Introduction to asymptotic techniques for stochastic systems with multiple time-scales

Eric Vanden-Eijnden

Courant Institute

Motivating examples

Consider the ODE

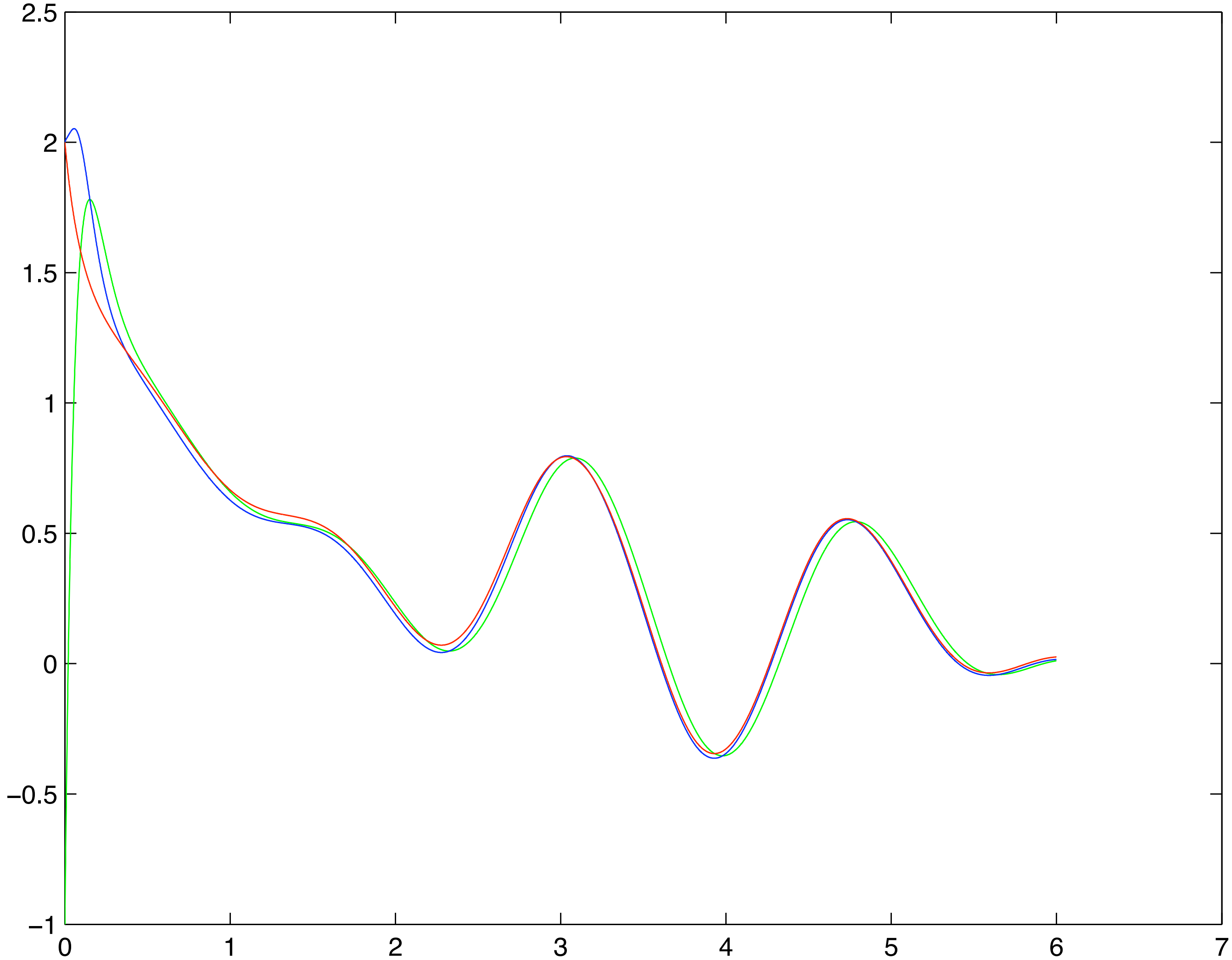
$$\begin{cases} \dot{X} = -Y^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t) & X(0) = x \\ \dot{Y} = -\varepsilon^{-1}(Y - X) & Y(0) = y. \end{cases} \tag{1}$$

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

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

Then the equation for X reduces to

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



The solution of (1) when $\varepsilon = 0.05$ and we took $X_{t=0} = 2$, $Y_{t=0} = -1$. X is shown in blue, and Y in green. Also shown in red is the solution of the limiting equation (2).

In contrast, consider the SDE

$$\begin{cases} \dot{X} = -Y^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), & X_{t=0} = x \\ \dot{Y} = -\varepsilon^{-1}(Y - X) + \varepsilon^{-1/2}\dot{W}, & Y_{t=0} = y. \end{cases} \tag{3}$$

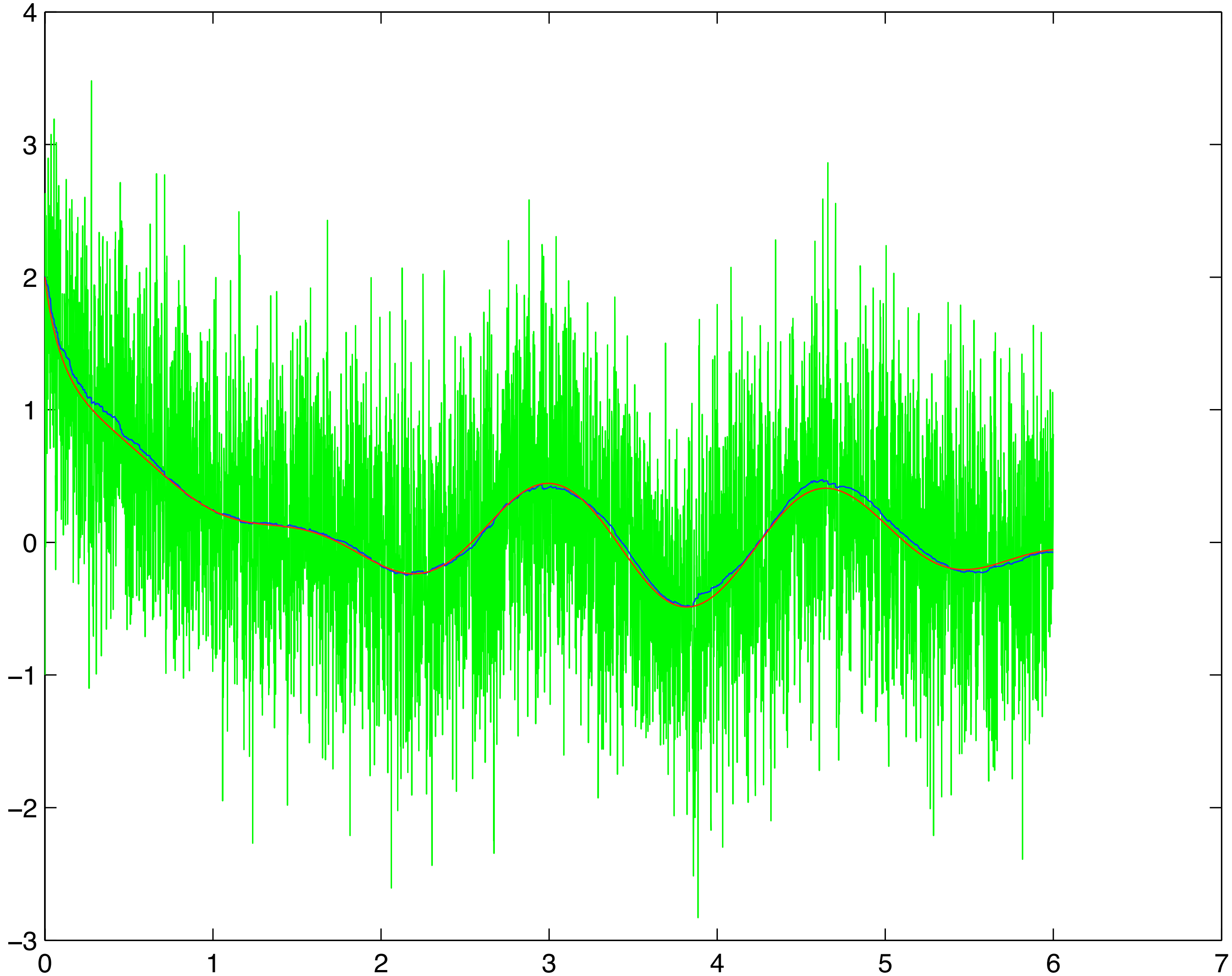
The equation for Y at fixed $X = x$ defines an Ornstein-Uhlenbeck process whose equilibrium probability density is

$$\rho(y|x) = \frac{e^{-(y-x)^2}}{\sqrt{\pi}}.$$

The limiting equation is obtained by averaging the right hand-side of the equation for X in (3) with respect to this density:

$$\dot{X} = -X^3 - \tfrac{3}{2}X + \sin(\pi t) + \cos(\sqrt{2}\pi t), \qquad X_0 = x. \tag{4}$$

Note the new term $-\frac{3}{2}\alpha^2 X$ due to the noise in (3).



The solution of (3) with $X_{t=0} = 2$, $Y_{t=0} = -1$ when $\varepsilon = 10^{-3}$. X is shown in blue, and Y in green. Also shown in red is the solution of the limiting equation (4). Notice how noisy Y is.

Singular perturbations techniques for Markov processes

Consider the system

$$\begin{cases} \dot{X} = f(X,Y), & X_{t=0} = x \in \mathbb{R}^n \\ \dot{Y} = \varepsilon^{-1} g(X,Y), & Y_{t=0} = y \in \mathbb{R}^m. \end{cases} \tag{5}$$

Assume that the equation for Y at fixed $X = x$ defines an Markov process which is ergodic for every x with respect to the probability distribution

$$d\mu_x(y)$$

If

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

then in the limit as $\varepsilon \to 0$ the evolution for X solution of (5) is governed by

$$\dot{X} = F(X) \qquad X(0) = x. \tag{6}$$

Derivation. Let L be the infinitesimal generator of the Markov process generated by (5)

$$\lim_{t\to 0^+} \frac{1}{t}\big(\mathbf{E}_{x,y}\phi(X_t, Y_t) - \phi(x, y)\big) = (L\phi)(x, y) \tag{7}$$

Then $L = L_0 + \varepsilon^{-1}L_1$, and $u(x, y, t) = \mathbf{E}_{x,y}\phi(Y, X)$ satisfies the backward Kolmogorov equation

$$\frac{\partial u}{\partial t} = L_0 u + \varepsilon^{-1}L_1 u, \qquad u|_{t=0} = \phi \tag{8}$$

Look for a solution in the form of

$$u = u_0 + \varepsilon u_1 + O(\varepsilon^2) \tag{9}$$

so that $\lim_{\varepsilon\to 0} u = u_0$.

Inserting (9) into (8) and equating equal powers in ε leads to the hierarchy of equations

$$\begin{cases} L_1 u_0 = 0, \\ L_1 u_1 = \dfrac{\partial u_0}{\partial t} - L_0 u_0, \\ L_1 u_2 = \cdots \end{cases} \tag{10}$$

The first equation tells that u_0 belong to the null-space of L_1. Assuming that for every x, L_1 is the generator of an ergodic Markov process with equilibrium distribution $\mu_x(y)$, this null-space is spanned by functions constant in y, i.e. $u_0 = u_0(x,t)$.

Since the null-space of L_1 is non-trivial, the next equations each requires a solvability condition, namely that their right hand-side belongs to the range of L_1.

To see what this solvability condition actually is, take the expectation of both sides of the second equation in (10) with respect to $d\mu_x(x)$. This gives

$$0 = \int_{\mathbb{R}^m} d\mu_x(y)\Big(\frac{\partial u_0}{\partial t} - L_0 u_0\Big) \tag{11}$$

Explicitly, (11) is

$$\frac{\partial u_0}{\partial t} = F(x) \cdot \nabla_x u_0 \tag{12}$$

where

$$F(x) = \int_{\mathbb{R}^m} f(x,y) d\mu_x(y) \tag{13}$$

(12) is the backward Kolmogorov equation of (6).

Remark: Computing the expectation with respect to $\mu_x(y)$ in practice. We have

$$(e^{L_1 t}\phi)(x,y) \to \int_{\mathbb{R}^m} \phi(x,y)d\mu_x(y) \quad \text{as } t \to \infty$$

In other words, if $u(x,y,t)$ satisfies

$$\frac{\partial u}{\partial t} = L_1 u, \qquad u|_{t=0} = \phi$$

so that formally $u(x,y,t) = (e^{L_1 t}\phi)(x,y)$, we have

$$\lim_{t\to\infty} u(x,y,t) = \int_{\mathbb{R}^m} \phi(x,y)d\mu_x(y)$$

But since $u(x,y,t) = \mathbf{E}_y \phi(Y^x)$ where

$$\dot{Y}^x = g(x, Y^x)$$

it follows that

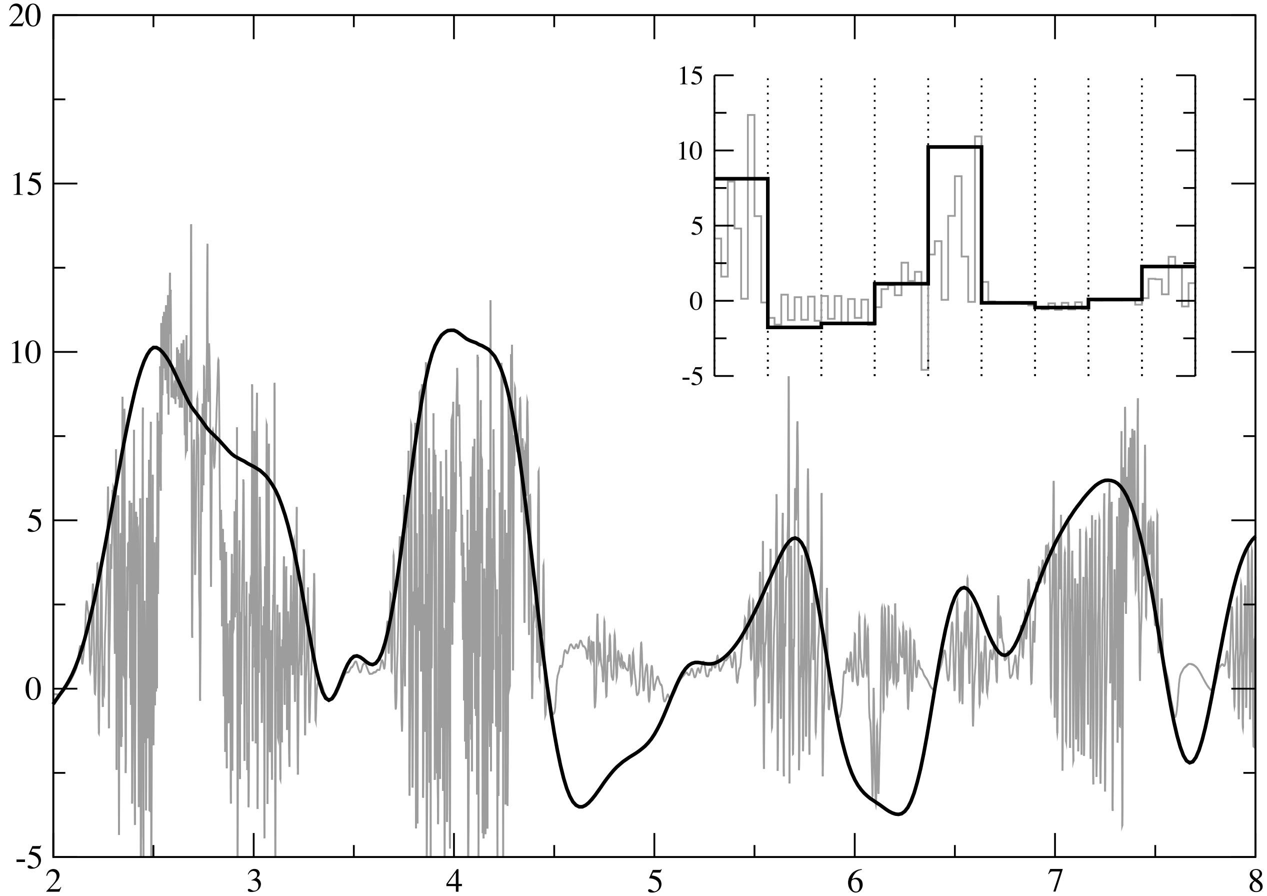
$$\begin{aligned} \int_{\mathbb{R}^m} \phi(x,y)d\mu_x(y) &= \lim_{t\to\infty} \mathbf{E}_y \phi(Y_t^x) \\ &= \lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(Y_t^x)dt \end{aligned} \tag{14}$$

Example: the Lorenz 96 (L96) model

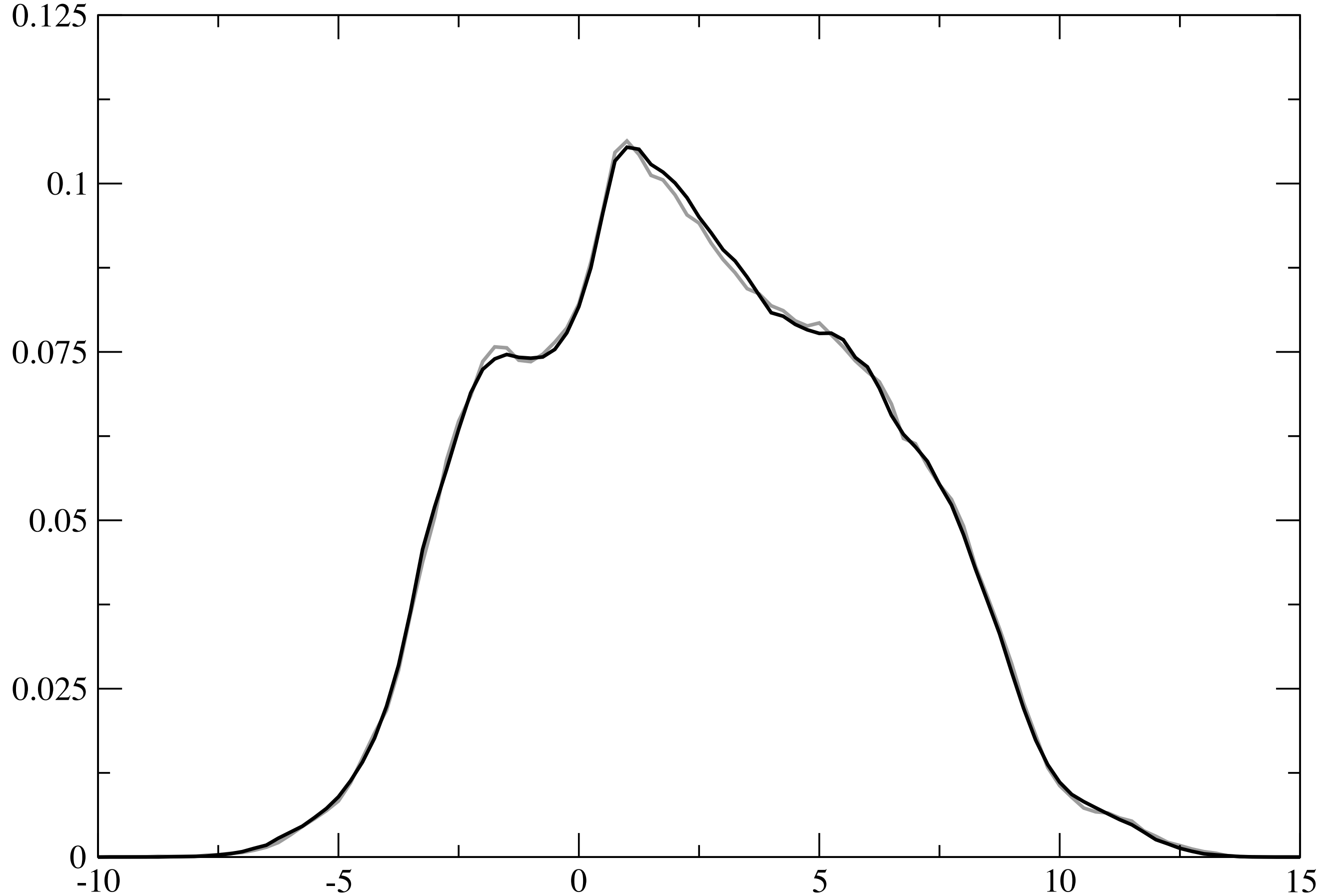
L96 consists of K slow variables X_k coupled to $J \times K$ fast variables $Y_{j,k}$ whose evolution is governed by

$$\begin{cases} \dot{X}_k = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F_x + \dfrac{h_x}{J} \displaystyle\sum_{j=1}^{J} Y_{j,k} \\ \dot{Y}_{j,k} = \dfrac{1}{\varepsilon} \left(-Y_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - Y_{j,k} + h_y X_k \right). \end{cases} \tag{15}$$

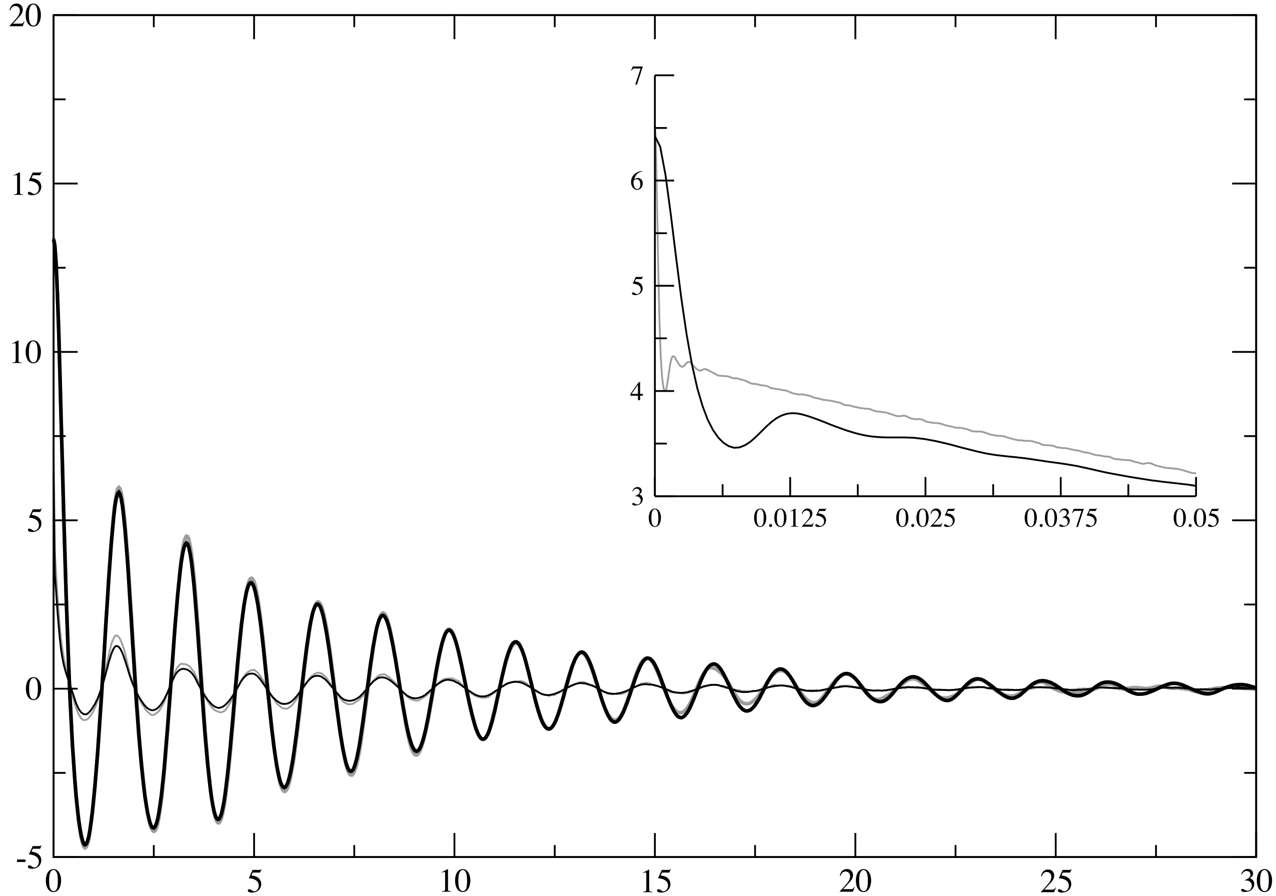
We will study (15) with $F_x = 10$, $h_x = -0.8$, $h_y = 1$, $K = 9$, $J = 8$, and two values of ε: $\varepsilon = 1/128$ and $\varepsilon = 1/1024$.



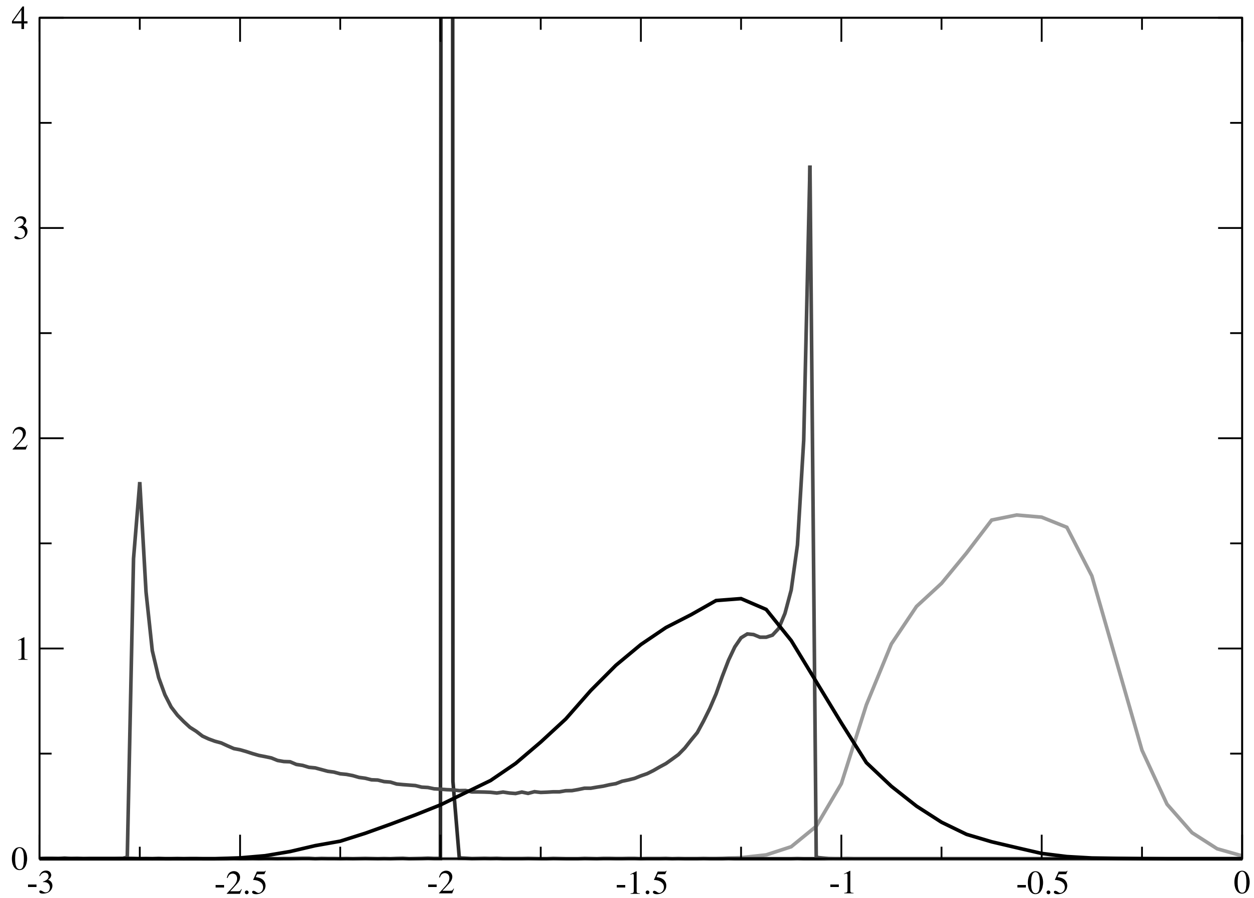
Typical time-series of the slow (black line) and fast (grey line) modes; $K = 9$, $J = 8$, $\varepsilon = 1/128$. The subplot displays a typical snapshot of the slow and fast modes at a given time.



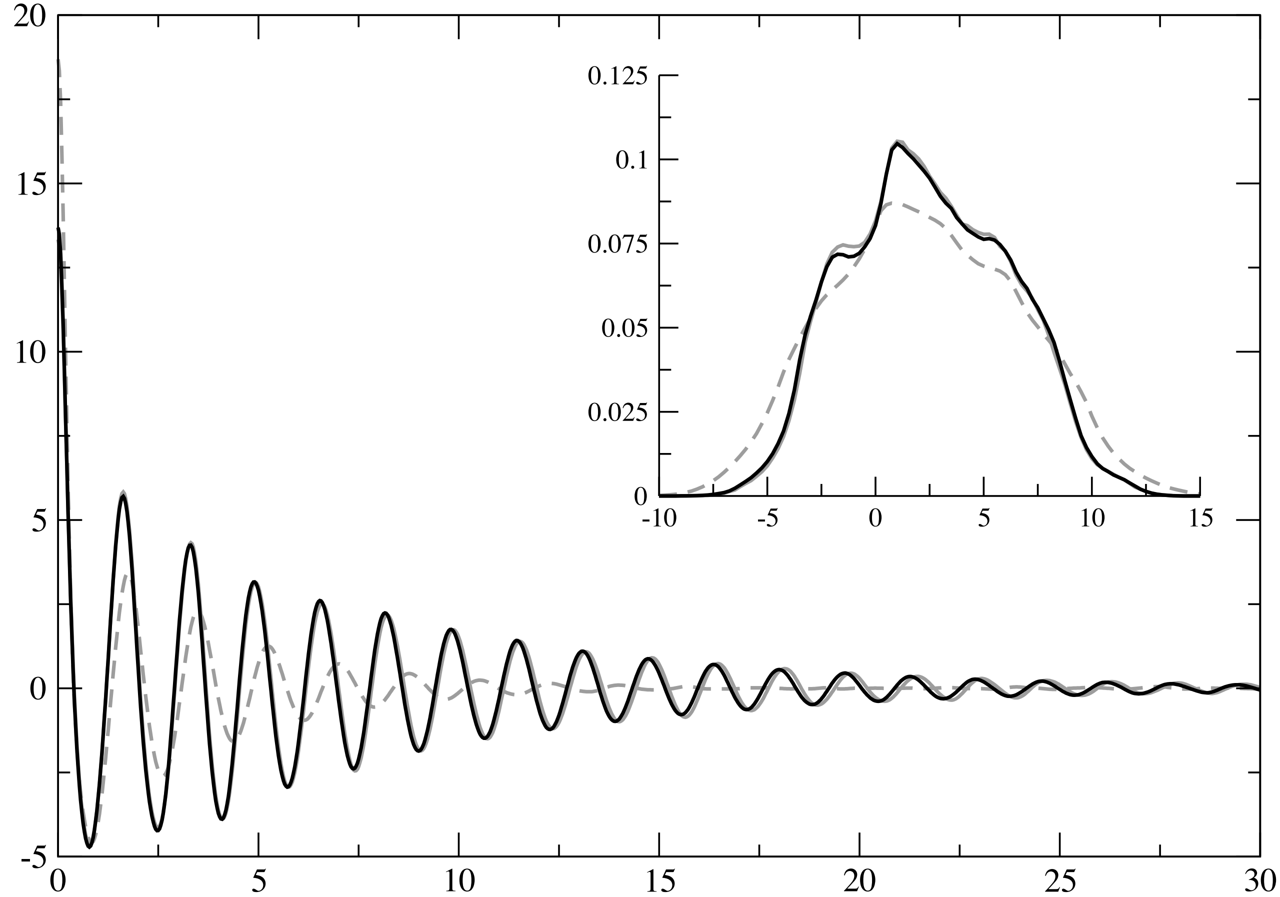
PDF of the slow variable; $K = 9$, $J = 8$, black line: $\varepsilon = 1/16$, grey line: $\varepsilon = 1/1024$. The insensitivity in ε of the PDFs indicates that the slow variables have already converged close to their limiting behavior when $\varepsilon = 1/16$.



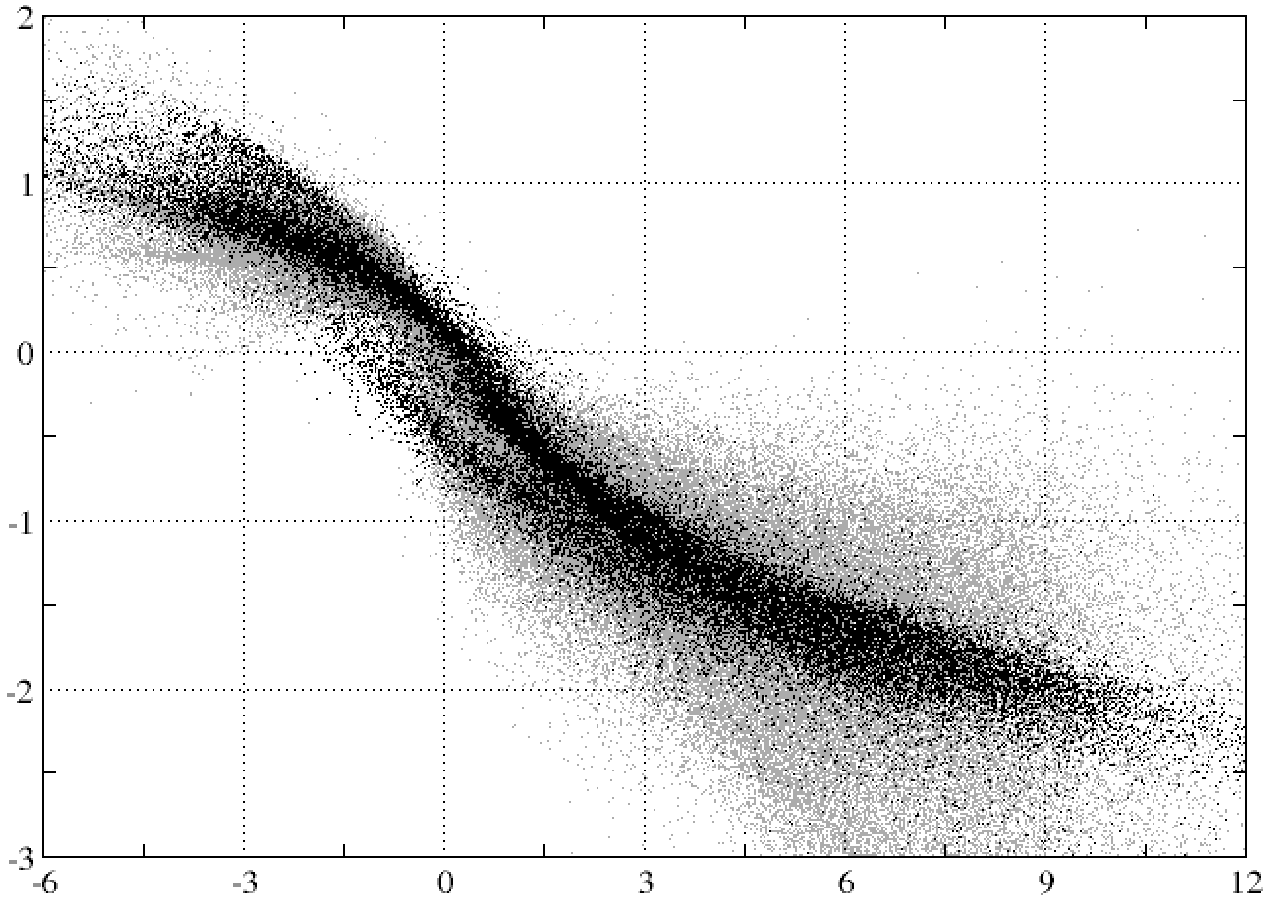
ACFs of the slow (thick line) and fast (thin line) variables; $\varepsilon = 1/128$, grey line: $\varepsilon = 1/1024$. The subplot is the zoom-in of the main graph which shows the transient decay of the ACFs of the fast modes becoming faster as ε is decreased: this is the only signature in the ACFs of the fact that the $Y_{j,k}$'s are faster.



Typical PDFs of the coupling term $(h_x/J)\sum_{j=1}^{J} Z_{j,k}(x)$ for various values of x. These PDFs are robust against variations in the initial conditions indicating that the dynamics of the fast modes conditional on the slow ones being fixed is ergodic.



Comparison between the ACFs and PDFs; black line: limiting dynamics; full grey line: $\varepsilon = 1/128$. Also shown in dashed grey are the corresponding ACF and PDF produced by the truncated dynamics where the coupling of the slow modes X_k with the fast ones, $Y_{j,k}$, is artificially switched off.



Black points: scatterplot of the forcing $F(x)$ in the limiting dynamics. Grey points: scatter plot of the bare coupling term $(h_x/J)\sum_{j=1}^{J} Y_{j,k}(x)$ when $\varepsilon = 1/128$.

Diffusive time-scales

Suppose that

$$F(x) = \int_{\mathbb{R}^m} f(x,y)d\mu_x(y) = 0. \tag{16}$$

Then the limiting equation on the $O(1)$ time-scale is trivial, $\dot{X} = 0$, and the interesting dynamics arises on the diffusive time-scale $O(\varepsilon^{-1})$.

Consider then

$$\begin{cases} \dot{X} = \varepsilon^{-1} f(X,Y), & X_{t=0} = x \in \mathbb{R}^n \\ \dot{Y} = \varepsilon^{-2} g(X,Y), & Y_{t=0} = y \in \mathbb{R}^m. \end{cases} \tag{17}$$

Proceeding as before we arrive at the following limiting equation when $\varepsilon \to 0$:

$$\dot{X} = b(X) + \sigma(X)\,\dot{W}_t, \tag{18}$$

where

$$\begin{aligned} (\bar{L}\phi)(x) &\equiv b(x)\cdot\nabla_x\phi(x) + \tfrac{1}{2}(\sigma\sigma^T)(x) : \nabla_x\nabla_x\phi(x) \\ &= \int_0^\infty dt \int_{\mathbb{R}^m} d\mu_x(y) f(x,y)\cdot\nabla_x\big(\mathbf{E}_y f(x,Y_t^x)\cdot\nabla_x\phi(x)\big) \end{aligned} \tag{19}$$

and

$$\dot{Y}^x = g(x, Y^x)$$

Derivation. The backward Kolmogorov equation for $u(x,y,t) = \mathbb{E}_{x,y} f(X)$ is now

$$\frac{\partial u}{\partial t} = \varepsilon^{-1} L_0 u + \varepsilon^{-2} L_1 u.$$

Inserting the expansion $u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + O(\varepsilon^2)$ (we will have to go one order in ε higher than before) in this equation now gives

$$\begin{cases} L_1 u_0 = 0, \\ L_1 u_1 = -L_0 u_0, \\ L_1 u_2 = \dfrac{\partial u_0}{\partial t} - L_0 u_1, \\ L_1 u_3 = \cdots \end{cases} \tag{20}$$

The first equation tells that $u_0(x,y,t) = u_0(x,t)$.

The solvability condition for the second equation is satisfied by assumption because of (16). Therefore this equation can be formally solved as

$$u_1 = -L_1^{-1} L_2 u_0.$$

Inserting this expression in the third equation in (20) and considering the solvability condition for this equation, we obtain the limiting equation for u_0:

$$\frac{\partial u_0}{\partial t} = \bar{L} u_0, \tag{21}$$

where

$$\bar{L} = \int_{\mathbb{R}^m} d\mu_x(y) L_0 L_1^{-1} L_0.$$

To see what this equation is explicitly, notice that $-L_1^{-1}g(y)$ is the steady state solution of

$$\frac{\partial v}{\partial t} = L_1 v + g(y).$$

The solution of this equation with the initial condition $v(y, 0) = 0$ can be represented by Feynman-Kac formula as

$$v(y,t) = \mathbb{E}_y \int_0^t g(Y_s^x) ds,$$

Therefore

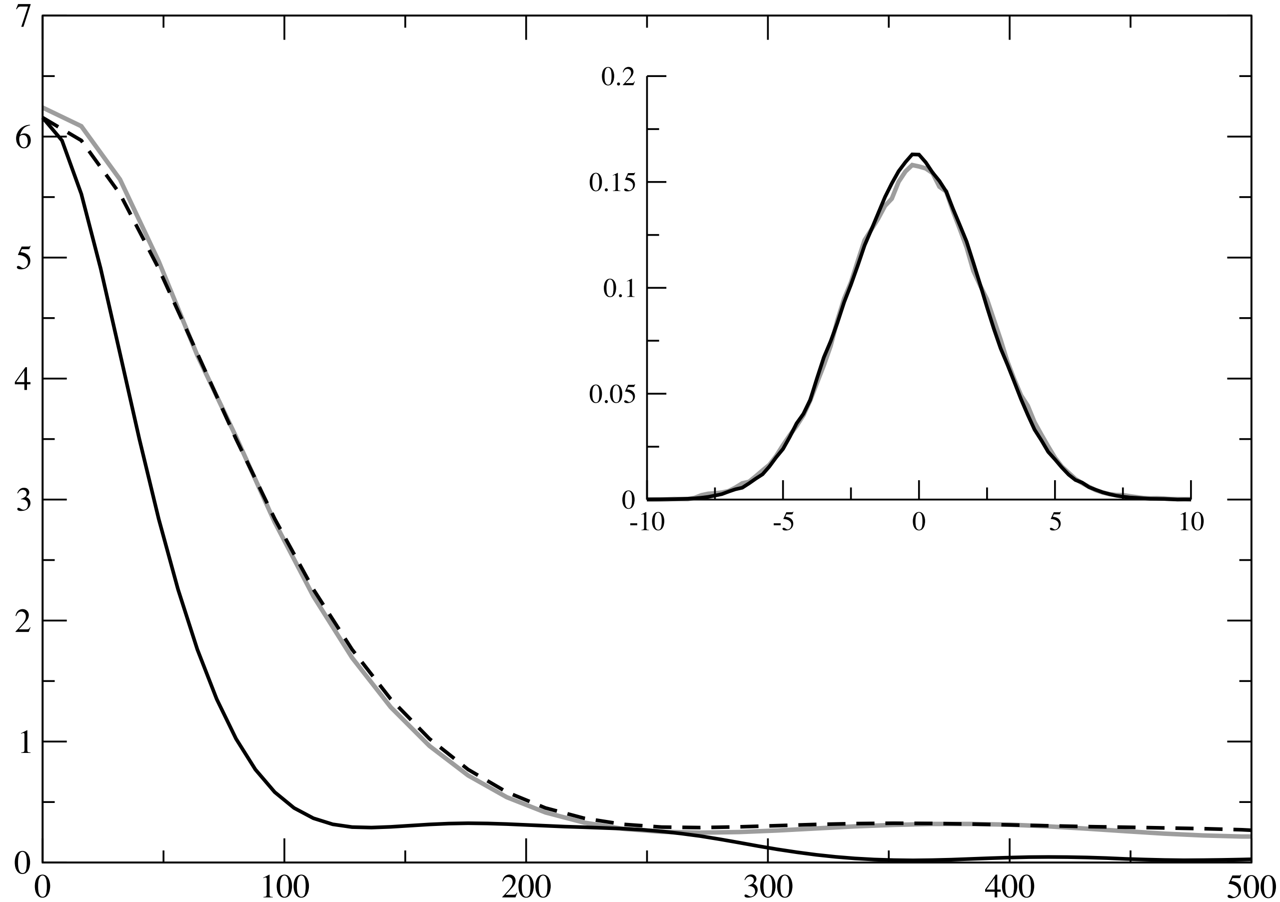
$$-L_1^{-1}g(y) = \mathbb{E}_y \int_0^\infty g(Y_t^x) dt,$$

and the operator $\bar{L}$ in the limiting backward Kolmogorov equation (21) is (19).

Example: the Lorenz 96 (L96) model

Consider

$$\begin{cases} \dot{X}_k = -\varepsilon \left(X_{k-1}(X_{k-2} - X_{k+1}) + X_k\right) + \dfrac{h_x}{J} \displaystyle\sum_{j=1}^{J} \left(Y_{j,k+1} - Y_{j,k-1}\right) \\ \dot{Y}_{j,k} = \dfrac{1}{\varepsilon} \left(-Y_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - Y_{j,k} + F_y\right) + h_y X_k, \end{cases} \tag{22}$$



ACFs and PDFs (subplot) of the slow variable X_k evolving under (22). Grey line: $\varepsilon = 1/256$; full black line: $\varepsilon = 1/128$; dashed black line: $\varepsilon = 1/128$ with time rescaled as $t \to 2t$. The near perfect match confirms that evolution of X_k converges to some limiting dynamics on the $O(1/\varepsilon)$ time-scale.

Coupled truncated Burgers-Hopf (TBH). I. Stable periodic orbits

$$\begin{cases} \dot{X}_1 = \dfrac{1}{\varepsilon} b_1 X_2 Y_1 + a Y_1 (R^2 - (X_1^2 + X_2^2)) - b X_2 (\alpha + (X_1^2 + X_2^2)), \\ \\ \dot{X}_2 = \dfrac{1}{\varepsilon} b_2 X_1 Y_1 + a x_2 (R^2 - (X_1^2 + X_2^2)) + b X_1 (\alpha + (X_1^2 + X_2^2)), \\ \\ \dot{Y}_k = -\mathrm{Re}\ \dfrac{ik}{2} \displaystyle\sum_{p+q+k=0} U_p^* U_q^* + \dfrac{1}{\varepsilon} b_3 \delta_{1,k} X_1 X_2, \\ \\ \dot{Z}_k = -\mathrm{Im}\ \dfrac{ik}{2} \displaystyle\sum_{p+q+k=0} U_p^* U_q^*, \end{cases}$$

where $U_k = Y_k + i Z_k$.

Truncated system: stable periodic orbit (limit cycle)

$(X_1(t), X_2(t)) = R\,(\cos \omega t, \sin \omega t)$ with frequency $\omega = b(\alpha + R^2)$.

NB: Truncated Burgers (Majda & Timofeyev, 2000)

Fourier-Galerkin truncation of the inviscid Burgers-Hopf equation, $u_t + \frac{1}{2}(u^2)_x = 0$:

$$\dot{U}_k = -\frac{ik}{2} \sum_{\substack{k+p+q=0 \\ |p|,|q| \le \Lambda}} U_p^* U_q^*, \qquad |k| < \Lambda$$

Features common with many complex systems. In particular:

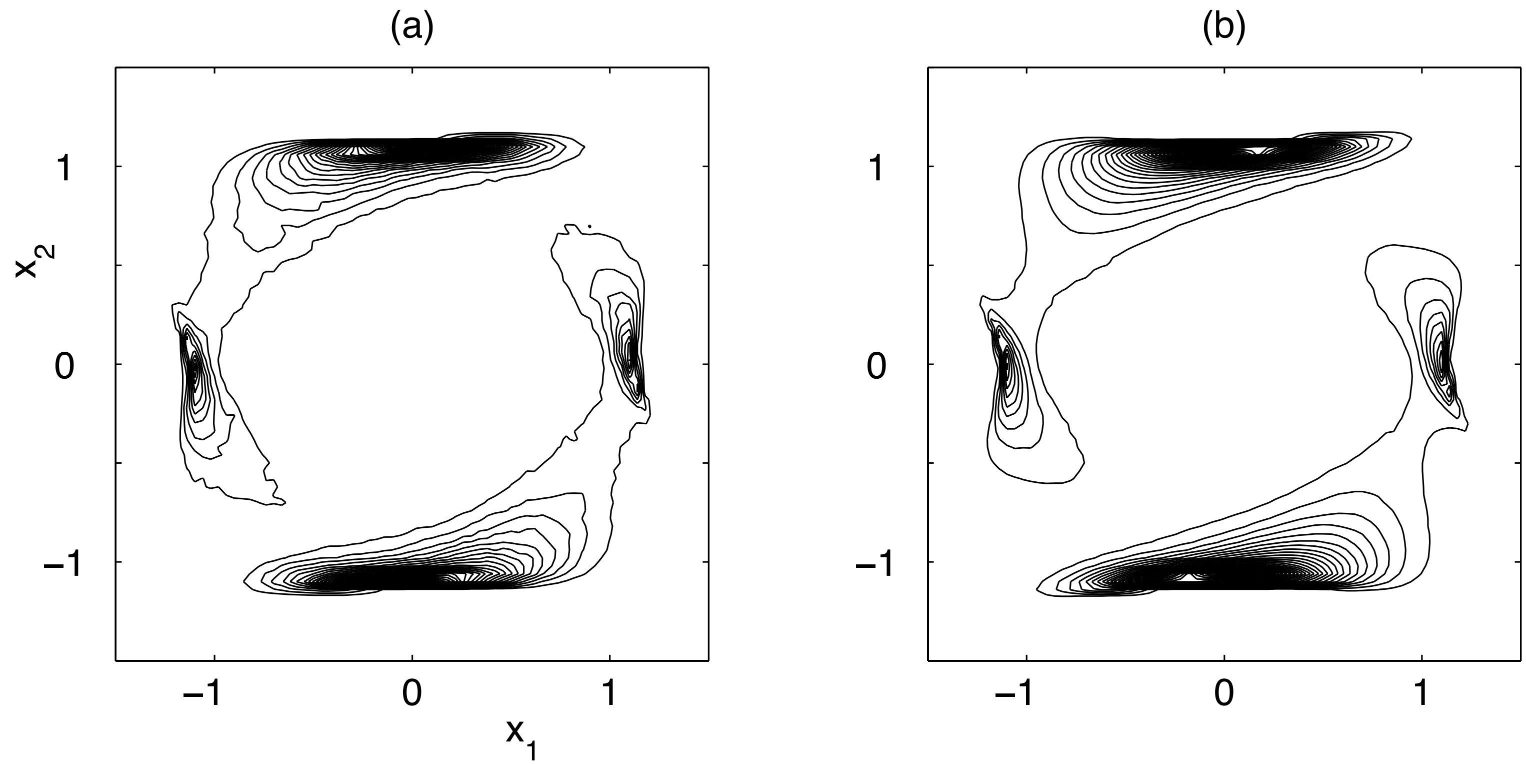
- Display deterministic chaos.
- Ergodic on $E = \sum_k |U_k|^2$.
- Scaling law for the correlation functions with $t_k \approx O(k^{-1})$.

Here: *Used as a model for unresolved modes.*
Couple truncated Burgers with two resolved variables.

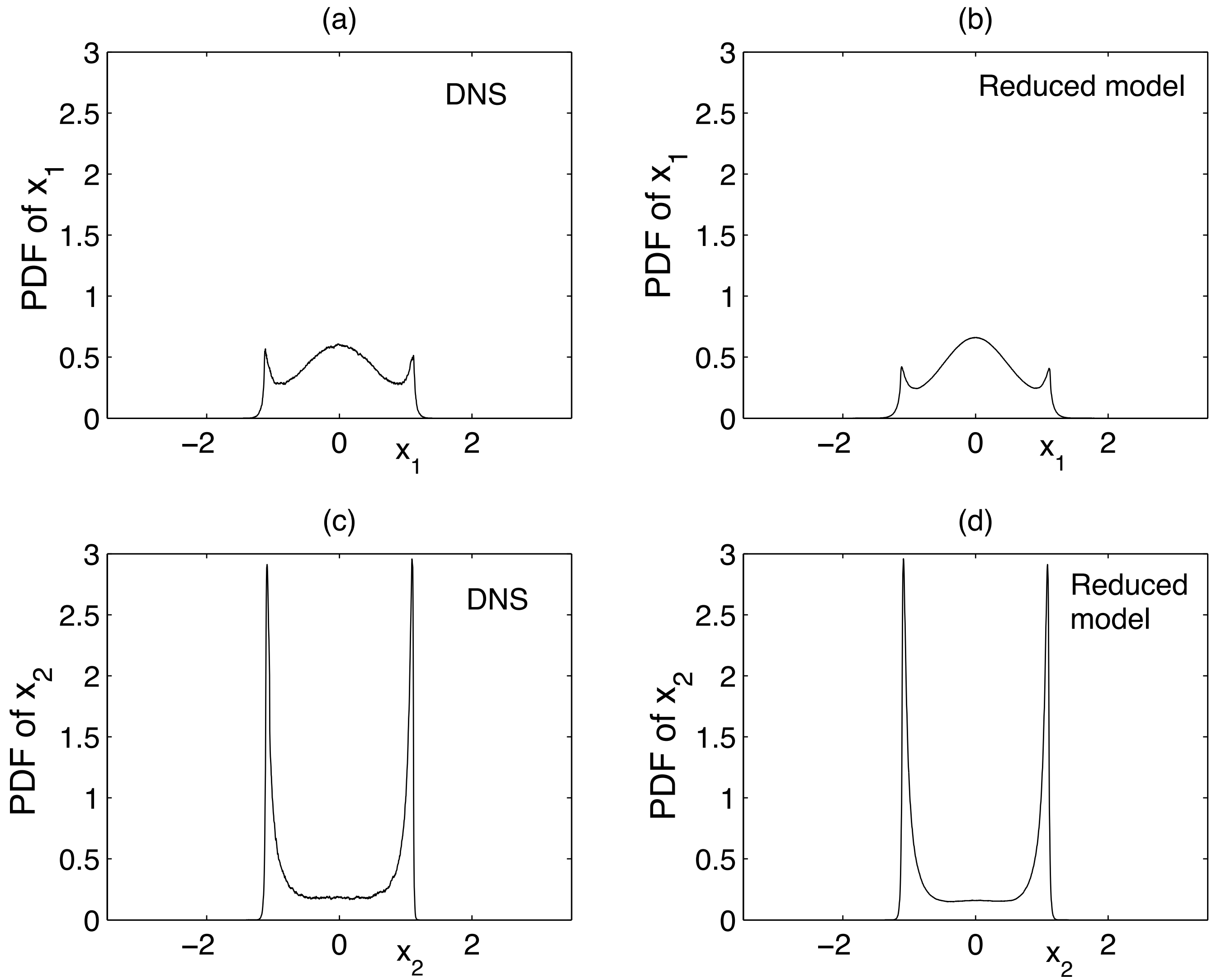
$$
\begin{cases}
\dot{X}_1 = \dfrac{1}{\varepsilon} b_1 X_2 Y_1 + a Y_1 (R^2 - (X_1^2 + X_2^2)) - b X_2 (\alpha + (X_1^2 + X_2^2)), \\[2ex]
\dot{X}_2 = \dfrac{1}{\varepsilon} b_2 X_1 Y_1 + a x_2 (R^2 - (X_1^2 + X_2^2)) + b X_1 (\alpha + (X_1^2 + X_2^2)), \\[2ex]
\dot{Y}_k = -\mathrm{Re}\, \dfrac{ik}{2} \displaystyle\sum_{p+q+k=0} U_p^* U_q^* + \dfrac{1}{\varepsilon} b_3 \delta_{1,k} X_1 X_2, \\[2ex]
\dot{Z}_k = -\mathrm{Im}\, \dfrac{ik}{2} \displaystyle\sum_{p+q+k=0} U_p^* U_q^*,
\end{cases}
$$

Limiting SDEs:

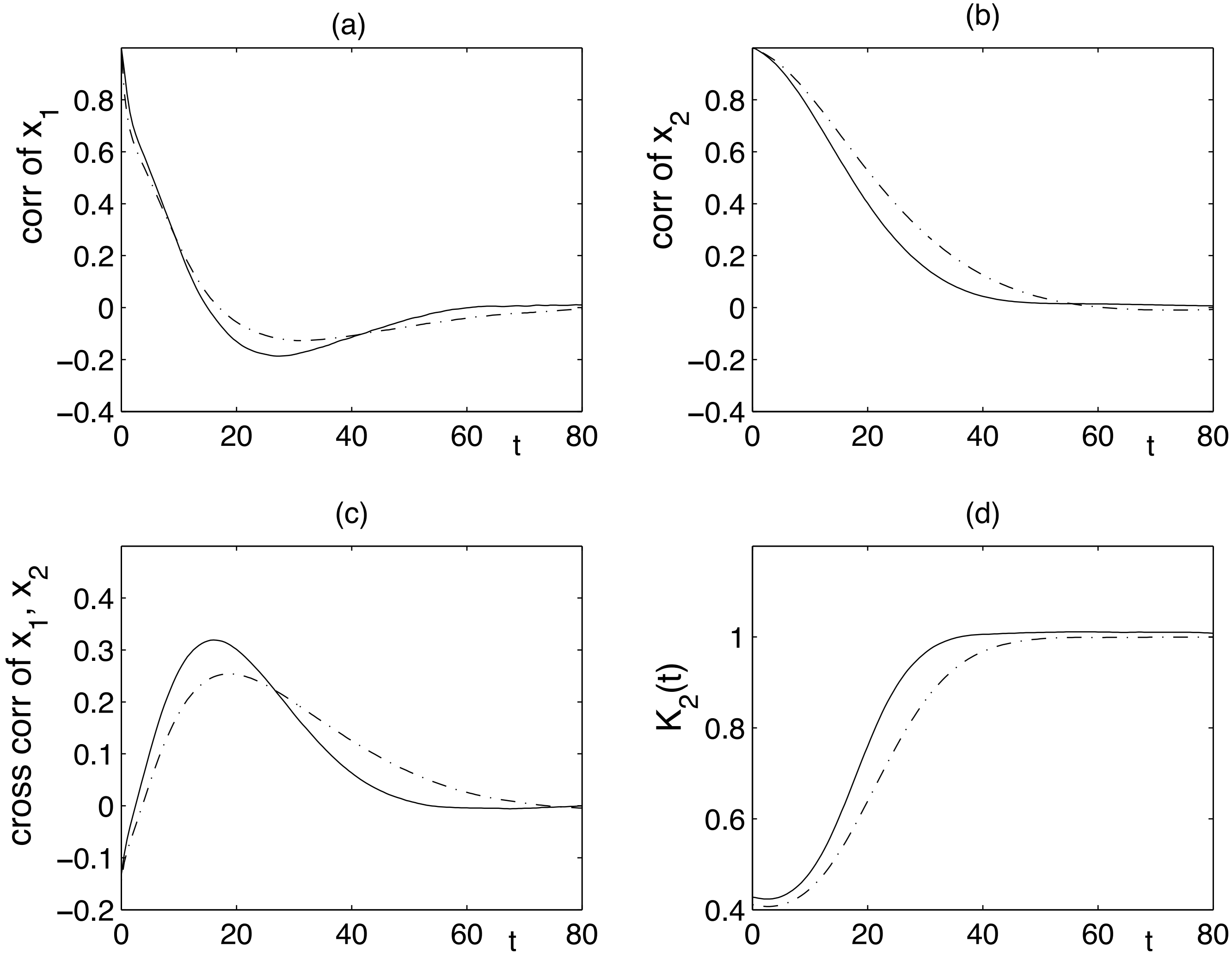
$$
\begin{cases}
\dot{X}_1 = b_1 b_2 X_1 + N_1 X_1 X_2^2 + \sigma_1 X_2 \dot{W}(t) \\
\qquad\quad + a X_1 (R^2 - (X_1^2 + X_2^2)) - b X_2 (\alpha + (X_1^2 + X_2^2)), \\[2ex]
\dot{X}_2 = b_1 b_2 X_2 + N_2 X_2 X_1^2 + \sigma_2 X_1 \dot{W}(t) \\
\qquad\quad + a X_2 (R^2 - (X_1^2 + X_2^2)) + b X_1 (\alpha + X_1^2 + X_2^2)),
\end{cases}
$$



Contour plots of the joint probability density for the climate variables X_1 and X_2; (a) deterministic system with 102 variables; (b) limit SDE. There only remains a ghost of the limit cycle.



Marginal PDFs of X_1 and X_2 for the simulations of the full deterministic system with 102 variables (DNS) and the limit SDE.



Two-point statistics for X_1 and X_2; solid lines - deterministic system with 102 degrees of freedom; dashed lines - limit SDE; (a), (b) correlation functions of X_1 and X_2, respectively; (c) cross-correlation function of X_1 and X_2; (d) normalized correlation of energy,

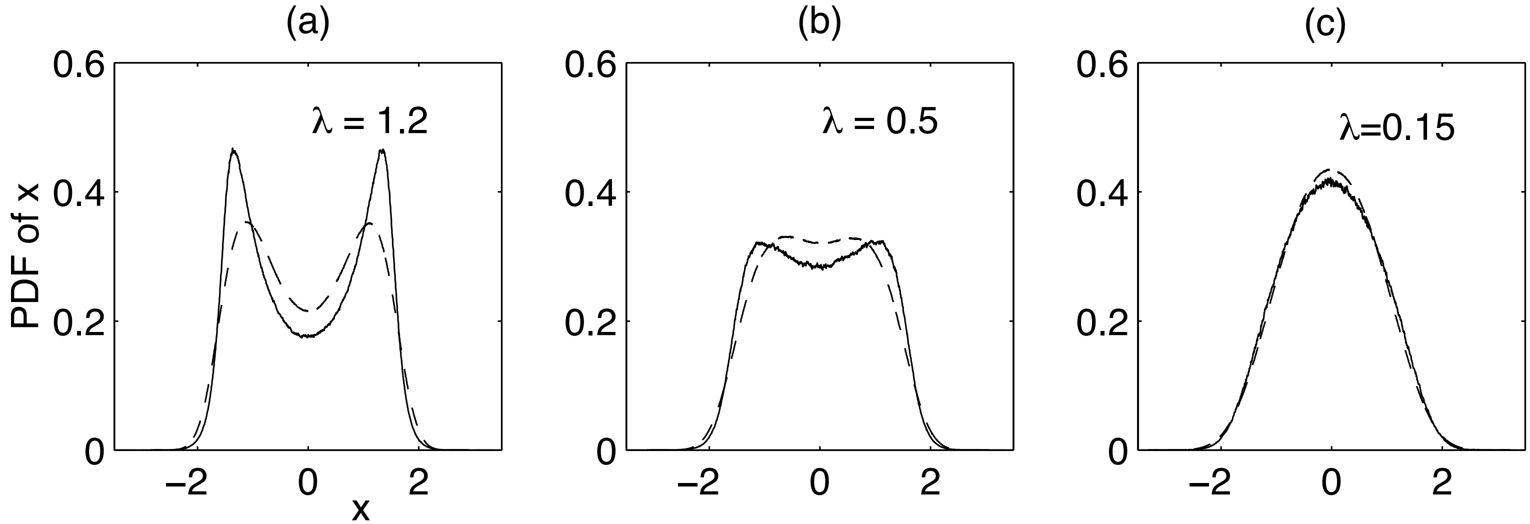
$$K_2(t) = \frac{\langle X_2^2(t) X_2^2(0)\rangle}{\langle X_2^2\rangle^2 + 2\langle X_2(t) X_2(0)\rangle^2}$$

Coupled truncated Burgers-Hopf (TBH). II. Multiple equilibria

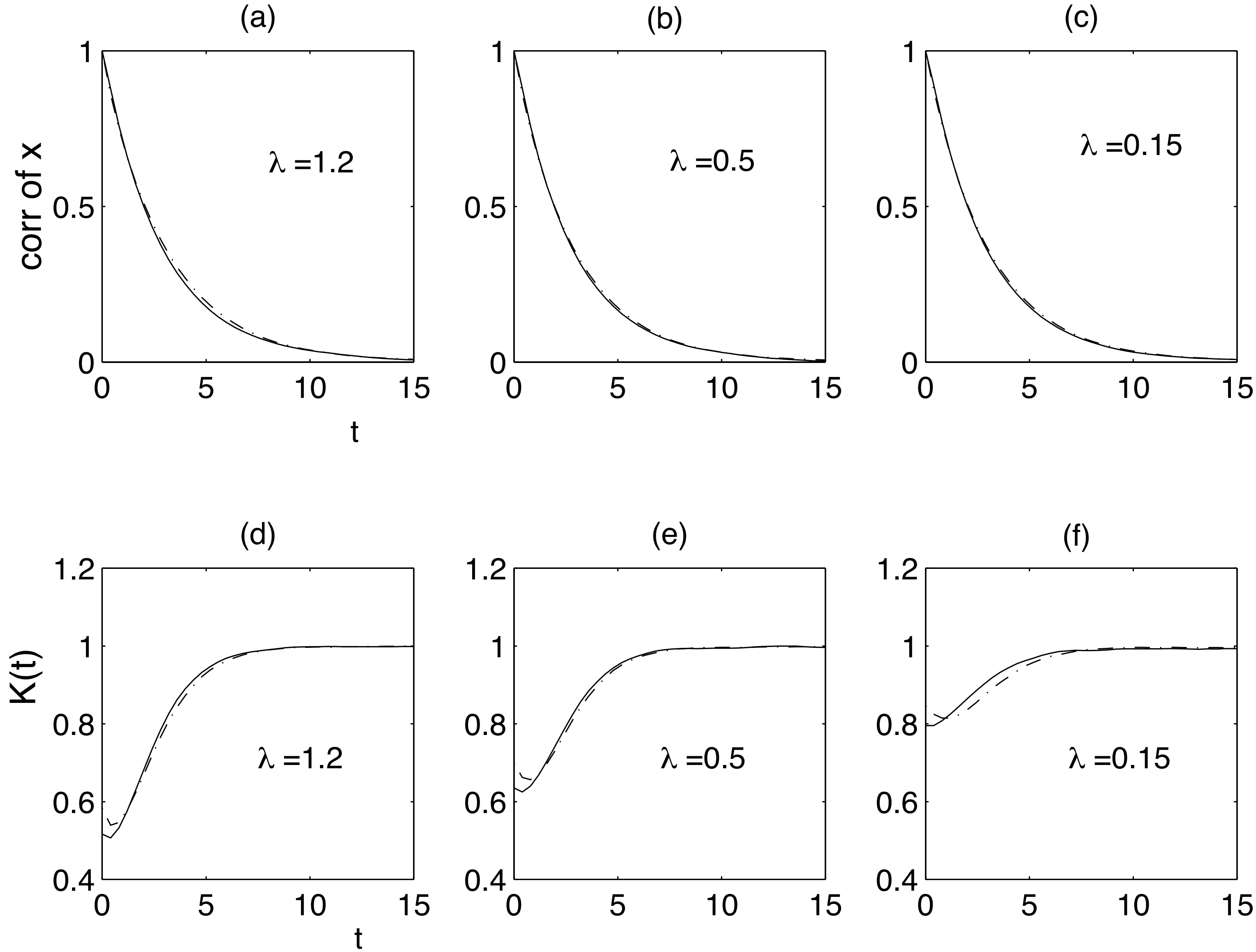
$$\begin{cases} \dot{X} = \frac{1}{\varepsilon} b_1 Y_1 Z_1 + \lambda(1 - \alpha x^2) x, \\ \dot{Y}_k = -\mathrm{Re}\, \frac{ik}{2} \sum_{p+q+k=0} U_p^* U_q^* + \frac{1}{\varepsilon} b_2 \delta_{1,k} X Z_k, \\ \dot{Z}_k = -\mathrm{Im}\, \frac{ik}{2} \sum_{p+q+k=0} u_p^* u_q^* + \frac{1}{\varepsilon} b_3 \delta_{1,k} X Y_k, \end{cases}$$

Limiting SDE:

$$\dot{X} = -\gamma X + \sigma \dot{W}(t) + \lambda(1 - \alpha X^2) X$$



PDF of X for the simulations of the deterministic model (solid lines) and the limit SDE (dashed lines) in three regimes, $\lambda = 1.2$, 0.5, 0.15.



Two-point statistics of X in three regimes, $\lambda = 1.2$, 0.5, 0.15 for the deterministic equations (solid lines) and limit SDE (dashed lines) (a), (b), (c) correlation function of x; (d), (e), (f) correlation of energy, $K(t)$, in x.

Generalizations

Let $Z_t \in S$ be the sample path of a continuous-time Markov process with generator

$$L = L_0 + \varepsilon^{-1} L_1$$

and assume that L_1 has several ergodic components indexed by $x \in S'$ with equilibrium distribution

$$\mu_x(z)$$

Then, as $\varepsilon \to 0$ there exists a limiting on S' with generator

$$\bar{L} = \mathbb{E}_{\mu_x} L_0$$

Similar results available on diffusive time-scales.

Can be applied to SDEs, Markov chains (i.e. discrete state-space), deterministic systems (periodic or chaotic), etc.

Other situations with limiting dynamics?

Consider

$$\begin{cases} \dot{X}_t = f(X_t, Y_t), \\ \dot{Y}_t = g(X_t, Y_t), \end{cases} \tag{23}$$

and denote by $\varphi(X_{[0,t]})$ the solution of the equation for Y at time t assuming that X_t is known on $[0, t]$

Observe that $\varphi(X_{[0,t]})$ is a functional of $\{X_s, s \in [0, t]\}$.

Then X_t satisfies the closed equation

$$\dot{X}_t = f(X_t, \varphi(X_{[0,t]})) \tag{24}$$

In general, X_t is not Markov!

It becomes Markov when Y_t is faster, or ...?

Other possibility: Weak coupling

The system

$$\begin{cases} \dot{X}_t = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} f(X_t, Y_t^n), \\ \dot{Y}_t^n = g(X_t, Y_t^n), \qquad \text{(all the } Y_t^n \text{ coupled only via } X_t) \end{cases}$$

may have a limit behavior as $N \to \infty$ which is the same as the limit behavior as

$$\begin{cases} \dot{X}_t = f(X_t, Y_t), \\ \dot{Y}_t = \dfrac{1}{\varepsilon} g(X_t, Y_t), \end{cases}$$

when $\varepsilon \to 0$.

Asymptotic techniques for singularly perturbed Markov processes:

R. Z. Khasminsky, On Stochastic Processes Defined by Differential Equations with a Small Parameter, *Theory Prob. Applications*, **11**:211–228, 1966.

R. Z. Khasminsky, A Limit Theorem for the Solutions of Differential Equations with Random Right-Hand Sides, *Theory Prob. Applications*, **11**:390–406, 1966.

T. G. Kurtz, Semigroups of conditioned shifts and approximations of Markov processes, *Annals of Probability*, **3**:618–642, 1975.

G. Papanicolaou, Some probabilistic problems and methods in singular perturbations, *Rocky Mountain J. Math*, **6**:653–673, 1976.

M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, 2nd edition, Springer-Verlag, 1998.

Effective dynamics in L96:

E. N. Lorenz, Predictability – A problem partly solved, pp. 1–18 in: ECMWF Seminar Proceedings on Predictability, Reading, United Kingdom, ECMWF, 1995.

C. Rödenbeck, C. Beck, H. Kantz, Dynamical systems with time scale-separation: averaging, stochastic modeling, and central limit theorems, pp. 187–209 in: Stochastic Climate Models, P. Imkeller, J.-S. von Storch eds., Progress in Probability **49**, Birkhäuser Verlag, Basel, 2001.

I. Fatkullin and E. Vanden-Eijnden, A computational strategy for multiscale systems with applications to Lorenz 96 model, *J. Comp. Phys.* **200**:605–638, 2004.

(Coupled) TBH:

A. J. Majda and I. Timofeyev, Remarkable statistical behavior for truncated Burgers-Hopf dynamics, *Proc. Nat. Acad. Sci. USA*, **97**:12413–12417, 2000.

A. J. Majda and I. Timofeyev, Statistical mechanics for truncations of the Burgers-Hopf equation: a model for intrinsic stochastic behavior with scaling, *Milan Journal of Mathematics*, **70**(1):39–96, 2002.

A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden, Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.*, **60**(14):1705–1722, 2003.

A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden, Stochastic models for selected slow variables in large deterministic systems, *Nonlinearity*, **19**:769–794, 2006.