Optimal closure for nonequilibrium statistical models

Bruce Turkington
University of Massachusetts Amherst

Collaborators:

Richard Kleeman (Courant Inst., NYU)
Qian-Yong Chen (UMass Amherst)
Simon Thalabard (postdoc)

Statistical closure via dynamical optimization

(1) Coarse-grain the underlying deterministic dynamics using a parametric statistical model,
(2) Quantify the lack-of-fit to the fully-resolved dynamics of any feasible path in the model parameter space,
(3) Derive the closed equations governing the reduced model from the optimality conditions for the best-fit path.

B.T., "An optimization principle for deriving nonequilibrium statistical models of Hamiltonian dynamics." *J. Stat. Phys.* 152:569–597 (2013).

- Optimal closure constructed for any Hamiltonian dynamics and any vector of relevant observables.

- The reduced equations have the "GENERIC" or "metriplectic" structure of nonequilibrium thermodynamics

An example: Two-dimensional fluid turbulence

B.T., Q.Y. Chen and S. Thalabard, "Coarse-graining two-dimensional turbulence via dynamical optimization." *arXiv* 1510.00341 (2015).

Microdynamics is the spectrally-truncated vorticity equation:

$$\zeta(x,t) = \sum_{k\in\Lambda_n} z_k(t)\, e^{ik\cdot x}\,, \qquad \Lambda_n \text{ is finite lattice of wavevectors}$$

$$\frac{dz_k}{dt} + \sum_{p+q=k} c(p,q) z_p z_q = 0\,, \qquad c(p,q) = \frac{1}{2}\, p\times q \left(\frac{1}{|p|^2} - \frac{1}{|q|^2} \right) .$$

Energy and enstrophy invariance determine Gibbs density,

$$\rho_{eq}(z) \propto \exp\left(- \sum_{k\in\Lambda_n} \beta_k |z_k|^2 \right) , \qquad \text{for } \beta_k = \alpha + \frac{\theta}{|k|^2} .$$

Parametric statistical model consists of Gaussian densities:

$$\tilde{\rho}(z; a, b) \propto \exp\left(- \sum_{k\in\Lambda_n} b_k |z_k - a_k|^2 \right) .$$

General framework

Hamiltonian microscopic dynamics (canonical or noncanonical)

$$\frac{dF}{dt} = \{F, H\} \qquad \text{for every observable } F \text{ on a phase space } \Gamma$$

Generic point $z = (z_1, \ldots, z_n) \in \Gamma$ is microstate. $n \gg 1$.

An *ensemble* of microscopic trajectories, $z(t)$, is described by a probability density, $\rho(z, t)$, which satisfies *Liouville's equation*:

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

Equivalently, for every observable F,

$$\frac{d}{dt}\langle F \,|\, \rho(\cdot, t)\rangle = \langle LF \,|\, \rho(\cdot, t)\rangle\,, \qquad \text{for} \quad \langle F \,|\, \rho\rangle \doteq \int_\Gamma F(z)\rho(z)\, dz\,.$$

But, the complexity of the exact solution, $\rho(\cdot, t) = e^{-tL}\rho(\cdot, 0)$, equals that of the underlying dynamics, making it too expensive to compute in practice.

Statistical-dynamical model

Replace the exact density $\rho(z,t)$ by an approximation $\tilde{\rho}(z;\lambda(t))$, in which the parameter vector, $\lambda = \lambda(t)$, evolves.

A classic choice is quasi-equilibrium densities associated to a vector of relevant (slow) observables, $A = (A_1, \ldots, A_m)$, $m \ll n$. This statistical model uses the exponential trial densities

$$\tilde{\rho}(z,\lambda) \,=\, \exp\left(\sum_{i=1}^{m} \lambda_i A_i(z) \,-\, \phi(\lambda)\right) \rho_{eq}(z)$$

$\tilde{\rho}(z,\lambda)$ maximizes entropy subject to the expectation of A:

$$s(a) = \max_{\rho} \langle -\log \frac{\rho}{\rho_{eq}} \,|\, \rho \,\rangle \quad \text{over} \quad \langle\, A \,|\, \rho \,\rangle = a\,, \quad \langle\, 1 \,|\, \rho \,\rangle = 1\,.$$

The parameter vector $\lambda = (\lambda_1, \ldots, \lambda_m)$ consists of the associated Lagrange multipliers, $\lambda_i = -\frac{\partial s}{\partial a_i}$.

Quantifying the lack-of-fit of the model

Evaluate the *Liouville residual* of the trial densities $\tilde{\rho}(z; \lambda(t))$ along any feasible parameter path $\lambda(t)$:

$$R = R(\,\cdot\,;\, \lambda(t), \dot{\lambda}(t)\,) \doteq \left(\frac{\partial}{\partial t} + L\right) \log \tilde{\rho}(\cdot\,; \lambda(t))$$

The ensemble-averaged evolution of any observable F along such a path satisfies

$$\frac{d}{dt}\langle F \,|\, \tilde{\rho}\,\rangle = \langle LF \,|\, \tilde{\rho}\,\rangle + \langle FR \,|\, \tilde{\rho}\,\rangle$$

For instance, in the quasi-equilibrium model,

$$R = \sum_{i=1}^{m} \dot{\lambda}_i(t)(A_i - a_i(t)) + \lambda_i(t)\{A_i\,, H\}\,.$$

Define the statistical *lack-of-fit* to be

$$\mathcal{L}(\lambda, \dot{\lambda}) = \frac{1}{2} \langle\, R(\cdot\,; \lambda, \dot{\lambda}\,)^2 \mid \tilde{\rho}(\cdot\,; \lambda)\,\rangle$$

$\mathcal{L}(\lambda, \dot{\lambda})$ represents the information loss rate due to coarse-graining:

$$\langle\, \log \frac{e^{-\Delta t\, L} \tilde{\rho}(\lambda(t))}{\tilde{\rho}(\lambda(t + \Delta t)\,)} \mid e^{-\Delta t\, L} \tilde{\rho}(\lambda(t))\,\rangle \;=\; (\Delta t)^2\, \mathcal{L}(\lambda, \dot{\lambda}) \;+\; O(\Delta t)^3$$

Incremental Kullback-Leibler distance between exact and model densities.

Characterization of optimal paths:

$$\min_{\lambda(t)} \int_0^\infty \mathcal{L}(\lambda, \dot{\lambda})\, dt \qquad \text{subject to } \lambda(0) = \lambda_0\,.$$

This optimization resembles a classical least action principle, BUT

(1) this "action" has units of entropy production,

(2) this "Lagrangian" $\mathcal{L}$ is a sum of positive-definite parts,

(3) these extremals relax to equilibrium as $t \to +\infty$.

Hamilton-Jacobi theory

Introduce the optimal cost function, or *value function*,

$$v(\lambda_0) = \min_{\lambda(0)=\lambda_0} \int_0^\infty \mathcal{L}(\lambda, \dot{\lambda})\, dt\,,$$

$v(\lambda)$ solves the stationary Hamilton-Jacobi equation

$$\mathcal{H}\left(\lambda, -\frac{\partial v}{\partial \lambda}\right) = 0\,,$$

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

$$\pi_i = \frac{\partial \mathcal{L}}{\partial \dot{\lambda}_i} = \cdots\,, \qquad \mathcal{H} = \sum_{i=1}^{m} \dot{\lambda}_i \pi_i - \mathcal{L} = \cdots\,.$$

Along an extremal path $\lambda(t)$, the conjugate vector $\pi(t)$ is

$$\pi_i(t) = -\frac{\partial v}{\partial \lambda_i}(\lambda(t))\,.$$

Replacing π by its expression in terms of $\dot{\lambda}$ and λ then produces the desired *closed system of first-order DEs* in λ.

Optimal closure equations for quasi-equilibrium models

The closed reduced equations expressed in terms of the mean resolved vector $a = \langle A \,|\, \tilde{\rho}(\lambda) \rangle$:

$$\frac{da}{dt} = J(a)\frac{\partial h}{\partial a} - \frac{\partial v}{\partial \lambda} \qquad \text{with} \qquad \lambda = -\frac{\partial s}{\partial a}$$

Reversible part: Hamiltonian-Poisson structure with

$$J_{ij} = \langle \{A_i, A_j\} \rangle = -J_{ji}\,, \qquad h(a) = \langle H \,|\, \tilde{\rho}(\lambda) \rangle$$

Irreversible part: Generalized gradient structure with "dissipation potential" $v(\lambda)$

Optimal closure then has the structure of *nonequilibrium thermodynamics:*

- "GENERIC" (General Equations of NonEquilibrium Reversible Irreversible Coupling) due to Grmela and Öttinger, and Beris and Edwards
- "Metriplectic dynamics" due to Morrison.

Statistical relaxation of truncated two-dimensional flow

Inviscid, unforced 2D Euler dynamics,

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = -\nabla p \,, \qquad \nabla \cdot u = 0 \,,$$

is equivalent to the self-advection of vorticity, $\zeta(x,t) = \nabla \times u(x,t)$. Spectrally truncate this dynamics onto $2n(n+1)$ Fourier modes,

$$\zeta(x,t) = \sum_{k \in \Lambda_n} z_k(t) \, e^{ik \cdot x} \,, \qquad \text{[doubly-periodic boundary conditions]}$$

using the lattice of wavevectors,

$$\Lambda_n = \{ \, k = (k_1, k_2) \neq (0,0) \, : \, k_1, k_2 = -n, \ldots, n \, \} \,,$$

The microdynamics has a typical quadratic nonlinearity:

$$\frac{dz_k}{dt} + \sum_{p+q=k} c(p,q) z_p z_q = 0 \,, \qquad c(p,q) = \frac{1}{2} \, p \times q \left(\frac{1}{|p|^2} - \frac{1}{|q|^2} \right) \,.$$

Two Gaussian nonequilibrium models

(1) As resolved variables use the modes with low wavenumbers:
$a_k = \langle z_k \rangle$ for $k \in \Lambda_m$, $m \ll n$; $\quad \langle |z_k - a_k|^2 \rangle = 1/\beta_k$ for all $k \in \Lambda_n$.

Trial densities are quasi-equilibrium with perturbed means a_k in the low modes, but equilibrium variance in *all* modes.

(2) Resolve both the means and variances of the low modes:
$a_k = \langle z_k \rangle$ and $1/b_k = \langle |z_k - a_k|^2 \rangle$ for $k \in \Lambda_m$.

Trial densities are Gaussian, and all *unresolved* modes have equilibrium means and variances:

$$\rho_{eq}(z) \propto \exp\left(- \sum_{k \in \Lambda_n} \beta_k |z_k|^2 \right) , \qquad \text{where } \beta_k \doteq \alpha + \frac{\theta}{|k|^2} \, .$$

Coarse-grained equations of motion

(1) Closure in terms of the low-mode means, $a_k = \langle\, z_k \,\rangle$:

$$\frac{da_k}{dt} + \sum_{p+q=k} [c(p,q) + d(p,q)]\, a_p a_q = -\sigma_k a_k \,,$$

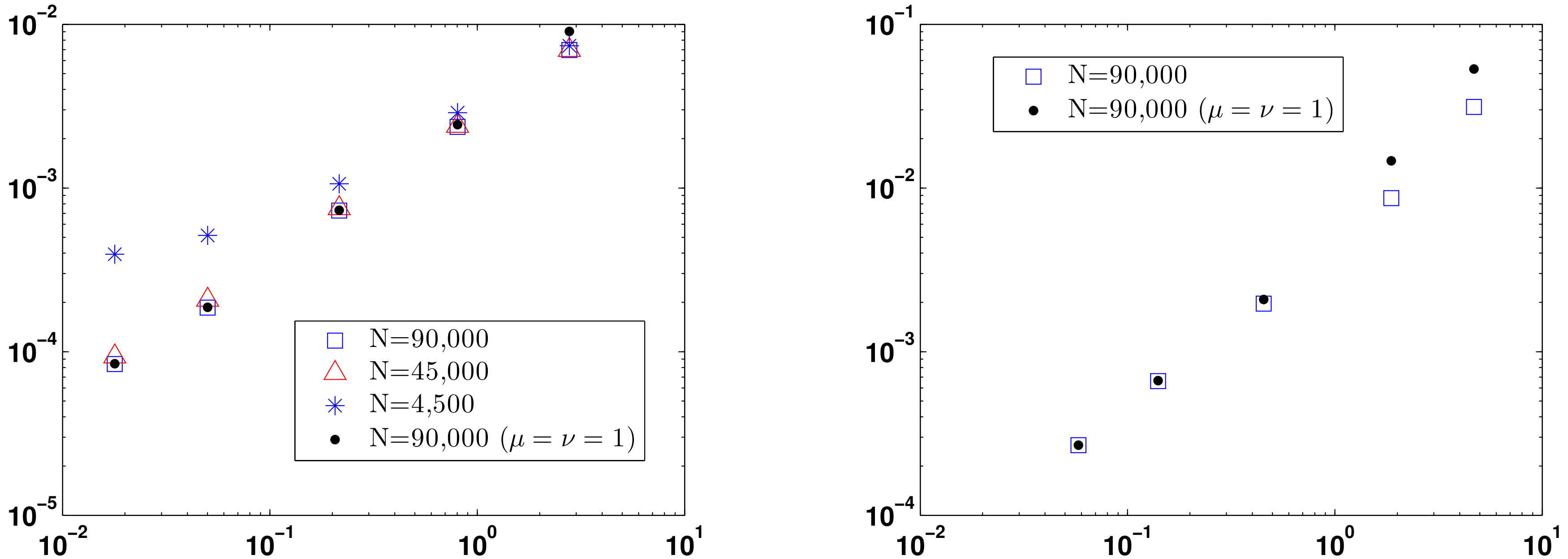
(2) Closure in terms of low-mode means, a_k, together with the (inverse) variances $b_k^{-1} = \langle\, |z_k - a_k|^2 \,\rangle$:

$$\frac{da_k}{dt} + \sum_{p+q=k} [c(p,q) + d(p,q)]\, a_p a_q = -\sigma_k a_k \left[1 + \xi_k (b_k - \beta_k)\right] \,,$$

$$\frac{db_k}{dt} = -\sigma_k \chi_k |a_k|^2 - \sigma_k' (b_k - \beta_k) \left[1 + \eta_k (b_k - \beta_k)\right]$$

Properties:
(a) Subdiffusive, nonlocal dissipation: $\sigma_k \sim |k| \log |k|$
(b) Modified nonlinear interactions, $d(p,q)$, between resolved modes.
(c) Variance relaxation coupled to mean, χ_k.

Validation of reduced model against ensemble DNS



Squared norm of fit $\bar{\Delta}_m(\rho^{dns} \,|\, \tilde{\rho})$ versus initial disturbance $\Delta_m(\rho^0 \,|\, \rho_{eq})$.

Left: Positive temperature; Right: Negative temperature.

Conclusions

- Coarse-graining and closure may be achieved by dynamical optimization using information-theoretic (relative entropy) metrics of statistical model fit.

- The optimal closure has a nonequilibrium thermodynamic structure and related properties.

- Practical implementation relies on perturbation analysis of the Hamilton-Jacobi equation (near equilibrium), or optimal control algorithms (beyond equilibrium).