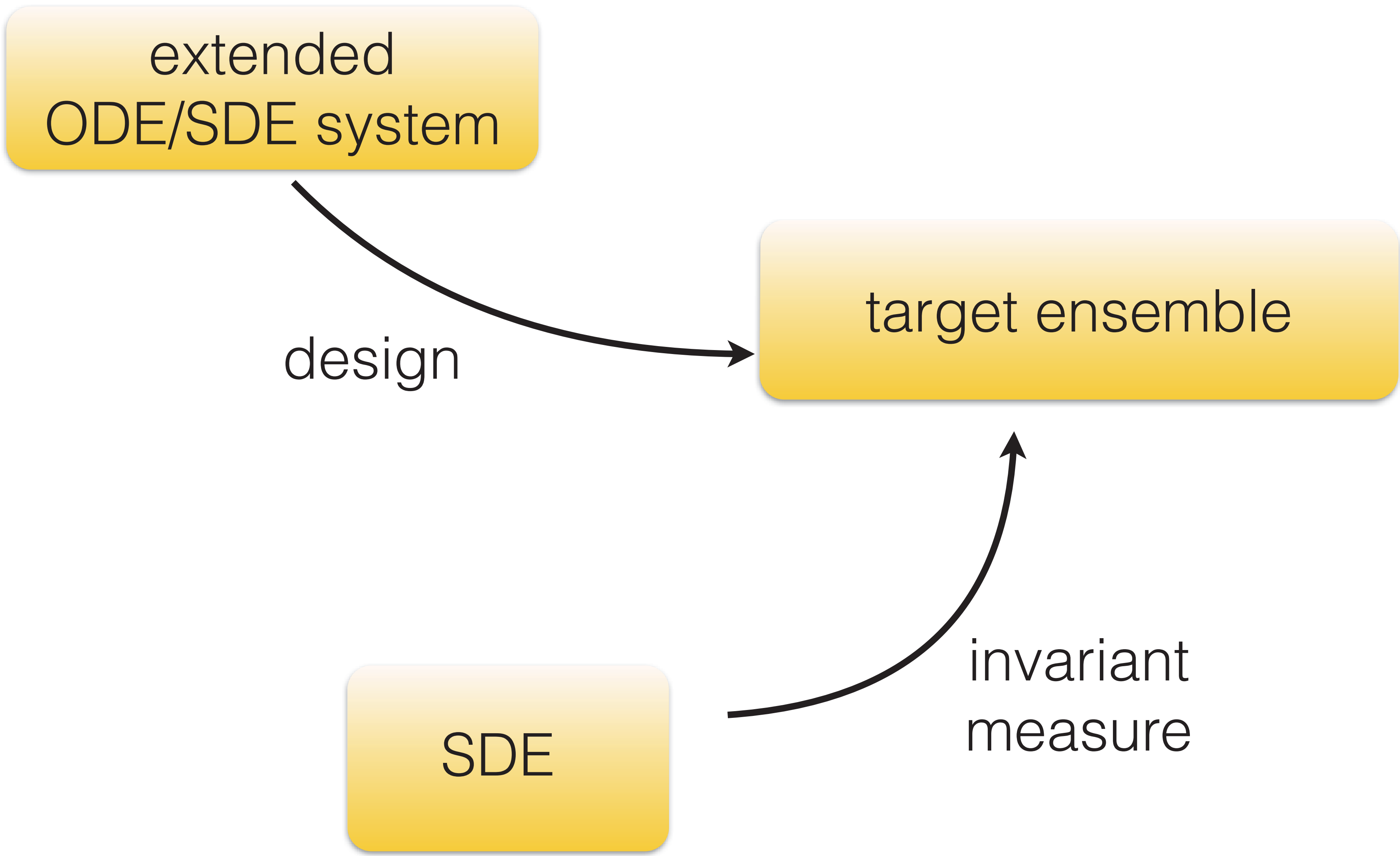
Algorithms for Ensemble Control

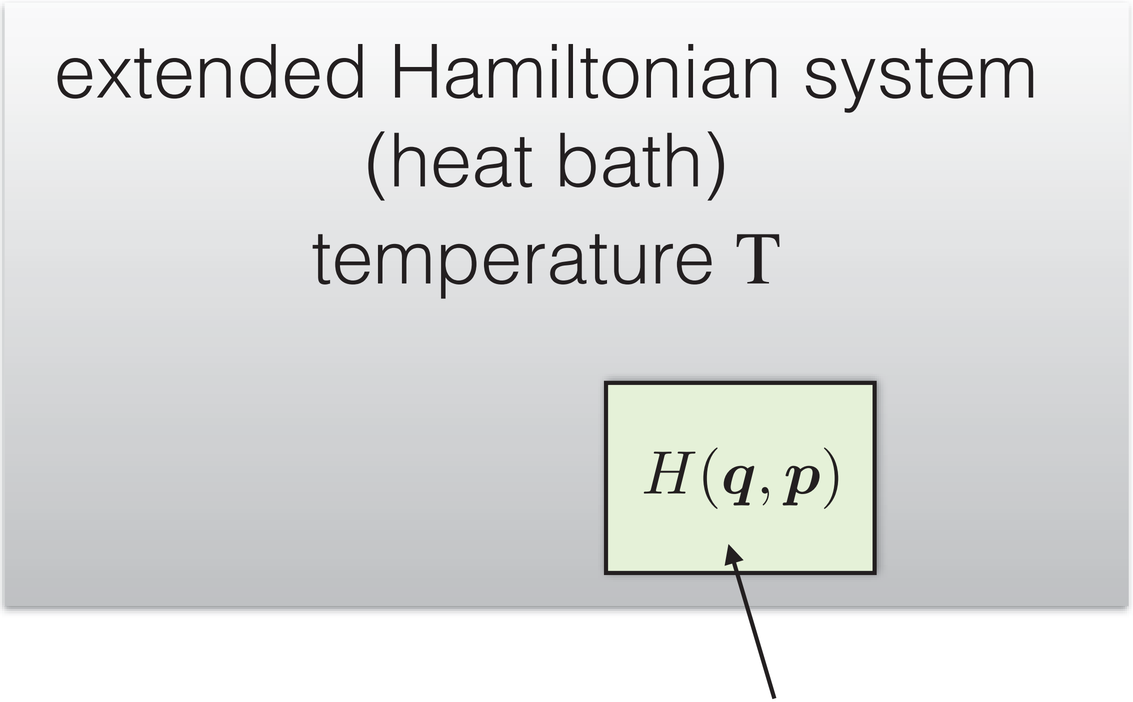
B Leimkuhler
University of Edinburgh

Model Reduction Across Disciplines • Leicester • 2014

(Stochastic) Ensemble Control



Thermostat



$$\rho_\beta \propto e^{-\beta H}, \qquad \beta = \frac{1}{\mathrm{k_B T}}$$

Thermostat = ODE/SDE with prescribed
unique invariant density
(typically Boltzmann-Gibbs)

Thermostats

Define:

$$A = \int a(z)\rho_{\mathrm{can}}(z)dz$$ stationary average

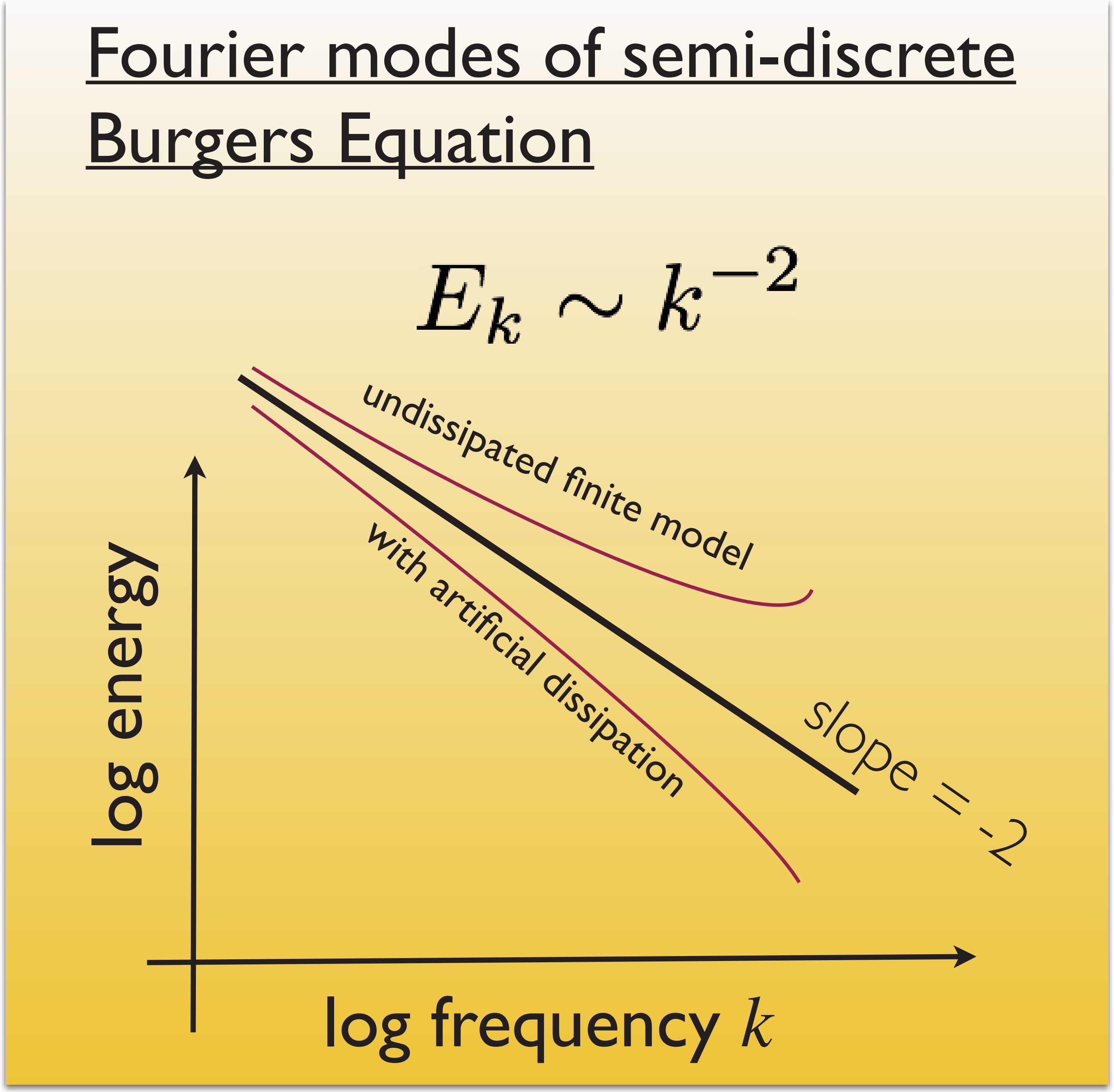
$$C(\tau) = \int \varphi_\tau(z)^T Bz\rho_{\mathrm{can}}(z)dz$$ autocorrelation function

A **thermostat** generates trajectories *z(t)* such that

$$\hat{A} := \lim_{t\to\infty} t^{-1}\int_0^t a(z(s))ds = A$$

$$\hat{C}(\tau) := \lim_{t\to\infty} t^{-1}\int_0^t z(s+\tau)^T Bz(s)ds \approx C(\tau) \quad \text{(T.L.)}$$

Control of "Model Error"



Can we correct the energy decay relation using a 'thermostat'-like device?

Some Questions

Many choices for reduced system with same invariant measure - how to design/choose?

Ergodicity? How to promote rapid mixing, convergence? How does the SDE approach equilibrium?

Role of **dynamics?** Relation to ‘natural’ timescales.

What is the effect of **numerical discretization?** (Invariant measures of numerical methods)
How does the SDE discretization approach equilibrium?

Can we **correct model error** using ensemble controls? (i.e. retroactively repair damaged models)

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 conditions
unique steady state of the Fokker-Planck equation:

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

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

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:

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

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

Under suitable conditions…

- 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

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

Hypoellipticity

A 2nd order differential operator with C^∞ coefficients is **hypoelliptic** if its zeros are C^∞

Let U be a compact, connected, invariant subset for an SDE.

$$\mathrm{d}x = X_0(x)\mathrm{d}t + \sum_{j=1}^{r} X_j(x)\mathrm{dW}_j$$

If the corresponding Kolmogorov operator is hypoelliptic on U, then the flow is ergodic on U.

Acknowledgement: *Hairer's Lecture Notes*

...Hörmander...Villani...Hairer...

Hörmander condition

The vector fields $X_0(x), \dots, X_r(x)$ satisfy a Hörmander condition if

$$Span\{X_0(x), \dots, X_r(x), [X_i, X_j](x),\ [X_i, [X_j, X_k]](x) \dots\} = \mathbb{R}^N$$

Theorem 1. *Let $U \subset \mathbb{R}^N$ be open. If $X_0, X_1 : U \to \mathbb{R}^N$ are two vectorfields that satisfy Hörmander's condition at every $z \in U$, then the operator L^* which is defined by*

$$L^*\rho := -\sum_{i=1}^{N} \frac{\partial}{\partial z_i}(\rho(z)X_{0,i}(z)) + \frac{1}{2}\sum_{i,j=1}^{N} \frac{\partial^2}{\partial z_i \partial z_j}(\rho(z)X_{1,i}(z)X_{1,j}(z))$$

is hypoelliptic.

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

Highly Degenerate Diffusions

Nose-Hoover

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

Newtonian dynamics

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -U'(q) \end{aligned}$$

preserves $\tilde{\rho} \propto \exp(-H/kT - \alpha\xi^2)$ **but not ergodic**

Nose-Hoover

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

Newtonian dynamics

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -U'(q) - \xi p \\ \mu\dot{\xi} &= p^2 - \theta \end{aligned}$$

preserves $\tilde{\rho} \propto \exp(-H/kT - \alpha\xi^2)$ **but not ergodic**

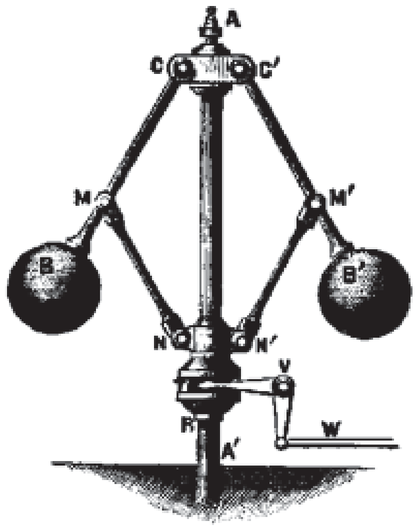
Nose-Hoover

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

Newtonian dynamics

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -U'(q) - \xi p \\ \mu\dot{\xi} &= p^2 - \theta \end{aligned}$$

'governor'



preserves $\tilde{\rho} \propto \exp(-H/kT - \alpha\xi^2)$ **but not ergodic**

Nose-Hoover

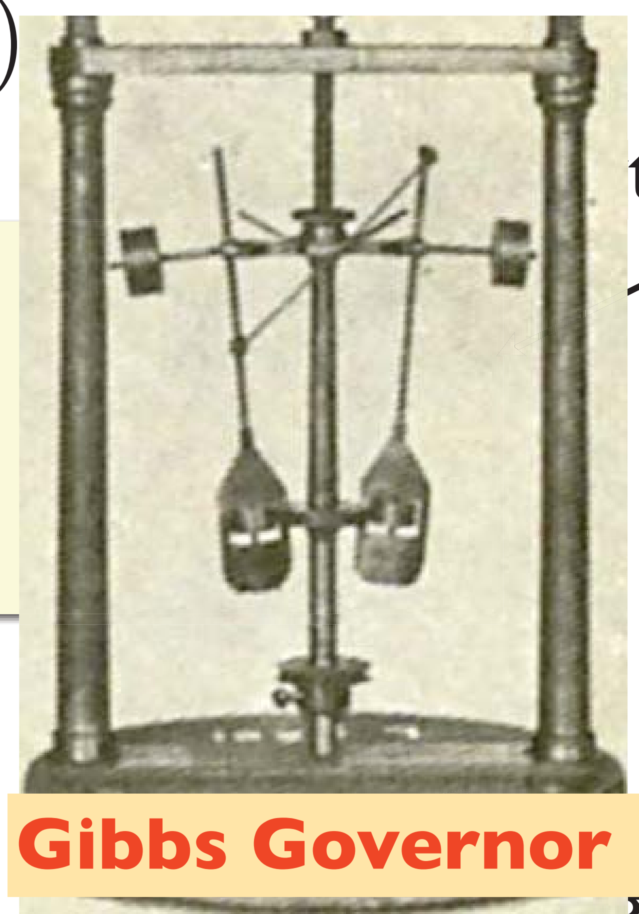
$$H = \frac{p^2}{2} + U(q)$$

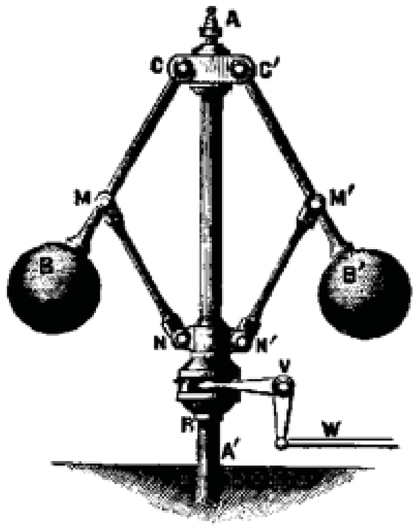
tonian dynamics

$$\dot q =$$

$$\dot p = \quad \xi p$$

$$\mu \dot\xi =$$

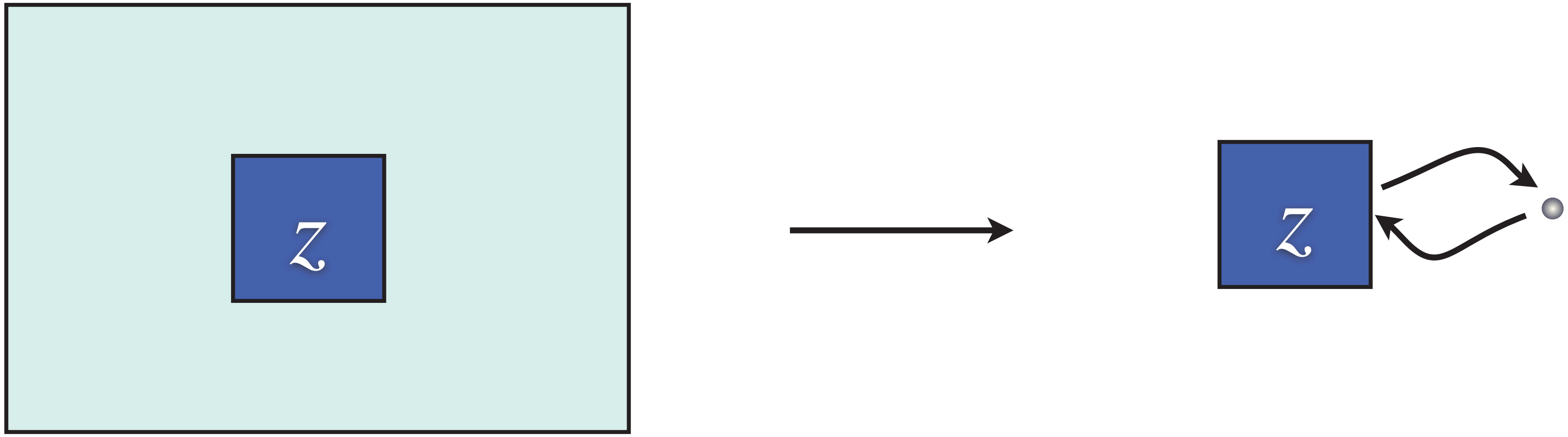




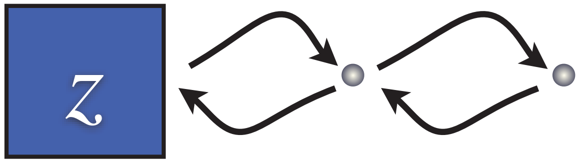
governor

preserves $\tilde\rho \propto \exp(-H/kT - \alpha\xi^2)$ **but not ergodic**

Need for Stochastics

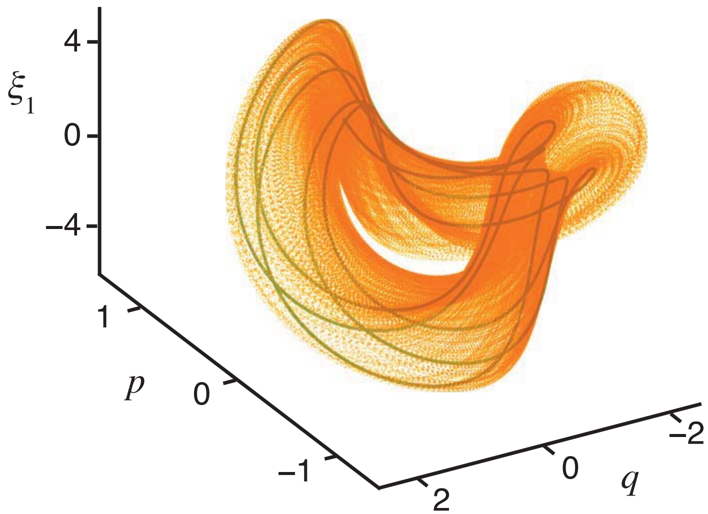


Chains:



$$
\begin{aligned}
\dot{q} &= p \\
\dot{p} &= -q - \xi p \\
\dot{\xi}_1 &= \mu_1^{-1}(p^2 - kT) - \xi_1 \xi_2 \\
\dot{\xi}_2 &= \mu_2^{-1}(\mu_1 \xi_1^2 - kT)
\end{aligned}
$$

$$\mu_1 = 0.2, \qquad \mu_2 = 1$$



Designer Diffusions

L. Noorizadeh, Theil JSP 2009,
L., Phys Rev E, 2010
L., Noorizadeh, Penrose JSP 2011

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

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

OU

- design to preserve extended Gibbs distribution

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

- 'weak' coupling to stochastic perturbation

Nose-Hoover-Langevin

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

$$\mathrm{d}p = -\nabla V - \xi p$$

$$\mathrm{d}\xi = \mu^{-1}[p^T p - n\mathrm{kT}]\mathrm{d}t - \gamma\xi\mathrm{d}t + \sqrt{2\mathrm{kT}\gamma/\mu}\,\mathrm{dW}$$

- Unification of Nosé-Hoover and Langevin thermostats
- Generalizes NH thermostat
- Includes kinetic energy regulator
- Single scalar stochastic variable

$$dX = f(X)dt + \xi g(X)dt$$

$$d\xi = h(X)dt - \gamma\xi dt + \sqrt{2\gamma}dW$$

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
f,g spans $\mathrm{T}\mathcal{M}$ at every point of $\mathcal{M}$

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

Ergodicity of NHL

$$\mathrm{d}q = p\mathrm{d}t$$
$$\mathrm{d}p = -\nabla V - \xi p$$
$$\mathrm{d}\xi = \mu^{-1}[p^T p - n\mathrm{kT}]\mathrm{d}t - \gamma\xi\mathrm{d}t + \sqrt{2\mathrm{kT}\gamma/\mu}\mathrm{dW}$$

Let the potential have the form

$$V = q^T B q$$

then, under a mild non-resonance assumption, the NHL equations are ergodic on a large set.

Proof: just check the Hörmander condition!

Ex: Nose-Hoover Langevin on a harmonic system

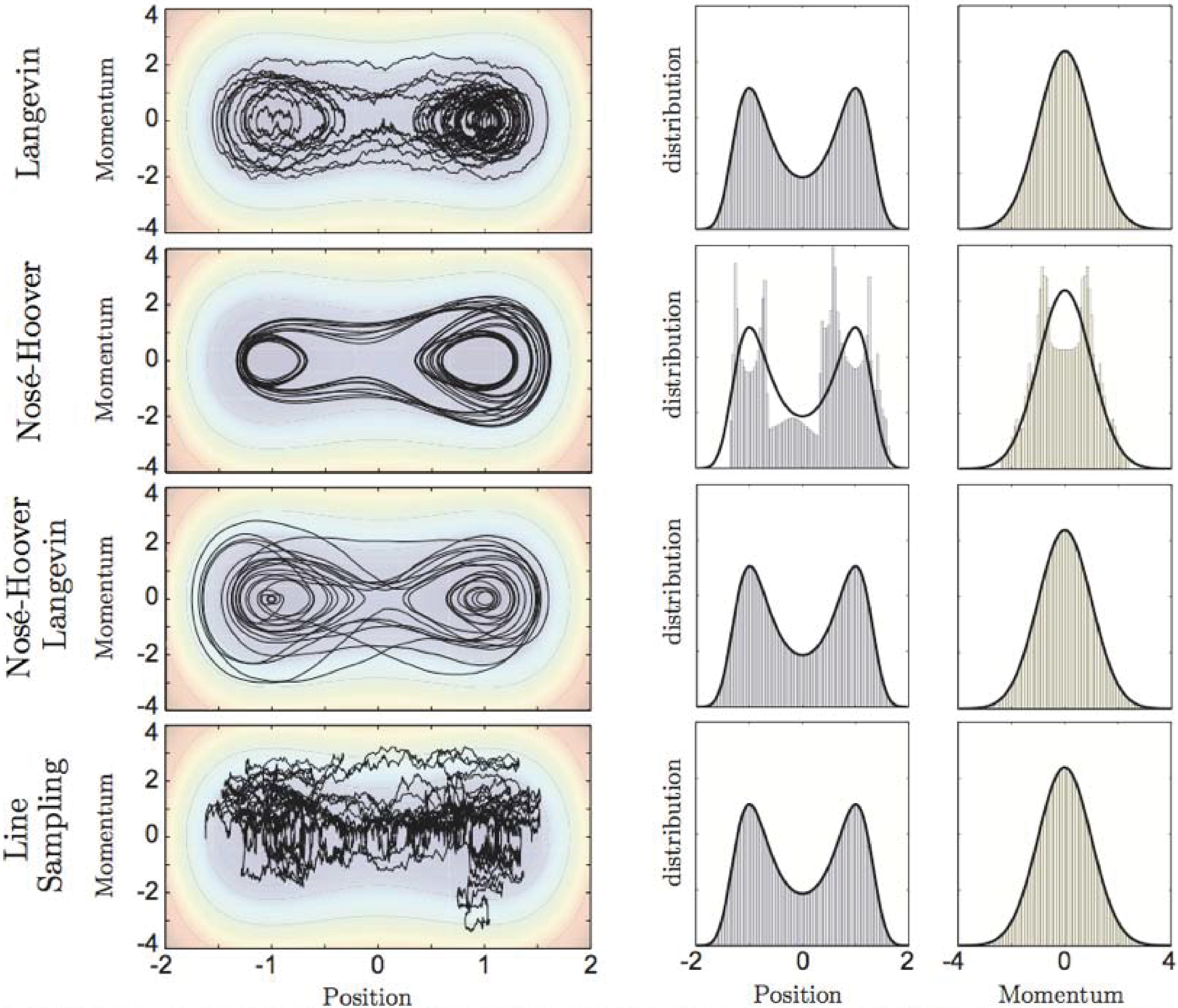
$$f = \begin{bmatrix} p \\ -Bq \end{bmatrix}, \qquad g = \begin{bmatrix} 0 \\ p \end{bmatrix}$$

$S\{f, g\}$ = Lie algebra (ideal) generated by f, g

Prop:

$$C_k = \begin{bmatrix} B^{k-1}p \\ B^k q \end{bmatrix} \in S\{f, g\}$$

$$D_k = \begin{bmatrix} B^k p \\ -B^k q \end{bmatrix} \in S\{f, g\}$$

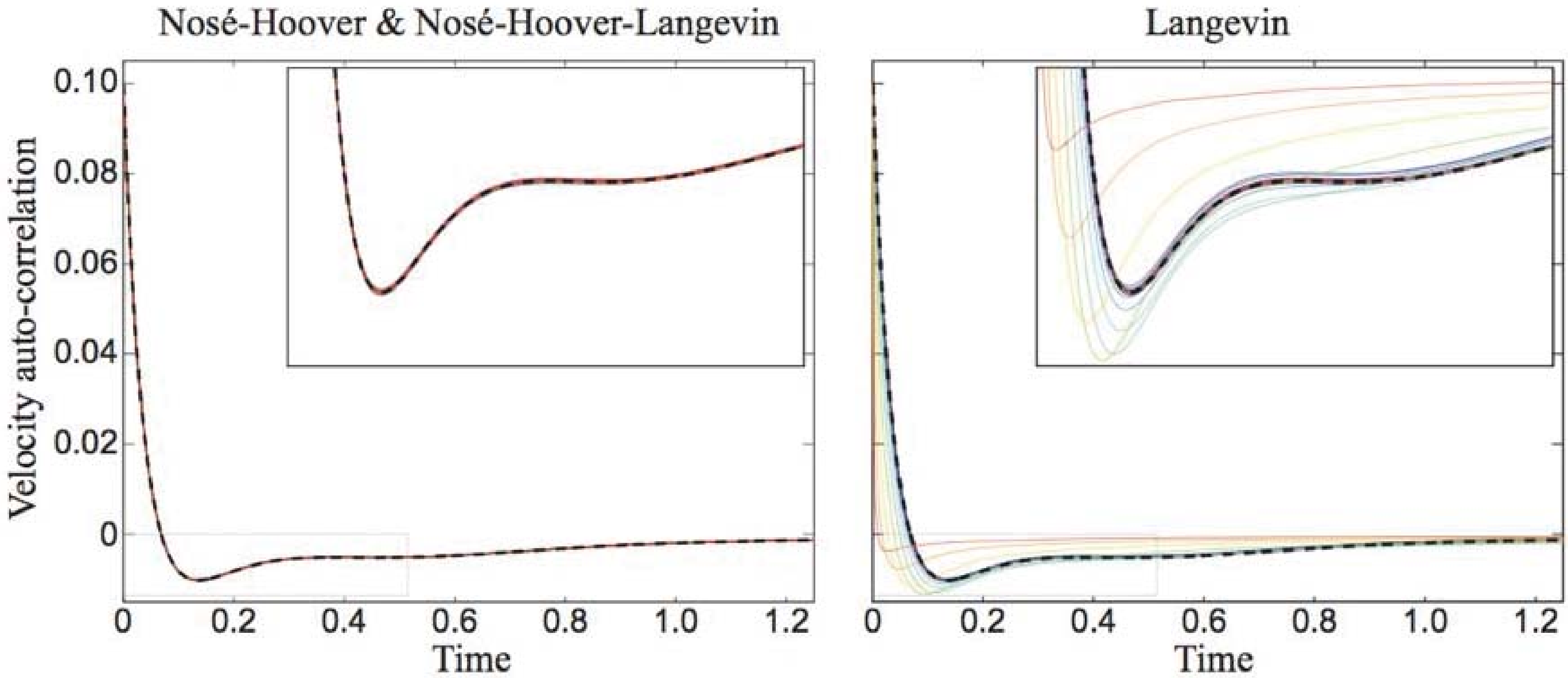


Langevin
Nosé-Hoover
Nosé-Hoover Langevin
Line Sampling
Momentum
distribution
Position
Momentum
4
2
0
-2
-4
-2
-1
0
1
2

Autocorrelation Functions

[L., Noorizadeh and Penrose, J. Stat. Phys. 2011]

quantify ‘efficiency’ of different thermostats
accumulation of error in dynamics vs convergence rate



parameter dependence

Vortex Method

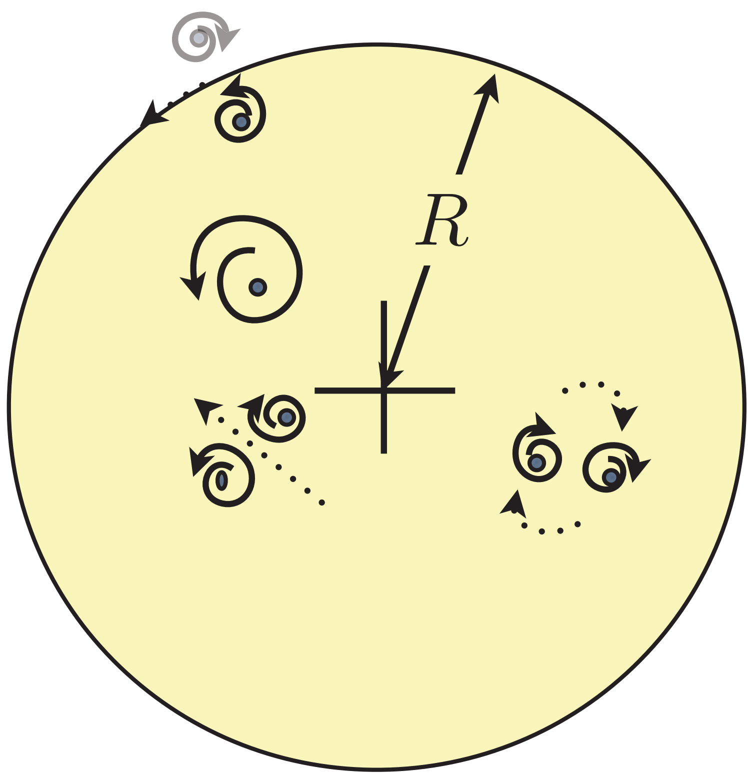
Point Vortices

[Dubinkina, Frank and L., SIAM MMS 2010]

A point vortex model for *N* vortices in a cylinder

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

$+$ *boundary terms*



$\longrightarrow$ $$\Gamma_i \dot{x}_i = J\nabla_{x_i} H$$

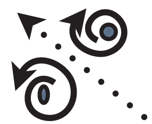
Onsager, 1949 “Statistical Hydrodynamics”
Oliver Bühler, 2002: a numerical study

Onsager's Prediction

"... vortices of the same sign will tend to cluster---preferably the strongest ones---so as to use up excess energy at the least possible cost in terms of degrees of freedom ... the weaker vortices, free to roam practically at random will yield rather erratic and disorganized contributions to the flow."

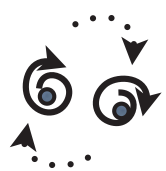
Positive temperatures:

Strong vortices of opposite sign tend to approach each other



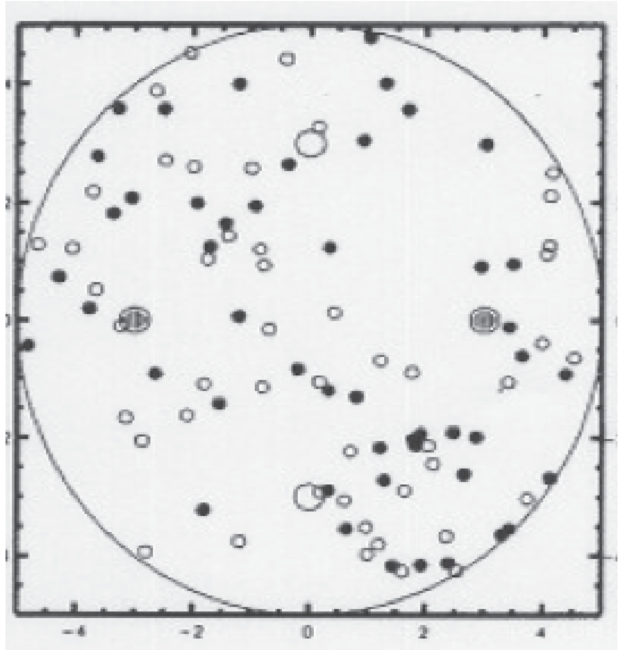
Negative temperatures:

Strong vortices of the same sign will cluster



Buhler (2002) Simulation

4 strong
96 weak vortices
sign indefinite,
0 net circulation in each group
fixed ang. mom.



Simulation results supported Onsager's predictions

Use *finite* bath - not the Gibbsian model

Modified Canonical Statistics

Assume the subsystem and reservoir variables decoupled in the Hamiltonian

$$H(X_A, X_B) = H_A(X_A) + H_B(X_B)$$

Notation: $\Omega(E) = \mathrm{vol}\{X \mid H(X) \in [E, E + dE)\}$

$$S(E) = \ln \Omega(E)$$

Then:
$$\begin{aligned}
\mathrm{Prob}\{X_A | H = E\} &\propto \Omega_B(E - H_A(X_A)) \\
&= \exp(S_B(E - H_A)) \\
&= \exp(S_B(E) - S'_B(E) H_A + S''_B(E) H_A^2 + \cdots) \\
&\propto \exp(-\beta H_A + \gamma H_A^2 + \cdots)
\end{aligned}$$

assume finite bath energy → Gibbs statistics

assume finite bath energy variance → Generalized Bath Model

Modified Gibbs

$$\rho_{\text{finite}} \propto e^{-\beta H - \gamma H^2}$$

Modified stochastic control law:

Gibbs:

$$\begin{aligned} \dot{X} &= J\nabla H(X) + \zeta s(X) \\ \dot{\zeta} &= \alpha^{-1}\left[\beta \nabla H \cdot s(X) - \nabla \cdot s(X)\right] + OU(\zeta) \end{aligned}$$

↓

modified Gibbs:

$$\dot{\zeta} = \alpha^{-1}\left[\beta(1 + \frac{\gamma}{\beta}H)\nabla H \cdot s(X) - \nabla \cdot s(X)\right] + OU(\zeta)$$

Allows direct comparison with Bühler's results

GBK thermostat gives a 100 → 5 model reduction

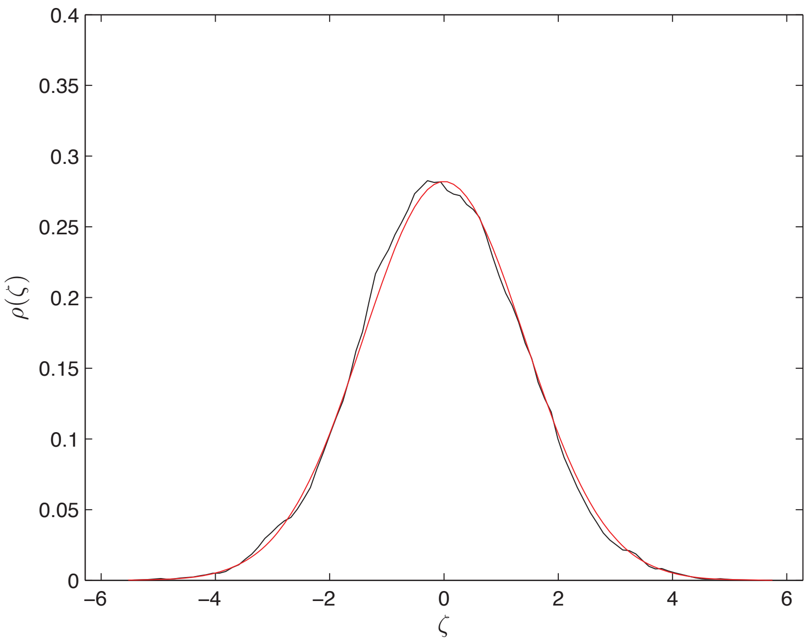
Experimental parameters

$\beta \in \{-0.006, -0.00055, 0.01\}$

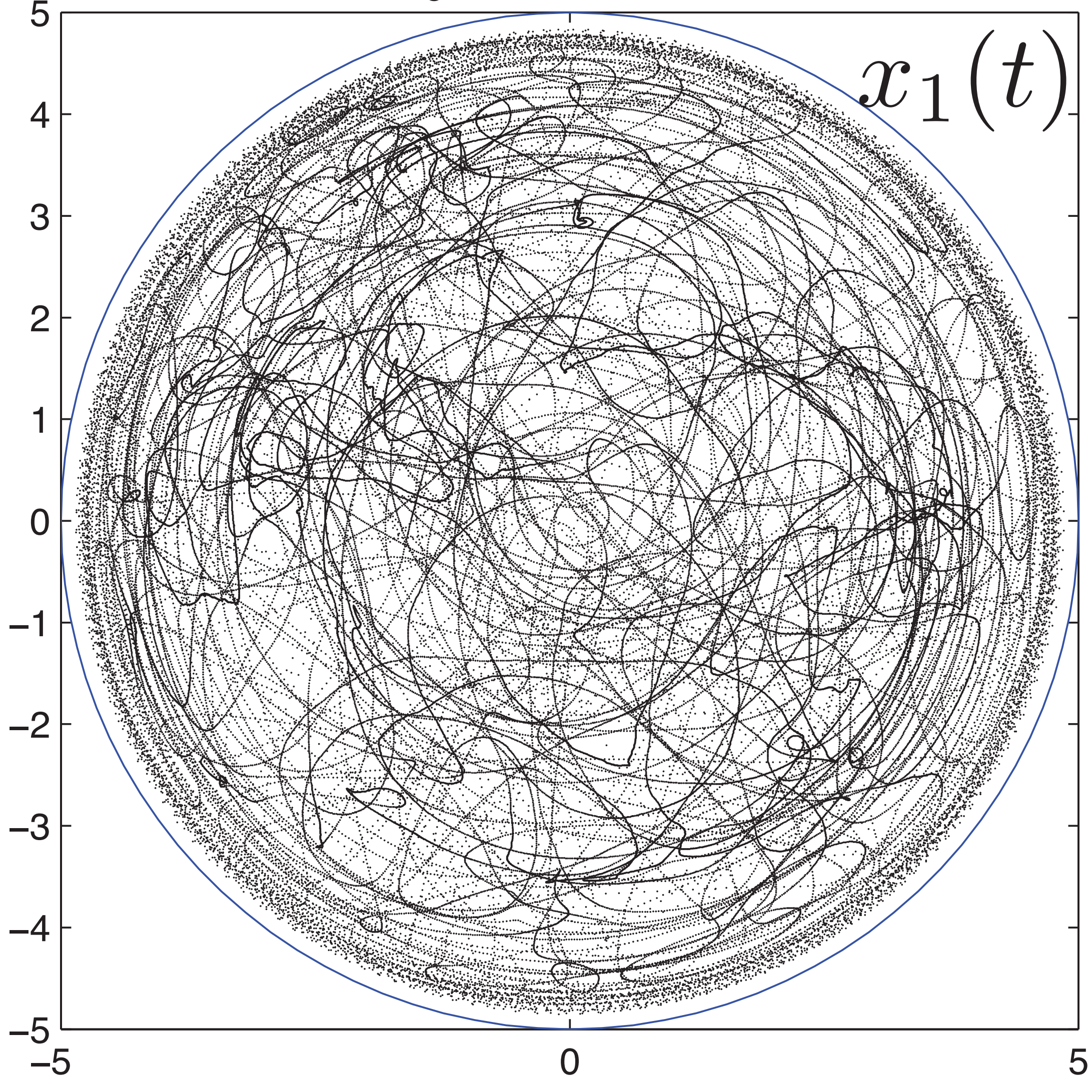
$\gamma = -\frac{\beta}{2E_0}, \quad E_0 \in \{628, 221, -197\}$

$\alpha = 0.5, \quad \sigma = \sqrt{0.4}$

$t \in [1500, 12000]$



ζ is Gaussian



$t \in [0, 1000] \qquad \beta = -0.00055$

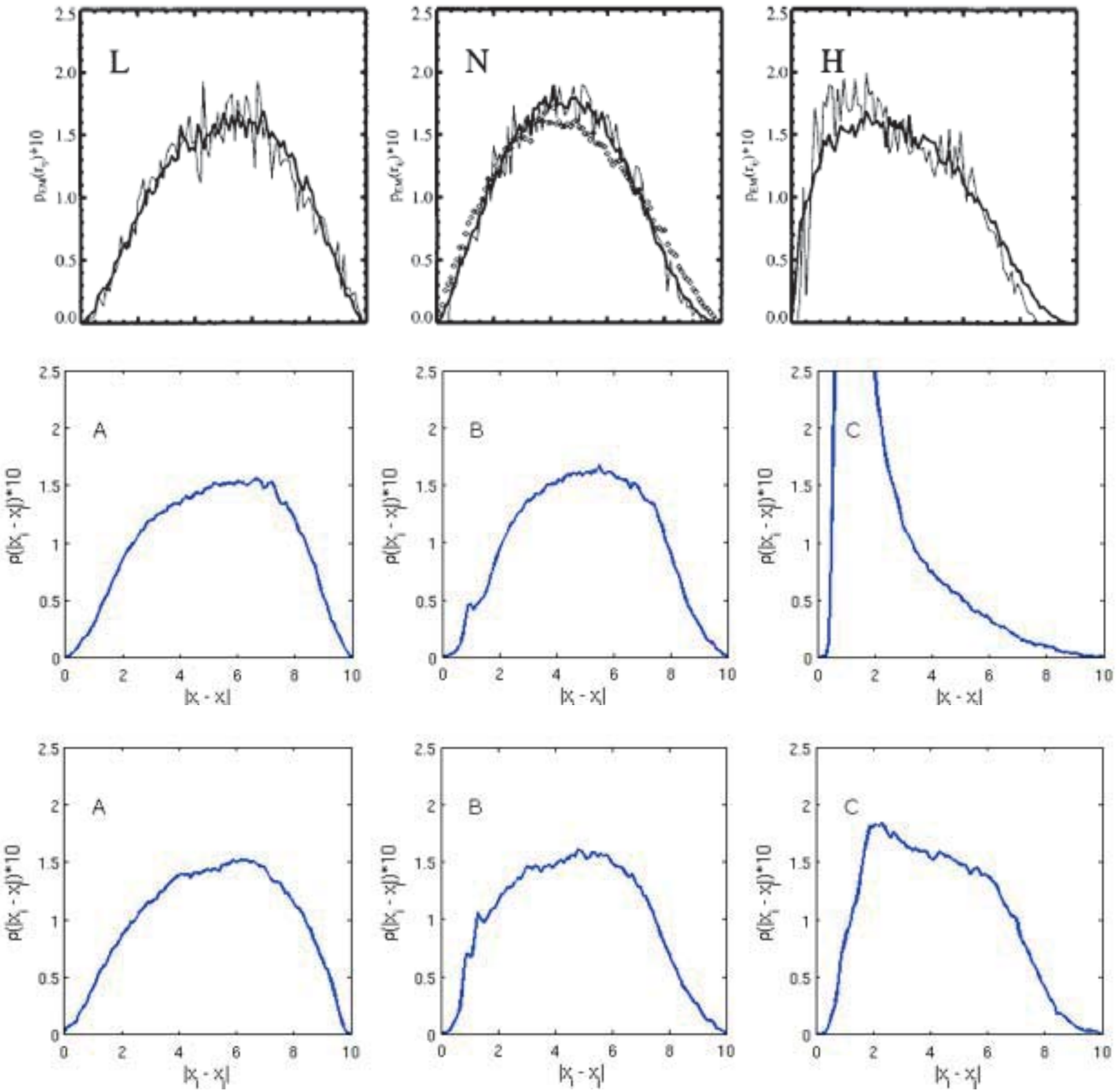
Distance between like signed vortices $|x_i - x_j|_{++}$

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

Buhler '02

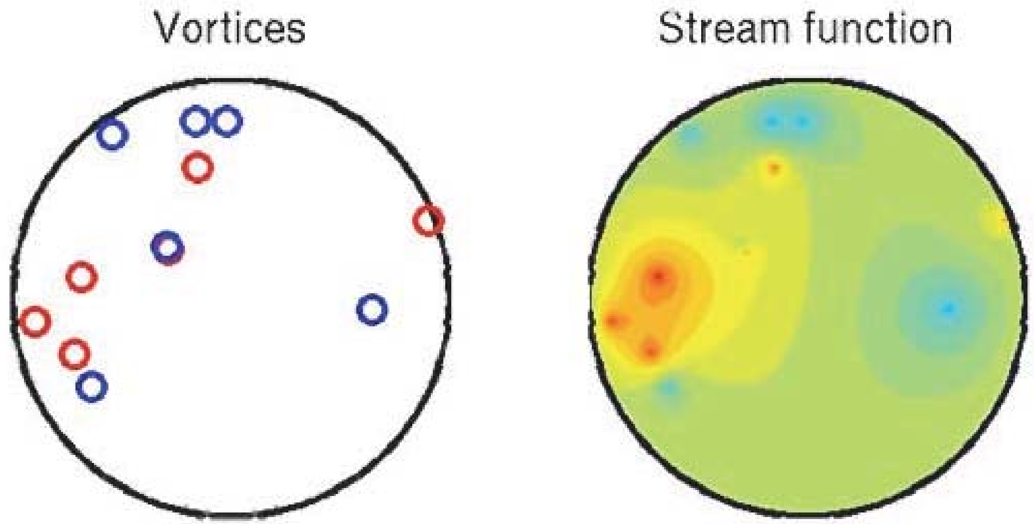
Gibbs

modified Gibbs

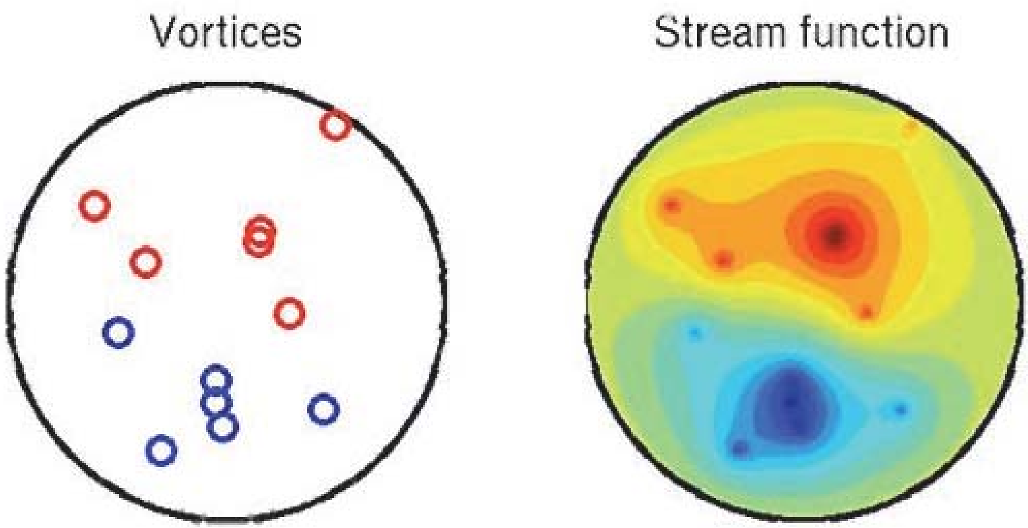


Vortex clustering, N=12

$\beta = 0.01$

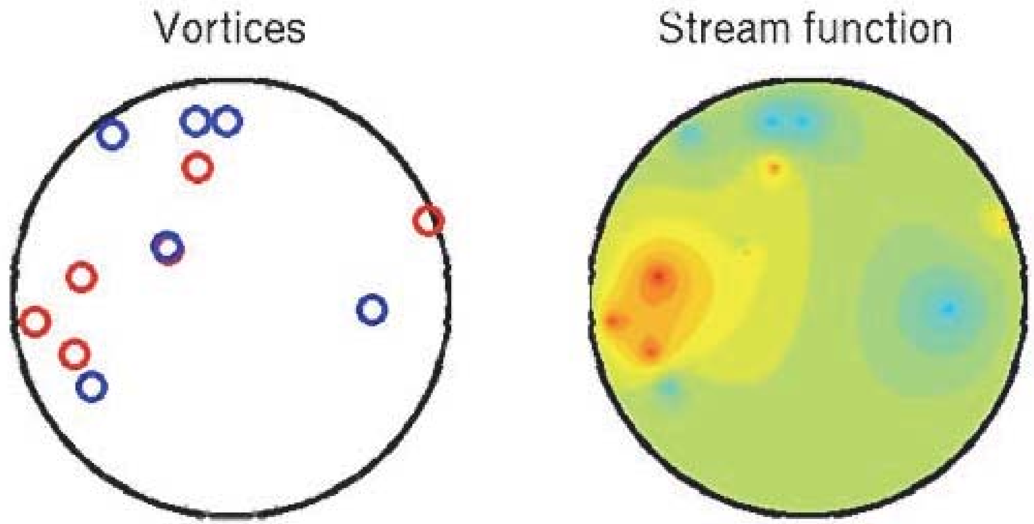


$\beta = -0.006$

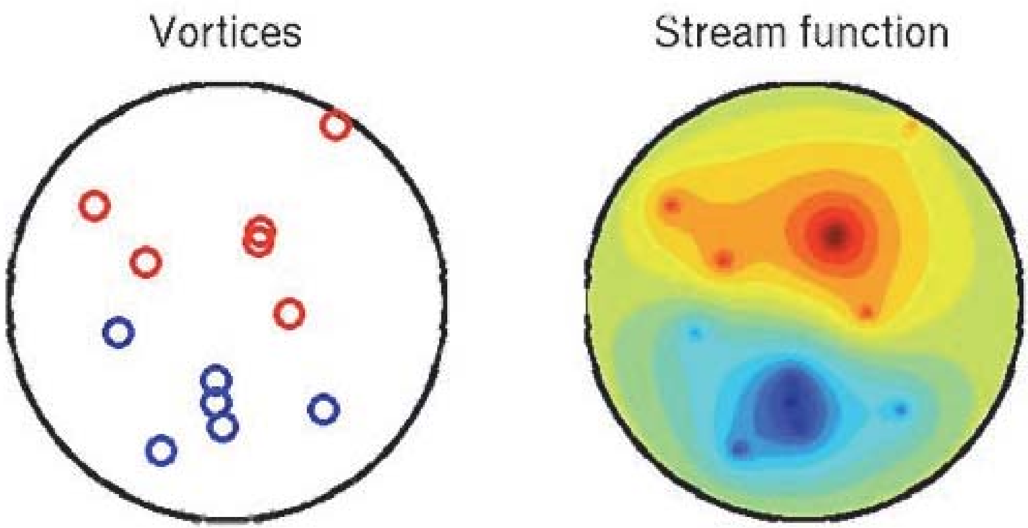


Vortex clustering, N=12

$\beta = 0.01$



$\beta = -0.006$



Burgers/KdV

Thermostat controls in Burgers-KdV

[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 used, artificial decrease

First steps: try to preserve a target equilibrium ensemble

Can we use a molecular ‘thermostat’ to control the ensemble in a semi-discrete Burgers/KdV model?

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

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

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

Truncated, discrete model

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

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

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

Two other first integrals

total momentum M, total kinetic energy E

Proposed ‘mixed’ distribution:

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

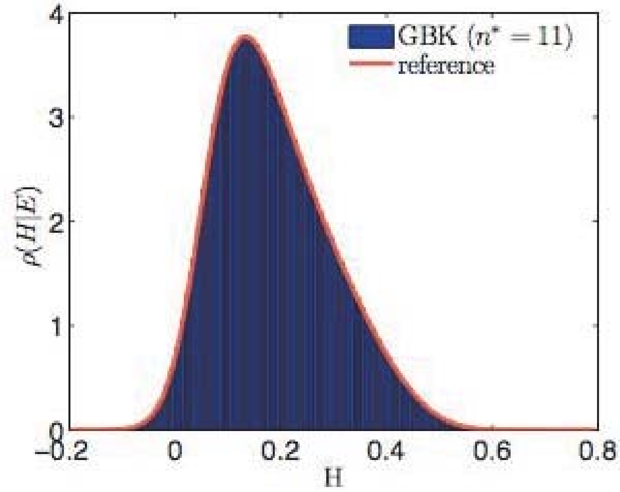
Now - design a highly degenerate thermostat

Notes:

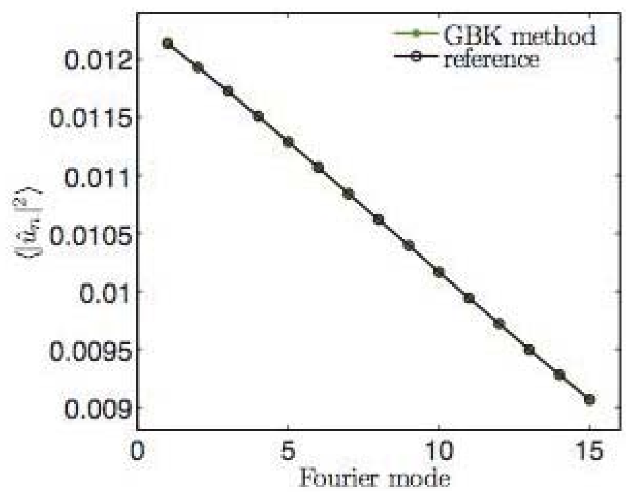
- The Hörmander condition is too hard for us to show
- we couple to the high wave numbers
and demonstrate ergodicity using numerics
- E and Mattingley - prove HC for coupling to **slow modes** (opposite of what we want)

Burgers

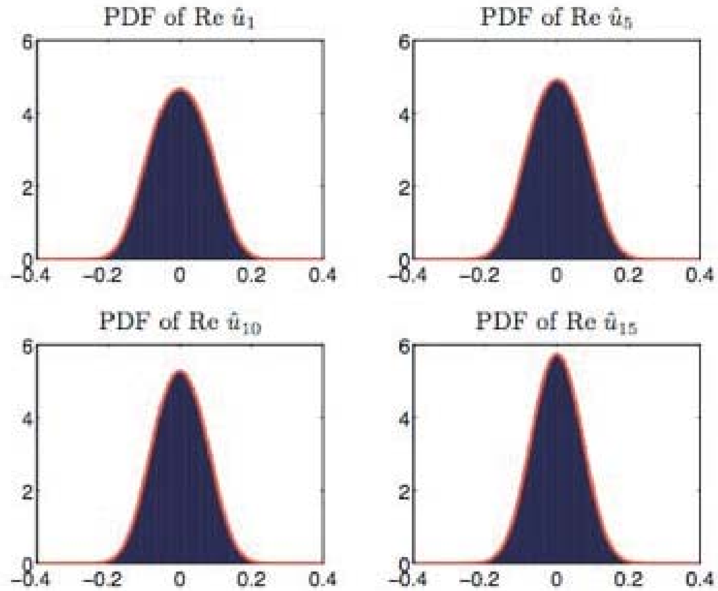
H Dist



Kinetic Energy



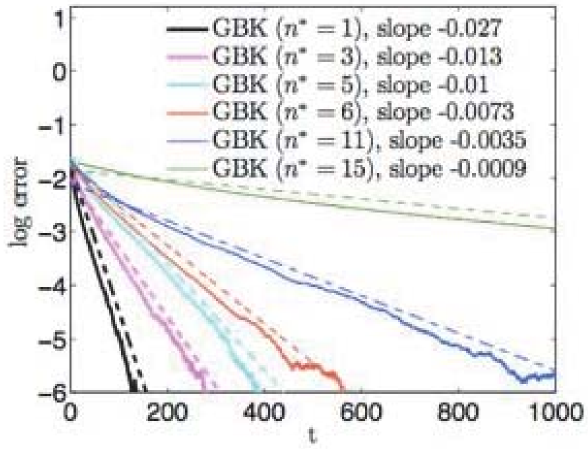
weak perturbation GBK(n^*=15)

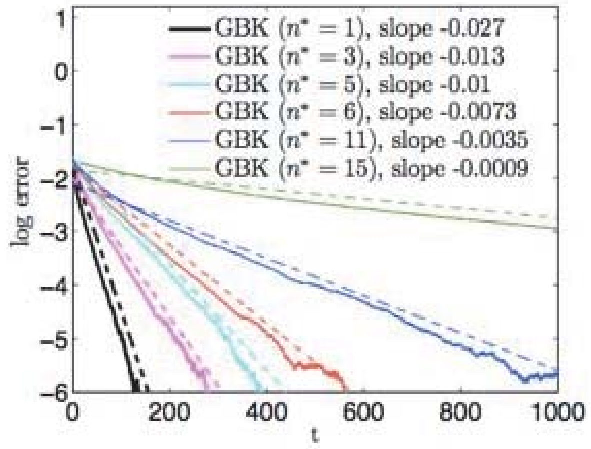


GBK(n=m)*: results using a **thermostat applied only to** modes *m…N*

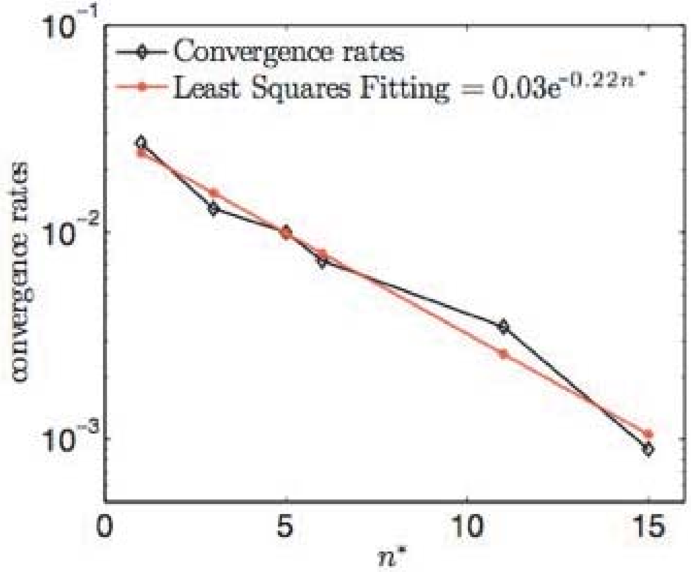
Burgers

Convergence of expected value of Hamiltonian

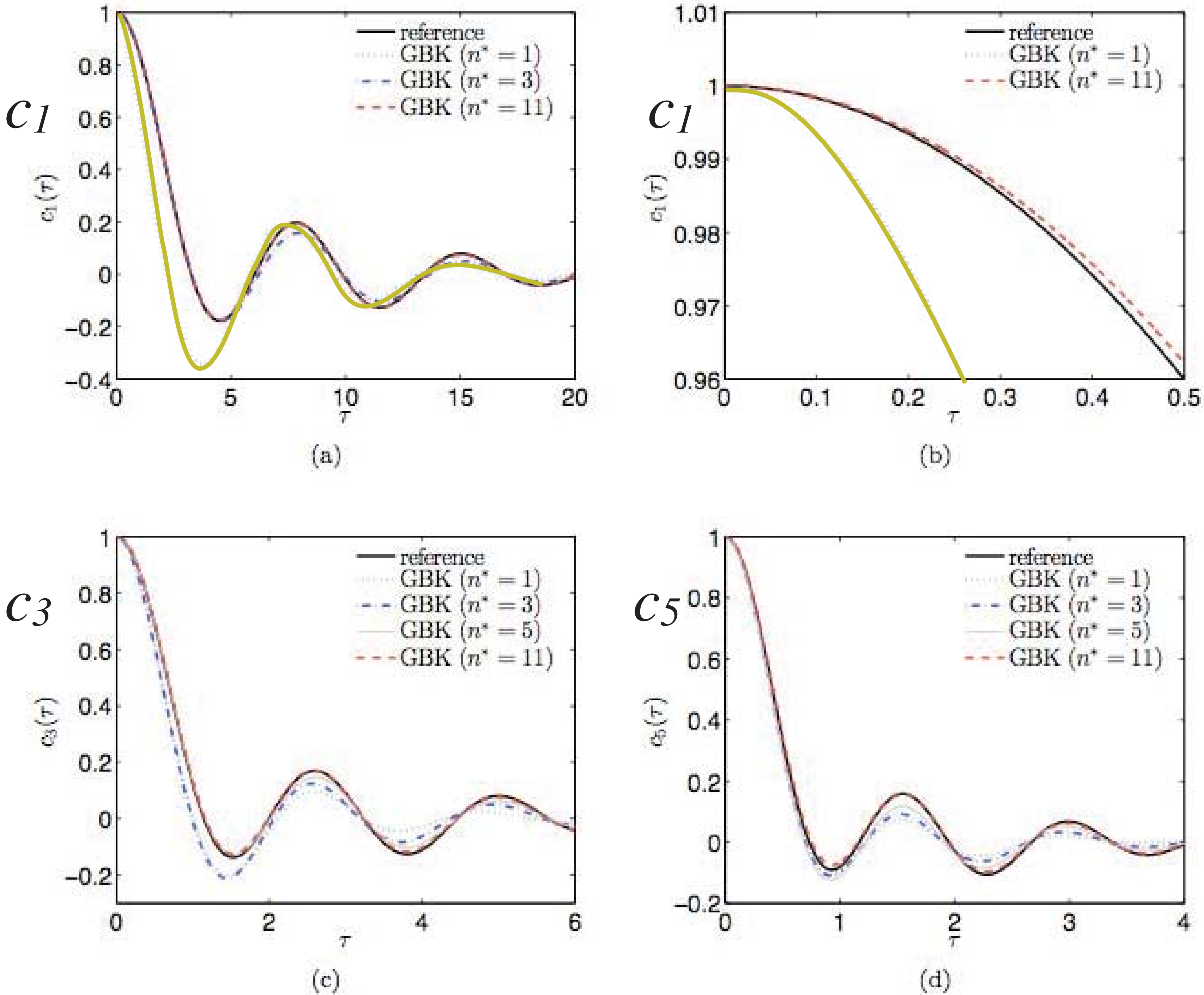




convergence of averages is observed in all cases, but is very slow for GBK($n^*=15$)



c_k : autocorrellation function for kth mode



2D Incompressible Navier Stokes
- 5 slides omitted.

Conclusions

1. SDE-based thermostats are versatile tools to approximate averages with respect to given density

2. Degenerate thermostats allow for efficient recovery, i.e., with small perturbation of dynamics

3. They can be applied beyond MD, e.g. in fluid dynamics (and more broadly)

4. Potentially valuable for model correction, data assimilation, etc., i.e. to restore properties of the equilibrium ideal system to a corrupted set of equations.