A new approach to statistical closure for complex deterministic dynamics

Bruce Turkington
Univ. of Massachusetts Amherst

Petr Plechac
Univ. of Tennessee / Oak Ridge National Lab.

Motivation

“Coarse-graining” and “subgrid-scale parameterization” abstracted to their mathematical essence:

- Choose a set of resolved variables that describe the macroscopic states of a complex dynamical system
- Identify macrostates with the ensemble-averaged resolved variables, and relegate unresolved variables to an appropriate statistical description
- Derive a closed set of equations for the evolution of those macrostates from the given microscopic dynamics

Statistical mechanics should furnish the natural tools for connecting microscopic dynamics to macroscopic behavior.
But a general theory of nonequilibrium statistical mechanics is not yet developed, and known techniques tend to focus on specially designed models.

For these reasons, we address the following generic problem:

* Consider a nonlinear Hamiltonian system with many degrees of freedom, and choose resolved dynamical variables. For instance,

- For a coupled system of oscillators, the coordinates of a few heavy particles interact with many light particles
- For a dispersive wave system, the energy spectrum in some selected bands of the many interacting modes

* Derive a statistical closure for the resolved dynamics that is consistent with the chaotic deterministic dynamics, without interposing a stochastic model of the unresolved variables.

Many multiscale problems that fall under the broad category of “turbulence” require such statistical closures, systematic model reductions and approximation schemes.

Background

The microstate is a point z in a *phase space* Γ:

$$z = (q_1, p_1, \ldots, q_n, p_n)^* \in \Gamma = R^{2n}, \qquad n \text{ is large}$$

The microscopic dynamics is *Hamiltonian*:

$$\frac{dz}{dt} = J \nabla_z H(z), \qquad \text{that is,} \quad \frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j},$$

Any *dynamical variable* $F : \Gamma \to R$ evolves under

$$\frac{dF}{dt} = \{F, H\} = (\nabla_z F)^* J \, \nabla_z H = \sum_j \frac{\partial F}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial F}{\partial p_j}\frac{\partial H}{\partial q_j}$$

The Hamiltonian dynamics defines a *phase flow* on Γ:

$$\Phi(t) : \Gamma \to \Gamma, \qquad \text{where} \qquad z(t) = \Phi(t)(z(0))$$

Probability on Γ flows with Φ: an *ensemble* of microscopic paths described by a probability measure $p(dz,t)$ satisfies

$$p(\Phi(t)(G),t) = p(G,0) \quad \text{for every} \;\; G \subset \Gamma.$$

In terms of probability density $\rho(z,t) = dp/dz$, the ensemble flow is governed by *Liouville's equation*:

$$\frac{\partial \rho}{\partial t} + L\rho = 0\,, \qquad \text{with } L\cdot = \{\,\cdot\,, H\}$$

The exact solution, $\rho(t) = e^{-tL}\rho(0)$, is extremely intricate and expensive to compute unless the dynamics is low-dimensional.

Instead we desire an approximation to $\rho(z,t)$ that is parametrized by the statistical averages of some *resolved variables*

$$A = (A_1, \ldots, A_m)^* \qquad\qquad (\text{ each } A_j : \Gamma \to R\,)$$

Equilibrium and nonequilibrium ensembles

Very long time average behavior is described by invariant densities such as the equilibrium canonical ensemble

$$\rho_{eq}(z) = \exp(-\beta[H(z) - f(\beta)])$$

with inverse temperature $\beta > 0$ and free energy

$$f(\beta) = -\beta^{-1} \log \int_\Gamma \exp(-\beta H(z))\, dz$$

For a *vector of resolved variables* $A : \Gamma \rightarrow R^m$ that are *not* invariants, a natural choice for a nonequilibrium ensemble is

$$\tilde{\rho}(z, \lambda) = \exp[\lambda^* A(z) - \phi(\lambda)]\, \rho_{eq}(z)$$

where

$$\phi(\lambda) = \log \int_\Gamma \exp(\lambda^* A(z)) \rho_{eq}(z)\, dz \qquad \text{for} \quad \lambda \in R^m$$

Provided that A contains the "slow" dynamical variables, we attempt to approximate the exact density $\rho(t)$ by the *trial density* $\tilde{\rho}(\lambda(t))$ with slowly evolving parameter vector $\lambda(t)$.

Two interpretations of $\tilde{\rho}(z,\lambda) = \exp[\lambda^* A(z) - \phi(\lambda)]\, \rho_{eq}(z)$:

Physics: Each $\tilde{\rho}$ is a *quasi-equilibrium ensemble*, which results from maximizing entropy

$$S(\tilde{\rho}) = -\int_\Gamma \tilde{\rho}(z) \log \frac{\tilde{\rho}(z)}{\rho_{eq}(z)}\, dz$$

given instantaneous mean values of the resolved variables

$$\int_\Gamma A(z)\tilde{\rho}(z)\, dz \doteq \langle A|\tilde{\rho}\rangle = a$$

The vector observable $a \in R^m$ is the macrostate in the coarse-grained description.

Statistics: $\tilde{\rho}$ is an *exponential family with natural parameter* λ. A is a minimal sufficient statistic for this parametric model.

Adiabatic closure

A naive way to evolve the parameter $\lambda(t) \in R^m$ is to impose $A_1, \ldots, A_m$ moments of the Liouville equation:

$$\frac{d}{dt}\langle A|\tilde{\rho}\,\rangle \; - \; \langle LA|\tilde{\rho}\,\rangle \; = \; \int_\Gamma A\,(\frac{\partial}{\partial t} + L)\,\tilde{\rho}\,dz \; = \; 0\,.$$

But this closed reduced dynamics conserves entropy exactly:

$$\frac{d}{dt}S(\tilde{\rho}) \; = \; -\frac{d}{dt}\left[\,\lambda^* a - \phi(\lambda)\,\right] \; = \; \int_\Gamma L\,e^{\lambda^* A - \phi(\lambda)}\,\rho_{eq}\,dz \; = \; 0\,.$$

The fundamental deficiency is that this closure is memoryless, whereas unresolved fluctuations influence macroscopic evolution over a finite time until they equilibriate.

Zubarev (1970s) proposed using "nonequilibrium statistical operators" with decaying memory

$$\rho(t) \; = \; \exp\left(\int_0^{+\infty} \lambda(t-\tau)\,A \circ \Phi(-\tau)\,e^{-\alpha\tau}\,\alpha d\tau \; - \; \psi[\lambda](t)\right)\,\rho_{eq}$$

Best-fit closure

Retain the trial densities $\tilde{\rho}$, and evaluate the *Liouville residual*:

$$R \doteq \left(\frac{\partial}{\partial t} + L\right) \log \tilde{\rho}(\lambda(t)) = \dot{\lambda}(t)^*(A - a(t)) + \lambda(t)^* LA$$

Two expressions for $R = R(z; \lambda, \dot{\lambda})$:

(1) At any microstate $z \in \Gamma$, consider the information for discrimination between the exact density and the trial density after a time increment θ:

$$\log \frac{e^{-\theta L}\tilde{\rho}(z, \lambda(t))}{\tilde{\rho}(z, \lambda(t+\theta))} = -\theta\, R(z) + O(\theta^2) \qquad \text{as } \theta \to 0.$$

(2) For any observable B, consider the evolution of the ensemble-averaged B with respect to the trial density:

$$\frac{d}{dt}\langle B \,|\, \tilde{\rho}(\lambda(t)) \rangle - \langle LB \,|\, \tilde{\rho}(\lambda(t)) \rangle = \langle BR \,|\, \tilde{\rho}(\lambda(t)) \rangle$$

The statistic R represents instantaneous “information loss rate” for $\tilde{\rho}(\lambda)$ along a parameter trajectory $\lambda(t)$.

Separate the Liouville residual into its orthogonal components along the resolved and unresolved subspaces of $L^2(\Gamma, \tilde{\rho}(\lambda))$:

$$\begin{aligned} R &= P_\lambda R + Q_\lambda R \\ &= [\, \dot{\lambda} - C(\lambda)^{-1} \langle LA | \tilde{\rho}(\lambda) \rangle \,]^* (A - a) + \lambda^* (Q_\lambda LA) \end{aligned}$$

where $C(\lambda) = \langle (A-a)(A-a)^* | \tilde{\rho}(\lambda) \rangle$ (Fisher information matrix).

Define the following lack-of-fit functional for this residual:

$$\int_{t_0}^{t_1} \mathcal{L}(\lambda, \dot{\lambda}) \, dt = \frac{1}{2} \int_{t_0}^{t_1} \langle \, (P_\lambda R)^2 | \, \tilde{\rho}(\lambda) \, \rangle + \epsilon^2 \langle \, (Q_\lambda R)^2 | \, \tilde{\rho}(\lambda) \, \rangle \, dt \, .$$

$\epsilon \in (0, 1]$ is an adjustable parameter in the “Lagrangian” $\mathcal{L}(\lambda, \dot{\lambda})$.

Our *best-fit closure* minimizes this “action” over trajectories $\lambda(t)$, $t_0 \le t \le t_1$, in the coarse-grained parameter space R^m.

The predicted macrostate at time t_1 corresponds to $\lambda(t_1)$.

Best-fit reduced dynamics is deduced from the *value function*:

$$v(\lambda_1, t_1) = \min_{\lambda(t_1)=\lambda_1} \; v_0(\lambda(t_0)) + \int_{t_0}^{t_1} \mathcal{L}(\lambda, \dot{\lambda})\, dt$$

This minimization is over parameter trajectories $\lambda(t) \in R^m$, $t_0 \leq t \leq t_1$ having terminal point λ_1 at terminal time t_1.

$v_0(\lambda)$ is a given penalty function on the initial parameter $\lambda(t_0)$.

$v(\lambda_1, t_1)$ assigns an optimal lack-of-fit to any feasible pair (λ_1, t_1).

The value function v quantifies the propagation of uncertainty due to coarse-graining the Hamiltonian microdynamics via the resolved vector A.

v depends on the *"relevance factor"* ϵ in the lack-of-fit metric.

An analysis using the Mori-Zwanzig memory kernel relates ϵ to the separation of relevant time scales:

$$\epsilon \sim \frac{\tau_{memory}}{\tau_{plateau}} \sim \frac{\tau_{plateau}}{\tau_{relaxation}}$$

The value function $v(\lambda, t)$ solves the *Hamilton-Jacobi equation*

$$\frac{\partial v}{\partial t} + \mathcal{H}\left(\lambda, \frac{\partial v}{\partial \lambda}\right) = 0\,, \qquad \text{with } \; v(\lambda, t_0) = v_0(\lambda),$$

where $\mathcal{H}(\lambda, \mu)$ is the Legendre transform of $\mathcal{L}(\lambda, \dot{\lambda})$,

$$\mu = \frac{\partial \mathcal{L}}{\partial \dot{\lambda}}\,, \qquad \mathcal{H} = \dot{\lambda}^* \mu - \mathcal{L}(\lambda, \dot{\lambda})\,.$$

$\mathcal{H}$ is explicitly calculable because $\mathcal{L}$ is quadratic in $\dot{\lambda}$.

The conjugate variable $\mu = \langle AR \,|\, \tilde{\rho} \rangle$ is the irreversible flux of A.

The best-fit prediction is the v-minimizing parameter trajectory

$$\widehat{\lambda}(t) = \arg\min_{\lambda} v(\lambda, t)\,.$$

The best-fit macrostate is

$$\widehat{a}(t) = \langle\, A \,|\, \tilde{\rho}(\widehat{\lambda}(t))\,\rangle = \frac{\partial \phi}{\partial \lambda}(\widehat{\lambda}(t))\,.$$

Summary of general closure scheme

Given a coarse-graining defined by the resolved vector A, and the associated family of trial densities $\tilde{\rho}(\lambda)$,

* Construct the lack-of-fit Lagrangian $\mathcal{L}$ and its Hamiltonian $\mathcal{H}$;

* Solve the Hamilton-Jacobi equation for the value function v;

*Derive the closed set of first-order differential equations for the best-fit parameter $\widehat{\lambda}(t)$, and hence the predicted macrostate $\widehat{a}(t)$.

The best-fit macroscopic dynamics is calculated without simulation of any microscopic solutions.

There is a single parameter ϵ that is tuned to produce the correct dissipation.

The uncertainty in the best-fit prediction is quantified by the value function.

Near equilibrium approximation

The best-fit reduced equations are linear relaxation equations near equilibrium.

Normalize A relative to equilibrium ρ_{eq}, $\langle A\rangle_{eq} = 0$, $\langle AH\rangle_{eq} = 0$, so that from any small initial disturbance $\lambda_0 \neq 0$ the system relaxes toward equilibrium ρ_{eq}: $\lambda(t) \to 0$ as $t \to \infty$.

Linear response approximation pertains to small $|\lambda|$:

$$\begin{aligned} \langle F \,|\, \tilde{\rho}(t)\rangle \;&=\; \langle\, F\,[\,1 + \lambda(t)^* A\,]\,\rangle_{eq} \;+\; O(|\lambda|^2) \\ &\approx\; \langle F A^*\rangle_{eq}\,\lambda(t) \end{aligned}$$

for any dynamical variable F with $\langle F\rangle_{eq} = 0$.

The Lagrangian (running cost) is a quadratic form

$$\mathcal{L}(\lambda,\dot{\lambda}) \;\approx\; \frac{1}{2}[\dot{\lambda} - C^{-1}J\lambda]^* C[\,\dot{\lambda} - C^{-1}J\lambda\,] \;+\; \frac{\epsilon^2}{2}\lambda^* D\lambda\,,$$

with coefficient matrices

$$C = \langle AA^*\rangle_{eq}\,, \quad J = \langle (LA)A^*\rangle_{eq}\,, \quad D = \langle\,(QLA)\,(QLA)^*\,\rangle_{eq}$$

Best-fit linear irreversible reduced dynamics

The predicted evolution of the macrostate is governed by

$$\frac{d\widehat{a}}{dt} = \Big[\, J \,-\, \epsilon^2 C G(t;\epsilon) D \,\Big]\, C^{-1}\widehat{a} \qquad \text{with} \quad \widehat{a}(0) = a_0\,,$$

where the matrix $G = G(t;\epsilon)$ solves the Riccati equation

$$\frac{dG}{dt} = JC^{-1}G - GC^{-1}J - \epsilon^2 GDG + C^{-1} \qquad \text{with} \quad G(0) = 0\,.$$

The value function is the time-varying quadratic form

$$v(\lambda,t) \;=\; \frac{1}{2}(\lambda - \widehat{\lambda}(t))^* G(t;\epsilon)^{-1}(\lambda - \widehat{\lambda}(t)) \;+\; \widehat{v}(t)$$

If the variables A are even under time reversal, then $J = 0$ and this closed reduced dynamics is analogous to *linear irreversible thermodynamics*:

- The matrix of transport coefficients $\epsilon^2 C G(t;\epsilon) D$ is positive-definite and symmetric (Onsager reciprocity relations).

- The entropy $\widehat{s} = -\frac{1}{2}\,\widehat{a}^* C^{-1}\widehat{a}$ is monotonic increasing in time.

Basic numerical test problem: Kac-Zwanzig model

Consider a coupled system of $n = 1 + N$ classical oscillators, having Hamiltonian

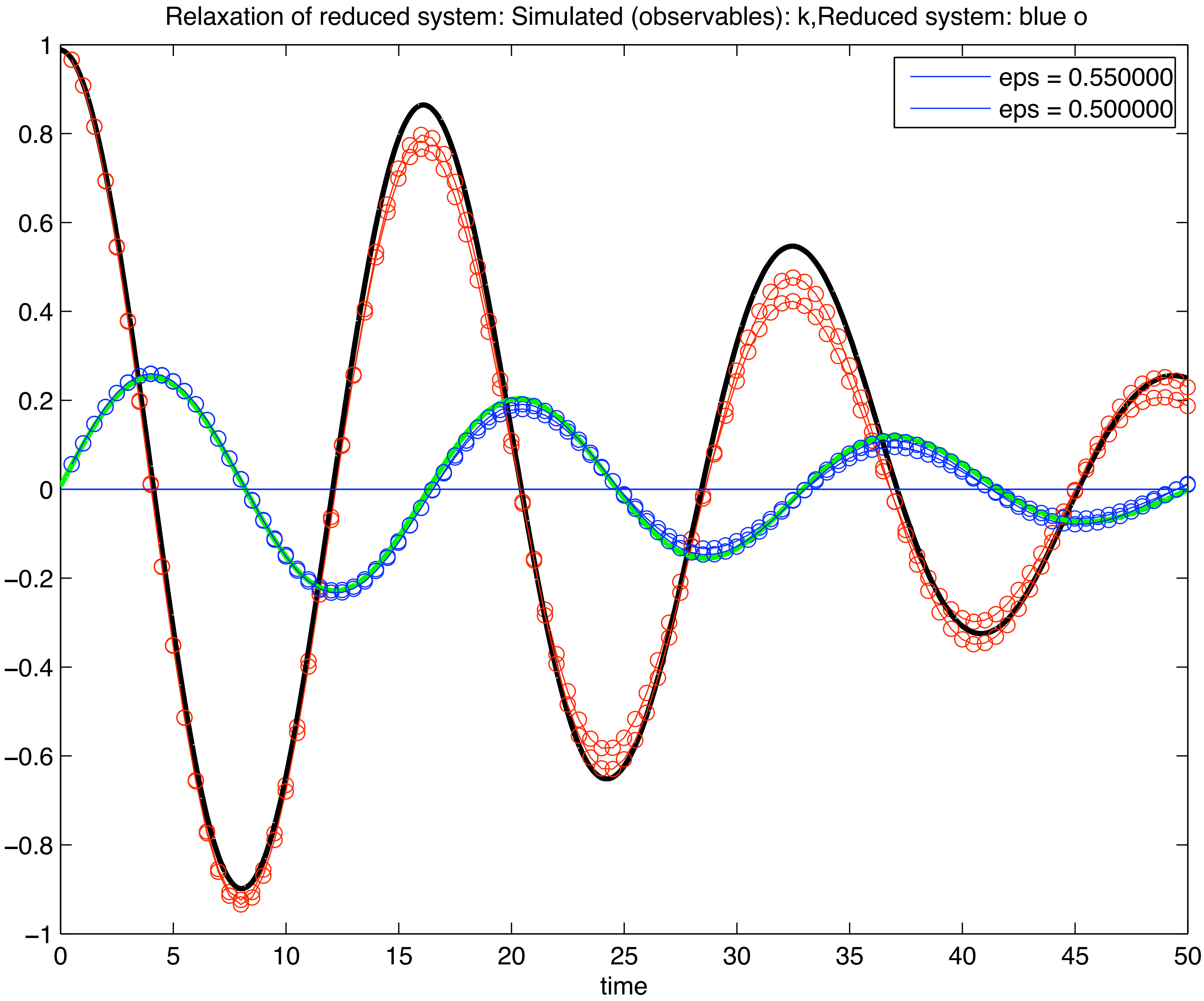
$$H = \frac{1}{2M}P^2 + \frac{1}{2}Q^2 + \frac{\eta}{4}Q^4 + \sum_{k=1}^{N} \frac{1}{2}p_k^2 + \frac{1}{2}\omega_k^2 \left(q_k - \frac{\gamma}{\omega_k^2}Q\right)^2$$

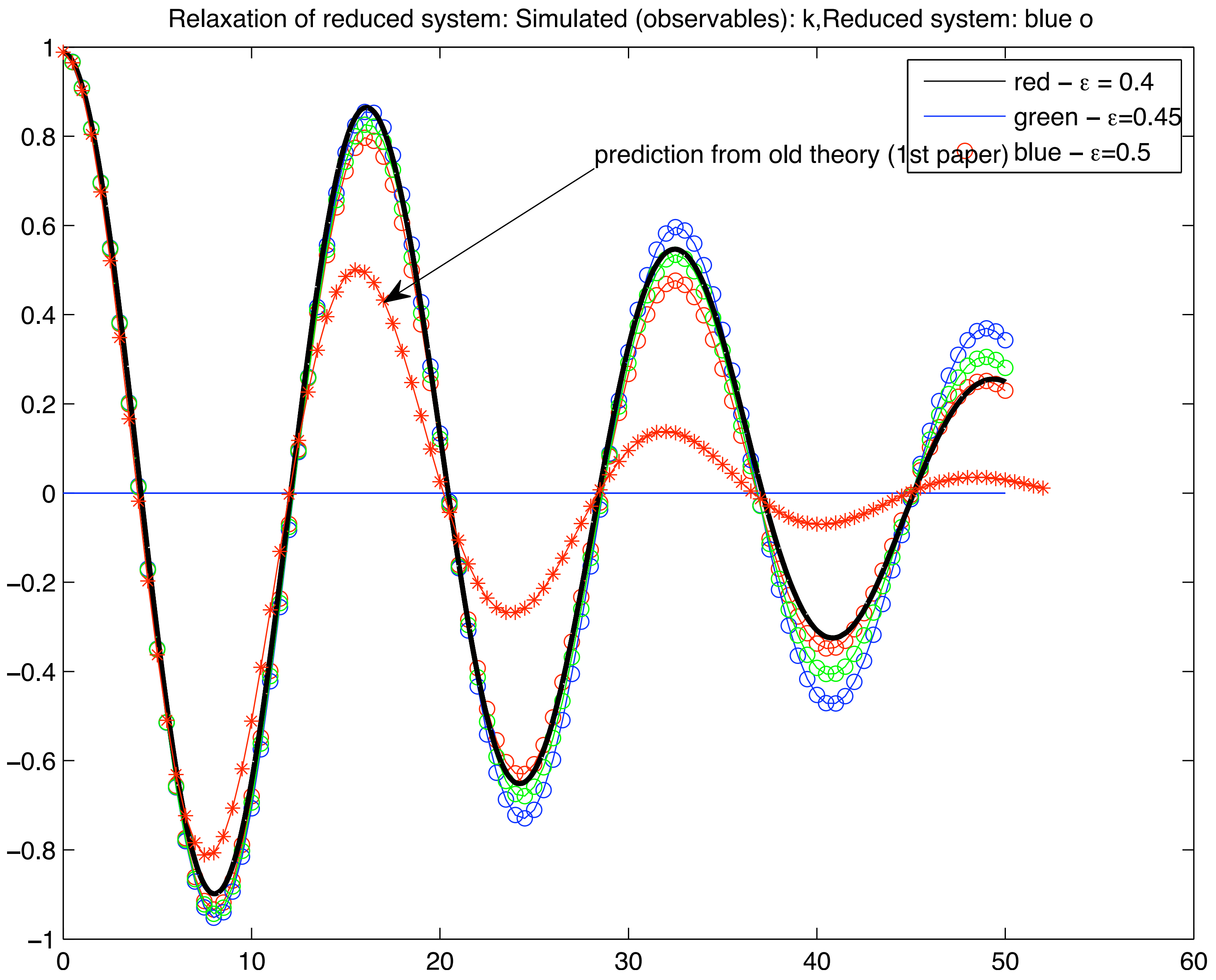
A single slow quartic oscillator is linearly coupled to N fast harmonic oscillators.

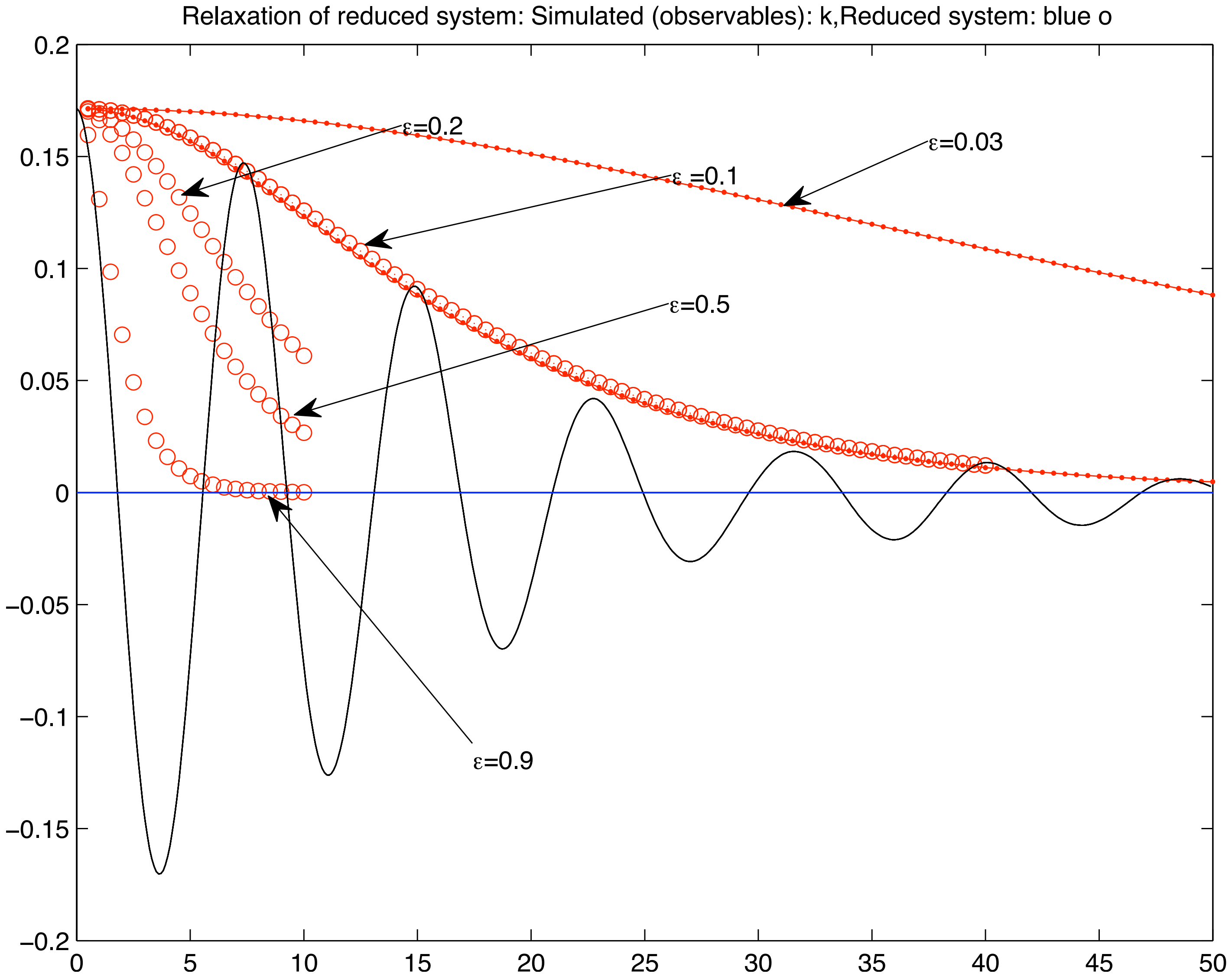
Select the resolved vector to be $A = (P, Q)^*$ for the single particle with mass $M > 1$. "Heat bath" of $N = 20$ particles with coordinates (p_k, q_k) and unit mass are unresolved.

In computed examples, the frequencies ω_k are sampled from a gamma distribution with mean ≈ 30 and shape factor 3; coupling $\gamma = 1$, and nonlinearity $\eta = 0.5$.

The equilibrium ensemble with $\beta = 1$ is simulated using an MCMC algorithm with between 10^4 to 10^5 samples.


Relaxation of reduced system: Simulated (observables): k,Reduced system: blue o
eps = 0.550000
eps = 0.500000
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0
5
10
15
20
25
30
35
40
45
50
time


Relaxation of reduced system: Simulated (observables): k,Reduced system: blue o
red – ε = 0.4
green – ε=0.45
blue – ε=0.5
prediction from old theory (1st paper)
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0
10
20
30
40
50
60


Relaxation of reduced system: Simulated (observables): k,Reduced system: blue o
0.2
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
−0.2
0
5
10
15
20
25
30
35
40
45
50
ε=0.2
ε=0.03
ε =0.1
ε=0.5
ε=0.9

Conclusions

* Best-fit closure using quasi-equilibrium ensembles offers a systematic alternative to the traditional closure schemes that truncate a moment hierarchy.

* Given a coarse-graining and a lack-of-fit metric, the best-fit prediction and its uncertainty are controlled by a value function, which propagates under a Hamilton-Jacobi PDE.

* In the near equilibrium regime, the theory produces linear relaxation equations whose transport matrices are derived from the underlying dynamics up to a single adjustable parameter.

* A future goal is to apply this approach to specific subgrid-scale parameterizations and turbulence problems, including geophysical fluid dynamics,