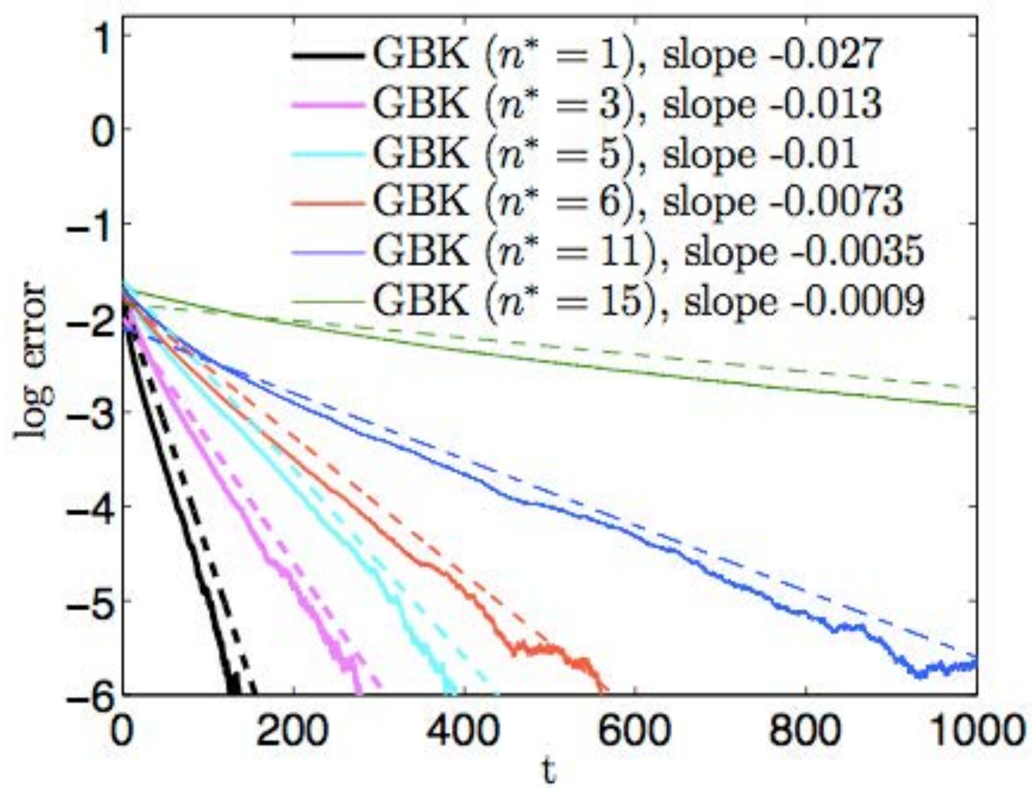
Spectrum

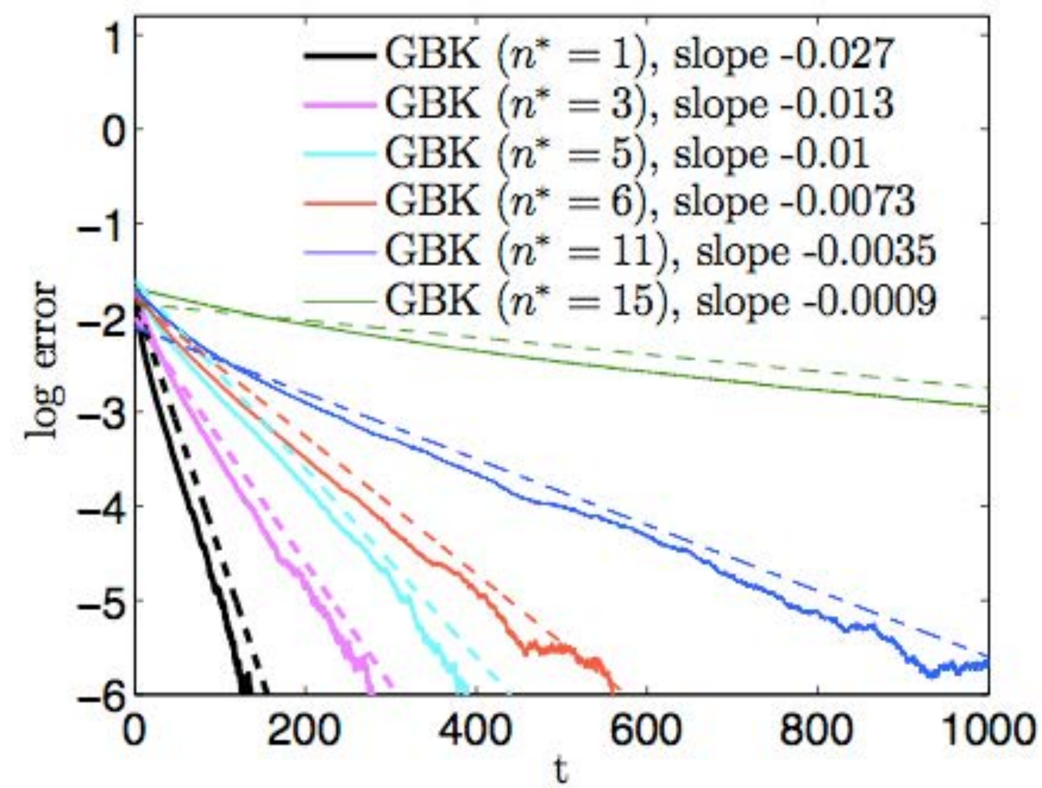


weak perturbation GBK(n*=15)

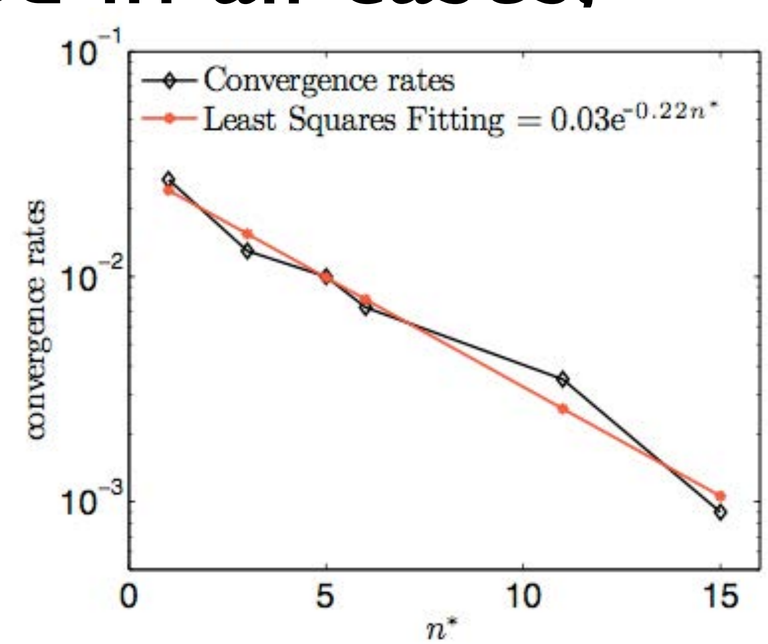


Convergence of expected value of Hamiltonian



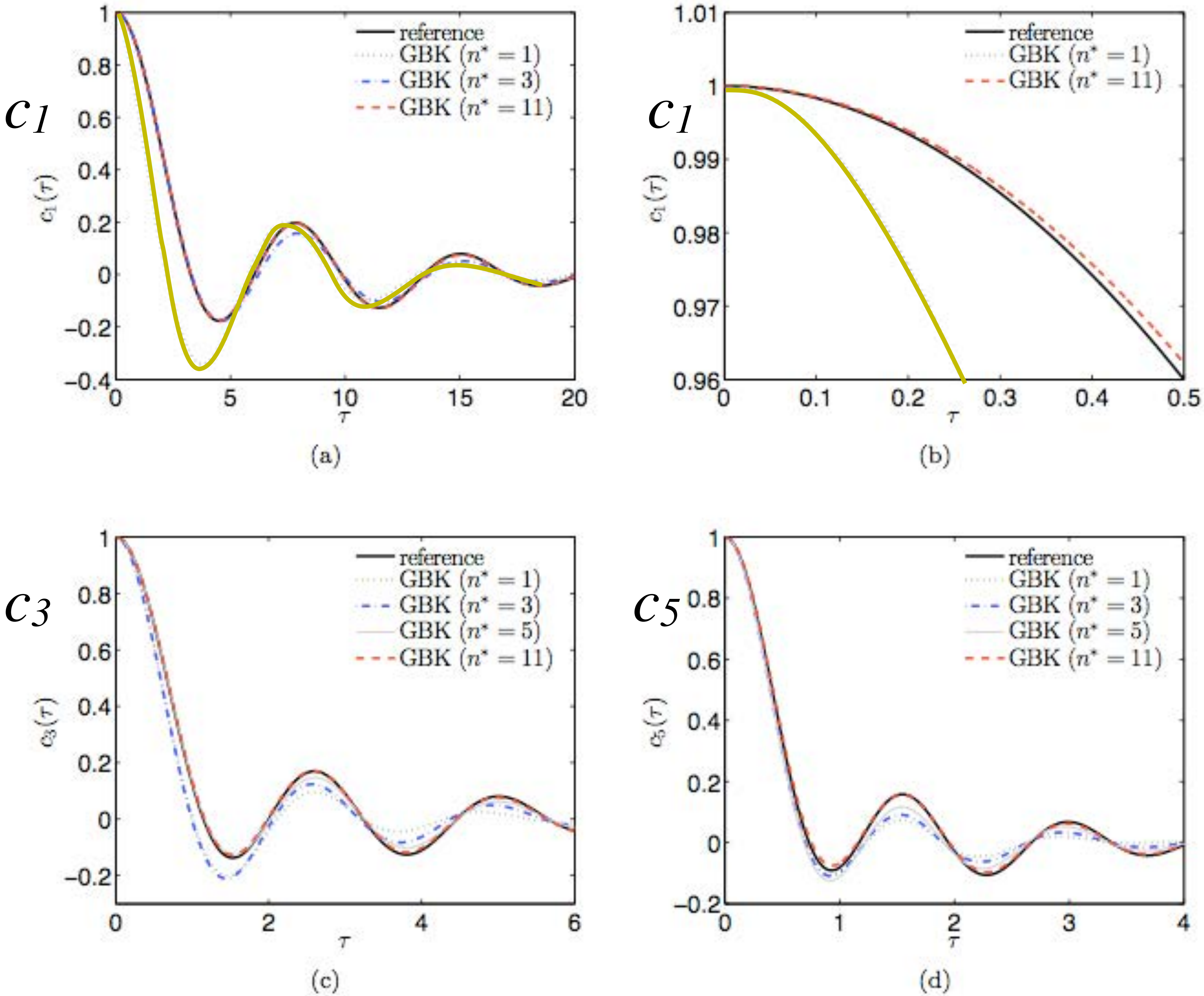


convergence of averages is observed in all cases, but is very slow for GBK (n^*=15)



exponential decay of the rate with n^*

c_k : autocorrelation function for kth mode



Observations

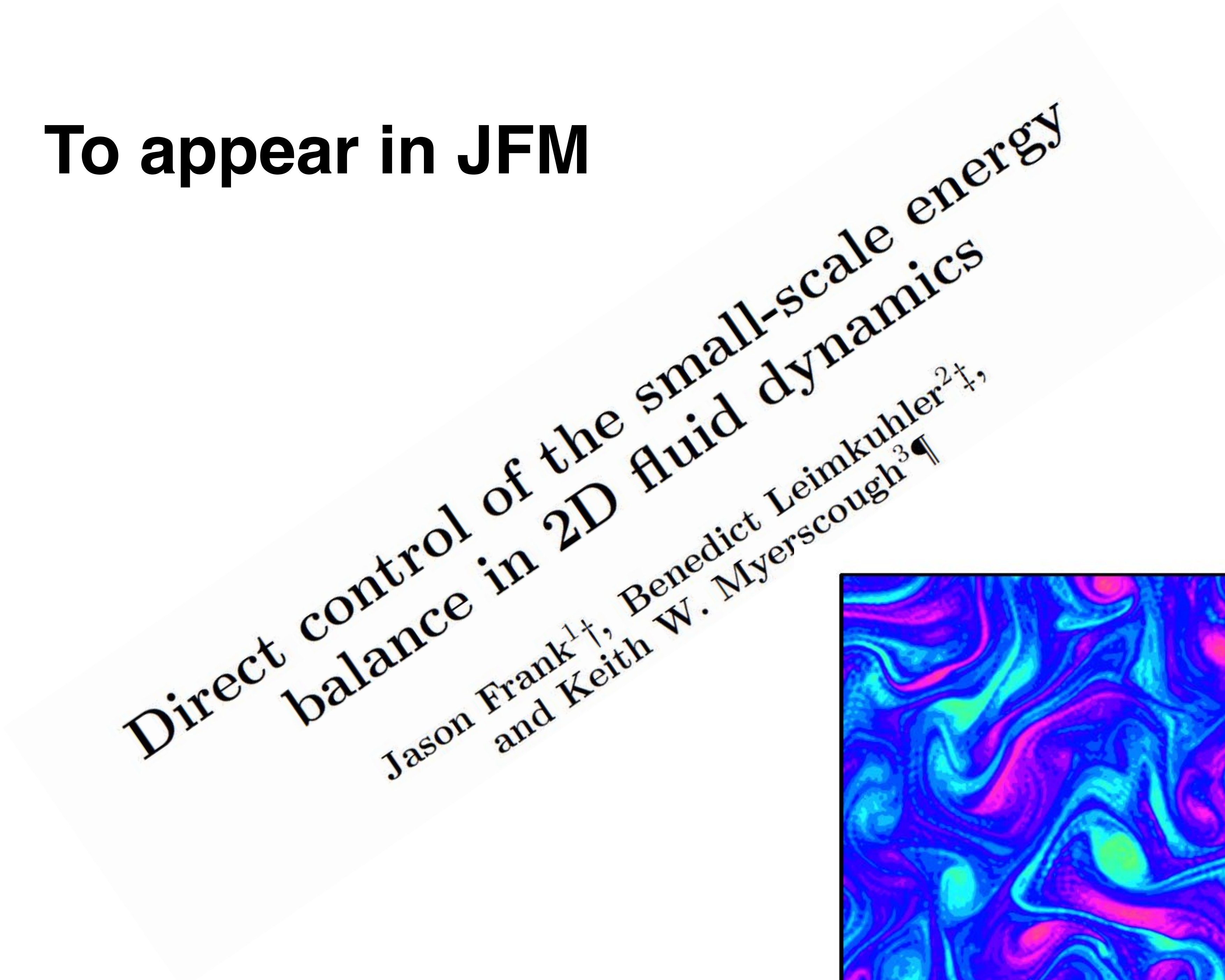
1. It is possible to design heat-bath models that contact arbitrary components in the Fourier spectral discretization

2. the methods appear to be ergodic

3. gentle thermostats are effective in the mixed density model, even if they only contact a single mode (although convergence is then quite slow)

4. all-mode thermostats (like Langevin dynamics) can distort the averaged dynamics of the system

To appear in JFM



Direct control of the small-scale energy balance in 2D fluid dynamics

Jason Frank[1]†, Benedict Leimkuhler[2]‡, and Keith W. Myerscough[3]¶