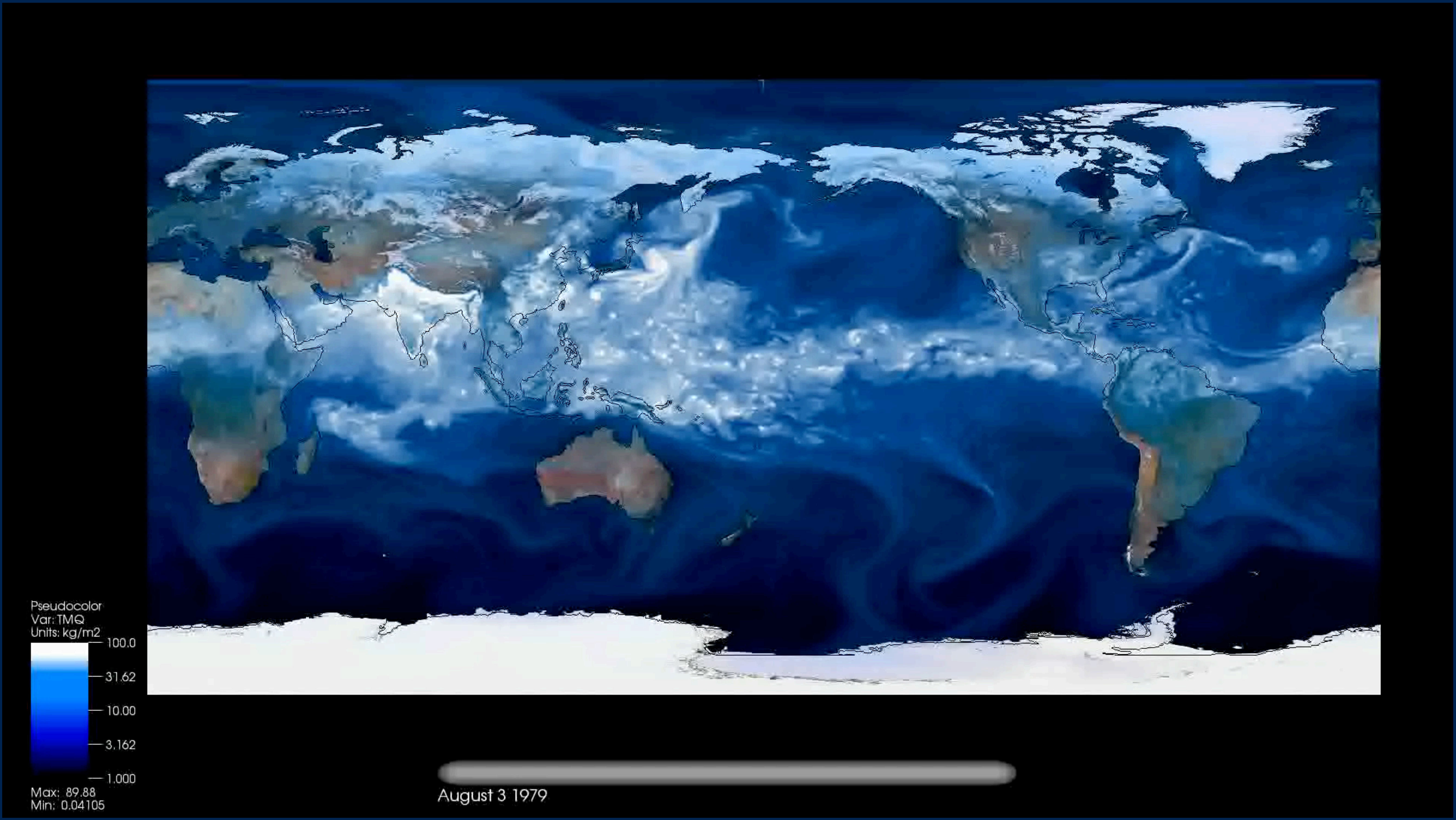
Stochastic parameterizations of deterministic dynamical systems: Theory, applications and challenges

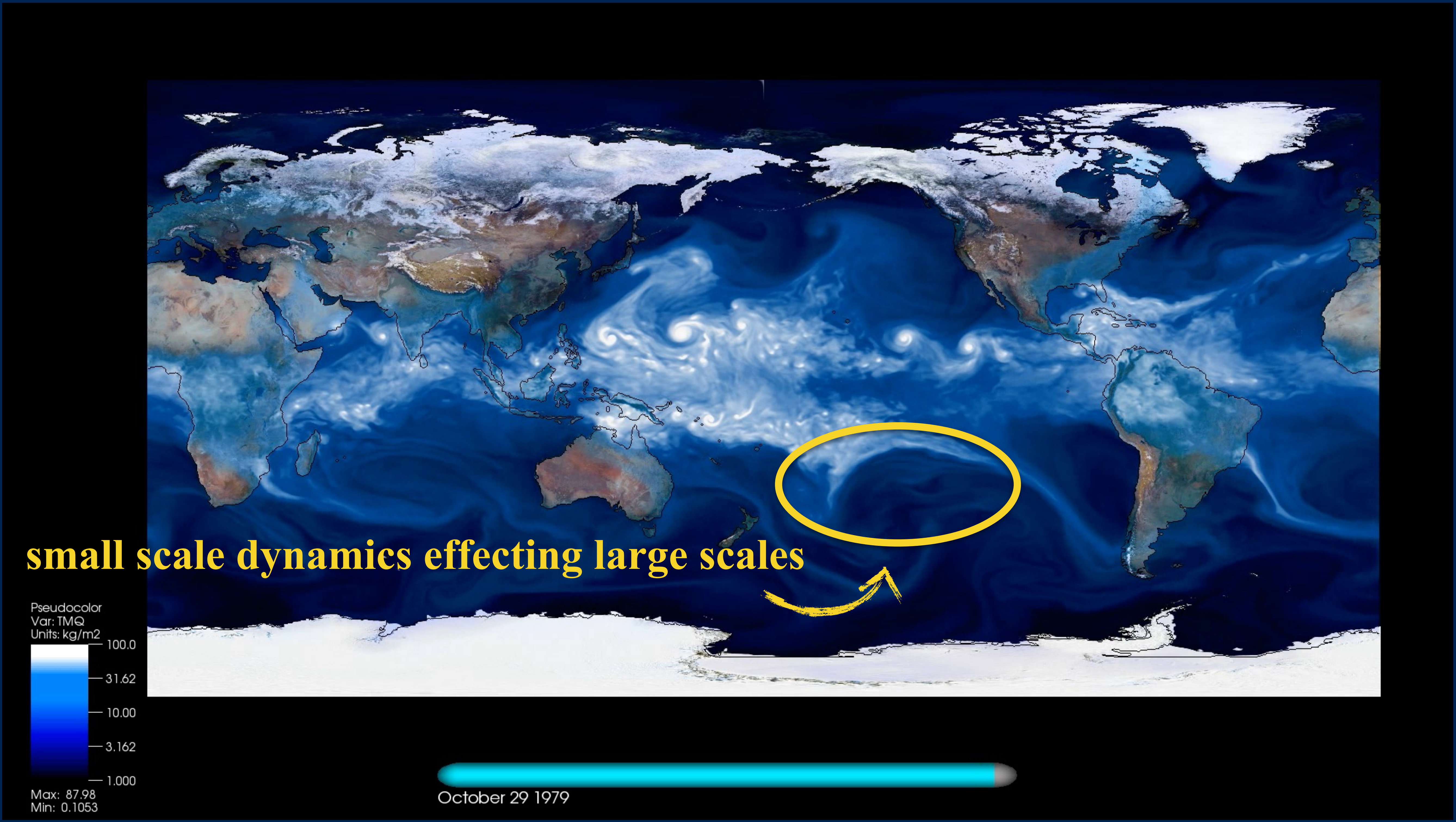
Georg Gottwald

joint work with Jeroen Wouters, Ian Melbourne, Jason Frank

THE UNIVERSITY OF SYDNEY

Geneva, January 30, 2017


Pseudocolor
Var: TMQ
Units: kg/m2
100.0
31.62
10.00
3.162
1.000
Max: 89.88
Min: 0.04105
August 3 1979


small scale dynamics effecting large scales
Pseudocolor
Var: TMQ
Units: kg/m2
100.0
31.62
10.00
3.162
1.000
Max: 87.98
Min: 0.1053
October 29 1979

Motivation for stochastic parametrisation:

- prediction: computational cost in running model

$$
\left.
\begin{aligned}
\dot{x} &= f(x,y) \\
\dot{y} &= \frac{1}{\varepsilon} g(x,y) \\
\\
x &\in \mathbb{R}^n \\
y &\in \mathbb{R}^{m} \\
\\
\varepsilon &\ll 1
\end{aligned}
\right\}
$$

$$dX = F(X)dt + \Sigma\, dW_t$$

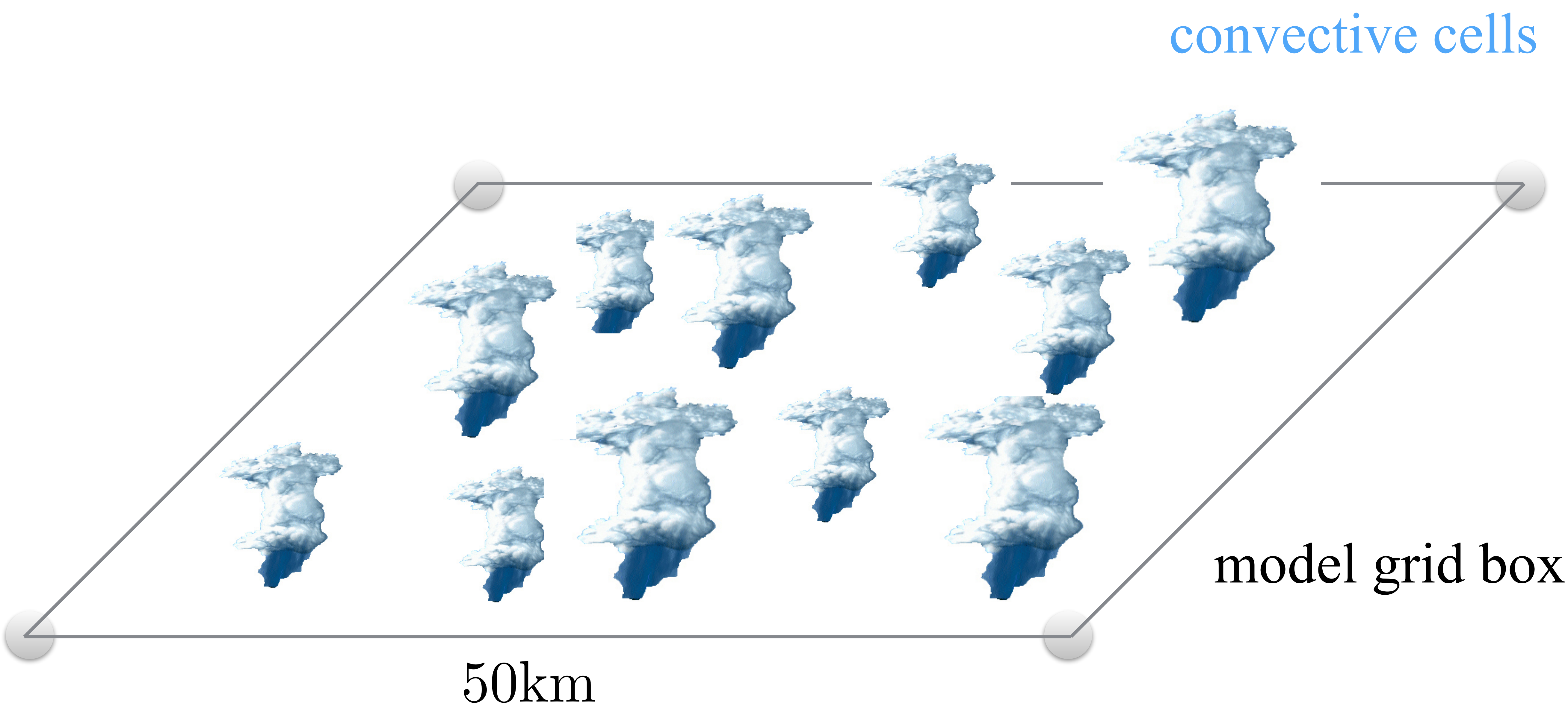
$$X \in \mathbb{R}^n$$

lower-dimensional stochastic problem

stiff high-dimensional deterministic
multi-scale problem

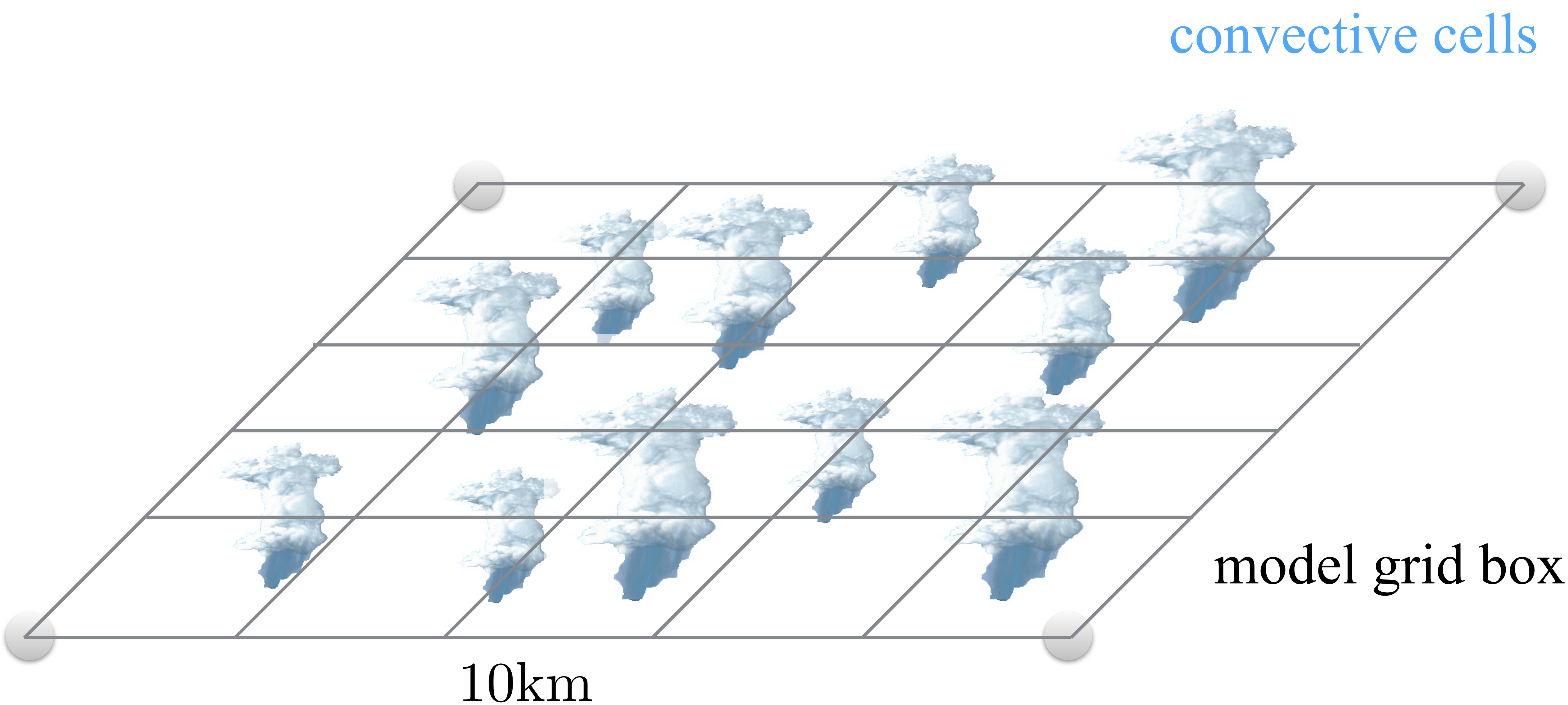
Motivation for stochastic parametrisation:

- prediction: computational cost in running model
- increase of resolution necessitates stochastic approach



Motivation for stochastic parametrisation:

- prediction: computational cost in running model
- increase of resolution necessitates stochastic approach



Motivation for stochastic parametrisation:

- prediction: computational cost in running model
- increase of resolution necessitates stochastic approach

- data assimilation/ensemble filters: use the reduced stochastic model as your forecast model (Mitchell and GAG, JAS (2012), GAG & Harlim, Proc Roy Soc A (2014))

combine limited observations with our knowledge of the laws of physics for optimal state estimation

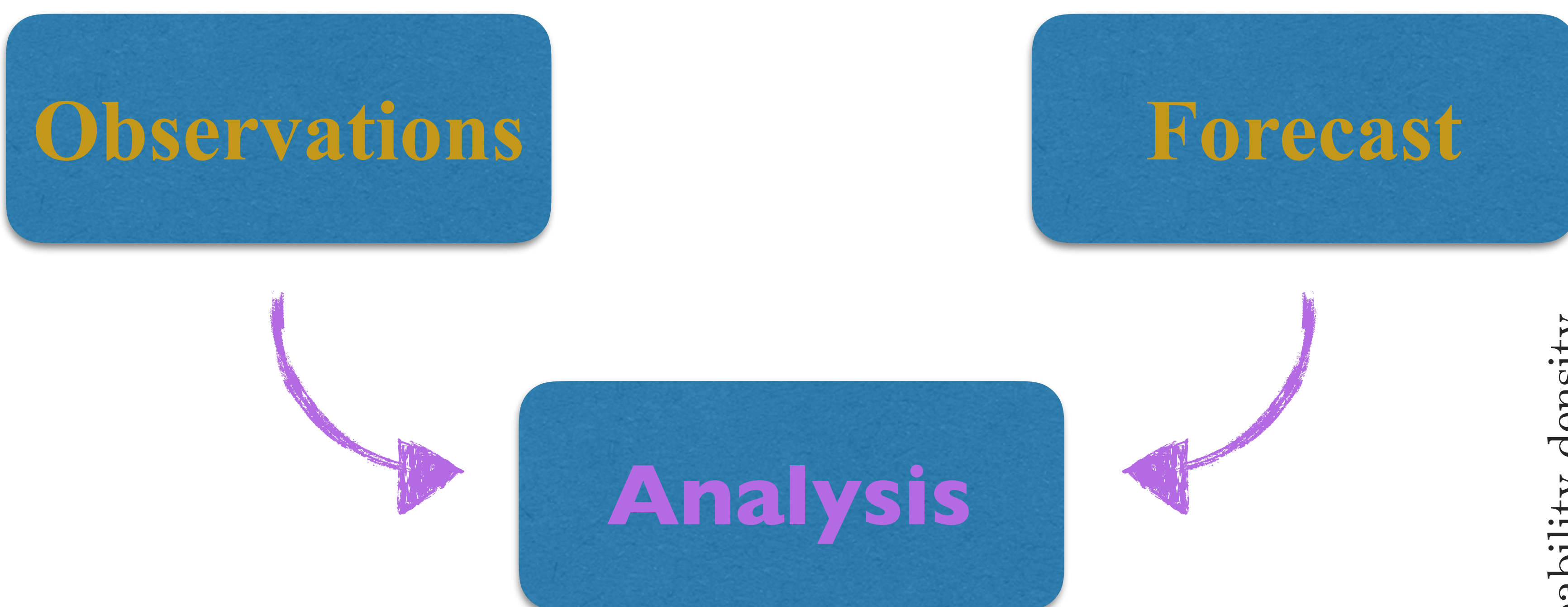
Optimal state estimate

Motivation for stochastic parametrisation:

- prediction: computational cost in running model
- increase of resolution necessitates stochastic approach

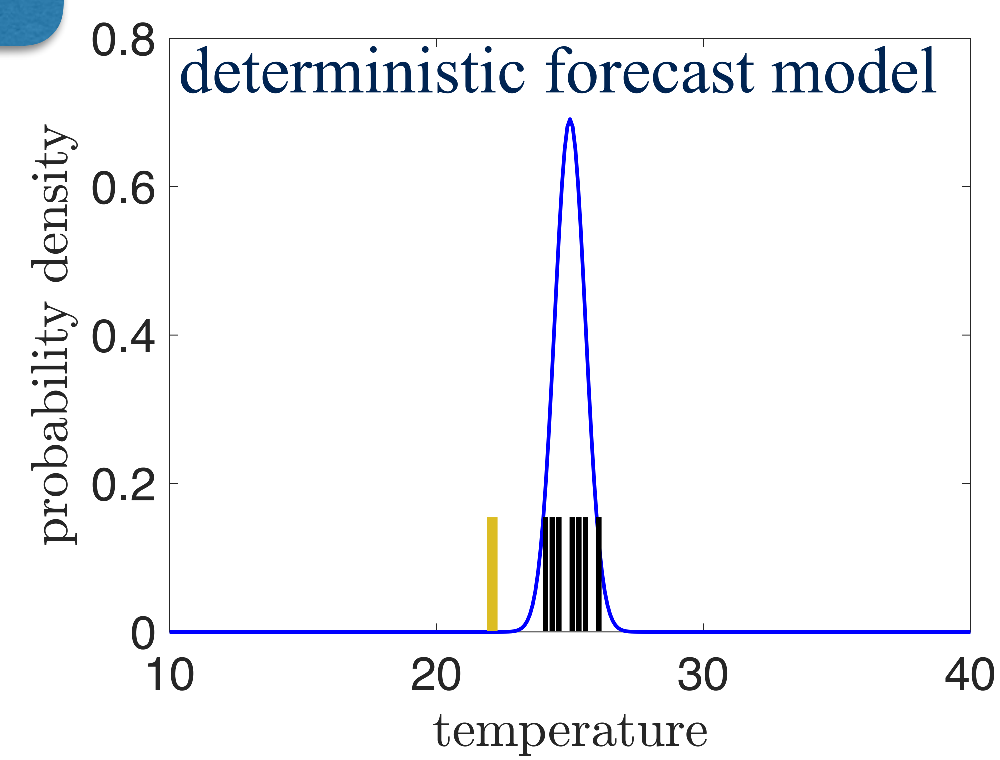
- data assimilation/ensemble filters: use the reduced stochastic model as your forecast model (Mitchell and GAG, JAS (2012), GAG & Harlim, Proc Roy Soc A (2014))

combine limited observations with our knowledge of the laws of physics for optimal state estimation



Optimal state estimate

What can go wrong?

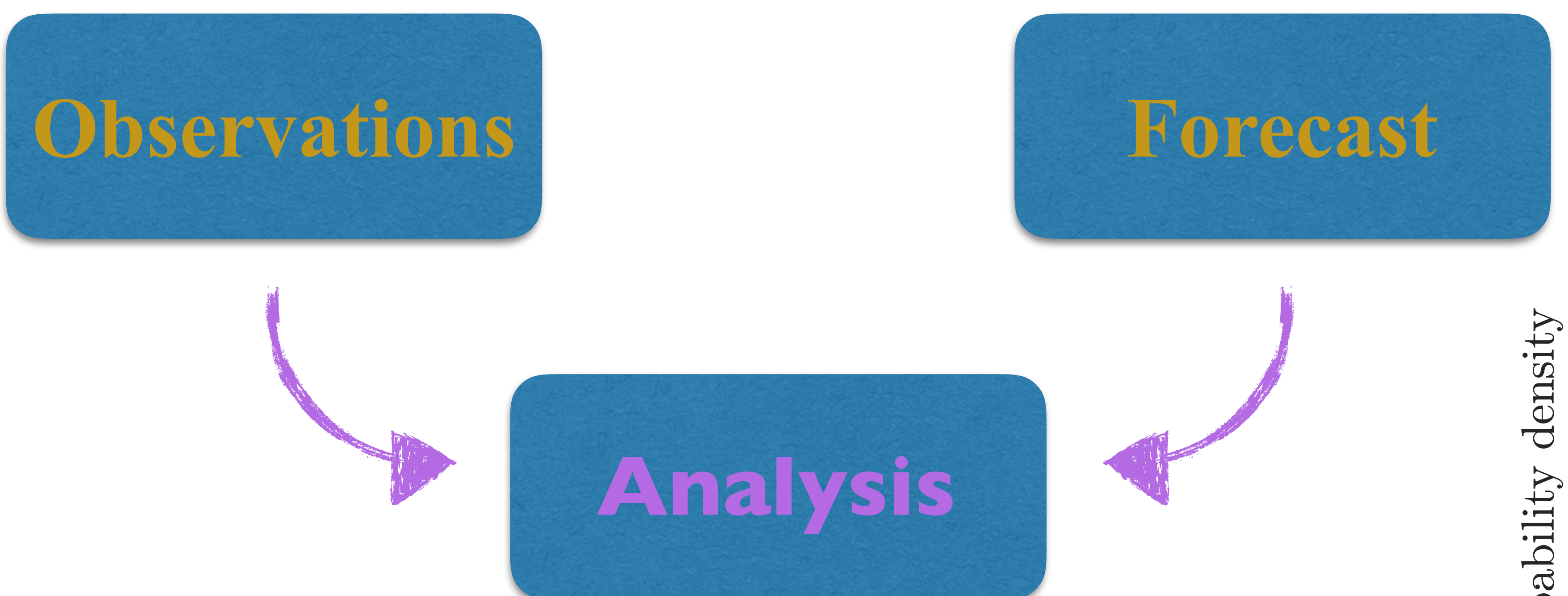


Motivation for stochastic parametrisation:

- prediction: computational cost in running model
- increase of resolution necessitates stochastic approach

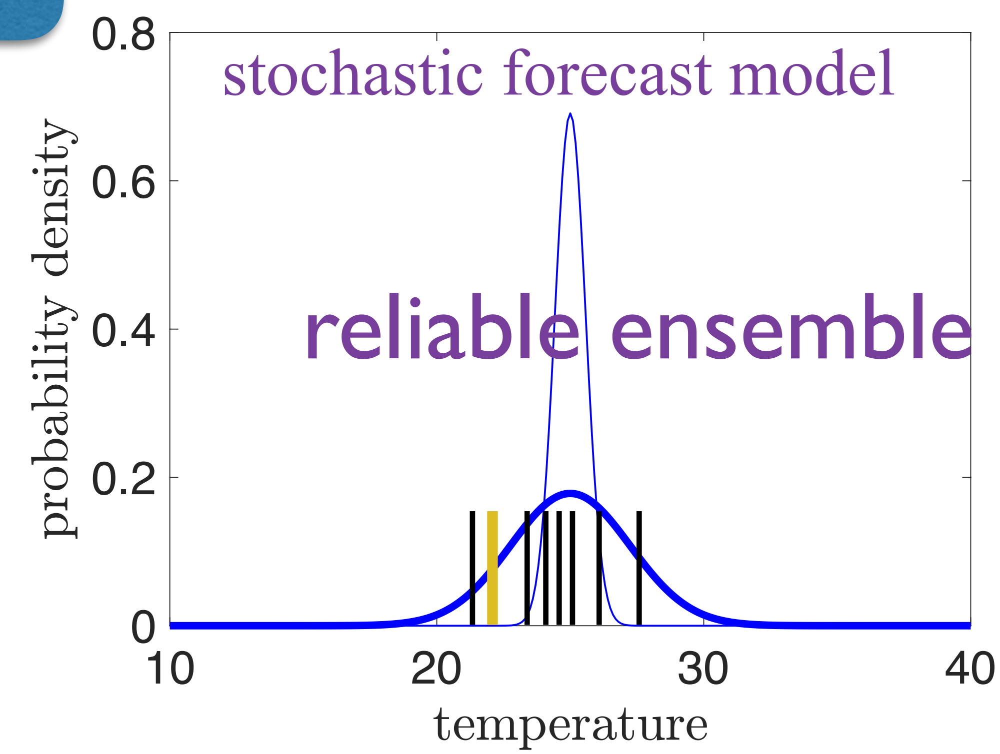
- data assimilation/ensemble filters: use the reduced stochastic model as your forecast model (Mitchell and GAG, JAS (2012), GAG & Harlim, Proc Roy Soc A (2014))

combine limited observations with our knowledge of the laws of physics for optimal state estimation



Optimal state estimate

What can go wrong?



Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\,dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\,dt + \frac{1}{\sqrt{\varepsilon}}\sigma\,dW_t$$

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\,dt$$
$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\,dt + \frac{1}{\sqrt{\varepsilon}}\sigma\,dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \int_0^t f(x^{(\varepsilon)}, y^{(\varepsilon)}(s))\,ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon}} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\,d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{n} \int_0^{nt} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\,d\tau \end{aligned}$$

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\, dt + \frac{1}{\sqrt{\varepsilon}} \sigma\, dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \int_0^t f(x^{(\varepsilon)}, y^{(\varepsilon)}(s))\, ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon}} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\, d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{n} \int_0^{nt} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\, d\tau \end{aligned}$$

Envoking Birkhoff's Ergodic Theorem

$$X(t) = X(0) + \int_0^t F(X(s))\, ds$$

$$F(X) = \int f(x, y)\, \mu(dy)$$

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\, dt + \frac{1}{\sqrt{\varepsilon}} \sigma\, dW_t$$

Integrate the slow equation

$$x^{(\varepsilon)}(t) = x^{(\varepsilon)}(0) + \int_0^t f(x^{(\varepsilon)}, y^{(\varepsilon)}(s))\, ds$$

$$= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon}} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\, d\tau$$

$$= x^{(\varepsilon)}(0) + \frac{1}{n} \int_0^{nt} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\, d\tau$$

Envoking Birkhoff's Ergodic Theorem

$$X(t) = X(0) + \int_0^t F(X(s))\, ds$$

$$F(X) = \int f(x, y)\, \mu(dy)$$

Averaged deterministic dynamics

law of large numbers

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\,dt$$
$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\,dt + \frac{1}{\sqrt{\varepsilon}}\sigma\,dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \int_0^t f(x^{(\varepsilon)}, y^{(\varepsilon)}(s))\,ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon}} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\,d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{n}\int_0^{nt} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\,d\tau \end{aligned}$$

Envoking Birkhoff's Ergodic Theorem

$$X(t) = X(0) + \int_0^t F(X(s))\,ds$$
$$F(X) = \int f(x,y)\,\mu(dy)$$

Averaged deterministic dynamics

law of large numbers

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\, dt + \frac{1}{\sqrt{\varepsilon}} \sigma\, dW_t$$

Integrate the slow equation

$$x^{(\varepsilon)}(t) = x^{(\varepsilon)}(0) + \int_0^t f(x^{(\varepsilon)}, y^{(\varepsilon)}(s))\, ds$$

$$= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon}} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\, d\tau$$

$$= x^{(\varepsilon)}(0) + \frac{1}{n} \int_0^{nt} f(x^{(\varepsilon)}, y^{(\varepsilon=1)}(\tau))\, d\tau$$

Envoking Birkhoff's Ergodic Theorem

$$X(t) = X(0) + \int_0^t F(X(s))\, ds$$

$$F(X) = \int f(x, y)\, \mu(dy)$$

Averaged deterministic dynamics

law of large numbers

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = f(x^{(\varepsilon)}, y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon} g(y^{(\varepsilon)})\, dt + \frac{1}{\sqrt{\varepsilon}} \sigma\, dW_t$$

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = \frac{1}{\varepsilon} f(y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon^2} g(y^{(\varepsilon)})\, dt + \frac{1}{\varepsilon} \sigma\, dW_t$$

go to long *diffusive* time scale

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = \frac{1}{\varepsilon} f(y^{(\varepsilon)})\, dt$$
$$dy^{(\varepsilon)} = \frac{1}{\varepsilon^2} g(y^{(\varepsilon)})\, dt + \frac{1}{\varepsilon}\sigma\, dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \frac{1}{\varepsilon}\int_0^t f(y^{(\varepsilon)}(s))\, ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f(y^{(\varepsilon=1)}(\tau))\, d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{\sqrt{n}} \int_0^{nt} f(y^{(\varepsilon=1)}(\tau))\, d\tau \end{aligned}$$

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = \frac{1}{\varepsilon} f(y^{(\varepsilon)})\, dt$$
$$dy^{(\varepsilon)} = \frac{1}{\varepsilon^2} g(y^{(\varepsilon)})\, dt + \frac{1}{\varepsilon} \sigma\, dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \frac{1}{\varepsilon} \int_0^t f(y^{(\varepsilon)}(s))\, ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f(y^{(\varepsilon=1)}(\tau))\, d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{\sqrt{n}} \int_0^{nt} f(y^{(\varepsilon=1)}(\tau))\, d\tau \end{aligned}$$

Assuming $\int f_0(y)\mu(dy) = 0$ and envoking the Central Limit Theorem

$$X(t) = X(0) + W_t$$
$$dX = dW_t$$

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = \frac{1}{\varepsilon} f(y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon^2} g(y^{(\varepsilon)})\, dt + \frac{1}{\varepsilon}\sigma\, dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \frac{1}{\varepsilon}\int_0^t f(y^{(\varepsilon)}(s))\, ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f(y^{(\varepsilon=1)}(\tau))\, d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{\sqrt{n}}\int_0^{nt} f(y^{(\varepsilon=1)}(\tau))\, d\tau \end{aligned}$$

Assuming $\int f_0(y)\mu(dy) = 0$ and envoking the Central Limit Theorem

$$X(t) = X(0) + W_t$$

$$dX = dW_t$$

Homogenised stochastic equation

central limit theorem

Heuristics for why the fast process can be replaced by noise

$$dx^{(\varepsilon)} = \frac{1}{\varepsilon} f(y^{(\varepsilon)})\, dt$$

$$dy^{(\varepsilon)} = \frac{1}{\varepsilon^2} g(y^{(\varepsilon)})\, dt + \frac{1}{\varepsilon}\sigma\, dW_t$$

Integrate the slow equation

$$\begin{aligned} x^{(\varepsilon)}(t) &= x^{(\varepsilon)}(0) + \frac{1}{\varepsilon}\int_0^t f(y^{(\varepsilon)}(s))\, ds \\ &= x^{(\varepsilon)}(0) + \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f(y^{(\varepsilon=1)}(\tau))\, d\tau \\ &= x^{(\varepsilon)}(0) + \frac{1}{\sqrt{n}} \int_0^{nt} f(y^{(\varepsilon=1)}(\tau))\, d\tau \end{aligned}$$

Assuming $\int f_0(y)\mu(dy) = 0$ and envoking the Central Limit Theorem

$$X(t) = X(0) + W_t$$

$$dX = dW_t$$

Homogenised stochastic equation

central limit theorem

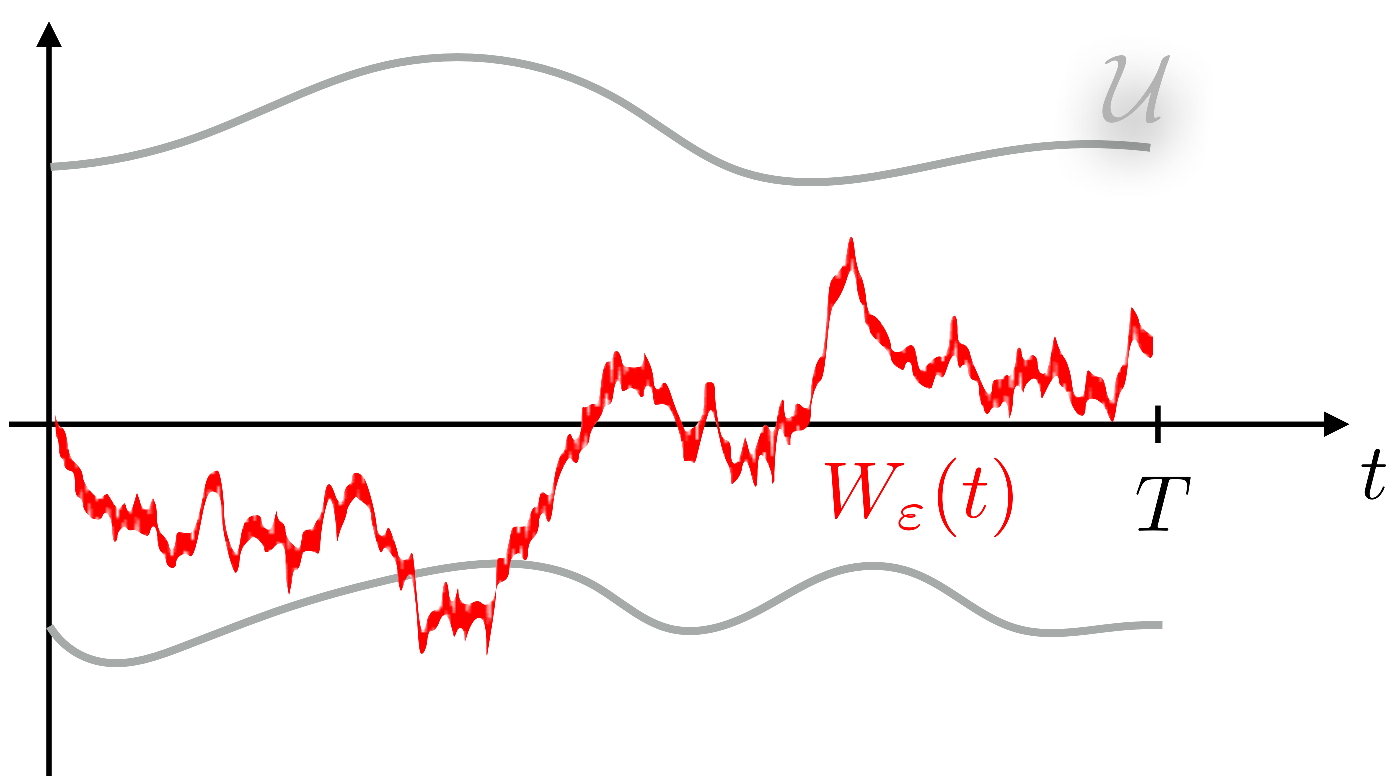
The Weak Invariance Principle

$$W_\varepsilon(t) = \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f_0(y(s))ds \longrightarrow_w W(t) \quad \text{as} \quad \varepsilon \to 0$$

weak convergence in $C([0,T],\mathbb{R})$

$$\mathbb{P}(W_\varepsilon \in \mathcal{U}) \longrightarrow \mathbb{P}(W \in \mathcal{U})$$

for suitable subsets of open collection of sample paths $\mathcal{U} \subset C([0,T],\mathbb{R})$



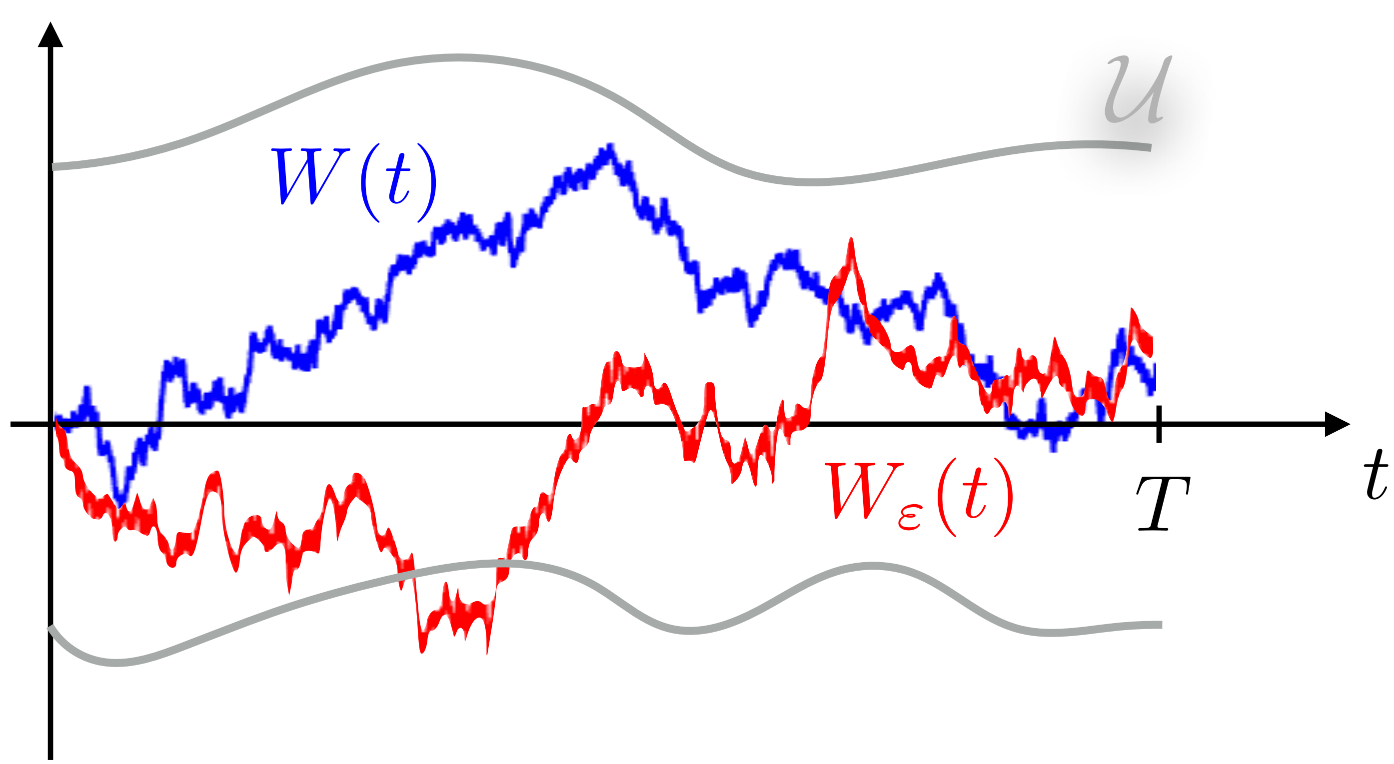
The Weak Invariance Principle

$$W_\varepsilon(t) = \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f_0(y(s))ds \longrightarrow_w W(t) \quad \text{as} \quad \varepsilon \to 0$$

weak convergence in $C([0,T],\mathbb{R})$

$$\mathbb{P}(W_\varepsilon \in \mathcal{U}) \longrightarrow \mathbb{P}(W \in \mathcal{U})$$

for suitable subsets of open collection of sample paths $\mathcal{U} \subset C([0,T],\mathbb{R})$



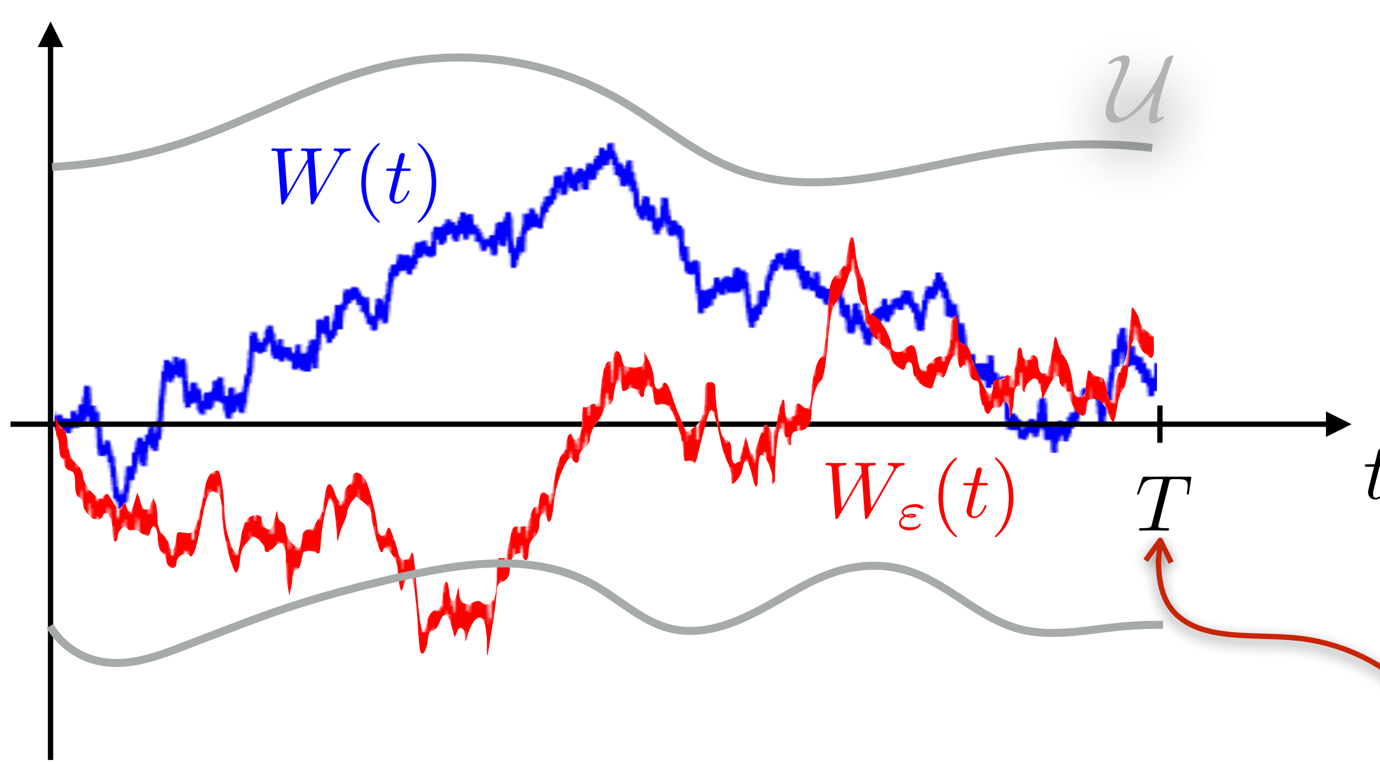
The Weak Invariance Principle

$$W_\varepsilon(t) = \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f_0(y(s))ds \longrightarrow_w W(t) \quad \text{as} \quad \varepsilon \to 0$$

weak convergence in $C([0,T],\mathbb{R})$

$$\mathbb{P}(W_\varepsilon \in \mathcal{U}) \longrightarrow \mathbb{P}(W \in \mathcal{U})$$

for suitable subsets of open collection of sample paths $\mathcal{U} \subset C([0,T],\mathbb{R})$

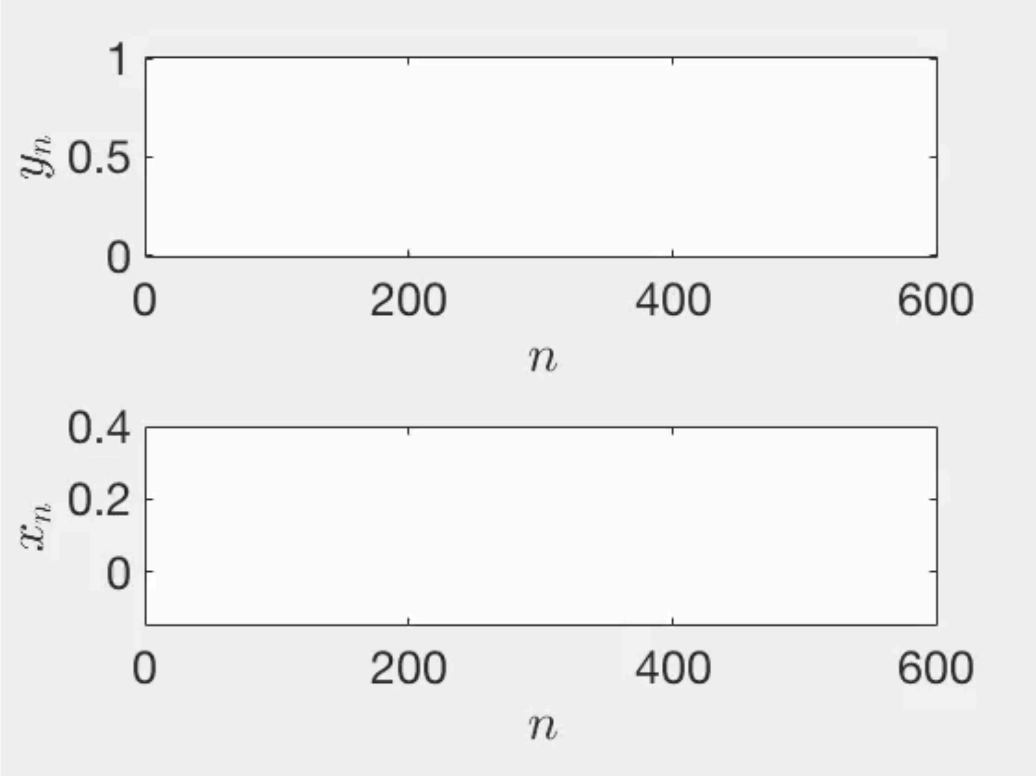


this implies the Central Limit Theorem at all times t ∈ [0,T]

Homogenisation in action

$$x_{n+1} = x_n + \varepsilon(y_n - \tfrac{1}{2})$$
$$y_{n+1} = 4y_n(1 - y_n)$$

strong chaos



Brownian motion

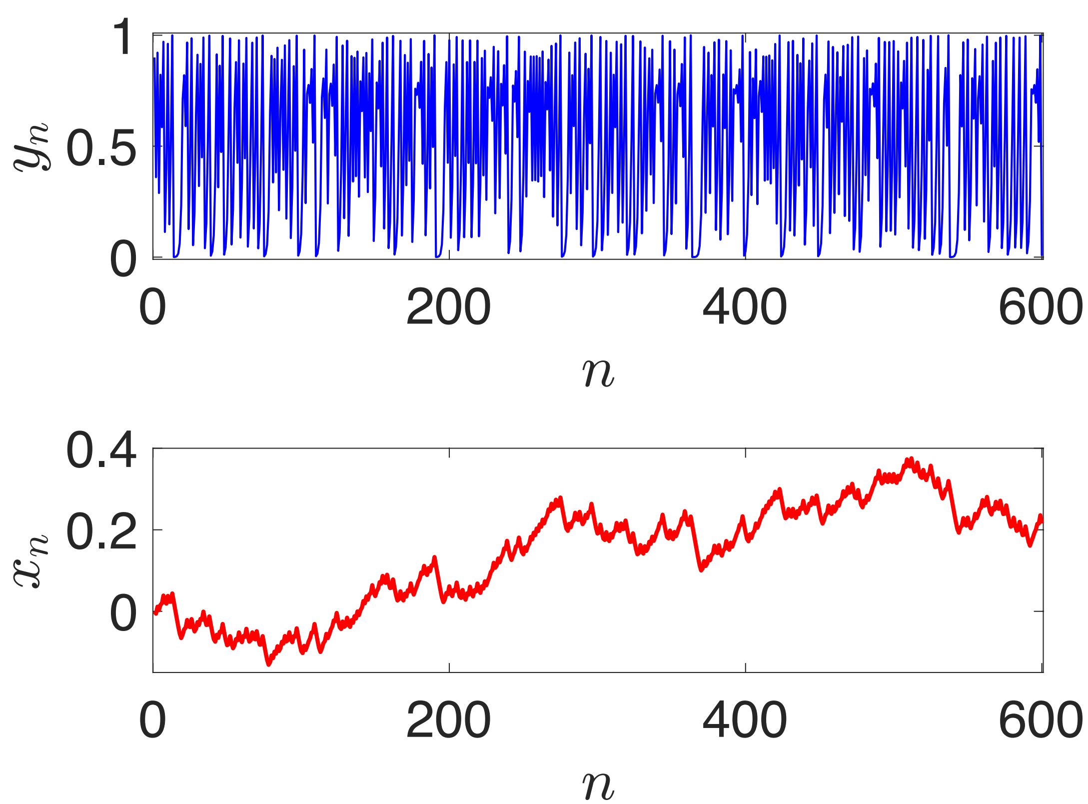
Homogenisation in action

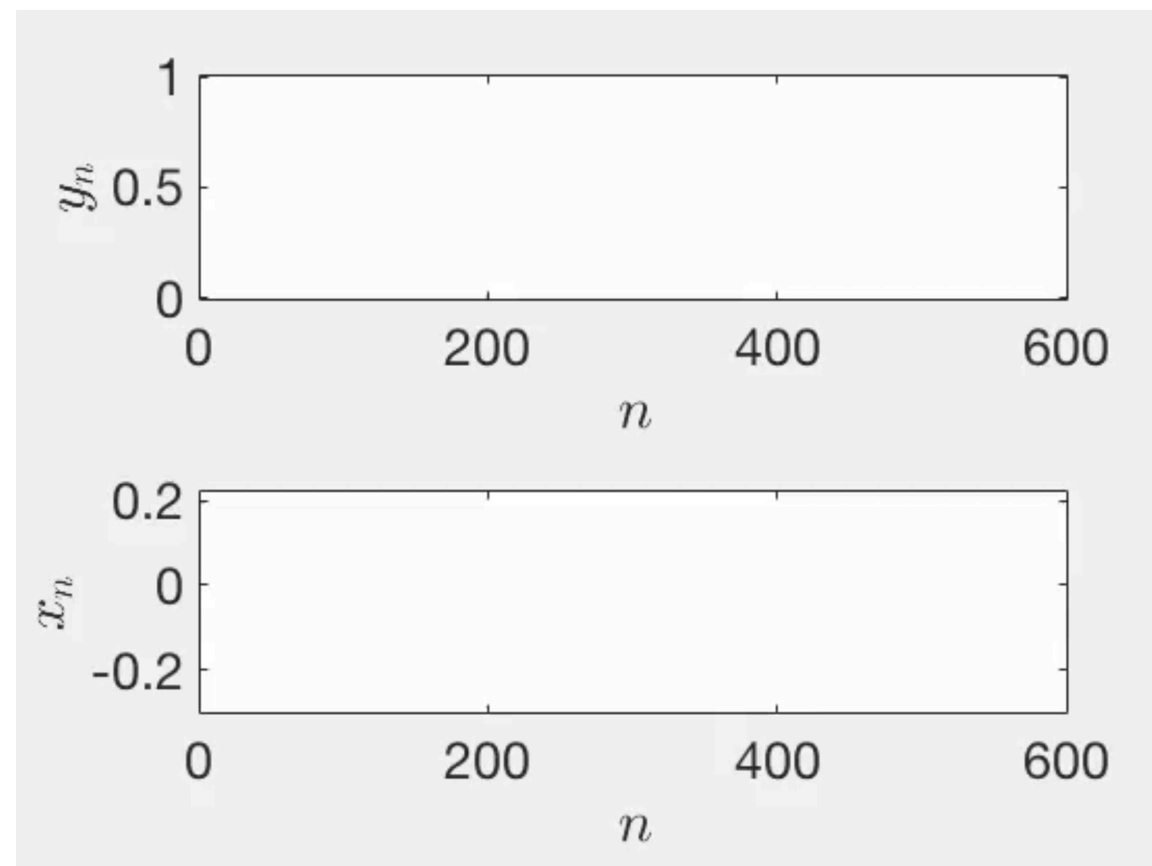
$$x_{n+1} = x_n + \varepsilon(y_n - \tfrac{1}{2})$$
$$y_{n+1} = 4y_n(1 - y_n)$$

strong chaos

$$x_{n+1} = x_n + \varepsilon(y^\star - y_n)$$
$$y_{n+1} = \begin{cases} y_n(1 + 2^\gamma y_n^\gamma) & 0 \le y_n \le \frac{1}{2} \\ 2y_n - 1 & \frac{1}{2} \le y_n \le 1 \end{cases}$$

weak chaos





Brownian motion

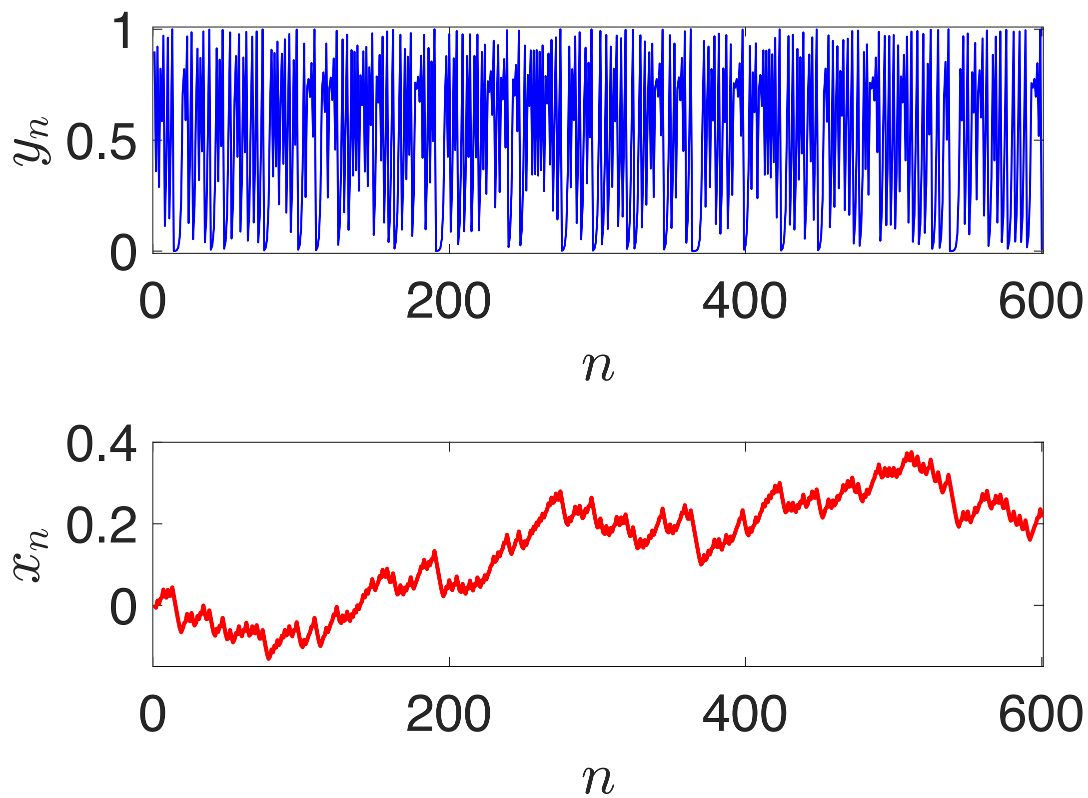
Homogenisation in action

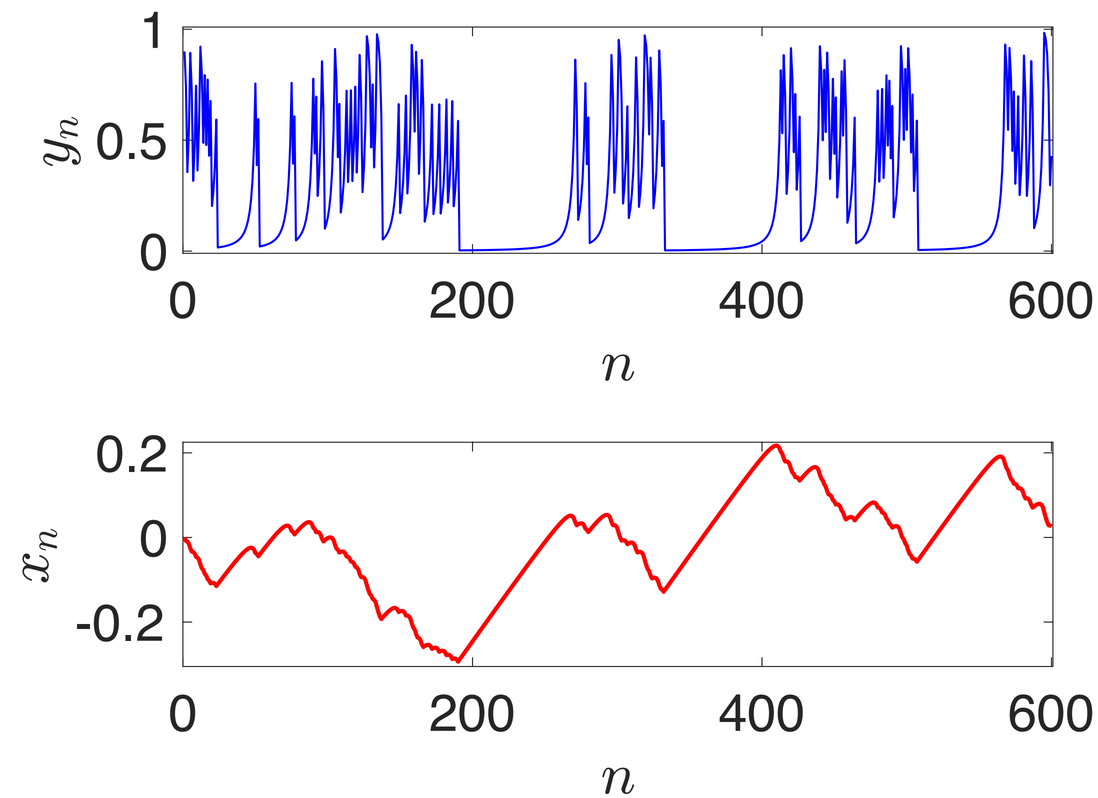
$$x_{n+1} = x_n + \varepsilon(y_n - \tfrac{1}{2})$$
$$y_{n+1} = 4y_n(1 - y_n)$$

strong chaos

$$x_{n+1} = x_n + \varepsilon(y^\star - y_n)$$
$$y_{n+1} = \begin{cases} y_n(1 + 2^\gamma y_n^\gamma) & 0 \le y_n \le \frac{1}{2} \\ 2y_n - 1 & \frac{1}{2} \le y_n \le 1 \end{cases}$$

weak chaos





Brownian motion

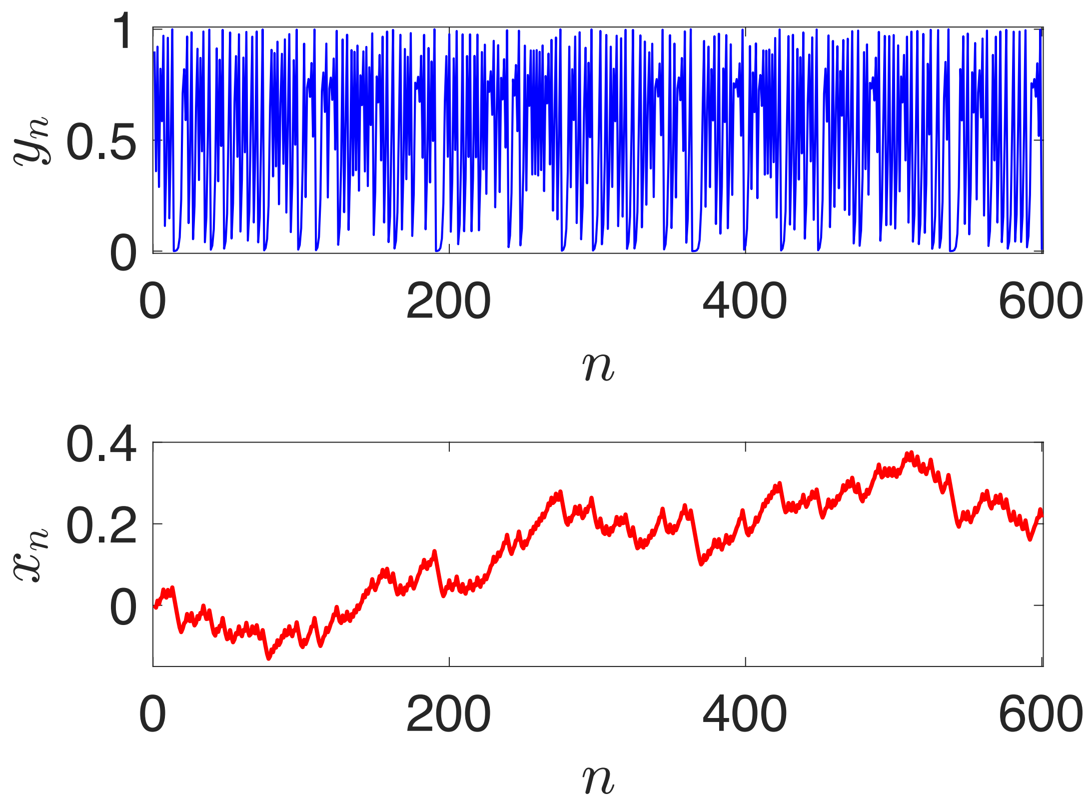
Homogenisation in action

$$x_{n+1} = x_n + \varepsilon(y_n - \tfrac{1}{2})$$
$$y_{n+1} = 4y_n(1 - y_n)$$

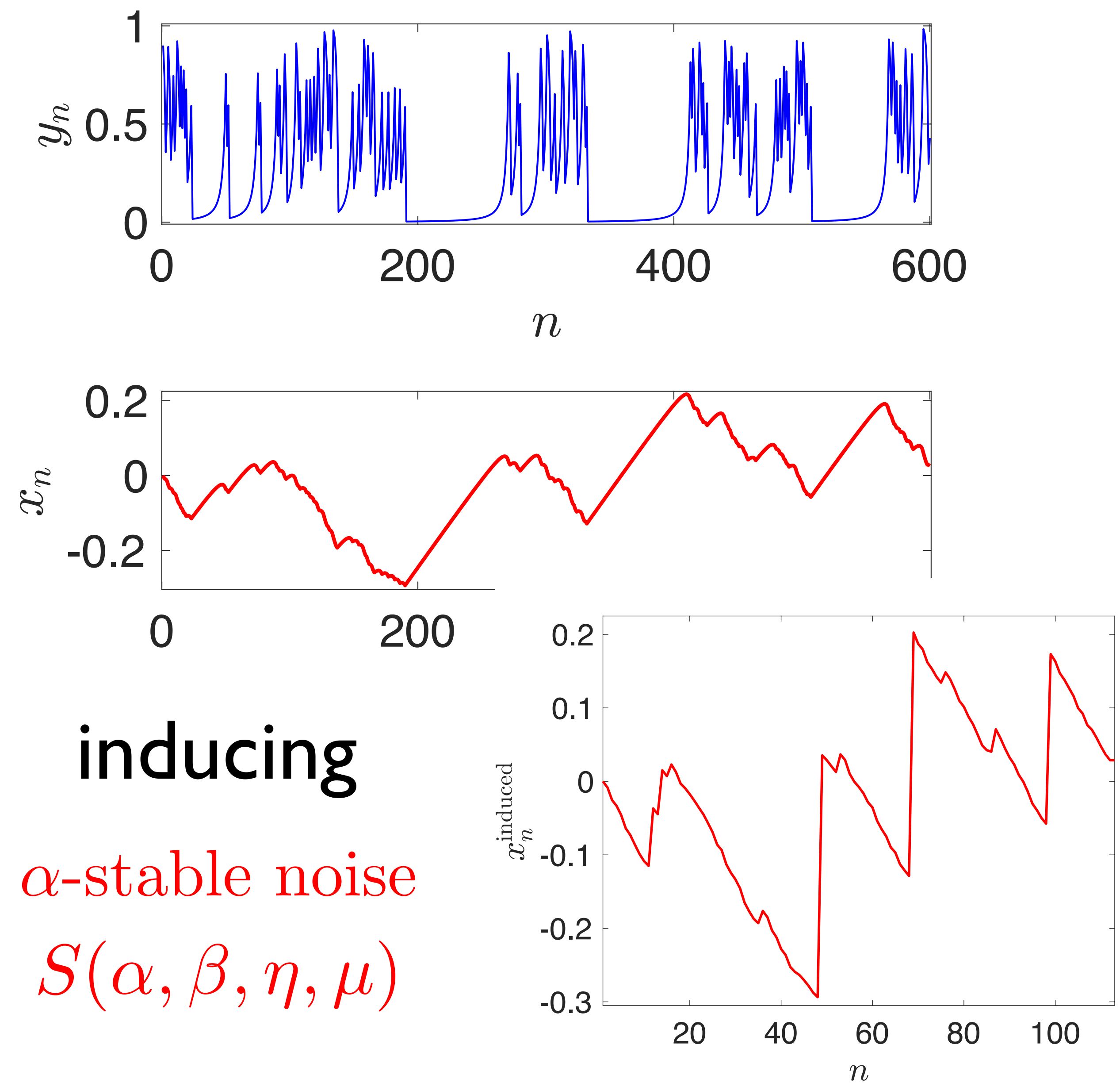
strong chaos

$$x_{n+1} = x_n + \varepsilon(y^\star - y_n)$$
$$y_{n+1} = \begin{cases} y_n(1 + 2^\gamma y_n^\gamma) & 0 \le y_n \le \frac{1}{2} \\ 2y_n - 1 & \frac{1}{2} \le y_n \le 1 \end{cases}$$

weak chaos



Brownian motion



inducing

α-stable noise

$S(\alpha, \beta, \eta, \mu)$

Homogenisation

resolved/slow: $dx = \frac{1}{\varepsilon} f_0(x,y)\,dt + f_1(x,y)\,dt$

unresolved/fast: $dy = \frac{1}{\varepsilon^2} g(x,y)\,dt + \frac{1}{\varepsilon}\sigma(x,y)\,dW_t$

Assumptions:

- fast y-process is ergodic with measure μ_x (mild chaoticity assumptions)
- $\int f_0(x,y) d\mu_x = 0$

Then, in the limit of $\varepsilon \to 0$, the statistics of the slow x-dynamics is approximated by

$$dX = F(X)\,dt + \Sigma(X)\,dW_t$$

where the diffusion matrix is given by a Green-Kubo formula

$$\frac{1}{2}\Sigma\Sigma^T = \int_0^\infty C(s) ds$$

with the auto-correlation matrix $C(t) = \mathbb{E}^{\mu_x}[f_0(x,y) f_0(x,y(t))]$ and

$$F(X) = \int f_1(x,y)\,d\mu_x + \int_0^\infty \int \nabla_x f_0(x,y(s)) \otimes f_0(x,y) d\mu_x \; ds$$

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

  skew product structure

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$

$$\dot{y} = \frac{1}{\varepsilon^2} g_0(\cancel{x}, y)$$

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

  skew product structure

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$

$$\dot{y} = \frac{1}{\varepsilon^2} g_0(\cancel{x}, y)$$

1. $f_0 = f_0(y)$ $\longrightarrow$ additive noise $dX = F(X)dt + \sigma dW$

   Melbourne & Stuart (2011)

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

  skew product structure

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g_0(x, y)$$

1. $f_0 = f_0(y)$ → additive noise $dX = F(X)dt + \sigma dW$

   Melbourne & Stuart (2011)

2. $f_0 = f_0(x, y)$ → multiplicative noise $dX = \tilde{F}(X)dt + \sigma(X)dW$

   GAG & Melbourne (2013)

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

  skew product structure

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g_0(\not{x}, y)$$

1. $f_0 = f_0(y)$ → additive noise $dX = F(X)dt + \sigma dW$

   Melbourne & Stuart (2011)

2. $f_0 = f_0(x, y)$ → multiplicative noise $dX = \tilde{F}(X)dt + \sigma(X)dW$

   GAG & Melbourne (2013)

Restrictions: $x \in \mathbb{R}^1$ or restrictive class of functions $f_0(x, y)$

What type of noise?
- ✱ strongly chaotic fast dynamics: Brownian noise
- ✱ weakly chaotic fast dynamics: α-stable noise

- ▷ continuous time: Stratonovich/Marcus (Wong-Zakai Theorem)
- ▷ discrete time: Ito or "neither" ← later

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

  skew product structure

$$\dot x = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot y = \frac{1}{\varepsilon^2} g_0(\cancel{x}, y)$$

1. $f_0 = f_0(y)$ → additive noise $dX = F(X)dt + \sigma dW$

   Melbourne & Stuart (2011)

3. $f_0 = f_0(x,y)$ → multiplicative noise $dX = \tilde F(X)dt + \sigma(X)dW$

   Melbourne & Kelly (2015)

No restriction on dimension of x

only the strongly chaotic case
leading to Stratonovich noise in line with the Wong-Zakai Theorem

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

  skew product structure

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$

$$\dot{y} = \frac{1}{\varepsilon^2} g_0(x, y)$$

back-coupling

What is known rigorously and what are the challenges?

- stochastic fast dynamics: *Khasminsky '66, Kurtz '73, Papanicolaou '76*
- deterministic fast dynamics:

skew product structure

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$

$$\dot{y} = \frac{1}{\varepsilon^2} g_0(x, y)$$

back-coupling

strongly chaotic case

weakly chaotic fast dynamics with $x \in \mathbb{R}^n$

allowing for multi-dimensional α-stable noise

Open problems from a modelling perspective

- slow dynamics couples back into the fast dynamics

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g_0(x,y)$$

What can go wrong?

If the fast invariant measure μ_x does not depend smoothly on x ("no linear response") even averaging does not "*work*"

$$F(X) = \underbrace{\int f_1(x,y)\mu_x(dy)}_{\text{non-Lipschitz}}$$

non-Lipschitz
uniqueness of solutions not guaranteed

Open problems from a modelling perspective

- slow dynamics couples back into the fast dynamics

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g_0(x,y)$$

What can go wrong?

If the fast invariant measure μ_x does not depend smoothly on x ("no linear response") even averaging does not "*work*"

$$F(X) = \underbrace{\int f_1(x,y)\mu_x(dy)}_{\text{non-Lipschitz}}$$

non-Lipschitz
uniqueness of solutions not guaranteed

How to detect failure of linear response in time series?
GAG, Wormell & Wouters (2016)

Open problems from a modelling perspective

- slow dynamics couples back into the fast dynamics
- finite time scale separation

Theory works in the limit $\varepsilon \to 0$
but in many physical applications ε is not so small

Where do we need the limit?

Averaging: Large deviation principle: $|\frac{1}{T}\int_0^T f_1(x, y(s))ds - F(x)|$

Homogenisation: Central Limit Theorem (Weak Invariance Principle)

$$W_\varepsilon(t) = \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f_0(y(s))ds \to_w W(t) \quad \text{as} \quad \varepsilon \to 0$$

Finite ε effects are finite size effects

The Central Limit Theorem and the Edgeworth expansion

The Central Limit Theorem

Assume X_i are *i.i.d.* random variables

$$S_n := \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{n}(X_j - \mu) \rightarrow_d \mathcal{N}(0,1)$$

where $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathbb{E}[X_i^2]$

For finite n there are *deviations* to the CLT

These are described by the **Edgeworth expansion**

$$\rho_n(x) = \Phi_{0,\sigma^2}(x) \times \left(1 + \frac{1}{6\sqrt{n}}\frac{\gamma}{\sigma^3}H_3(x/\sigma)\right) + o(\frac{1}{\sqrt{n}})$$

where $H_3(x) = x^3 - 3x$ is the third Hermite polynomial

and γ/σ^3 is the skewness of X_i

The Central Limit Theorem and the Edgeworth expansion

The Central Limit Theorem

Assume X_i are *i.i.d.* random variables

$$S_n := \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{n}(X_j - \mu) \to_d \mathcal{N}(0,1)$$

where $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathbb{E}[X_i^2]$

For finite n there are *deviations* to the CLT

These are described by the **Edgeworth expansion**

$$\rho_n(x) = \Phi_{0,\sigma^2}(x) \times \left(1 + \frac{1}{6\sqrt{n}}\frac{\gamma}{\sigma^3}H_3(x/\sigma)\right) + o(\frac{1}{\sqrt{n}})$$

where $H_3(x) = x^3 - 3x$ is the third Hermite polynomial

and γ/σ^3 is the skewness of X_i

this is not a density!

The Central Limit Theorem and the Edgeworth expansion

The Central Limit Theorem

Assume X_i are *i.i.d.* random variables

$$S_n := \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} (X_j - \mu) \to_d \mathcal{N}(0,1)$$

where $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathbb{E}[X_i^2]$

For finite n there are *deviations* to the CLT

These are described by the **Edgeworth expansion**

$$\rho_n(x) = \Phi_{0,\sigma^2}(x) \times \left(1 + \frac{1}{6\sqrt{n}} \frac{\gamma}{\sigma^3} H_3(x/\sigma)\right) + o\left(\frac{1}{\sqrt{n}}\right)$$

where $H_3(x) = x^3 - 3x$ is the third Hermite polynomial

and γ/σ^3 is the skewness of X_i

can be pushed to any order involving higher-order moments

The Central Limit Theorem and the Edgeworth expansion

The Central Limit Theorem

Assume X_i are stationary *weakly dependent* random variables

$$S_n := \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} (X_j - \mu) \to_d \mathcal{N}(0,1)$$

where $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathbb{E}[X_i^2] + 2\sum_{j=1}^{\infty} \mathbb{E}[X_1 X_{j+1}]$

For finite n there are *deviations* to the CLT

These are described by the **Edgeworth expansion**

$$\rho_n(x) = \Phi_{0,\sigma^2+\delta\sigma^2/n}(x) \times \left(1 + \frac{1}{\sqrt{n}} \delta\kappa H_3(x/\sigma)\right) + o(\frac{1}{\sqrt{n}})$$

where H_3 is the third Hermite polynomial and $\delta\sigma^2$ and $\delta\kappa$ are integrals of correlation functions of X_i *(Götze & Hipp (1983))*

Stochastic Parametrisation using the Edgeworth expansion

Given a multi-scale dynamical system

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

(I) determine the Edgeworth expansion coefficients $\sigma^2_{\rm GK},\ \delta\kappa$ associated with $f_0(x, y)$

Stochastic Parametrisation using the Edgeworth expansion

Given a multi-scale dynamical system

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

(I) determine the Edgeworth expansion coefficients $\sigma^2_{\mathrm{GK}},\ \delta\kappa$ associated with $f_0(x,y)$

(II) model the multi-scale system by the surrogate stochastic process

$$\dot{X} = \frac{1}{\varepsilon} A(\eta) + F(X)$$
$$d\eta = -\frac{1}{\varepsilon^2}\gamma\eta\, dt + \frac{1}{\sqrt{\varepsilon}} dW_t$$

with $A(\eta) = a\eta^2 + b\eta + c$

1d Ornstein-Uhlenbeck process

where the parameters $a,\ b,\ c,\ \gamma$ are determined such that the Edgeworth expansion coefficients associated with $A(\eta)$ match $\sigma^2_{\mathrm{GK}},\ \delta\kappa$

Stochastic Parametrisation using the Edgeworth expansion

Given a multi-scale dynamical system

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

(I) determine the Edgeworth expansion coefficients $\sigma^2_{\mathrm{GK}},\ \delta\kappa$ associated with $f_0(x,y)$

(II) model the multi-scale system by the surrogate stochastic process

$$\dot{X} = \frac{1}{\varepsilon} A(\eta) + F(X)$$
$$d\eta = -\frac{1}{\varepsilon^2}\gamma\eta\, dt + \frac{1}{\sqrt{\varepsilon}} dW_t$$

with $A(\eta) = a\eta^2 + b\eta + c$

1d Ornstein-Uhlenbeck process

where the parameters $a,\ b,\ c,\ \gamma$ are determined such that the Edgeworth expansion coefficients associated with $A(\eta)$ match $\sigma^2_{\mathrm{GK}},\ \delta\kappa$

Remark: By construction the homogenized limit system of the original and the surrogate system are the same!

How to calculate the Edgeworth coefficients?

The three time scales of multi-scale systems

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

How to calculate the Edgeworth coefficients?

The three time scales of multi-scale systems

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

nontrivial fast dynamics

trivial slow dynamics $x(t) = x_0$

How to calculate the Edgeworth coefficients?

The three time scales of multi-scale systems

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

nontrivial fast dynamics

trivial slow dynamics $x(t) = x_0$

fast dynamics has equilibrated

trivial slow dynamics $x(t) = x_0$

How to calculate the Edgeworth coefficients?

The three time scales of multi-scale systems

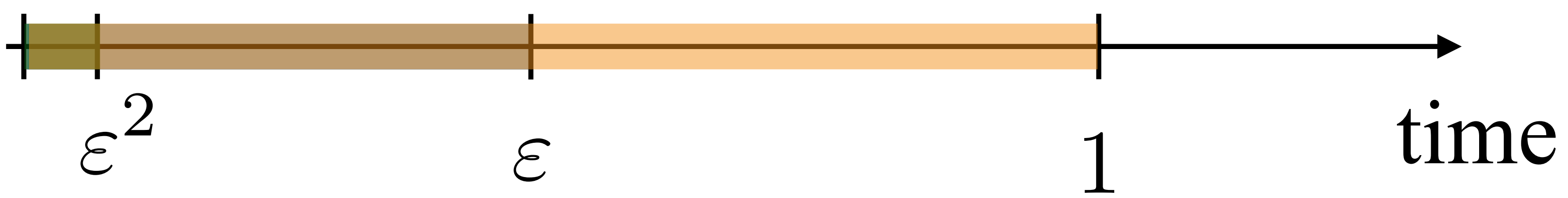
$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

nontrivial fast dynamics

trivial slow dynamics $x(t) = x_0$

diffusive time scale: CLT

$$dX = F(X)\,dt + \sigma(X) \circ dW_t$$



fast dynamics has equilibrated

trivial slow dynamics $x(t) = x_0$

How to calculate the Edgeworth coefficients?

The three time scales of multi-scale systems

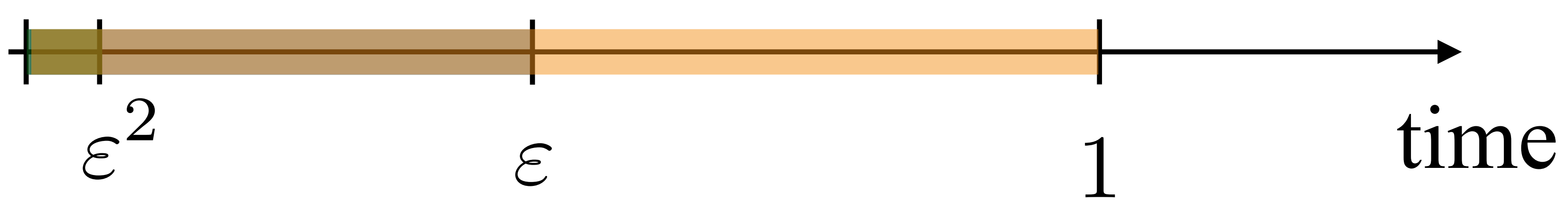
$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$
$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

nontrivial fast dynamics

trivial slow dynamics $x(t) = x_0$

diffusive time scale: CLT

$$dX = F(X)\,dt + \sigma(X) \circ dW_t$$



fast dynamics has equilibrated

trivial slow dynamics $x(t) = x_0$

expect deviations of CLT on timescale $t = \varepsilon$

$$\frac{x(t) - x_0}{\sqrt{t}} \to \sigma(x_0) W_t$$

How to calculate the Edgeworth coefficients?

Consider $\rho_t(x(t)|x(0) = x_0) = \int dxdy\, e^{\mathcal{L}t} \delta_{x_0}(x)\mu(dy)$ for $t = \varepsilon$

transfer operator

$$\dot{x} = \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y)$$

$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

$$\mathcal{L} = \frac{1}{\varepsilon^2}\mathcal{L}_0 + \frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_2$$

$$\mathcal{L}_0\rho = -\partial_y\left(g(y)\rho\right),\ \mathcal{L}_1\rho = -\,\partial_x\left(f_0(x,y)\rho\right),\ \mathcal{L}_2\rho = -\partial_x\left(f_1(x,y)\rho\right)$$

How to calculate the Edgeworth coefficients?

Consider $\rho_t(x(t)|x(0) = x_0) = \int dxdy\, e^{\mathcal{L}t} \delta_{x_0}(x)\mu(dy)$ for $t = \varepsilon$

transfer operator

$$\dot{x} = \frac{1}{\varepsilon} f_0(x, y) + f_1(x, y)$$

$$\dot{y} = \frac{1}{\varepsilon^2} g(y)$$

$$\mathcal{L} = \frac{1}{\varepsilon^2} \mathcal{L}_0 + \frac{1}{\varepsilon} \mathcal{L}_1 + \mathcal{L}_2$$

$$\mathcal{L}_0 \rho = -\partial_y \left(g(y)\rho \right), \ \mathcal{L}_1 \rho = -\partial_x \left(f_0(x, y)\rho \right), \ \mathcal{L}_2 \rho = -\partial_x \left(f_1(x, y)\rho \right)$$

Calculate asymptotically, using successive applications of the Duhamel-Dyson formula, up to $\mathcal{O}(\varepsilon^n)$:

$$\frac{\mathbb{E}[x(\varepsilon) - x_0]}{\sqrt{\varepsilon}} = \sqrt{\epsilon}\, \xi = \sqrt{\epsilon} \langle f_1(x_0) \rangle$$

$$\frac{\mathbb{E}[\hat{x}^2]}{\varepsilon} = \sigma_{\text{GK}}^2 - 2\varepsilon \int_0^{\frac{t}{\varepsilon^2}} ds \left(s \langle f_0 e^{\mathcal{L}_0 s} f_0 \rangle - \langle f_0 e^{\mathcal{L}_0 s} f_1 \rangle \right) + \cdots$$

$\hat{x} = x - \mathbb{E}[x]$

$$\frac{\mathbb{E}[\hat{x}^3]}{\varepsilon^{\frac{3}{2}}} = \sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon^2}} ds_1\, ds_2 \, \langle f_0 e^{\mathcal{L}_0 s_1} f_0 e^{\mathcal{L}_0 s_2} f_0 \rangle$$

Example I Diffusive limit of a deterministic multi-scale system

$$x_{j+1}^{(\varepsilon)} = x_j^{(\varepsilon)} + \varepsilon f_0(y_j) + \varepsilon^2 f_1(x_j^{(\varepsilon)})$$
$$y_{j+1} = p\, y_j \pmod 1$$

Homogenisation

$$dX = f(X)dt + \sigma_{\mathrm{GK}}\, dW$$

GAG & Melbourne (2013)

Edgeworth expansion

$$X_{j+1}^{(\varepsilon)} = X_j^{(\varepsilon)} + \varepsilon A(\eta_j) + \varepsilon^2 f_1(X_j^{(\varepsilon)})$$
$$A(\eta) = a_s \eta^2 + b_s \eta + c_s$$
$$\eta_{j+1} = \phi \eta_j + N_j \qquad N_j \sim \mathcal{N}(0,1)$$
$$\sigma_{\mathrm{GK}}^2 \text{ and } \delta\kappa_3$$

Example I Diffusive limit of a deterministic multi-scale system

$$x_{j+1}^{(\varepsilon)} = x_j^{(\varepsilon)} + \varepsilon f_0(y_j) + \varepsilon^2 f_1(x_j^{(\varepsilon)})$$
$$y_{j+1} = p\, y_j \pmod 1$$

Homogenisation

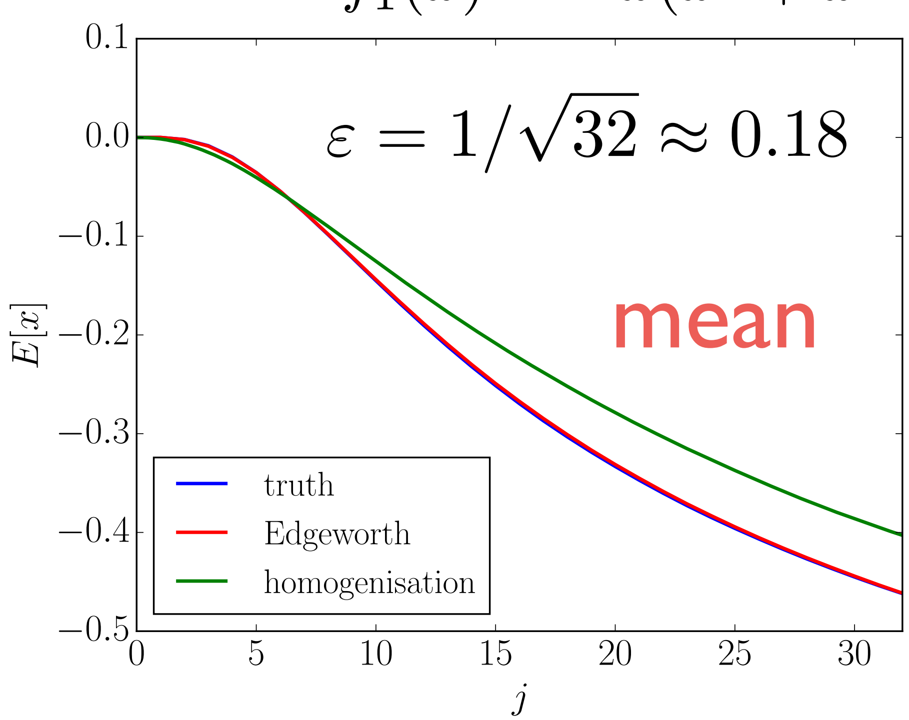
$$dX = f(X)dt + \sigma_{\rm GK}\, dW$$

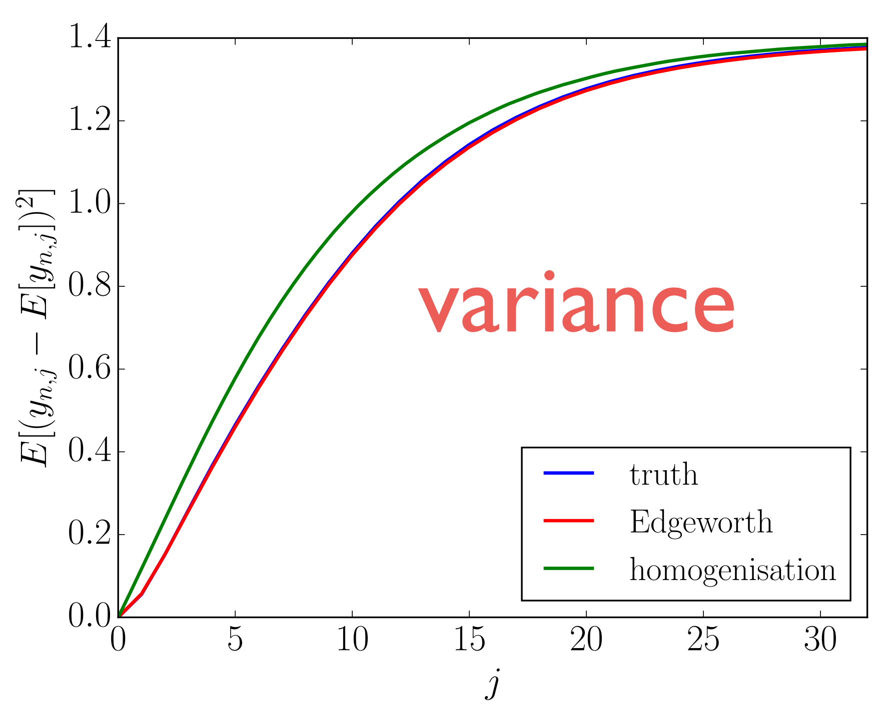
GAG & Melbourne (2013)

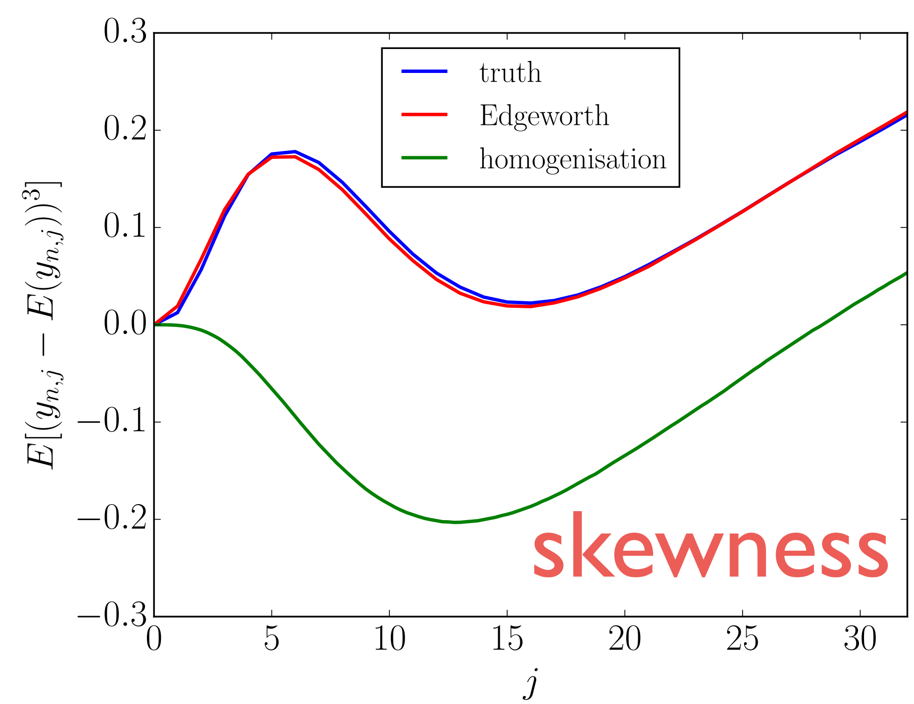
Edgeworth expansion

$$X_{j+1}^{(\varepsilon)} = X_j^{(\varepsilon)} + \varepsilon A(\eta_j) + \varepsilon^2 f_1(X_j^{(\varepsilon)})$$
$$A(\eta) = a_s\eta^2 + b_s\eta + c_s$$
$$\eta_{j+1} = \phi\eta_j + N_j \qquad N_j \sim \mathcal{N}(0,1)$$
$$\sigma^2_{\rm GK} \text{ and } \delta\kappa_3$$

$$p = 3 \qquad f_0(y) = y^5 + y^4 + y^3 + y^2 + y - \frac{29}{20}$$
$$f_1(x) = -x(x^2 + x - 1)$$







Example I Diffusive limit of a deterministic multi-scale system

$$x_{j+1}^{(\varepsilon)} = x_j^{(\varepsilon)} + \varepsilon f_0(y_j) + \varepsilon^2 f_1(x_j^{(\varepsilon)})$$
$$y_{j+1} = p\, y_j \pmod 1$$

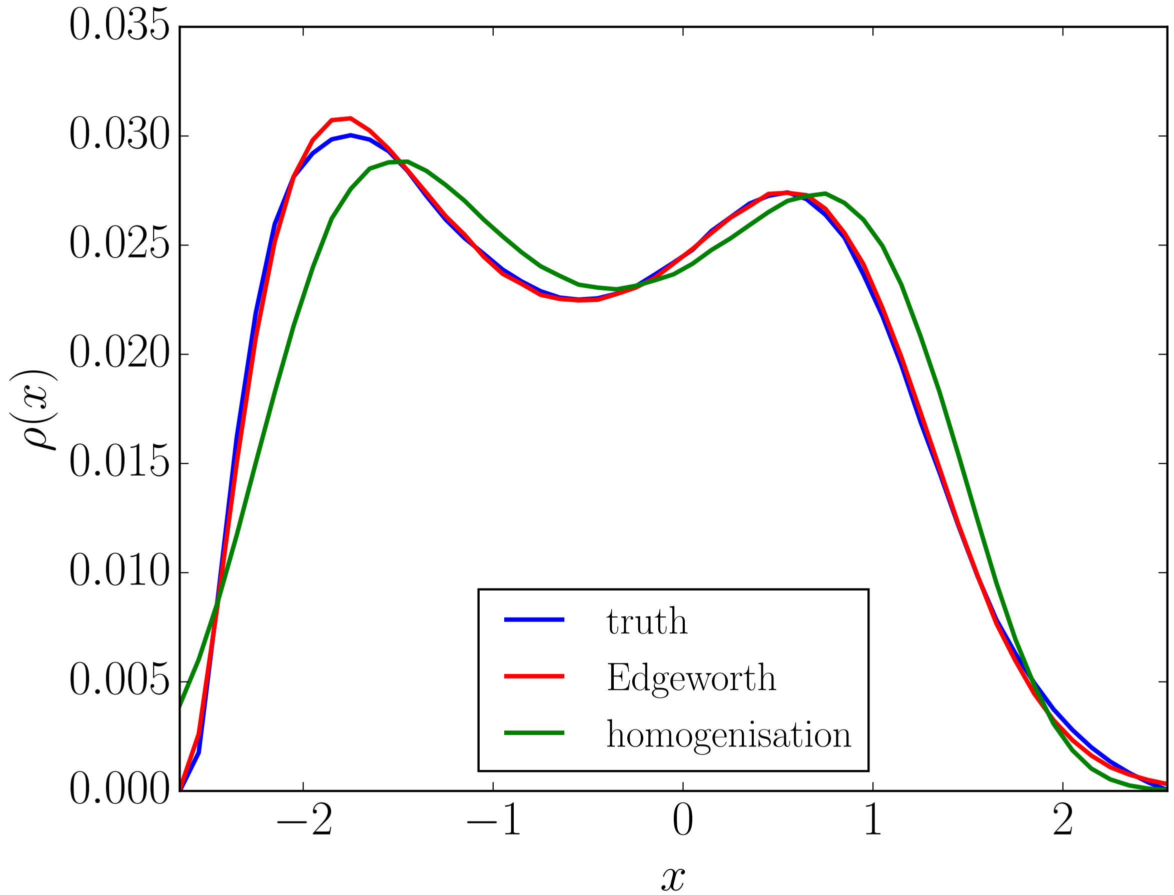
Homogenisation

$$dX = f(X)dt + \sigma_{\mathrm{GK}}\, dW$$

GAG & Melbourne (2013)

Edgeworth expansion

$$X_{j+1}^{(\varepsilon)} = X_j^{(\varepsilon)} + \varepsilon A(\eta_j) + \varepsilon^2 f_1(X_j^{(\varepsilon)})$$
$$A(\eta) = a_s \eta^2 + b_s \eta + c_s$$
$$\eta_{j+1} = \phi \eta_j + N_j \qquad N_j \sim \mathcal{N}(0,1)$$
$$\sigma_{\mathrm{GK}}^2 \text{ and } \delta\kappa_3$$



empirical density

Example II Diffusive limit of a triad system

$$\dot{x} = \frac{1}{\varepsilon} B_0 y_1 y_2$$
$$\dot{y}_1 = \frac{1}{\varepsilon} B_1 x y_2 - \frac{1}{\varepsilon^2}\gamma_1 y_1 - \frac{1}{\varepsilon}\sigma_1 \dot{W}_1$$
$$\dot{y}_2 = \frac{1}{\varepsilon} B_2 x y_1 - \frac{1}{\varepsilon^2}\gamma_2 y_2 - \frac{1}{\varepsilon}\sigma_2 \dot{W}_2$$

Triad system

Majda et al (2001)

backcoupling

$$\dot{X} = \frac{1}{\varepsilon} A(\eta)$$
$$\dot{\eta} = \frac{1}{\varepsilon}\alpha X - \frac{1}{\varepsilon^2}\eta - \frac{1}{\varepsilon}\sigma \dot{W}$$
$$A(\eta) = a_s \eta^2 + b_s \eta + c_s$$

Example II Diffusive limit of a triad system

Triad system

$$\sigma^2_{\mathrm{GK}} = \int_0^\infty C(\tau)d\tau = \frac{B_0^2\sigma^2_{1\infty}\sigma^2_{2\infty}}{\gamma_1+\gamma_2}$$

$$\delta\kappa_3 = 0$$

$$\mu = \int_0^\infty R(\tau)d\tau = x\frac{B_0}{\gamma_1+\gamma_2}(B_1\sigma^2_{2\infty} + B_2\sigma^2_{1\infty})$$

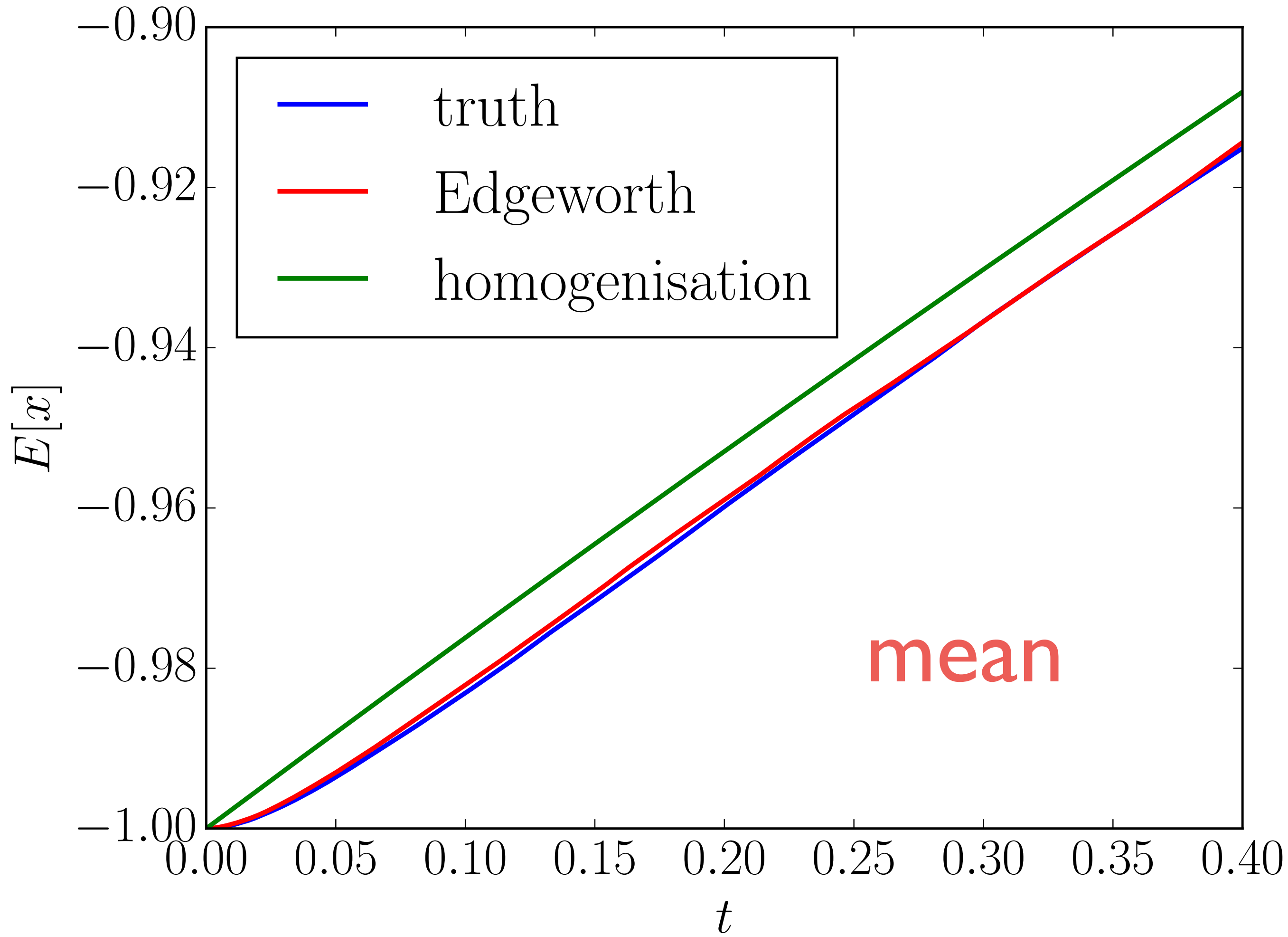
$$\delta\mu = \int_0^\infty \tau R(\tau)d\tau = x\frac{B_0}{(\gamma_1+\gamma_2)^2}(B_1\sigma^2_{2\infty} + B_2\sigma^2_{1\infty})$$

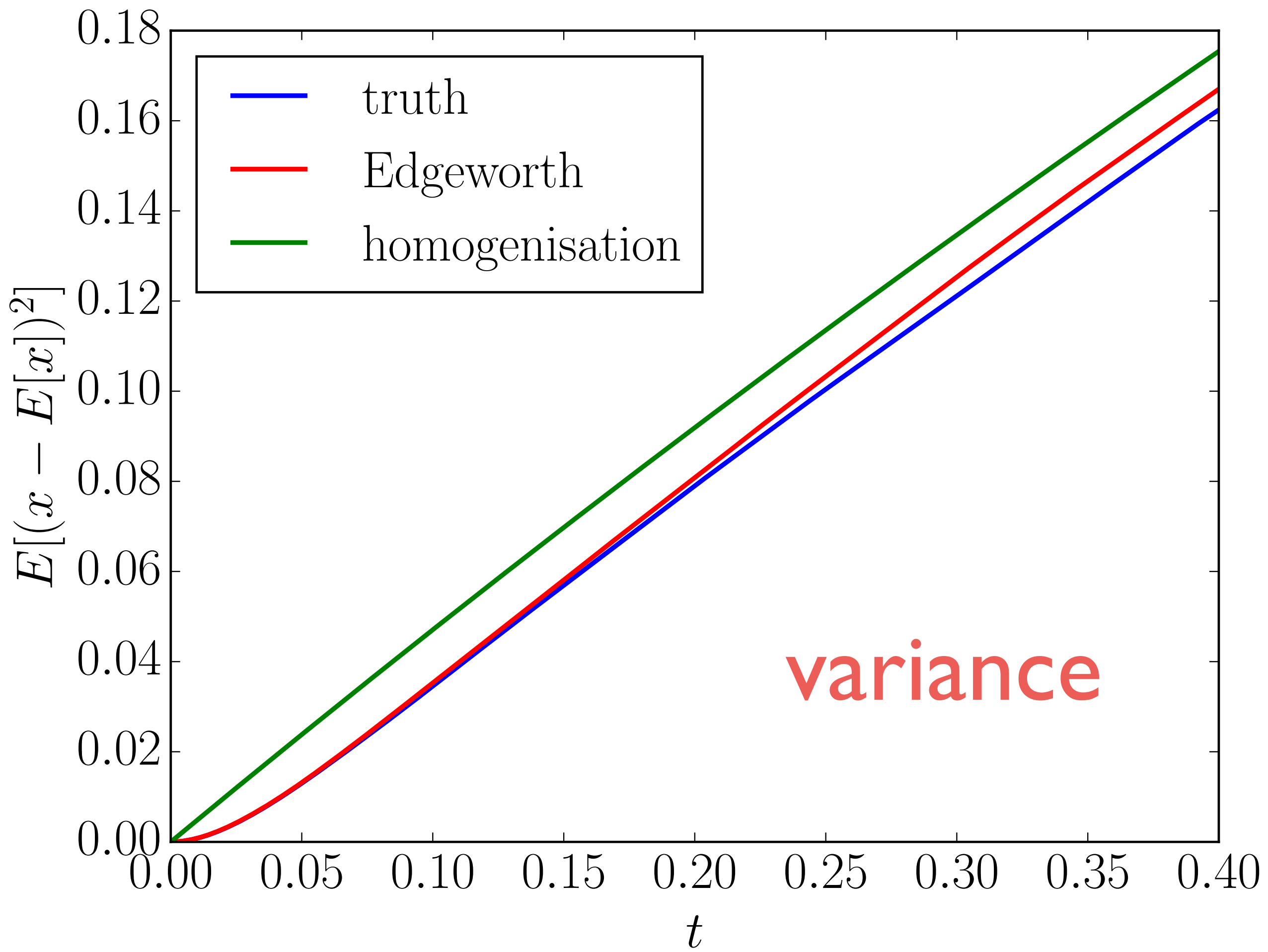
Surrogate system

$$\int_0^\infty C(\tau)d\tau = \frac{\sigma^2_\infty}{\gamma}(a^2\sigma^2_\infty + b^2)$$

$$\int_0^\infty R(\tau)d\tau = x\frac{\alpha b}{\gamma}$$

$$\int_0^\infty \tau R(\tau)d\tau = x\frac{\alpha b}{\gamma^2}$$





Statistical consistency of numerical integrators for deterministic multi-scale systems

Question:

How does the numerical time integrator affect the statistical behaviour of the simulation?

*E*xample: Influence of conservation laws on invariant measure in Hamiltonian systems
Dubinkina & Frank (2007)

Statistical consistency of numerical integrators for deterministic multi-scale systems

Question:

How does the numerical time integrator affect the statistical behaviour of the simulation?

*E*xample: Influence of conservation laws on invariant measure in Hamiltonian systems
Dubinkina & Frank (2007)

What are minimal requirements for a numerical scheme to assure that the statistics of the numerical simulations match those of the original continuous-time system?

Take Home Message:

Avoid first-order time-stepping when simulating deterministic multi-scale systems

Statistical consistency of numerical integrators for deterministic multi-scale systems

Question:

How does the numerical time integrator affect the statistical behaviour of the simulation?

*E*xample: Influence of conservation laws on invariant measure in Hamiltonian systems
Dubinkina & Frank (2007)

What are minimal requirements for a numerical scheme to assure that the statistics of the numerical simulations match those of the original continuous-time system?

Take Home Message:

Avoid first-order time-stepping when simulating deterministic multi-scale systems

Tools:

- Homogenisation
- Backward error analysis

Homogenisation

The statistical behaviour of deterministic multi-scale systems is well described by **homogenisation** (modulo Edgeworth corrections)

The statistical behaviour of the slow dynamics of the deterministic multi-scale system

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

$x \in \mathbb{R}^n$
$y \in \mathbb{R}^m,\ m \geq 3$

with chaotic fast dynamics and fast invariant measure μ, and $\int_\Lambda f_0\, d\mu = 0$

is (in the limit $\varepsilon \to 0$) described by the homogenised SDE

Melbourne & Stuart (2011)
GAG & Melbourne (2013)
Kelly & Melbourne (2015)

$$dX = F(X)\, dt + \sigma h(X) \circ dW_t$$

where

$$F(X) = \int_\Lambda f(X,y)\, d\mu$$

flow map of fast dynamics

$$\frac{1}{2}\sigma^2 = \int_0^\infty \mathbb{E}[f_0(y) f_0(\varphi^t y)]\, dt$$

Green-Kubo formula

Homogenisation

The statistical behaviour of deterministic multi-scale systems is well described by **homogenisation** (modulo Edgeworth corrections)

The statistical behaviour of the slow dynamics of the deterministic multi-scale system

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2 f(x,y)$$
$$\dot{y} = g(y)$$

$x \in \mathbb{R}^n$
$y \in \mathbb{R}^m,\ m \geq 3$

with chaotic fast dynamics and fast invariant measure μ, and $\int_\Lambda f_0\, d\mu = 0$

is (in the limit $\varepsilon \to 0$) described by the homogenised SDE

$$dX = F(X)\,dt + \sigma h(X) \circ dW_t$$

☆ noise is Stratonovich á la Wong-Zakai Theorem:
"approximate a rough noise by smooth functions"

$$W_\varepsilon(t) = \varepsilon \int_0^{\frac{t}{\varepsilon^2}} f_0(y(s))ds \longrightarrow_w W(t)$$

The forward Euler scheme for the slow variables of

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x, y)$$
$$\dot{y} = g(y)$$

is

$$x_{n+1} = x_n + \Delta t \varepsilon\, h(x_n) f_0(y_n) + \Delta t \varepsilon^2\, f(x_n, y_n)$$

The forward Euler scheme for the slow variables of

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

is

$$x_{n+1} = x_n + \Delta t \varepsilon\, h(x_n) f_0(y_n) + \Delta t \varepsilon^2\, f(x_n, y_n)$$

This map has the homogenised limit system

GAG & Melbourne (2013)

$$dX = \left(F(X) - \frac{1}{2} \Delta t\, h(X) h'(X)\, \mathbb{E}[f_0^2] \right) dt + \sqrt{\Delta t}\, \hat{\sigma} h(X) \circ d\tilde{W}_t$$

where

$$F(X) = \int_\Lambda f(X,y)\, d\mu$$

$$\hat{\sigma}^2 = \mathbb{E}[f_0^2] + 2 \sum_{n=1}^{\infty} \mathbb{E}[f_0(y) f_0(\Phi^n\, y)]$$

Discretisation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

$$x_{n+1} = x_n + \Delta t \varepsilon\, h(x_n) f_0(y_n) + \Delta t \varepsilon^2\, f(x_n, y_n)$$

Discretisation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2 f(x,y)$$
$$\dot{y} = g(y)$$

$$x_{n+1} = x_n + \Delta t \varepsilon\, h(x_n) f_0(y_n) + \Delta t \varepsilon^2 f(x_n, y_n)$$

Homogenisation

$$dX = F(X)\,dt + \sigma h(X) \circ dW_t$$
$$F(X) = \int_\Lambda f(X,y)\,d\mu$$
$$\frac{1}{2}\sigma^2 = \int_0^\infty \mathbb{E}[f_0(y) f_0(\varphi^t y)]\,dt$$

Homogenisation

$$dX = \Big(F(X) - \frac{1}{2}\Delta t\, h(X) h'(X)\, \mathbb{E}[f_0^2]\Big)\,dt + \sqrt{\Delta t}\,\hat{\sigma} h(X) \circ d\tilde{W}_t$$
$$F(X) = \int_\Lambda f(X,y)\,d\mu$$
$$\hat{\sigma}^2 = \mathbb{E}[f_0^2] + 2\sum_{n=1}^{\infty} \mathbb{E}[f_0(y) f_0(\Phi^n\, y)]$$

Discretisation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

$$x_{n+1} = x_n + \Delta t \varepsilon\, h(x_n) f_0(y_n) + \Delta t \varepsilon^2\, f(x_n, y_n)$$

Homogenisation

Homogenisation

$$dX = F(X)\, dt + \sigma h(X) \circ dW_t$$
$$F(X) = \int_\Lambda f(X,y)\, d\mu$$
$$\frac{1}{2}\sigma^2 = \int_0^\infty \mathbb{E}[f_0(y) f_0(\varphi^t y)]\, dt$$

$$dX = \left(F(X) - \frac{1}{2}\Delta t\, h(X) h'(X)\, \mathbb{E}[f_0^2]\right) dt + \sqrt{\Delta t}\, \hat{\sigma} h(X) \circ d\tilde{W}_t$$
$$F(X) = \int_\Lambda f(X,y)\, d\mu$$
$$\hat{\sigma}^2 = \mathbb{E}[f_0^2] + 2\sum_{n=1}^{\infty} \mathbb{E}[f_0(y) f_0(\Phi^n y)]$$

Remarks: $\hat{\sigma}^2\, \Delta t \to \sigma^2$ for $\Delta t \to 0$

Discretisation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

$$x_{n+1} = x_n + \Delta t \varepsilon\, h(x_n) f_0(y_n) + \Delta t \varepsilon^2\, f(x_n, y_n)$$

Homogenisation

Homogenisation

$$dX = F(X)\,dt + \sigma h(X) \circ dW_t$$
$$F(X) = \int_\Lambda f(X,y)\,d\mu$$
$$\frac{1}{2}\sigma^2 = \int_0^\infty \mathbb{E}[f_0(y) f_0(\varphi^t y)]\,dt$$

$$dX = \left(F(X) - \frac{1}{2}\Delta t\, h(X) h'(X)\, \mathbb{E}[f_0^2]\right) dt + \sqrt{\Delta t}\,\hat{\sigma} h(X) \circ d\tilde{W}_t$$
$$F(X) = \int_\Lambda f(X,y)\,d\mu$$
$$\hat{\sigma}^2 = \mathbb{E}[f_0^2] + 2\sum_{n=1}^{\infty} \mathbb{E}[f_0(y) f_0(\Phi^n\, y)]$$

Remarks: $\hat{\sigma}^2\,\Delta t \to \sigma^2$ for $\Delta t \to 0$

noise is neither Stratonovich nor Itô

$$\mathbf{E} := -\frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\Delta}\mathbf{t}\, \mathbf{h}(\mathbf{X})\mathbf{h}'(\mathbf{X})\, \mathbb{E}[\mathbf{f}_0^2]$$

for $i.i.d.$ fast dynamics, i.e. $\hat{\sigma}^2 = \mathbb{E}[f_0^2]$, the noise is Itô
(dynamics is already rough on time scale of $\mathcal{O}(\Delta t)$)

but it is never Stratonovich!

Question:

The only difference between the two homogenised equations is

$$\mathbf{E} := -\frac{1}{2}\Delta\mathbf{t}\,\mathbf{h}(\mathbf{X})\mathbf{h}'(\mathbf{X})\,\mathbb{E}[\mathbf{f_0^2}]$$

How can we interpret this extra drift term in the homogenised equation of the discretisation?

Can the extra term be significant? It is only $\mathcal{O}(\Delta t)$

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Solutions of the modified equation can be Taylor expanded as

$$z(t + \Delta t) = z(t) + \Delta t\, \tilde{b} + \frac{\Delta t^2}{2} D\tilde{b}\, \tilde{b} + \mathcal{O}(\Delta t^3)$$

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Solutions of the modified equation can be Taylor expanded as

$$\begin{aligned} z(t+\Delta t) &= z(t) + \Delta t\, \tilde{b} + \frac{\Delta t^2}{2} D\tilde{b}\, \tilde{b} + \mathcal{O}(\Delta t^3) \\ &= z(t) + \Delta t\, b + \Delta t^2 \left[b_1 + \frac{1}{2} Db\, b \right] + \mathcal{O}(\Delta t^3) \end{aligned}$$

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Solutions of the modified equation can be Taylor expanded as

$$\begin{aligned} z(t+\Delta t) &= z(t) + \Delta t\,\tilde{b} + \frac{\Delta t^2}{2} D\tilde{b}\,\tilde{b} + \mathcal{O}(\Delta t^3) \\ &= z(t) + \Delta t\, b + \Delta t^2 \left[b_1 + \frac{1}{2} Db\, b \right] + \mathcal{O}(\Delta t^3) \end{aligned}$$

Example 1: Forward Euler $z_{n+1} = z_n + \Delta t\, b(z_n)$

consistency up to $\mathcal{O}(\Delta t^2)$ → first-order scheme

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Solutions of the modified equation can be Taylor expanded as

$$\begin{aligned} z(t+\Delta t) &= z(t) + \Delta t\, \tilde{b} + \frac{\Delta t^2}{2} D\tilde{b}\, \tilde{b} + \mathcal{O}(\Delta t^3) \\ &= z(t) + \Delta t\, b + \Delta t^2 \left[b_1 + \frac{1}{2} Db\, b \right] + \mathcal{O}(\Delta t^3) \end{aligned}$$

Example 1: Forward Euler $z_{n+1} = z_n + \Delta t\, b(z_n)$

consistency up to $\mathcal{O}(\Delta t^2)$ → first-order scheme

However, it would be a second-order scheme for the modified equation

$$\dot{z} = b - \frac{\Delta t}{2} Db\, b$$

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Solutions of the modified equation can be Taylor expanded as

$$\begin{aligned} z(t+\Delta t) &= z(t) + \Delta t\, \tilde{b} + \frac{\Delta t^2}{2} D\tilde{b}\, \tilde{b} + \mathcal{O}(\Delta t^3) \\ &= z(t) + \Delta t\, b + \Delta t^2 \left[b_1 + \frac{1}{2} Db\, b \right] + \mathcal{O}(\Delta t^3) \end{aligned}$$

Example II: **Second-order Runge-Kutta method**

$$z_{n+1} = z_n + \frac{\Delta t}{2} \left[b(z_n) + b(z_n + \Delta t b(z_n)) \right]$$

Backward error analysis

Hairer, Lubich & Wanner

A numerical integrator $z_{n+1} = \Phi(z_n)$ does not approximate solutions of the original system $\dot{z} = b(z)$ but the solution of a so called *modified equation*

$$\dot{z} = \tilde{b}(z) = b(z) + \Delta t\, b_1(z) + \Delta t^2\, b_2(z) + \cdots$$

Solutions of the modified equation can be Taylor expanded as

$$\begin{aligned} z(t+\Delta t) &= z(t) + \Delta t\, \tilde{b} + \frac{\Delta t^2}{2} D\tilde{b}\, \tilde{b} + \mathcal{O}(\Delta t^3) \\ &= z(t) + \Delta t\, b + \Delta t^2 \left[b_1 + \frac{1}{2} Db\, b \right] + \mathcal{O}(\Delta t^3) \end{aligned}$$

Example II: **Second-order Runge-Kutta method**

$$\begin{aligned} z_{n+1} &= z_n + \frac{\Delta t}{2} \left[b(z_n) + b(z_n + \Delta t b(z_n)) \right] \\ &= z_n + \frac{\Delta t}{2} \left[b(z_n) + b(z_n) + \Delta t Db(z_n)\, b(z_n) + \mathcal{O}(\Delta t^2) \right] \end{aligned}$$

so it approximates the modified equation $\dot{z} = b(z) + \mathcal{O}(\Delta t^2)$ $\quad (b_1 \equiv 0)$

Backward error analysis

Back to our deterministic multi-scale system

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

The forward Euler discretisation of the slow dynamics yields as a modified equation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y) - \frac{\Delta t}{2}\left(\varepsilon^2 \partial_x h(x) h(x) f_0^2(y) + \varepsilon h(x) \partial_y f_0(y)\, g(y)\right) + \mathcal{O}(\varepsilon^3 \Delta t)$$

Backward error analysis

Back to our deterministic multi-scale system

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

The forward Euler discretisation of the slow dynamics yields as a modified equation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y) - \frac{\Delta t}{2}\left(\varepsilon^2 \partial_x h(x) h(x) f_0^2(y) + \varepsilon h(x) \partial_y f_0(y)\, g(y)\right) + \mathcal{O}(\varepsilon^3 \Delta t)$$

which has the same homogenisation limit as the forward Euler map ✓

Backward error analysis

Back to our deterministic multi-scale system

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y)$$
$$\dot{y} = g(y)$$

The forward Euler discretisation of the slow dynamics yields as a modified equation

$$\dot{x} = \varepsilon\, h(x) f_0(y) + \varepsilon^2\, f(x,y) - \frac{\Delta t}{2}\left(\varepsilon^2 \partial_x h(x) h(x) f_0^2(y) + \varepsilon h(x) \partial_y f_0(y)\, g(y)\right) + \mathcal{O}(\varepsilon^3 \Delta t)$$

which has the same homogenisation limit as the forward Euler map ✓

Remark: The additional drift term $E := -\frac{1}{2}\Delta t\, h(x)\partial_x h(x)\, f_0^2$ would be absent in a numerical scheme of at least second order

For a second-order time-stepping method the homogenized modified equation therefore agrees with the homogenized equation of the full multi-scale system up to $\mathcal{O}(\Delta t^3)$

Numerical results

$$\dot{x} = \varepsilon\sqrt{x}y + \varepsilon^2 b(c-x)y^2$$

$$\begin{cases} \dot{\xi} = -\eta - \zeta \\ \dot{\eta} = \xi + r\eta \\ \dot{\zeta} = s + (\xi - u)\zeta \end{cases}$$

$y = \eta + \zeta$

Rössler system

$r = s = 0.25$
$u = 7$

discretise → Second-order Runge-Kutta

discretise ↓ First-order forward Euler

Numerical results

$$\dot{x} = \varepsilon\sqrt{x}y + \varepsilon^2 b(c-x)y^2$$

$$\begin{cases} \dot{\xi} = -\eta - \zeta \\ \dot{\eta} = \xi + r\eta \\ \dot{\zeta} = s + (\xi - u)\zeta \end{cases}$$

$y = \eta + \zeta$

Rössler system

$r = s = 0.25$
$u = 7$

discretise → Second-order Runge-Kutta

homogenisation

$$\sigma^2 = 2\int_0^\infty \mathbb{E}[(\varphi^t y)y]\,dt$$

$$\beta = c + \frac{\sigma^2 a^2}{8\alpha b}$$

discretise → First-order forward Euler

Cox-Ingersoll-Ross model

$$dX = \sigma a\sqrt{X}\,dW + 2\alpha b(\beta - X)\,dt$$

$$\alpha = \frac{1}{2}\,\mathbb{E}[y^2]$$

Numerical results

$$\dot{x} = \varepsilon\sqrt{x}y + \varepsilon^2 b(c-x)y^2$$

$$y = \eta + \zeta$$

$$\begin{cases} \dot{\xi} = -\eta - \zeta \\ \dot{\eta} = \xi + r\eta \\ \dot{\zeta} = s + (\xi - u)\zeta \end{cases}$$

Rössler system

$r = s = 0.25$
$u = 7$

discretise → **Second-order Runge-Kutta**

homogenisation

$$\sigma^2 = 2\int_0^\infty \mathbb{E}[(\varphi^t y)y]\,dt$$

$$\beta = c + \frac{\sigma^2 a^2}{8\alpha b}$$

discretise → **First-order forward Euler**

homogenisation

$$\hat{\sigma}^2 = \mathbb{E}[y^2] + 2\sum_{n=1}^{\infty}\mathbb{E}[(\Phi^n y)y]$$

$$\beta = c + \frac{\Delta t \hat{\sigma}^2 a^2}{8\alpha b} - \frac{a^2 \Delta t}{4b}$$

Cox-Ingersoll-Ross model

$$dX = \sigma a\sqrt{X}\,dW + 2\alpha b(\beta - X)\,dt$$

$$\alpha = \frac{1}{2}\mathbb{E}[y^2]$$

Numerical results

parameters for
continuous-time ODE

$$\sigma^2 = 2\int_0^\infty \mathbb{E}[(\varphi^t y)y]\,dt$$

$$\beta = c + \frac{\sigma^2 a^2}{8\alpha b}$$

parameters for
discrete-time map

$$\hat\sigma^2 = \mathbb{E}[y^2] + 2\sum_{n=1}^{\infty} \mathbb{E}[(\Phi^n\, y)y]$$

$$\beta = c + \frac{\Delta t \hat\sigma^2 a^2}{8\alpha b} - \frac{a^2 \Delta t}{4b}$$

Cox-Ingersoll-Ross model

$$dX = \sigma a\sqrt{X}\,dW + 2\alpha b(\beta - X)\,dt$$

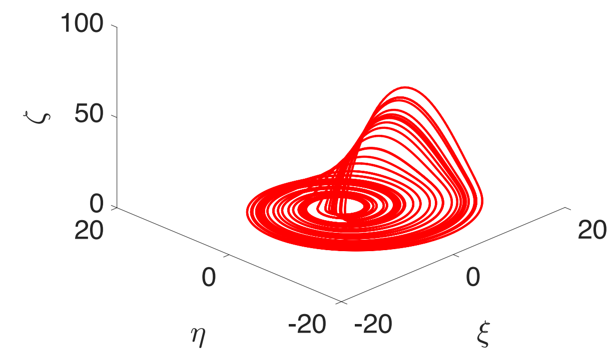
$$\alpha = \frac{1}{2}\,\mathbb{E}[y^2]$$

Numerical results

When is the difference significant?

$$a^2/b \gg 1$$

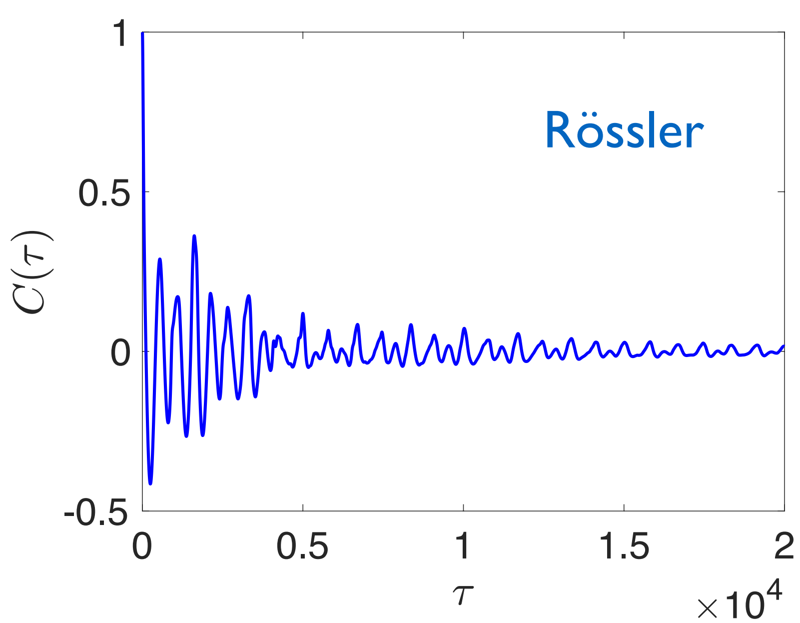
$$\sigma^2/4\alpha = \frac{\int_0^\infty \mathbb{E}[y(\varphi^t y)]\,dt}{\mathbb{E}[y^2]} \ll 1$$



parameters for
continuous-time ODE

$$\sigma^2 = 2\int_0^\infty \mathbb{E}[(\varphi^t y)y]\,dt$$

$$\beta = c + \frac{\sigma^2 a^2}{8\alpha b}$$



parameters for
discrete-time map

$$\hat{\sigma}^2 = \mathbb{E}[y^2] + 2\sum_{n=1}^{\infty} \mathbb{E}[(\Phi^n\, y)y]$$

$$\beta = c + \frac{\Delta t \hat{\sigma}^2 a^2}{8\alpha b} - \frac{a^2 \Delta t}{4b}$$

Cox-Ingersoll-Ross model

$$dX = \sigma a \sqrt{X}\,dW + 2\alpha b(\beta - X)\,dt$$

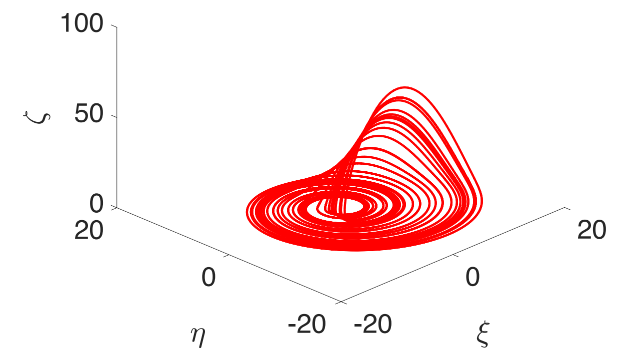
$$\alpha = \frac{1}{2}\mathbb{E}[y^2]$$

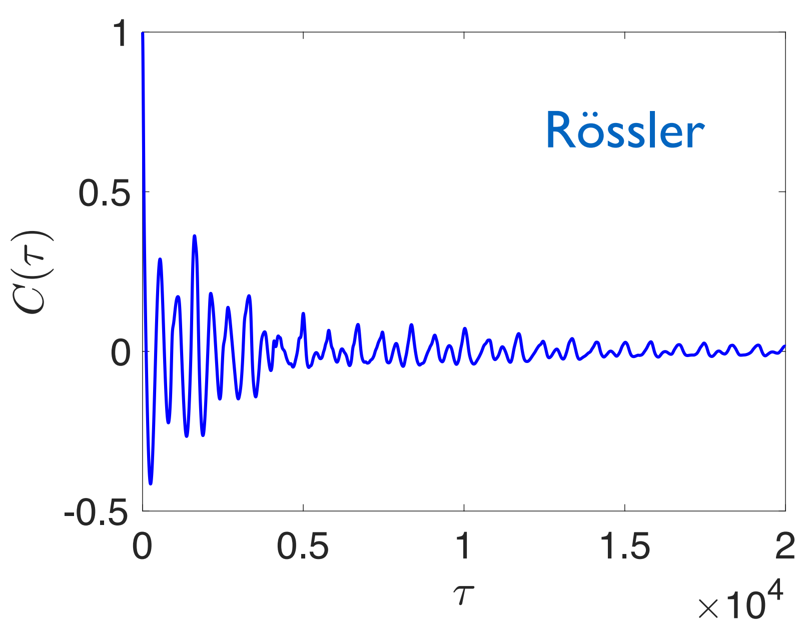
Numerical results

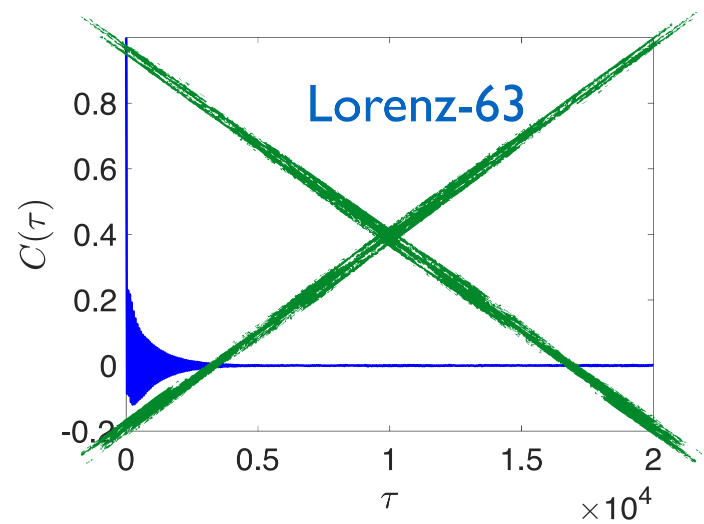
When is the difference significant?

$$a^2/b \gg 1$$

$$\sigma^2/4\alpha = \frac{\int_0^\infty \mathbb{E}[y(\varphi^t y)]\,dt}{\mathbb{E}[y^2]} \ll 1$$







parameters for continuous-time ODE

$$\sigma^2 = 2\int_0^\infty \mathbb{E}[(\varphi^t y)y]\,dt$$

$$\beta = c + \frac{\sigma^2 a^2}{8\alpha b}$$

parameters for discrete-time map

$$\hat{\sigma}^2 = \mathbb{E}[y^2] + 2\sum_{n=1}^{\infty}\mathbb{E}[(\Phi^n\, y)y]$$

$$\beta = c + \frac{\Delta t\hat{\sigma}^2 a^2}{8\alpha b} - \frac{a^2\Delta t}{4b}$$

Cox-Ingersoll-Ross model

$$dX = \sigma a\sqrt{X}\,dW + 2\alpha b(\beta - X)\,dt$$

$$\alpha = \frac{1}{2}\,\mathbb{E}[y^2]$$

The Cox-Ingersoll-Ross model has an exact solution

$$X(t) = c(t)H(t) \quad \text{with} \quad c(t) = \frac{\sigma^2}{4\alpha}(1 - e^{-\alpha t})$$

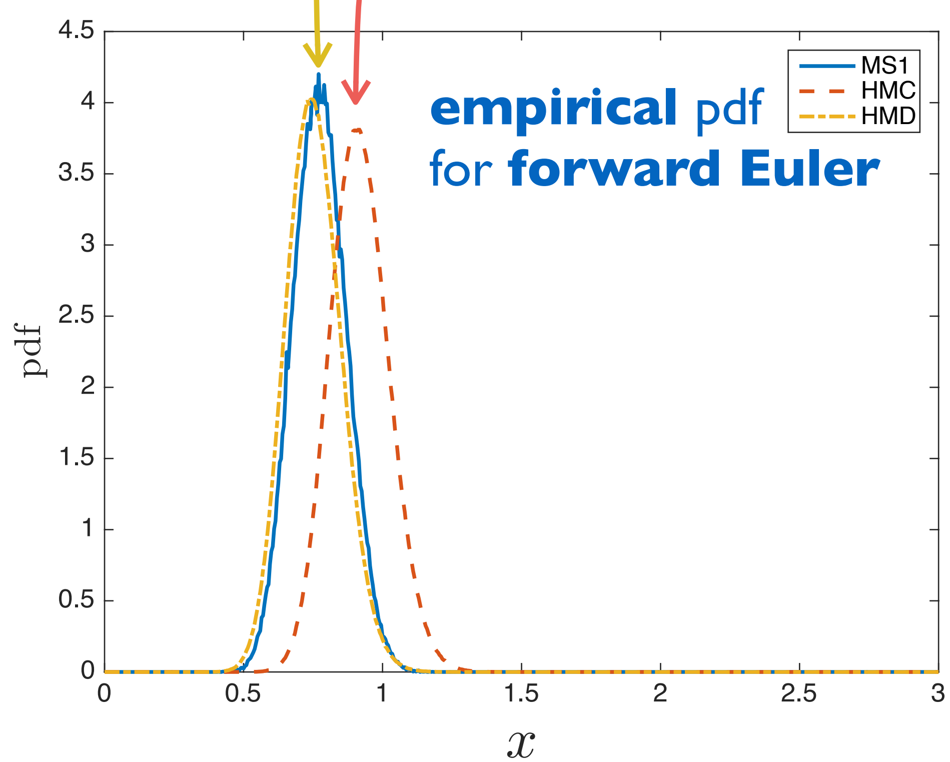
noncentral χ-squared distribution

$4\alpha\beta/\sigma^2$ degrees of freedom

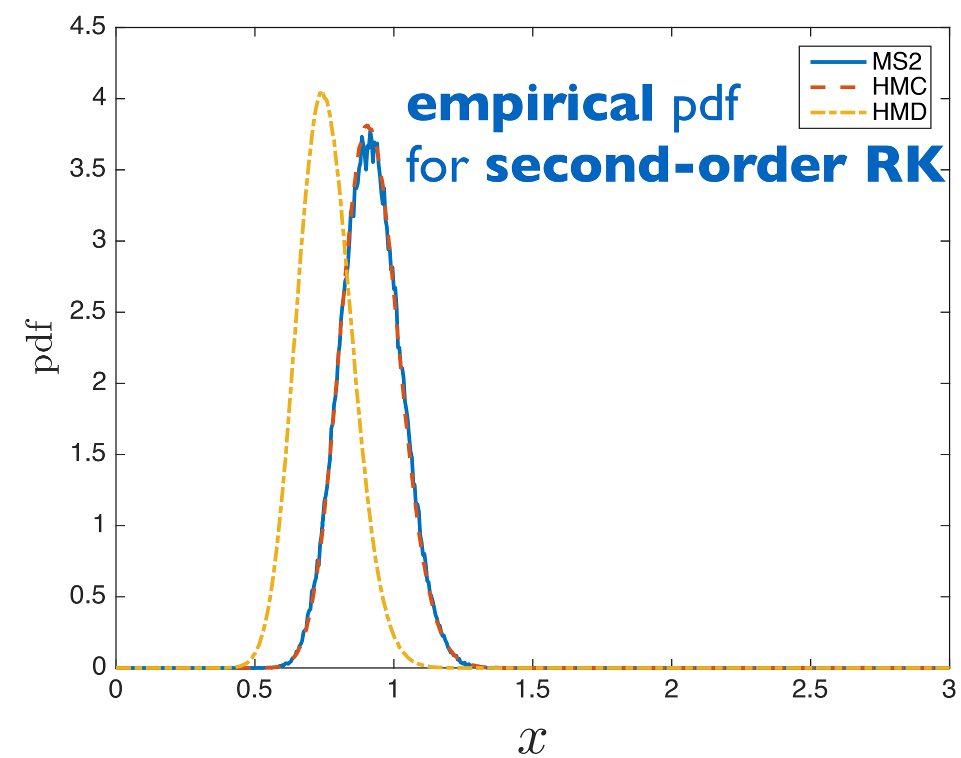
noncentrality parameter $c(t)^{-1}e^{-\alpha t}\xi$

analytical pdf of homogenised equation for full **discrete**-time multi-scale system

analytical pdf of homogenised equation for full **continuous**-time multi-scale system



$\varepsilon = 0.1$



15.6% error in mean!

Summary

We have used the Edgeworth expansion to push stochastic model reduction past the limit of infinite time scale separation, going beyond the Central Limit Theorem

We have developed a machinery to calculate the Edgeworth corrections for continuous time deterministic systems

The fast dynamics are replaced by a stochastic surrogate process, the parameters of which are tuned to match the Edgeworth expansion corrections of the full multi-scale system

Summary

We have used the Edgeworth expansion to push stochastic model reduction past the limit of infinite time scale separation, going beyond the Central Limit Theorem

We have developed a machinery to calculate the Edgeworth corrections for continuous time deterministic systems

The fast dynamics are replaced by a stochastic surrogate process, the parameters of which are tuned to match the Edgeworth expansion corrections of the full multi-scale system

Outlook:

- Use the strategy for the triad system to apply Edgeworth expansion to the barotropic vorticity equation
- Use Edgeworth expansions in a data-driven approach
- Prove the corrections rigorously (start with stochastic fast dynamics)

Summary

- We have resolved the discrepancy between the homogenized equations for a continuous-time fast-slow system and its first-order discretization using backward error analysis

Take Home Message:

Avoid first-order time-stepping when simulating deterministic multi-scale systems
In particular, when the system is far from *i.i.d.*