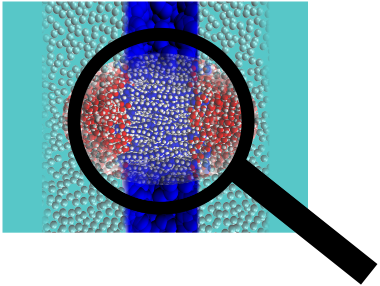
Coarse-graining of complex systems with parameter uncertainties

Carsten Hartmann (BTU Cottbus-Senftenberg),
with Hafida Bouanani (Saida) & Omar Kebiri (Cottbus)

SFB 1294 Data Assimilation, U Postdam

Parametric multiscale problems in science and engineering



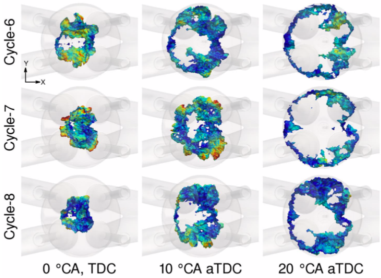
Modelling of lipid bilayers (source: Luigi Delle Site)

Atomistic-to-continuum models

- large biomolecular systems
- coarse grain models of materials

Stochastic turbulence models

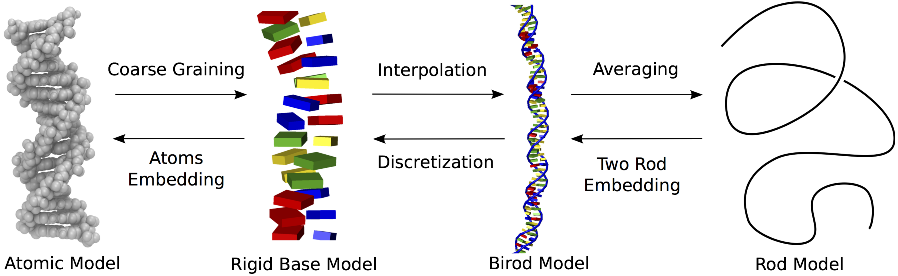
- atmosphere ocean interaction
- combustion models



Flame front propagation (source: Michael Oevermann)

Parametric multiscale problems in science and engineering, cont'd

- microscale model not fully parametrised (e.g. due to unobservable DOFs)
- coarse-graining technique may be sensitive to parameter uncertainties
- macroscale model may exhibit parameter-dependent regime changes



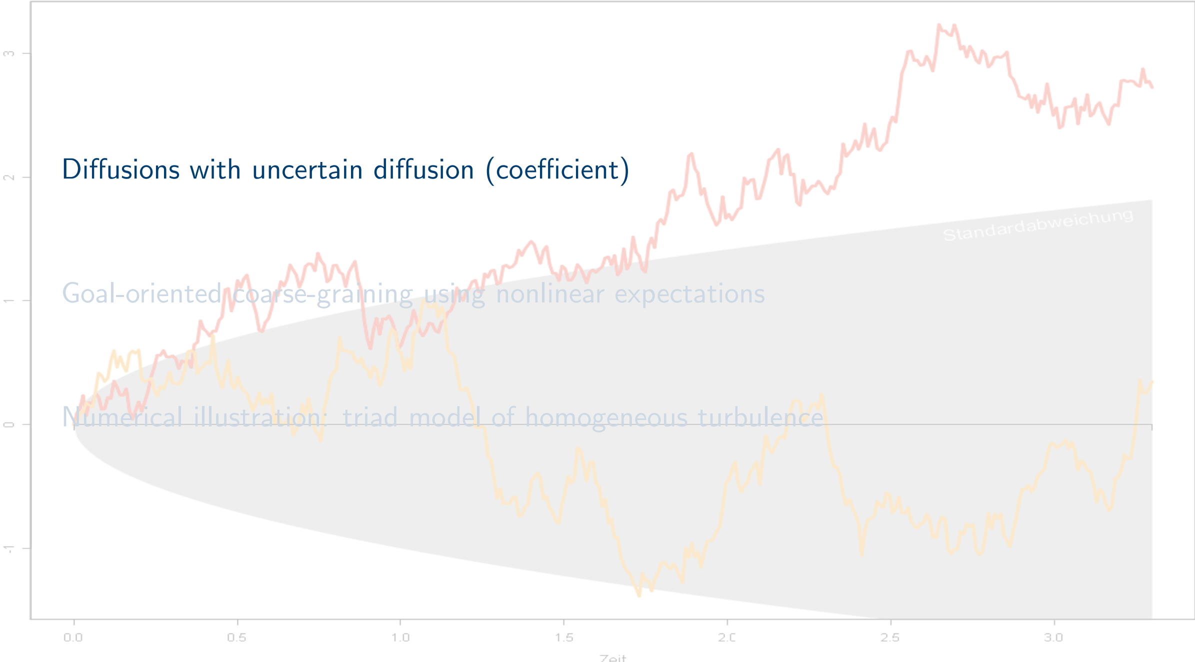
Sequence-dependent mechanical stiffness of coarse-grained B-DNA models (source: John Maddocks)

Outline

Diffusions with uncertain diffusion (coefficient)

Goal-oriented coarse-graining using nonlinear expectations

Numerical illustration: triad model of homogeneous turbulence



Diffusions with uncertain diffusion (coefficient)

Goal-oriented coarse-graining using nonlinear expectations

Numerical illustration: triad model of homogeneous turbulence

Illustrative example I: opinion dynamics

Cobb suggested the following Wright-Fisher-type diffusion model

$$dX_t = \alpha(\mu - X_t)dt + \theta\sqrt{X_t(1 - X_t)}\, dW_t\,, \quad X_0 = x$$

for the political opinion $X_t \in [0, 1]$ of an individual at time $t > 0$ where

- 0 and 1 represent the extreme ends of the political spectrum
- $\mu \in (0, 1)$ denotes the average ("mainstream") opinion
- $\alpha > 0$ is the rate at which the average opinion becomes mainstream, i.e.

$$\mathbb{E}[X_t] - \mu = (x - \mu)\exp(-\alpha t)$$

- $\theta > 0$ is the strength of the random perturbations around the average opinion.

[Cobb, Behav Sci, 1978]; cf. [Higham & Mantzaris, Chaos, 2020]

Illustrative example I, cont'd

Here is again the opinion dynamics model,

$$dX_t = \alpha(\mu - X_t)dt + \theta\sqrt{X_t(1 - X_t)}\, dW_t\,, \quad X_0 = x\,,$$

and the resulting ODE for the average opinion: $m' = \alpha(\mu - m)$.

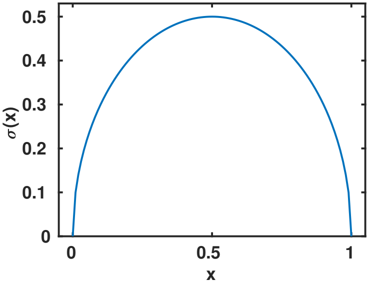
Observations:

- fluctuations are smaller for people holding extreme views, since

$$\sigma(x) = \sqrt{x(1 - x)}$$

  vanishes at $x = 0$ and $x = 1$

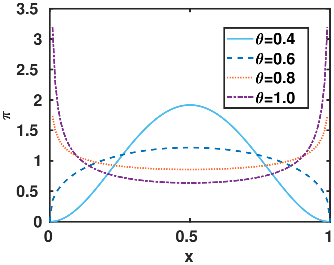
- there is no fixed point for $\epsilon > 0$



Illustrative example I, cont'd

Stationary long-term opinion distribution (approached exponentially fast)

$$\pi(x) = C_{\mu,\lambda} x^{\mu/\lambda - 1}(1-x)^{(1-\mu)/\lambda - 1} \quad (\lambda = \theta^2/\alpha)$$



Remark: Don't try to simulate the SDE with Euler-Maruyama.

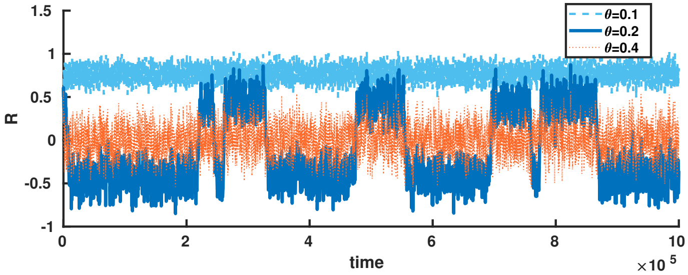
[Furioli et al, Ann Inst Henri Poincare C, 2019]; cf. [Jenkins & Sanõ, Ann Appl Probab, 2017]

Illustrative example II: slow-fast system

Now consider the 2-dimensional slow-fast system

$$dR_t = (R_t - U_t^3)dt\,, \quad R_0 = r$$
$$dU_t = \frac{1}{\epsilon}(R_t - U_t)dt + \sqrt{\frac{2\theta}{\epsilon}}dW_t\,, \quad U_0 = u\,.$$

with **unknown** $\theta \in [0, 1]$. Here are some realisations for $\epsilon = 0.01$:



Illustrative example II, cont'd

The fast dynamics conditional on $R_t = r$,

$$dU_t^r = (r - U_t^r)dt + \sqrt{2\theta}dW_t\,, \quad U_0 = u\,.$$

has the **unique invariant measure**

$$\mu_r = \begin{cases} \mathcal{N}(r,\theta)\,, & \theta \in (0,1] \\ \delta_r\,, & \theta = 0\,. \end{cases}$$

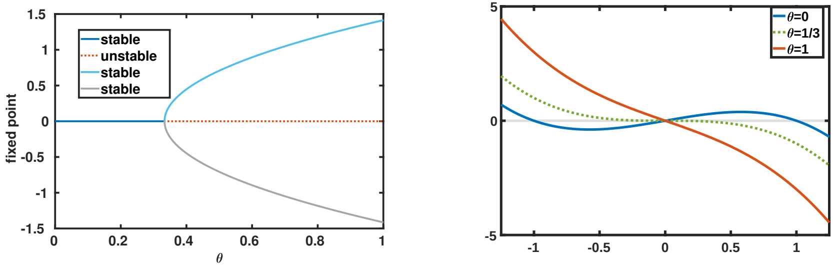
By the averaging principle, $\lim_{\epsilon \to 0} \mathbb{E}_\theta[\|R_t - \overline{R}_t\|_{[0,T]}] = 0$, where $\overline{R}$ solves the ODE

$$\frac{d\overline{R}}{dt} = -\overline{R}^3 + \overline{R}(1 - 3\theta)\,, \quad \overline{R}(0) = r\,.$$

[Freidlin & Wentzell, Springer, 1998], [Kifer, Cambridge University Press, 2004]

Illustrative example II, cont'd

For $0 \leq \theta \leq 1$ fixed, the averaged dynamics undergoes a supercritical pitchfork bifurcation at $\theta = 1/3$ (see green dotted curve of right panel):

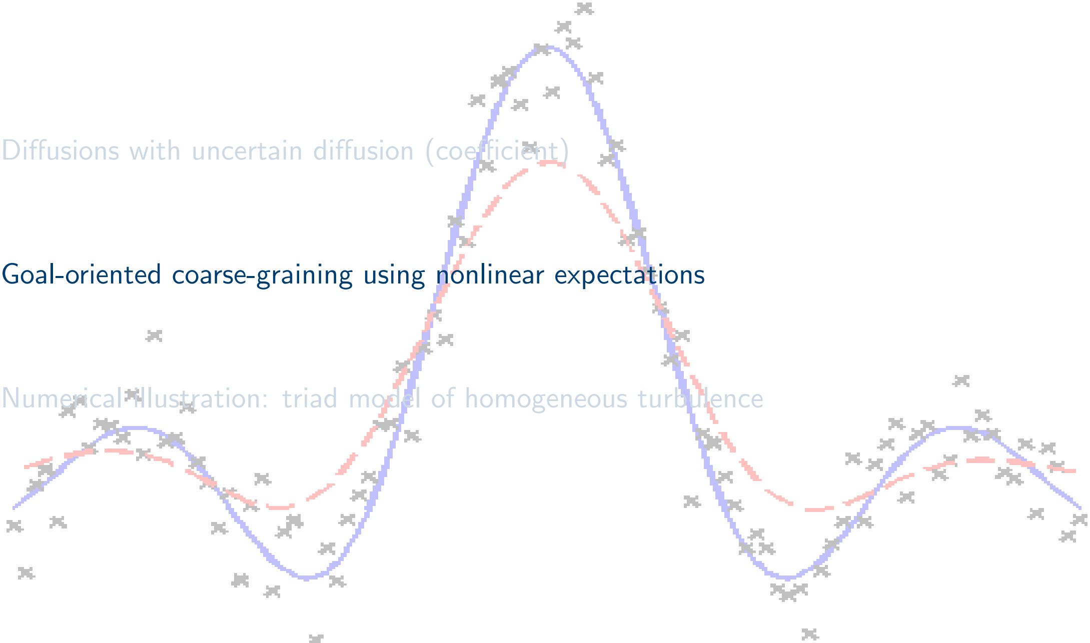


By construction, the averaged dynamics $\bar{R}$ is the **best approximation** of R as $\epsilon \to 0$. Even though the limit dynamics continuously depends on θ, the notion of a best approximation remains ambigous, since very little is known about θ.

Intermediate summary

- From the perspective of **SDE parameter estimation**, estimating diffusion coefficients is more difficult than estimating the drift.
- The **mean** of an SDE solution (or any stochastic process) may be misleading, and you better know the diffusion coefficient.
- Asymptotic properties (e.g. $t \to \infty$ or $\epsilon \to 0$) and the convergence rates can sensitively depend on unobservable and unknown parameters of the underlying dynamics; in some cases these **parameters may fluctuate** (e.g. in diffusing diffusivity or Heston-type models used in physics and finance).
- If macroscopic quantities (e.g. moments, etc.) depend on unknown parameters, it makes sense to consider a **worst-case scenario** for some functional $\varphi(\cdot)$, e.g.

$$\hat{\mathbb{E}}[|\varphi(R) - \varphi(\overline{R})|] := \sup\{\mathbb{E}_\theta[|\varphi(R) - \varphi(\overline{R})|] \colon 0 \le \theta \le 1\}$$



Diffusions with uncertain diffusion (coefficient)

Goal-oriented coarse-graining using nonlinear expectations

Numerical illustration: triad model of homogeneous turbulence

Related work (non-exhaustive)

Model reduction of parametric systems

- Parametric model order reduction (Gramians, moment-matching, . . .)

  Baur & Benner; Bond & Daniel; Drohmann, Haasdonk & Ohlberger; Rowley & Marsden; Willcox; ...

- Surrogate modelling (regression, interpolation, response surfaces, . . .)

  Box & Wilson; Draper; Constantine & Wang; Hardin & Sloane; Simpson & Mistree; ...

- Equation-free modelling (projective integration, gap-tooth schemes, . . .)

  Gear, Kevrekidis & Samaey; (Tony) Roberts; Dietrich, Reich & Kevrikidis et al.; ...

- Variational inference (relative entropy, Bayesian approaches, . . .)

  Katsoulakis & Plechac; Koumoutsakos; Majda & Harlim; Turkington; Reich & Opper; Freitag et al; ...

- Parameter estimation for multiscale systems (MLE, regression-based, . . .)

  Abdulle; Crommelin & Vanden-Eijnden; Krumscheid; Pavliotis & Stuart; Spiliopoulos; Timofeyev; ...

Set-up: parametric multiscale diffusions

We consider **slow-fast systems** of the form

$$
\begin{aligned}
dR_t &= \left(f_0(R_t, U_t) + \frac{1}{\sqrt{\epsilon}} f_1(R_t, U_t) \right) dt + \alpha(R_t, U_t) dW_t^r \\
dU_t &= \frac{1}{\epsilon} g(R_t, U_t; \theta) dt + \frac{1}{\sqrt{\epsilon}} \beta(R_t, U_t; \theta) dW_t^u ,
\end{aligned}
$$

with sufficiently smooth drift and diffusion coefficients, and

- $R = R^\epsilon \in \mathbb{R}^{n_s}$ slow variables (resolved)
- $U = U^\epsilon \in \mathbb{R}^{n_f}$ fast variables (unresolved)
- $\theta \in \Theta$ unknown parameter ($\Theta \subset \mathbb{R}^p$ compact)

Aim: Find a closed equation for the best approximation of R as $\epsilon \to 0$.

An worst-case averaging principle

Given the **slow-fast system**

$$dR_t = \left(f_0(R_t, U_t) + \frac{1}{\sqrt{\epsilon}} f_1(R_t, U_t) \right) dt + \alpha(R_t, U_t) dW_t^r$$
$$dU_t = \frac{1}{\epsilon} g(R_t, U_t; \theta) dt + \frac{1}{\sqrt{\epsilon}} \beta(R_t, U_t; \theta) dW_t^u \,,$$

for fixed $\theta \in \Theta$ and its **averaged (or: homogenised)** equation

$$d\overline{R} = b(\overline{R}_t; \theta)\, dt + \sigma(\overline{R}_t; \theta) dW_t^r$$

we say that some quantity of interest $\varphi = \varphi(R)$ converges to in the sense of **sublinear expectations** $\hat{\mathbb{E}}[\cdot] := \sup_\theta \mathbb{E}_\theta[\cdot]$ if

$$\lim_{\epsilon \to 0} \left| \hat{\mathbb{E}}[\varphi(R)] - \hat{\mathbb{E}}[\varphi(\overline{R})] \right| = 0$$

A short excurse . . .

Stochastic calculus under uncertainty: Peng's sublinear expectation

Let $\mathcal{H}$ be a space of random variables.

Definition (Peng's sublinear expectation)

The functional $\hat{\mathbb{E}}\colon \mathcal{H} \to \mathbb{R}$ is called a sublinear expectation if for all $X, Y \in \mathcal{H}$:

(a) $X \geq Y$ implies $\hat{\mathbb{E}}[X] \geq \hat{\mathbb{E}}[Y]$ (monotonicity)
(b) $\hat{\mathbb{E}}[c] = c$ for all constants $c \in \mathbb{R}$ (preservation of constants)
(c) $\hat{\mathbb{E}}[X + Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$ (sublinearity)
(d) $\hat{\mathbb{E}}[\lambda X] = \lambda \mathbb{E}[X]$ for all $\lambda > 0$ (positive homogeneity).

By properties (c)–(d), sublinear expectations are convex and admit the following **characterisation:** there exists a family $\{\mathbb{E}_\theta\colon \theta \in \Theta\}$, such that

$$\hat{\mathbb{E}}[X] = \sup\{\mathbb{E}_\theta[X]\colon \theta \in \Theta\}, \quad X \in \mathcal{H}.$$

[Peng, Springer, 2019]

Stochastic calculus under uncertainty: G-Brownian motion

Recall that any standard Gaussian random variable $X \sim \mathcal{N}(0,1)$ with expectation

$$\mathbb{E}[\varphi(X)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(x) e^{-\frac{x^2}{2}}\, dx\,, \quad \varphi \in \mathrm{Lip}(\mathbb{R})$$

for an arbitrary Lipschitz function φ can be characterised by $\mathbb{E}[\varphi(X)] = u(0,1)$, where $u = u(x,t)$ is the solution to the heat equation with initial condition $u(x,0) = \varphi(x)$.

We can characterise the G-**normal distribution** $\mathcal{N}(0,[\underline{\sigma}^2,\overline{\sigma}^2])$ by $\hat{\mathbb{E}}[\varphi(X)] = v(0,1)$, where v is the **viscosity solution to the G-heat equation**

$$\frac{\partial v}{\partial t} = G\left(\frac{\partial^2 v}{\partial x^2}\right), \quad v(x,0) = \varphi(x)\,,$$

with $G(w) = \dfrac{1}{2}\max\{cw : \underline{\sigma}^2 \le c \le \overline{\sigma}^2\}$.

[Peng, Stochastic Analysis and Applications, 2007]

G-Brownian motion, cont'd

The right hand side of the G-heat equation is the (nonlinear) infinitesimal generator of a stochastic process B, called **G-Brownian motion**, with the properties

$$\hat{\mathbb{E}}[B_t] = 0\,, \quad \hat{\mathbb{E}}[B_t^2] = \overline{\sigma}^2 \quad \text{and} \quad -\hat{\mathbb{E}}[-B_t^2] = \underline{\sigma}^2$$

for any $t > 0$.

Noting that the G-heat equation can be recast as a Hamilton-Jacobi-Bellman equation for a **diffusion-controlled SDE**, the G-Brownian motion B admits the representation

$$B_t = \int_0^t C_s\, dW_s$$

where W denotes standard Brownian motion and $C \in [\underline{\sigma}^2, \overline{\sigma}^2]$ is adapted.

Remark: The generalisation to the multidimensional case is straightforward.

[Peng, Stochastic Analysis and Applications, 2007]; [Denis & Hu, Potential Analysis, 2011]

Back to our problem . . .

Goal-oriented coarse-graining of multiscale systems

We consider the nonlinear HJB-type Kolmogorov backward equation

$$-\frac{\partial v^\epsilon}{\partial t} = \sup_{\theta\in\Theta}\left\{\frac{1}{2}a^\epsilon : \nabla^2 v^\epsilon + b^\epsilon \cdot \nabla v^\epsilon\right\} + \psi(x)\,, \quad v^\epsilon(T,x) = \phi(r)$$

associated with our **slow-fast system** that we compactly write as

$$dX_t = b^\epsilon(X_t;\theta)\,dt + \sigma^\epsilon(X_t;\theta)dW_t$$

with $X_t = (R_t, U_t)$ and $a^\epsilon = \sigma\sigma^T$.

As a our **quantity of interest (QoI)**, we consider the continuous functional

$$v^\epsilon(t,x) = \hat{\mathbb{E}}\left[\int_t^T \psi(R_s)\,ds + \phi(R_T)\middle| X_t = x\right].$$

[Mezdoud, H, Remita & Kebiri, arXiv:2108:06965, 2021]

A worst-case averaging principle, cont'd

Theorem (Bouanani, H & Kebiri, 2021)

Technical details aside, we have

$$v^\epsilon \to \bar{v}\,, \quad \nabla v^\epsilon \to \nabla \bar{v}$$

as $\epsilon \to 0$ where the convergence of v^ϵ is uniform on any compact subset of $[0, T] \times \mathbb{R}^n$ and pointwise for ∇v^ϵ for all $(t, x) \in [0, T] \times \mathbb{R}^n$, with $n = n_s + n_f$, and

$$\bar{v}(r, t) = \hat{\mathbb{E}}\left[\int_t^T \psi(\bar{R}_s)\, ds + \phi(\bar{R}_T) \middle| \bar{R}_t = r\right].$$

denotes the unique classical solution to the nonlinear Kolmogorov backward equation associated with the averaged (or: homogenised) equation.

[Bouanani, H & Kebiri, arXiv:2102.04908, 2021]; cf. [Kebiri, Neureither & H, Computation, 2018]

Some remarks

- The proof relies on a **G-FBSDE representation** of the backward equation.
- It uses a stability result of Zhang & Chen together with **Gronwall and BDG-type estimates for G-Brownian motion**, assuming uniform Lipschitz conditions for the drift coefficients and constant diffusion coefficients (i.e. additive noise).
- We have moreover proved convergence for an **optimally controlled SDE**, with a parameter uncertainty that is sitting only in the diffusion coefficient, while the control is acting on the drift part of the SDE, i.e. we consider

$$v^{\epsilon}(t,x) = \min_{\alpha \in \mathcal{A}} \hat{\mathbb{E}}\left[\int_t^T \psi(R_s, \alpha_t)\, ds + \phi(R_T) \middle| X_t = x\right].$$

  with α controlling the drift. If the control is acting on the uncertain parts of the SDE, extra saddle-point conditions are necessary.

[Cheridito, Soner, Touzi & Victoir, Commun Pure Appl Math, 2007], [Zhang & Chen, An St Univ Ovidius Constanta Ser Mat, 2011]

Diffusions with uncertain diffusion (coefficient)

Goal-oriented coarse-graining using nonlinear expectations

Numerical illustration: triad model of homogeneous turbulence

A simplified stochastic turbulence model

We consider the **bilinear triad interaction model**

$$dX(t) = \frac{1}{\epsilon^2} L(X(t))dt + \frac{1}{\epsilon} B(X(t), X(t))dt + \frac{1}{\epsilon}\Sigma\, dW_t\,, \quad X^{\epsilon}(0) = x,$$

where $X(t) = (R_1(t), R_2(t), U(t)) \in \mathbb{R}^3$ and

$$L(x) = -\begin{pmatrix} 0 \\ 0 \\ u \end{pmatrix}, \quad B(x,x) = \begin{pmatrix} A_1 r_2 u \\ A_2 r_1 u \\ A_3 r_1 r_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix},$$

Here $0 < \epsilon \ll 1$, the coefficients A_1, A_2, A_3 with $A_1 + A_2 + A_3 = 0$ are fixed, and

$$\lambda \in [\underline{\sigma}, \overline{\sigma}]$$

denotes the **unknown diffusion coefficient**.

[Majda, Timofeyev & Vanden-Eijnden, PNAS, 1999]

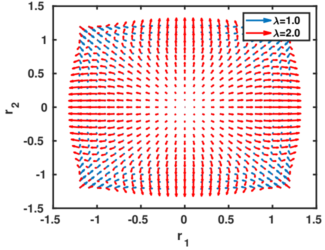
Coarse-grained triad interaction model

Homogenised limit equation

$$dR(t) = b(R(t);\lambda)dt + \sigma(R(t);\lambda)dW_t\,, \quad R(0) = r,$$

for climate variables $r = (r_1, r_2)$ as $\epsilon \to 0$, with

$$b = \begin{pmatrix} A_1 r_1 (A_3 r_2^2 + \frac{\lambda^2}{2} A_2) \\ A_2 r_2 (A_3 r_1^2 + \frac{\lambda^2}{2} A_1) \end{pmatrix}, \; \sigma = \frac{\lambda}{\gamma} \begin{pmatrix} A_1 r_2 \\ A_2 r_1 \end{pmatrix}.$$



Limit vector field b for $A_1 = A_2 = 1$ and $A_3 = -2$ and different noise parameters λ (invariant manifolds: hyperbola).

Using Itô's formula, it readily follows that

$$I(r_1, r_2) = A_1 r_2^2 - A_2 r_1^2$$

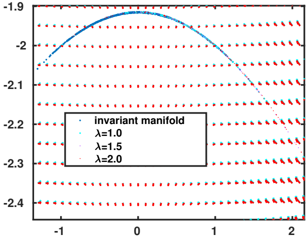
is a **conserved quantity** for both the reduced and the original system.

Qualitative behaviour of invariant manifolds and equilibria

Equilbria: The origin is an unstable hyperbolic equilibrium. The rays that connect the origin with any of the four equilibria

$$r^*_{\pm,\pm} = \left(\pm\lambda\sqrt{\frac{A_1}{2|A_3|}}, \pm\lambda\sqrt{\frac{A_2}{2|A_3|}}\right), \quad A_1,\ A_2 > 0\,.$$

are (locally hyperbolically unstable) invariant sets.



Independent realisations for different $\lambda \in [1,\ 2]$ and fixed $T = 0.5$, all starting from the same initial value $r = (1, -2)$.

Observation: The invariant manifolds are independent of λ, but the dynamics *on* the invariant manifolds changes as λ is varied.

QoI and worst-case parameter

Quantity of interest (QoI):

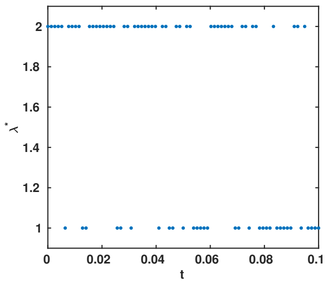
$$v^\epsilon = \hat{\mathbb{E}}[X_1(T)|X(t) = x]\,, \;\; \bar{v} = \hat{\mathbb{E}}[R_1(T)|R(t) = r]$$

Example 1: $A_1 = A_2 = 1$, $A_3 = -2$, $\lambda \in [0.8,\ 1.2]$, $T = 0.1$, $\epsilon = 0.2$, $x = (r, u) = (1, -2, -2)^T$:

$$v^\epsilon(0, x) = 0.9291\,,\;\; v(0, r) = 0.9326$$

Example 2: $A_1 = 1, A_2 = 2, A_3 = -3$, $\lambda \in [0.6,\ 1.2]$, $T = 0.5$, ϵ and x as before:

$$v^\epsilon(0, x) = 1.3202\,,\;\; v(0, r) = 1.3549$$



Parameter λ^* that maximises the nonlinear generator G over the noise coefficient $\lambda \in [1, 2]$.

[Beck, E & Jentzen, J Nonl Sci, 2019]

Outlook: variational parameter estimation

Given an observation time series $\hat{R} = \{\hat{R}_t : t \in [0, T]\}$, we may use the stochastic control ansatz to minimise a **tracking-type functional**

$$J(\theta; r, u, t) = \mathbb{E}\left[\int_0^T |R_s - \hat{R}_s|^2 + |\theta_s|^2 ds \,\middle|\, R_t = r,\ U_t = u\right]$$

over the unknown θ. Ideally there will be a unique **value function**

$$v(r, u, t) = \inf_\theta J(\theta; r, u, t)$$

solving an HJB equation, with θ becoming a **time-dependent feedback control**

$$\theta_s = c(X_s, Y_s, s)$$

that can be expressed in terms of v (starting point for systematic approximations).

[de Wiljes, Majda & Horenko, SIAM MMS, 2013], [Boyko, Krumscheid, Vercauteren, arXiv2102.12395, 2021]

Take-home message

- The sublinear expectation and related G-Brownian framework is a versatile tool for goal-oriented **coarse-graining with worst-case scenarios**.
- The **worst-case scenario may not correspond to a single parameter value**.
- The underlying nonlinear PDEs can be efficiently solved using the regression-type approximation schemes of Beck, E & Jentzen for **second-order BSDEs** .
- **Open problems:** strong convergence *in* sublinear expectation, multiplicative noise, infinite time horizon, ...

Thank you for your attention!

Acknowledgement:

Hafida Bouanani
Omar Kebiri
Lara Neureither
Lorenz Richter
Markus Strehlau
Upanshu Sharma
Wei Zhang

German Science Foundation (Math+, SFB 1114)
Einstein Center for Mathematics Berlin (ECMath)