Multiscale Dynamics and Information: Dimensional Reduction and Data Assimilation

N. Sri Namachchivaya

Joint work with:

Peter Imkeller (HU-B), Nishanth Lingala, Nicolas Perkowski (HU-B) and Hoong Chieh Yeong

Departments of Aerospace Engineering & Applied Mathematics
Information Trust Institute
University of Illinois at Urbana-Champaign
IL, USA

January 8-11, 2014

Workshop on Nonlinear filtering and data assimilation

TIFR Centre for Applicable Mathematics, Bangalore, India.

Research Supported by: NSF EFRI -1024772, CMMI -1030144, and AFOSR FA9550-12-1-0390

Data Assimilation in High-Dimensional Chaotic Systems

Objectives are to develop efficient methods and algorithms for

- the assimilation of data in high-dimensional multi-scale systems, and
- steering the measurement process towards "information rich" data locations.

↓↓↓: efficient methods for assimilation : ↓↓↓

Dimensional reduction and Nonlinear filtering

- Dimensional reduction in nonlinear filtering: – issues in high-dimensions
- Construct and validate homogenized particle filters with Optimal nudging

↓↓↓: information rich data locations :↓↓↓

Sensor Control

- Dynamically steer the measurement processes (use Finite-Time LE)
- Control that minimize the loss attributed to the estimation procedure

Data Assimilation in High-Dimensional Chaotic Systems

Objectives are to develop efficient methods and algorithms for

- the assimilation of data in high-dimensional multi-scale systems, and
- steering the measurement process towards "information rich" data locations.

↓↓↓: efficient methods for assimilation : ↓↓↓

Dimensional reduction and Nonlinear filtering

- Dimensional reduction in nonlinear filtering: – issues in high-dimensions
- Construct and validate homogenized particle filters with Optimal nudging

↓↓↓: information rich data locations :↓↓↓

Sensor Control

- Dynamically steer the measurement processes (use Finite-Time LE)
- Control that minimize the loss attributed to the estimation procedure

Outline: Dimensional Reduction and Data Assimilation

1. MOTIVATION: DECADAL CLIMATE PREDICTION – *NCAR Community Climate System Model*

2. DIMENSIONAL REDUCTION IN A MULTI-SCALE ENVIRONMENT – *N. Sri Namachchivaya and R. Sowers, Rigorous Stochastic Averaging at a Center With Additive Noise, MECCANICA, Vol. 37(2), 2002*, *Kristjan Onu and N. Sri Namachchivaya, 'Stochastic Averaging of Surface Waves.* Proceedings of the Royal Society, A, *Vol. 466, 2010*.

Estimation and prediction in Earth (climate) system models

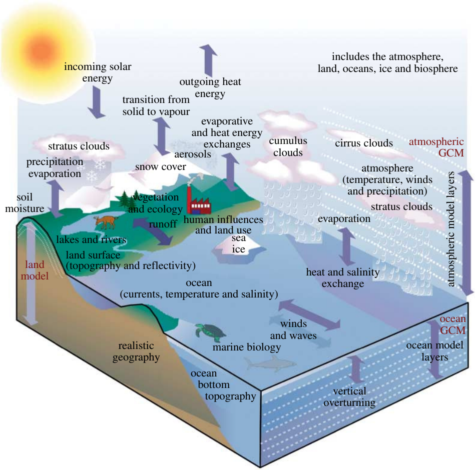


Figure: NCAR Community Climate System Model

W. M. Washington, L. Buja, A. Craig, The computational future for climate and Earth system models: on the path to petaflop and beyond, *Phil. Trans. R. Soc. A*, 367, 2009

Earth system/Global Circulation Models (GCMs)

- multiple timescales (≥ 2)
- high dimensions, $\geq \mathcal{O}(10^4)$ grid points
- uncertain initial conditions, inaccurate/changing model parameters

Estimation and prediction in Earth (climate) system models

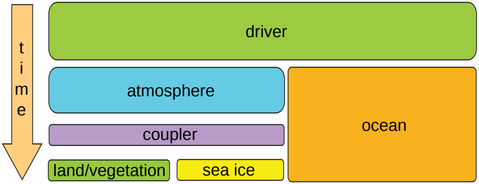


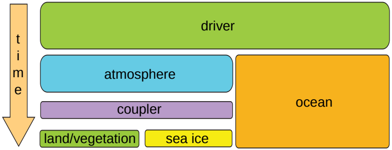
Figure: Coupling components in climate model [NCAR]

Earth system/Global Circulation Models (GCMs)

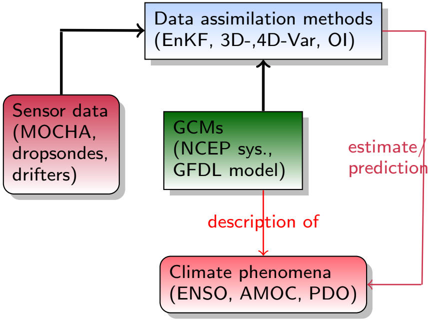
- multiple timescales (≥ 2)
- high dimensions, $\geq \mathcal{O}(10^4)$ grid points
- uncertain initial conditions, inaccurate/changing model parameters

W. M. Washington, L. Buja, A. Craig, The computational future for climate and Earth system models: on the path to petaflop and beyond, *Phil. Trans. R. Soc. A*, 367, 2009

Estimation in Global Circulation Models



Coupling components in climate models.
W. M. Washington, L. Buja, A. Craig, The computational future for climate and Earth system models: on the path to petaflop and beyond, Phil. Trans. R. Soc. A, 367, 2009



Global Circulation Models (GCMs)
– multiple timescales (≥ 2)
– high dimensions, on $\geq \mathcal{O}(10^4)$ grid points
– uncertain initial conditions, inaccurate model parameters

Climate phenomena (e.g. El Niño Southern Oscillation, Atlantic Meridional Overturning Circulation, Pacific Decadal Oscillation) can be studied through estimation/prediction using GCMs

Problem: Decadal Climate Prediction

The "new field" of decadal prediction focuses on 10 to 30 years in the future that is of interest to policymakers.

North Atlantic climate exhibits fluctuations on multidecadal timescales:

- large influence on sea surface-temperature (SST);
- large influence on hurricane activity in the Atlantic;
- large influence on surface-temperature and rainfall variations over North America, Europe and northern Africa.

The Atlantic Meridional Overturning Circulation (AMOC) plays an important role in driving the multidecadal SST variations.

Understanding Climate Variability Using a Hierarchy of Models and Observations

Bottom-up: From conceptual to realistic climate models

$\downarrow\downarrow\downarrow$

Climate models are valuable tools – however, current climate model projections are still considerably uncertain.

This uncertainty can be traced back to:

- the inability of models to capture important physical processes – e.g., from grid size of 1 deg $\sim$ 150km to 1m resolution $\Longrightarrow$ stochastic parameterization;
- inadequate observations used to constrain and initialize the models – e.g., $p_0(x) = \nu(x) \Longrightarrow$ how to use the measurements to reduce the uncertainty and obtain the best possible state estimates;
- uncertainties in future anthropogenic forcings – e.g., IPCC report $\Longrightarrow$ anthropogenic aerosols that include black carbon, organic carbon and sulfate aerosols.

Understanding and quantifying the effects, interactions, and magnitudes of these uncertainties is crucial:

- to improving climate predictions;
- for analyzing risk; and
- for instituting sound decision making strategies under uncertainty.

TOP-DOWN: ANALYSIS AND ASSIMILATION OF OBSERVATIONS

$\downarrow\downarrow\downarrow$

This requires information rich observations

observations that are aggregated into globally gridded data products using well-validated interpolation techniques

.

Multiscale Information:

Atmospheric data from weather stations, aircraft observations and satellites, that has been aggregated into globally gridded data products using well-validated interpolation techniques.

Currently ∼3000 Argo floats covering the majority of the global ocean, providing information about temperature, salinity, and currents within the upper 2 km of the ocean.

TASK

Improving decadal-scale projections depends critically on assimilation of observations from the atmosphere, land and ocean, to reduce initial conditions uncertainty and evaluate model performance.

In complex systems, non-linearities of the governing physical processes allow energy transfer between different scales. Many aspects of this complex behavior can be represented by stochastic models.

For example, Large Eddy Simulation (LES) seeks to directly solve large spatial scales, while modeling the smaller scales (smaller scales have been found to be more universal, and hence are more easily modeled).

The main challenge in multiscale modeling is to recognize their simplicity. Take advantage of the multi-scales in the problem to see how information interacts with these complex structures and scales.

Simple or "conceptual" coupled atmosphere - ocean model

Signal Process:

$$\varepsilon \dot{Z}_t^\varepsilon = g(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon),\ Z_0^\varepsilon = z \in \mathbb{R}^m,\ \text{core atmosphere model}$$
$$\dot{X}_t^\varepsilon = f(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon),\ X_0^\varepsilon = x \in \mathbb{R}^n,\ \text{core ocean model}$$

where, z represent the general circulation of the atmosphere; x are the ocean components such as the density gradients, angular momentum, etc.. ξ^ε represents the unmodeled dynamics of the system or an additive noise.

Observation process:

For example, Meridional Overturning Circulation and Heatflux Array (MOCHA): Current Meters, CTDs (salinity and temperature every hour or so for a period of up to two years), buoys, acoustic releases. Deployed along 26.5° N in the Atlantic.

The observation process is a function of the signal process corrupted by noise

$$Y_t^\varepsilon = \int_0^t h^\varepsilon(X_s^\varepsilon, Z_s^\varepsilon)ds + V_t,$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is called the sensor function.

Signal Process:

$$\varepsilon \dot{Z}_t^\varepsilon = g(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon), \; Z_0^\varepsilon = z \in \mathbb{R}^m, \text{ core atmosphere model}$$
$$\dot{X}_t^\varepsilon = f(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon), \; X_0^\varepsilon = x \in \mathbb{R}^n, \text{ core ocean model}$$

where, z represent the general circulation of the atmosphere; x are the ocean components such as the density gradients, angular momentum, etc.. ξ^ε represents the unmodeled dynamics of the system or an additive noise.

Observation process:

For example, Meridional Overturning Circulation and Heatflux Array (MOCHA): Current Meters, CTDs (salinity and temperature every hour or so for a period of up to two years), buoys, acoustic releases. Deployed along 26.5° N in the Atlantic.

The observation process is a function of the signal process corrupted by noise

$$Y_t^\varepsilon = \int_0^t h^\varepsilon(X_s^\varepsilon, Z_s^\varepsilon) ds + V_t,$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is called the sensor function.

Signal Process:

$$\varepsilon \dot{Z}_t^\varepsilon = g(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon),\ Z_0^\varepsilon = z \in \mathbb{R}^m,\ \text{core atmosphere model}$$
$$\dot{X}_t^\varepsilon = f(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon),\ X_0^\varepsilon = x \in \mathbb{R}^n,\ \text{core ocean model}$$

where, z represent the general circulation of the atmosphere; x are the ocean components such as the density gradients, angular momentum, etc.. ξ^ε represents the unmodeled dynamics of the system or an additive noise.

Observation process:

For example, Meridional Overturning Circulation and Heatflux Array (MOCHA): Current Meters, CTDs (salinity and temperature every hour or so for a period of up to two years), buoys, acoustic releases. Deployed along 26.5° N in the Atlantic.

The observation process is a function of the signal process corrupted by noise

$$Y_t^\varepsilon = \int_0^t h^\varepsilon(X_s^\varepsilon, Z_s^\varepsilon)ds + V_t,$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is called the sensor function.

Signal Process:

$$\varepsilon \dot{Z}_t^\varepsilon = g(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon),\ Z_0^\varepsilon = z \in \mathbb{R}^m,\ \text{core atmosphere model}$$
$$\dot{X}_t^\varepsilon = f(X_t^\varepsilon, Z_t^\varepsilon, \xi_t^\varepsilon),\ X_0^\varepsilon = x \in \mathbb{R}^n,\ \text{core ocean model}$$

where, z represent the general circulation of the atmosphere; x are the ocean components such as the density gradients, angular momentum, etc.. ξ^ε represents the unmodeled dynamics of the system or an additive noise.

Observation process:

For example, Meridional Overturning Circulation and Heatflux Array (MOCHA): Current Meters, CTDs (salinity and temperature every hour or so for a period of up to two years), buoys, acoustic releases. Deployed along 26.5° N in the Atlantic.

The observation process is a function of the signal process corrupted by noise

$$Y_t^\varepsilon = \int_0^t h^\varepsilon(X_s^\varepsilon, Z_s^\varepsilon)ds + V_t,$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is called the sensor function.

Mathematical methodology for multiscale estimation

Multiscale process

$$(\text{slow})\ \dot{X}_t^\varepsilon = B(X_t^\varepsilon, Z_t^\varepsilon, \theta, \xi_t), \quad X_0^\varepsilon = x \in \mathbb{R}^m$$
$$(\text{fast})\ \dot{Z}_t^\varepsilon = F(\varepsilon, X_t^\varepsilon, Z_t^\varepsilon, \theta, \zeta_t), \quad Z_0^\varepsilon = z \in \mathbb{R}^n$$
$$(\text{obs})\ Y_t^\varepsilon = H(X_t^\varepsilon, Z_t^\varepsilon, \chi_t), \quad Y_0^\varepsilon = 0 \in \mathbb{R}^d$$

Nonlinear filtering
Compute conditional expectation of signal process, given observations:
$\pi_t^\varepsilon(\varphi) \stackrel{\text{def}}{=} \mathbb{E}_{\mathbb{P}^\varepsilon}\left[\varphi(X_t^\varepsilon, Z_t^\varepsilon) \middle| Y_{0:t}^\varepsilon\right]$

Maximum likelihood estimation
Maximize likelihood of signal, given parameter value:
$\theta^* \stackrel{\text{def}}{=} \arg\max_\theta \log \mathbb{E}_{\mathbb{P}^\varepsilon_\alpha}\left[\frac{d\mathbb{P}^\varepsilon_\theta}{d\mathbb{P}^\varepsilon_\alpha} \middle| Y_{0:T}^\varepsilon\right]$

Difficulty: $\mathbb{R}^{m+n}$ is high dimensional state-space; numerical computation issues

Question: for each $t \geq 0$, find the *conditional law* of X_t (only the ocean state) given $\mathscr{Y}_t$ – the information available at time $t > 0$. $\implies$ only interested in estimating X^ε dynamics

We develope methods that combine techniques of model reduction (from RDS) and nonlinear filtering (from SPDE), that enable more effective data assimilation and prediction of complex systems with multiple scales.

If only interested in estimating X^ε dynamics $\implies$ make use of process X^0

Question: for each $t \geq 0$, find the *conditional law* of X_t (only the ocean state) given $\mathscr{Y}_t$ – the information available at time $t > 0$. $\Longrightarrow$ only interested in estimating X^ε dynamics

We develope methods that combine techniques of model reduction (from RDS) and nonlinear filtering (from SPDE), that enable more effective data assimilation and prediction of complex systems with multiple scales.

If only interested in estimating X^ε dynamics $\Longrightarrow$ make use of process X^0

Lecture 1: Some Basic Stochastics – SDE's

Consider a stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^n, \tag{1}$$

and let $X \stackrel{\text{def}}{=} \{X_t, t \geq 0\}$ be the solution of (1). Denote by $\mathbf{P}_x$ the probability measure under which X satisfies (1) with $X_0 = x$, that is, $\mathbf{P}_x$ be a family of probability measures on some probability space, one for each possible initial point x. Then we know from basic stochastic process, that the SDE (1) defines a Markov process $(X, \mathbf{P}_x)$. Let $\mathbb{E}_x$ be the expectation under the probability measure $\mathbf{P}_x$.

The integral form (correct) of (1) is

$$X_t(\omega) = x + \int_0^t b(X_s(\omega))ds + \underbrace{\int_0^t \sigma(X_s(\omega))dW_s(\omega)}_{Ito}, \tag{2}$$

Theorem

For a homogeneous diffusive Markov process, with continuous drift $b(x)$ and diffusion $a(x)(\stackrel{def}{=} \sigma(x)\sigma^T(x))$ coefficients the infinitesimal generator is given by

$$(\mathscr{L}\varphi)(x) \stackrel{def}{=} \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2\varphi}{\partial x_i \partial x_j}(x) + \sum_i b_i(x)\frac{\partial\varphi}{\partial x_i}(x) \tag{3}$$

for each φ in the domain of the generator $\mathscr{D}_{\mathscr{L}} \stackrel{def}{=} C^2(\mathbb{R}^n)$.

Itô formula for $\varphi \in C^2(\mathbb{R})$

$$\varphi(X_t) = \varphi(x) + \int_0^t \varphi'(X_s)\, dX_s + \frac{1}{2}\int_0^t \varphi''(X_s)\, \langle dX, dX \rangle_s \,, \tag{4}$$

where $\langle X \rangle_s$ represents the quadratic variation of X_t is

$$\langle X, X \rangle_t \stackrel{\text{def}}{=} \int_0^t \sigma^2(X_s) ds.$$

In general, for a $\varphi(x, t)$

$$\begin{aligned} d\varphi(X_t, t) &= \frac{\partial \varphi}{\partial t}(X_t, t)dt + \frac{\partial \varphi}{\partial x}(X_t, t)dX_t + \frac{1}{2}\frac{\partial^2 \varphi}{\partial x^2}(X_t, t)\, \langle dX, dX \rangle_t \\ &= \frac{\partial \varphi}{\partial t}(X_t, t)dt + (\mathscr{L}\varphi)(X_t, t)dt + \frac{\partial \varphi}{\partial x}(X_t, t)\sigma(X_t)dW_t, \end{aligned}$$

For future reference, let's define some operators.

Definition (Generator and Symbol)

For each φ and ψ in $C^2(\mathbb{R}^2)$, define

$$(\mathscr{L}\varphi)(x,t) \stackrel{\text{def}}{=} \frac{1}{2}\sum_{i,j} a_{ij}(x,t)\frac{\partial^2\varphi}{\partial x_i \partial x_j}(x) + \sum_i b_i(x,t)\frac{\partial \varphi}{\partial x_i}(x)$$

$$\begin{aligned}\langle d\varphi, d\psi\rangle_t (x,t) &\stackrel{\text{def}}{=} \mathscr{L}(\varphi\psi)(x,t) - \varphi(x)(\mathscr{L}\psi)(x,t) - \psi(x)(\mathscr{L}\varphi)(x,t) \\ &= \sum_{i,j} a_{ij}(x,t)\frac{\partial \varphi}{\partial x_i}(x_1,x_2)\frac{\partial \psi}{\partial x_j}(x_1,x_2)\end{aligned}$$

for all $x = (x_1, x_2) \in \mathbb{R}^2$ and $t \geq 0$, where $a_{ij}(x,t) \stackrel{\text{def}}{=} (\sigma(x,t)\sigma^T(x,t))_{ij}$. The operator $\langle \cdot, \cdot \rangle$ is known as the *symbol* of $\mathscr{L}$.

Classical Stochastic Averging

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the underlying probability space. Consider the equation in standard form:

$$\frac{dX_t}{dt} = \varepsilon F(X_t, t, \omega, \varepsilon) \stackrel{\text{def}}{=} \varepsilon F(X_t, t, \xi(\omega, t), \varepsilon), \qquad X_0 = x, \qquad \omega \in \Omega, \quad (5)$$

where, for simplicity, we shall assume that $F(x, t, \omega, \varepsilon)$ is periodic in t with period T and have bounded continuous first and second partial derivatives.

Assume that the stochastic process $\xi(\omega, t)$ satisfies the strong mixing condition or, loosely speaking, ξ satisfies the condition of weak dependence (the dependence between $\xi(\omega, t)$ and $\xi(\omega, t+\tau)$ becomes weaker in some sense with increasing τ).

In the sequel, we shall state the essence of the two basic stochastic averaging theorems without proof.

Classical Stochastic Averging

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the underlying probability space. Consider the equation in standard form:

$$\frac{dX_t}{dt} = \varepsilon F(X_t, t, \omega, \varepsilon) \stackrel{\text{def}}{=} \varepsilon F(X_t, t, \xi(\omega, t), \varepsilon), \qquad X_0 = x, \qquad \omega \in \Omega, \quad (5)$$

where, for simplicity, we shall assume that $F(x, t, \omega, \varepsilon)$ is periodic in t with period T and have bounded continuous first and second partial derivatives.

Assume that the stochastic process $\xi(\omega, t)$ satisfies the strong mixing condition or, loosely speaking, ξ satisfies the condition of weak dependence (the dependence between $\xi(\omega, t)$ and $\xi(\omega, t + \tau)$ becomes weaker in some sense with increasing τ).

In the sequel, we shall state the essence of the two basic stochastic averaging theorems without proof.

It was shown that the solution of Eq. (5) in an interval of time of order $\mathcal{O}(1/\varepsilon)$ can be uniformly approximated by the solution of:

$$\frac{d\bar{X}_t}{dt} = \varepsilon \bar{F}(\bar{X}_t), \qquad \bar{X}_0 = x, \qquad \bar{F}(\bar{X}_t) = M_t(\mathbb{E}[F(X_t, t, \omega, 0)]), \qquad (6)$$

where $\mathbb{E}[\cdot]$ is the expectation, and M_t is the time-averaging operator, and the difference $X_t^\varepsilon - \bar{X}_t$ between the solutions of Eq. (5) and Eq. (6) is of order $\sqrt{\varepsilon}$.

Furthermore, it was shown by Khasminskii that the normalized difference,

$$\zeta_t^\varepsilon = \frac{(X_t^\varepsilon - \bar{X}_t)}{\sqrt{\varepsilon}},$$

not only converges to a Gaussian distribution for every fixed t as $\varepsilon \to 0$, but also converges in the sense of weak convergence to a Gaussian Markov process.

Theorem (Khasminskii (1966a))

Let the limits

$$M_t\left(\beta_i(x,t)\right) = \bar{F}_i(x), \qquad M_t\left(\int_{-\infty}^{\infty} \alpha_{kj}(x,t,s)ds\right) = [\sigma(x)\sigma^T(x)]_{kj}, \quad (7)$$

where

$$\begin{aligned} \beta_i(x,t) &= \mathbb{E}\left[F_i(x,t,\xi_t)\right], \\ \alpha_{kj} &= \mathbb{E}\left[\{F_k(x,t,\xi_t) - \beta_k(x,t)\}\{F_j(x,s,\xi_s) - \beta_j(x,s)\}\right], \\ \bar{\varphi}(x) &= M_t\left(\varphi(x,t)\right) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(x,t)dt \end{aligned} \quad (8)$$

exists uniformly in $x \in \mathbf{R}^n$.

Theorem (continued)

Then under certain conditions on $\beta(x,t)$ and its x-derivatives up to second order as $\varepsilon \to 0$ the process $\zeta_t^\varepsilon = (X_t^\varepsilon - \bar{X}_t)/\sqrt{\varepsilon}$ (normalized difference) converges weakly on the time interval of order $\mathcal{O}(1/\varepsilon)$ to a Gaussian Markov process ζ_t^0 satisfying the system of linear Itô stochastic differential equations

$$d\zeta_t^0 = B(\bar{X}_t)\zeta_t^0 dt + \sigma(\bar{X}_t)dw_t, \qquad \zeta_0^0 = 0, \tag{9}$$

where

$$B(x) = \frac{\partial \bar{F}}{\partial x}(x), \tag{10}$$

W_t is an m-dimensional Wiener process, and $X_t^\varepsilon, \bar{X}_t$ are the solutions of Eq. (5) and Eq. (6) respectively.

Extension in Higher Dimension

Illustrative example (singularly perturbed system in advective time scale).

$$\dot{Z}_t^\varepsilon = -\frac{1}{\varepsilon}(Z_t^\varepsilon - X_t^\varepsilon) + \frac{\sigma}{\sqrt{\varepsilon}}\dot{W}_t, \quad Z_0^\varepsilon = z \tag{11a}$$

$$\dot{X}_t^\varepsilon = -(Z_t^\varepsilon)^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), \quad X_0^\varepsilon = x \tag{11b}$$

For a fixed $X_t^\varepsilon = x$, (11a) becomes the Ornstein-Uhlenbeck process in which stationary density is

$$\mu(z|x) = \frac{1}{\sqrt{\pi}\sigma}\exp\{\frac{-(z-x)^2}{\sigma^2}\}.$$

As $\varepsilon \to 0$, one can show that $X_t^\varepsilon \to X_t^0$ and X_t^0 satisfies

$$\dot{X}_t^0 = b(X_t^0), \quad X_0^0 = x, \tag{12}$$

where

$$b(x) = -x^3 - \frac{3}{2}\sigma^2 x + \sin(\pi t) + \cos(\sqrt{2}\pi t).$$

For completeness, consider the Ornstein-Uhlenbeck process, for a fixed x

$$\dot{Z}_t = -(Z_t - x) + \sigma \dot{W}_t, \quad Z_0 = z.$$

The particular solution can be obtained by applying the Itô formula to $e^t Z_t$.

$$d\left(e^t Z_t\right) = e^t Z_t dt + e^t dZ_t = xe^t dt + \sigma e^t dW_t$$

Integrating both sides, yields

$$\left(e^t Z_t\right)_0^t = xe^t + \sigma \int_0^t e^s dW_s \tag{13a}$$

$$Z_t = ze^{-t} + x + \sigma \int_0^t e^{-(t-s)} dW_s \tag{13b}$$

Hence the mean and variance of the Z_t process are

$$\mathbb{E}[Z_t] = ze^{-t} + x \tag{14a}$$

$$\mathbb{E}\left[(Z_t - \mathbb{E}[Z_t])^2\right] = \sigma^2 \mathbb{E}\left[\left(\int_0^t e^{-(t-s)} dW_s\right)^2\right] = \sigma^2 \int_0^t \left(e^{-(t-s)}\right)^2 ds \tag{14b}$$

$$= \frac{\sigma^2}{2}\left[1 - e^{-2t}\right] \tag{14c}$$

Thus as $t \to \infty$,

$$Z_t \xrightarrow{d} N(x, \frac{\sigma^2}{2}).$$

For a fixed $X_t^\varepsilon = x$, (11a) becomes the Ornstein-Uhlenbeck process in which stationary density is

$$\mu(z|x) = \frac{1}{\sqrt{\pi}\sigma} \exp\{\frac{-(z-x)^2}{\sigma^2}\}.$$

The averaged drift is given by

$$b(x) \stackrel{\text{def}}{=} -\int_{-\infty}^{\infty} z^3 \mu(z|x) dz + \sin(\pi t) + \cos(\sqrt{2}\pi t)$$

and

$$\int_{-\infty}^{\infty} z^3 \mu(z|x) dz = \int_{-\infty}^{\infty} z^3 \frac{1}{\sqrt{\pi}\sigma} \exp\{\frac{-(z-x)^2}{\sigma^2}\} dz = x^3 + \frac{3}{2}\sigma^2 x$$

$$\boxed{\dot{X}_t^0 = -(X_t^0)^3 - \frac{3}{2}\sigma^2 X_t^0 + \sin(\pi t) + \cos(\sqrt{2}\pi t), \quad X_0^0 = x.} \tag{15}$$

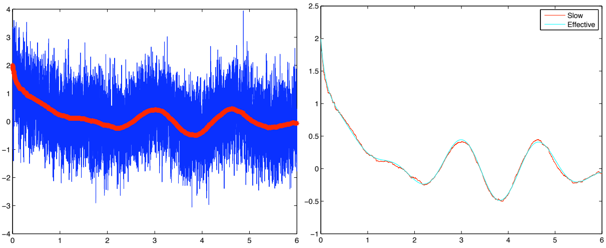


Figure: Original Signal processes and Averaged process

For the case $F(\bar{x}) \equiv 0$, the above theorem implies that in a time interval of order $\mathcal{O}(1/\varepsilon)$ the solution of Eq. (5), X_t^ε, converges to zero in probability as $\varepsilon \to 0$ because

$$\lim_{\varepsilon \to 0} \mathbf{P} \left\{ \sup_{0 \leq t \leq T} |X_t^\varepsilon - \bar{X}_t| > \delta \right\} = 0, \qquad \text{for} \quad T > 0, \quad \delta > 0 \tag{16}$$

and $\bar{X}_t$ is identically zero.

However, Stratonovich called attention to the fact that a nontrival limit distribution exists at a still slower time $\tau = \varepsilon^2 t$ if $\bar{F}(x) = 0$. On a physically rigorous level, he established that under certain conditions, the family of processes x_t^ε converges to a diffusion process, and he computed the characteristics of the limit process. A mathematically rigorous proof if this result was given by Khasminskii.

Theorem (Khasminskii (1966b))

Let $F_i(x, t, \xi_t, \varepsilon)$ *be written as*

$$F_i(x, t, \xi_t, \varepsilon) = F_i(x, t, \xi_t) + \varepsilon G(x, t) + \mathcal{O}(\varepsilon) \tag{17}$$

where the G_i *are deterministic terms, and suppose the limits*

$$\begin{aligned} b_i(x) &\stackrel{def}{=} M_t \left\{ G_i(x, t) + \int_{-\infty}^{0} \mathbb{E}\left[\frac{\partial F_i}{\partial x_j}(x, t, \xi_t)\, F_j(x, t+\tau, \xi_{t+\tau}) \right] d\tau \right\}, \\ a_{kj}(x) &\stackrel{def}{=} M_t \left\{ \int_{-\infty}^{\infty} \mathbb{E}\left[F_k(x, t, \xi_t) F_j(x, t+\tau, \xi_{t+\tau}) \right] d\tau \right\} \end{aligned} \tag{18}$$

exists uniformly in $x \in \mathbf{R}^n$.

Theorem (continued)

Then as $\varepsilon \to 0$, the process X_t^ε, the solution of Eq. (5) (with the form given in Eq. (17)) converges weakly *on a time interval of order $1/\varepsilon^2$ to a diffuse Markov process $\bar{X}_t$ with the infinitesimal generator*

$$\mathcal{L}(\cdot) = b_i(x)\frac{\partial}{\partial x_i}(\cdot) + \frac{1}{2}a_{ij}(x)\frac{\partial^2(\cdot)}{\partial x_i \partial x_j}. \tag{19}$$

One of the preeminent modern frameworks for considering convergence of the laws of Markov processes is that of the *martingale problem* (Stroock-Varadan). Papanicolaou-Stroock-Varadan were the first to make use of the martingale approach in proving limit theorems for stochastic averaging.

In the presence of bifurcations in the orbits of the fast dynamics, we need to use ideas from martingale problem to characterize the Markov process on a lower dimensional space.

Freidlin & Weber [1998], Namachchivaya & Sowers [2001] Freidlin & Wentzell [2004], Onu & Namachchivaya [2009]

Consider a SDE

$$dZ_t^{\varepsilon} = \{\frac{1}{\varepsilon}\bar{\nabla} H(Z_t^{\varepsilon}) + b(Z_t^{\varepsilon})\}dt + \sigma(Z_t^{\varepsilon})dW_t, \quad Z_0^{\varepsilon} = x \in \mathbb{R}^2. \tag{20}$$

It is clear that Z_t^{ε} defined in Eq. (20) is a Markov process on $\mathbb{R}^2$ whose generator is

$$\mathscr{L}^{\varepsilon} \stackrel{\text{def}}{=} \mathscr{L} + \frac{1}{\varepsilon}\bar{\nabla} H \tag{21}$$

with domain $\mathscr{D}(\mathscr{L}^{\varepsilon}) = \mathbf{C}^2(\mathbb{R}^2)$.

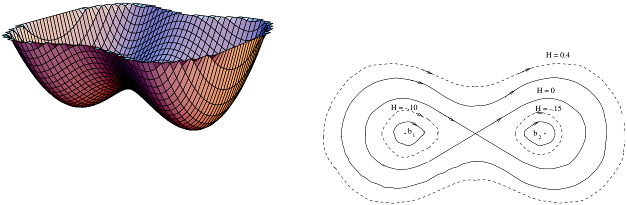


Figure: Deterministic fast processes and associated phase portraits

By moving to the canonical setting with event space $\Omega = \mathbf{C}([0,\infty);\mathbb{R}^2)$, the coordinate functions $X_t(\omega) = \omega(t)$ for all $t \geq 0$ and all $\omega \in \Omega$, Eq. (20) and Eq. (21) can be interpreted using the martingale formulation as follows:

Then according to the martingale problem if we fix $f \in \mathscr{D}(\mathscr{L}^{\varepsilon})$, any $0 \leq s < t$, any $0 \leq r_1 < r_2 \cdots < r_n \leq s$ and any $\{\varphi_j; j = 1, 2 \ldots n\} \subset \mathbf{C}_b^2(\mathbb{R}^2)$, then

$$\mathbb{E}_x^{\varepsilon}\left[\left\{f(X_t) - f(X_s) - \int_s^t (\mathscr{L}^{\varepsilon} f)(X_s)ds\right\}\prod_{i=1}^{n}\varphi_j(X_{r_j})\right] = 0. \tag{22}$$

where $\mathbb{E}_x^{\varepsilon}$ is the expectation operator corresponding to $\mathbb{P}_x^{\varepsilon}$.

The purpose is to replace, in some limiting regime, Eq. (20) by a *simpler, constructive, and rational* approximation – low-dimensional model of the dynamical system.

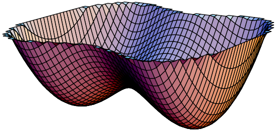
Since H is conserved, by Itô lemma

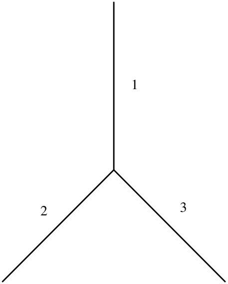
$$H(Z_t^\varepsilon) = H(z) + \int_0^t \mathscr{L}^\varepsilon H(Z_s^\varepsilon) ds + \int_0^t (\nabla H(Z_s^\varepsilon), \sigma(Z_s^\varepsilon) dW_s) \tag{23}$$

We would like to approximate the Hamiltonian $H(Z_t^\varepsilon)$ itself as a Markov process in the limit as $\varepsilon \to 0$. We achieve the model-reduction from a two-dimensional Markov process in Z_t^ε to a one-dimensional process in H through *stochastic averaging*.

The underpinning of stochastic averaging is a separation of scales; under $\mathbb{P}_x^\varepsilon$, the process X_t runs around the level sets of H very quickly, and thus a coarse-grained description of the process records only $H(X_t)$, and the $\mathbb{P}_x^\varepsilon$-dynamics of $H(X_t)$ depends only on $H(X_t)$ itself.

Mathematically, this can be described by an equivalence relation; we say that any two points x and y in $\mathbb{R}^2$ are equivalent, i.e., $x \sim y$, if $H(x) = H(y)$ and they are in the same connected component of $H^{-1}(H(x)) = H^{-1}(H(y))$; as usual, we denote by $[x]$ the equivalence class of x for any $x \in \mathbb{R}^2$. In other words, $[x]$ denotes the projection from points in $\mathbb{R}^2$ to its equivalence class under $\sim$. The $\mathbb{P}^\varepsilon_x$-law of $\varrho_t \overset{\text{def}}{=} \{[X_t];\, t \geq 0\}$ should converge to that of a Markov process on the state space of equivalence classes. We define the quotient space $\Gamma \overset{\text{def}}{=} \mathbb{R}^2/\sim$ where all points in $H^{-1}(h)$ are equivalent for $h \in \mathbb{R}$, so Γ is a 1-dimensional manifold. As $\varepsilon \to o$ we want to show that in the appropriate topology, the $\mathbb{P}^\varepsilon_x$ law of $\varrho_t = \{[X^\varepsilon_t];\, t \geq 0\}$ tends to the solution, $\mathbb{P}^0_x$, of the martingale problem for some averaged elliptic operator $\mathscr{L}_{\text{ave}}$ on Γ, i.e, $\mathscr{D}(\mathscr{L}_{\text{ave}}) = \mathbf{C}^2(\Gamma)$.





In order to prove the martingale problem, and we want to show that for fix $f \in \mathscr{D}^\dagger$ and any $0 \le r_1 < r_2 \cdots < r_n \le s < t$, and $\{\varphi_j^\dagger; j = 1, 2 \ldots n\} \subset C(\bar{\Gamma})$,

$$\mathbb{E}^\dagger \left[\left\{ f(X_t^\dagger) - f(X_s^\dagger) - \int_s^t (\mathscr{L}^\dagger f)(X_u^\dagger) du \right\} \prod_{j=1}^n \varphi_j^\dagger(X_{r_j}^\dagger) \right] = 0. \tag{24}$$

Making use of the fact $\mathbb{P}^\dagger \overset{\text{def}}{=} \lim_{\varepsilon \to 0} \mathbb{P}^{\varepsilon,\dagger}$ in weak topology of probability measures on $C([0,\infty); \bar{\Gamma})$, we can rewrite (24) as

$$\lim_{\varepsilon \to 0} \mathbb{E}^{\varepsilon,\dagger} \left[\left\{ f(X_t^\dagger) - f(X_s^\dagger) - \int_s^t (\mathscr{L}^\dagger f)(X_u^\dagger) du \right\} \prod_{j=1}^n \varphi_j^\dagger(X_{r_j}^\dagger) \right] = 0. \tag{25}$$

Revering back to the original canonical space, <u>we want to show</u>

$$\lim_{\varepsilon \to 0} \mathbb{E}^{\varepsilon} \left[\left\{ f(H(X_t)) - f(H(X_s)) - \int_s^t \underbrace{(\mathscr{L}^\dagger f)(H(X_u)) du}_{A} \right\} \prod_{j=1}^n \varphi_j(H(X_{r_j})) \right] = 0. \tag{26}$$

for fix $f \in \mathscr{D}^\dagger$ and any $0 \le r_1 < r_2 \cdots < r_n \le s < t$, and $\{\varphi_j^\dagger; j = 1, 2 \ldots n\} \subset C(\bar{\Gamma})$, But in the original canonical space, the given martingale problem (22) tells us that

$$\mathbb{E}^{\varepsilon}\left[\left\{f(H(X_t)) - f(H(X_s)) - \int_s^t \underbrace{L^{\varepsilon}(X_u)du}_{B}\right\}\prod_{j=1}^{n}\varphi_j(X_{r_j})\right] = 0. \qquad (27)$$

for all $\varepsilon > 0$, where L^{ε} is

$$\begin{aligned} L^{\varepsilon}(x) \stackrel{\text{def}}{=} \left(\mathscr{L}^{\varepsilon}\left(f \circ H\right)\right)(x) &= \frac{1}{\varepsilon}\dot{f}\left(H(x)\right)\left(\bar{\nabla} H, \nabla H\right)(x) \\ &+ \dot{f}\left(H(x)\right)(\mathscr{L} H)(x) + \frac{1}{2}\ddot{f}\left(H(x)\right)\langle dH, dH\rangle(x) \end{aligned} \qquad (28)$$

for all $x \in \mathbb{R}^2$ and $t \geq 0$.

Martingale problem (26) which we need to prove tells us that as ε goes to zero, $H(X_t)$ by itself a Markov process. However the Martingale problem (27) which we know to be true tells us that for $\varepsilon > 0$, $H(X_t)$ and X_t together form a Markov process.

It thus suffices to show that the integrands A and B match as ε goes to zero, or

$$\lim_{\varepsilon \to 0} \mathbb{E}^{\varepsilon} \left[\left| \int_0^t \left\{ L^{\varepsilon}(X_u) - \left(\mathscr{L}^{\dagger} f \right) (\varrho_u) \right\} du \right| \right] = 0.$$

We have stated heuristic arguments as to how to prove the martingale problem.

Non Standard Reduction: Illustrative example

– Surface gravity waves

Wikimedia Commons, NASA

- waves at the interface of two immiscible fluids (liquid/gas, liquid/liquid)
- Application: g-jitter

Physical model: J. W. Miles, JFM, Vol 149, 1984.

- Conservation of mass: $\nabla^2 \varphi = 0$
- Impermeable walls: $n \cdot \nabla \varphi = 0$
- Kinematic surface boundary condition

$$\frac{\partial \eta}{\partial t} + \nabla \eta \cdot \nabla \varphi = \frac{\partial \varphi}{\partial z}$$

Relevant two mode dynamics of surface-gravity waves

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k} + \zeta_k p_k + \varepsilon \dot{W}_t$$

- Horizontal forcing (W) is a Brownian motion (representing subscale behavior)
- Viscosity re-introduced, ad-hoc, as linear damping (ζ)
- Damping added to momentum, but not coordinate equations

Physical model: J. W. Miles, JFM, Vol 149, 1984.

- Conservation of mass: $\nabla^2 \varphi = 0$
- Impermeable walls: $n \cdot \nabla \varphi = 0$
- Kinematic surface boundary condition

$$\frac{\partial \eta}{\partial t} + \nabla \eta \cdot \nabla \varphi = \frac{\partial \varphi}{\partial z}$$

Relevant two mode dynamics of surface-gravity waves

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k} + \zeta_k p_k + \varepsilon \dot{W}_t$$

- Horizontal forcing (W) is a Brownian motion (representing subscale behavior)
- Viscosity re-introduced, ad-hoc, as linear damping (ζ)
- Damping added to momentum, but not coordinate equations

Non Standard Reduction: Illustrative example
- Surface gravity waves

Relevant two mode dynamics of surface-gravity waves

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k} + \zeta_k p_k + \varepsilon \dot{W}_t$$

- Horizontal forcing (W) is a Brownian motion (representing subscale behavior)
- Viscosity re-introduced, ad-hoc, as linear damping (ζ)
- Damping added to momentum, but not coordinate equations

Kristjan Onu and N. Sri Namachchivaya, "Stochastically forced water waves in a circular basin," *Proceedings of the Royal Society, A*, Vol. 466, 2010

Our *starting point* is that the process $X_t^\varepsilon \stackrel{\text{def}}{=} \{q_1(t), p_1(t), q_2(t), p_2(t)\}$ on $\mathbb{R}^4$ is a Markov process with the generator $\mathscr{L}^\varepsilon$.

For the original Markov process in $\mathbb{R}^4$ with the generator $\mathscr{L}^\varepsilon$, we know that for all test functions $f \in C^2(\mathbb{R}^4)$,

$$f(X_t) = f(X_0) + \int_0^t (\mathscr{L}^\varepsilon f)(X_u)du + M_t^{f,\varepsilon}$$

where $M^{f,\varepsilon}$ is a martingale, i.e., $\mathbb{E}^\varepsilon[M_t^{f,\varepsilon}|X_r : 0 \le r \le s] = M_s^{f,\varepsilon}$. $M^{f,\varepsilon}$ is a generalization of a zero-mean process.

Our *starting point* is that the process $X_t^\varepsilon \stackrel{\text{def}}{=} \{q_1(t), p_1(t), q_2(t), p_2(t)\}$ on $\mathbb{R}^4$ is a Markov process with the generator $\mathscr{L}^\varepsilon$.

For the original Markov process in $\mathbb{R}^4$ with the generator $\mathscr{L}^\varepsilon$, we know that for all test functions $f \in C^2(\mathbb{R}^4)$,

$$f(X_t) = f(X_0) + \int_0^t (\mathscr{L}^\varepsilon f)(X_u)du + M_t^{f,\varepsilon}$$

where $M^{f,\varepsilon}$ is a martingale, i.e., $\mathbb{E}^\varepsilon[M_t^{f,\varepsilon}|X_r : 0 \leq r \leq s] = M_s^{f,\varepsilon}$. $M^{f,\varepsilon}$ is a generalization of a zero-mean process.

State Space Reduction

We want to reduce the state space ($\mathbb{R}^4$) by using the fast (deterministic) dynamics in a way which respects noise.

Correct way is *chain equivalence*.

Informally, we say that two points x and x' in $\mathbb{R}^4$ are equivalent, i.e., $x \sim x'$, if there is a way to get from x to x' by means of a finite number of trajectories of the deterministic flow and a finite number of arbitrarily small jumps (Conley).

The reduced space is then

$$\mathfrak{M} \stackrel{\text{def}}{=} \mathbb{R}^4 / \sim .$$

Note:

- in the absence of bifurcations in the orbits, chain equivalence is the same as standard orbit equivalence
- in the presence of bifurcations, chain equivalence represents the allowable effects of noise.
- $\mathfrak{M}$ is a "stratified" space, which consists of a number of manifolds pasted together at the edges.

Freidlin & Weber [1998], Namachchivaya & Sowers [2001]

State Space Reduction

We want to reduce the state space ($\mathbb{R}^4$) by using the fast (deterministic) dynamics in a way which respects noise.

Correct way is *chain equivalence*.

Informally, we say that two points x and x' in $\mathbb{R}^4$ are equivalent, i.e., $x \sim x'$, if there is a way to get from x to x' by means of a finite number of trajectories of the deterministic flow and a finite number of arbitrarily small jumps (Conley).

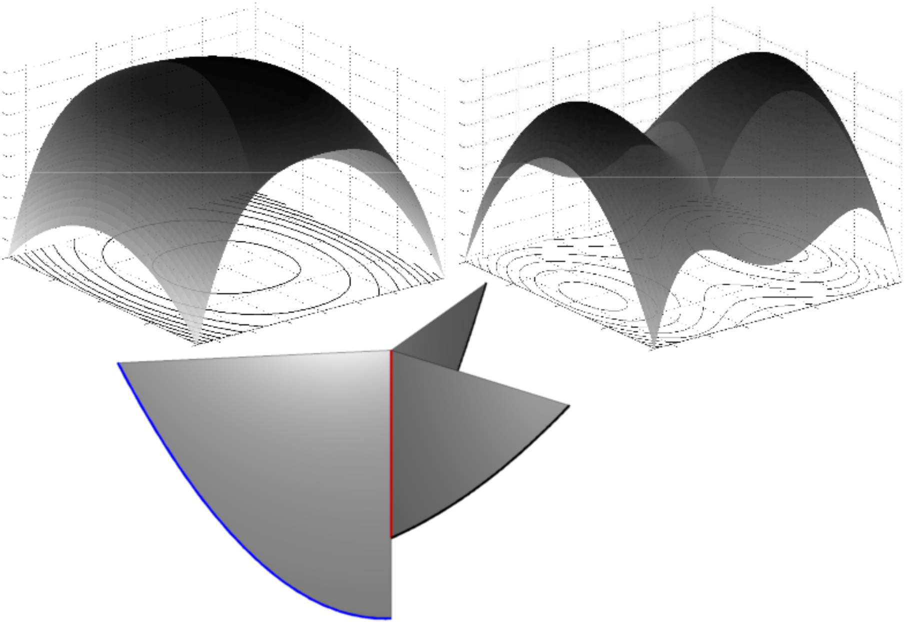
The reduced space is then

$$\mathfrak{M} \stackrel{\text{def}}{=} \mathbb{R}^4 / \sim .$$

Note:

- in the absence of bifurcations in the orbits, chain equivalence is the same as standard orbit equivalence
- in the presence of bifurcations, chain equivalence represents the allowable effects of noise.
- $\mathfrak{M}$ is a "stratified" space, which consists of a number of manifolds pasted together at the edges.

Freidlin & Weber [1998], Namachchivaya & Sowers [2001]



The dimension of $\mathfrak{M}$ is 2 (smaller than the dimension of the original state space).

If we define the slowly-varying quantity $Z : \mathbb{R}^4 \to \mathbb{R}^2$ by

$$Z(x) \stackrel{\text{def}}{=} (H(x), I(x)), \qquad x \in \mathbb{R}^4 \tag{29}$$

Our goal is to show that as ε tends to zero, the dynamics of $Z_t^\varepsilon = Z(X_t^\varepsilon)$ tends to a lower-dimensional Markov process and to identify its limiting generator $\mathscr{L}^\dagger$.

We prove that the $\mathbb{P}^\varepsilon$-law of $\{Z_t;\ t \geq 0\}$ tends to a $\mathfrak{M}$-valued Markov process with an identifiable generator. We define the limiting generator

$$(\mathscr{L}^\dagger f)(z) \stackrel{\text{def}}{=} \sum_{j=1}^{2} \mathfrak{b}_i(z)\frac{\partial f}{\partial z_j}(z) + \frac{1}{2}\sum_{j,k=1}^{2} \mathfrak{a}_{jk}(z)\frac{\partial^2 f}{\partial z_j \partial z_j}(z) \tag{30}$$

for all $z = (z_1, z_2) \in \mathcal{I}$. At the bifurcation points, *glueing conditions* define the behavior of the process.

If we define the slowly-varying quantity $Z : \mathbb{R}^4 \to \mathbb{R}^2$ by

$$Z(x) \stackrel{\text{def}}{=} (H(x), I(x)), \qquad x \in \mathbb{R}^4 \tag{29}$$

Our goal is to show that as ε tends to zero, the dynamics of $Z_t^\varepsilon = Z(X_t^\varepsilon)$ tends to a lower-dimensional Markov process and to identify its limiting generator $\mathscr{L}^\dagger$.

We prove that the $\mathbb{P}^\varepsilon$-law of $\{Z_t;\ t \geq 0\}$ tends to a $\mathfrak{M}$-valued Markov process with an identifiable generator. We define the limiting generator

$$(\mathscr{L}^\dagger f)(z) \stackrel{\text{def}}{=} \sum_{j=1}^{2} \mathfrak{b}_i(z) \frac{\partial f}{\partial z_j}(z) + \frac{1}{2} \sum_{j,k=1}^{2} \mathfrak{a}_{jk}(z) \frac{\partial^2 f}{\partial z_j \partial z_j}(z) \tag{30}$$

for all $z = (z_1, z_2) \in \mathcal{I}$. At the bifurcation points, *glueing conditions* define the behavior of the process.

If we define the slowly-varying quantity $Z : \mathbb{R}^4 \to \mathbb{R}^2$ by

$$Z(x) \stackrel{\text{def}}{=} (H(x), I(x)), \qquad x \in \mathbb{R}^4 \tag{29}$$

Our goal is to show that as ε tends to zero, the dynamics of $Z_t^\varepsilon = Z(X_t^\varepsilon)$ tends to a lower-dimensional Markov process and to identify its limiting generator $\mathscr{L}^\dagger$.

We prove that the $\mathbb{P}^\varepsilon$-law of $\{Z_t;\ t \geq 0\}$ tends to a $\mathfrak{M}$-valued Markov process with an identifiable generator. We define the limiting generator

$$(\mathscr{L}^\dagger f)(z) \stackrel{\text{def}}{=} \sum_{j=1}^{2} \mathfrak{b}_i(z) \frac{\partial f}{\partial z_j}(z) + \frac{1}{2} \sum_{j,k=1}^{2} \mathfrak{a}_{jk}(z) \frac{\partial^2 f}{\partial z_j \partial z_j}(z) \tag{30}$$

for all $z = (z_1, z_2) \in \mathcal{I}$. At the bifurcation points, *glueing conditions* define the behavior of the process.

The limiting domain $\mathscr{D}^\dagger$ for the graph valued process is

$$\mathscr{D}^\dagger_{\mathfrak{M}} = \Bigg\{ f \in C(\mathfrak{M}) \cap C^2(\cup_{i=1}^3 \mathcal{I}_i) : \lim_{z \searrow (H(\mathfrak{c}_i), I(\mathfrak{c}_i))} (\mathscr{L}_i f_i)(h) \quad \text{exists} \quad \forall i,$$

$$\lim_{z_2 \nearrow I^*} (\mathscr{L}_i f_i)(z) = 0 \quad \forall i, \quad \underbrace{\sum_{i=1}^{3} \{\pm\} \sum_{j=1}^{2} \left\{ \sum_{k=1}^{2} \overset{\circ}{\mathfrak{a}}{}^i_{jk}(z) \frac{\partial f_i}{\partial z_k}(z) \right\} \cdot \nu_j \Big|_{z=\mathcal{O}} = 0}_{\text{glueing conditions}} \Bigg\}$$

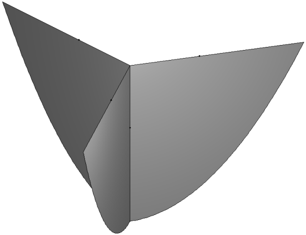


Figure: Diffusion with *glueing conditions* on the reduced space.

Heuristically "coin flips" to decide behavior at a vertex.

Glueing conditions

- Are similar to Dirichlet or Neumann boundary conditions (from the perspective of PDE's)
- Are similar to reflection or absorption (from the perspective of a diffusing particle).
- Heuristically correspond to "coin flips" to decide behavior at a vertex.

We prove that the $\mathbb{P}^\varepsilon$-law of $\{Z_t;\, t \geq 0\}$ tends to a $\mathfrak{M}$-valued Markov process with an identifiable generator.

Our main result is that the $\mathbb{P}^\varepsilon$'s tend to the unique solution $\mathbb{P}^\dagger$ of the martingale problem with generator $\mathscr{L}^\dagger_{\mathfrak{M}}$, with domain $\mathscr{D}^\dagger_{\mathfrak{M}}$ and with initial condition δ_z.

Dimensional Reduction for Multi-scale RDS [Summary]

Developed rigorous methods to replace the original n - dimensional signal process X_t by a simpler m - dimensional Markov process Z_t on a stratified reduced space with the generator $\mathscr{L}^{\dagger}_{\mathfrak{M}}$,

$$\mathscr{L}^{\dagger}_{\mathfrak{M}} = b_i \frac{\partial}{\partial z_i} + \frac{1}{2} a_{ij} \frac{\partial^2}{\partial z_i \partial z_j} \quad z \in \mathbb{R}^m, \quad m << n, \quad \text{formulas for } \mathrm{a_{ij}} \text{ and } \mathrm{b_i} \text{ were derive}$$

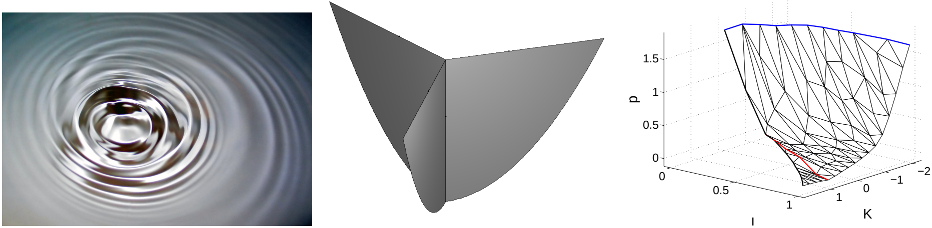


Figure: Two mode dynamics of surface-gravity waves, Stratified reduced space $\mathfrak{M}$ & PDF on reduced space

Dimensional Reduction for Multi-scale RDS

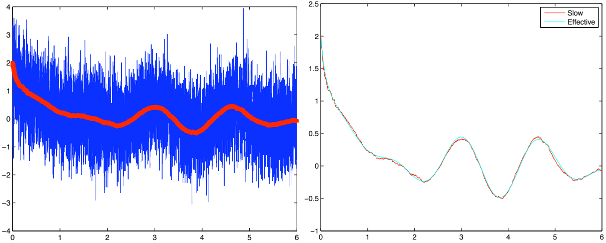


Figure: Original Signal processes and Averaged process

Original Signal processes

$$\dot{X}_t^{\varepsilon} = f(X_t^{\varepsilon}, Z_t^{\varepsilon}, \xi_t^{\varepsilon}), \quad \varepsilon \dot{Z}_t^{\varepsilon} = g(X_t^{\varepsilon}, Z_t^{\varepsilon}, \xi_t^{\varepsilon})$$

Averaged process

$$d\overline{X}_t = b(\overline{X}_t) + \sigma(\overline{X}_t) dW_t$$