# Quantifying Uncertainty in Differential Equation Models: Manifolds, Metrics and Russian Roulette

Mark Girolami

Department of Statistical Science
University College London

October 9, 2013

Enabling Quantification of
EQUIP
Uncertainty for Inverse Problems

## Talk Outline

- Why statistical inference for mechanistic models is hugely challenging

# Talk Outline

- Why statistical inference for mechanistic models is hugely challenging
- One approach is to attack the problem exploiting intrinsic geometry of mechanistic dynamics

## Talk Outline

- Why statistical inference for mechanistic models is hugely challenging
- One approach is to attack the problem exploiting intrinsic geometry of mechanistic dynamics
- Mechanistic dynamics described via PDE and ODE systems with non-analytic solutions

## Talk Outline

- Why statistical inference for mechanistic models is hugely challenging
- One approach is to attack the problem exploiting intrinsic geometry of mechanistic dynamics
- Mechanistic dynamics described via PDE and ODE systems with non-analytic solutions
- How to account for uncertainty induced by implicit definition of dynamics via PDE and ODE representation

## Talk Outline

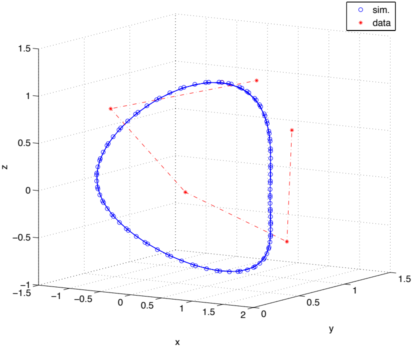
- Why statistical inference for mechanistic models is hugely challenging
- One approach is to attack the problem exploiting intrinsic geometry of mechanistic dynamics
- Mechanistic dynamics described via PDE and ODE systems with non-analytic solutions
- How to account for uncertainty induced by implicit definition of dynamics via PDE and ODE representation
- Oftentimes under fine spatial mesh refinement computing a likelihood exactly may be infeasible

## Talk Outline

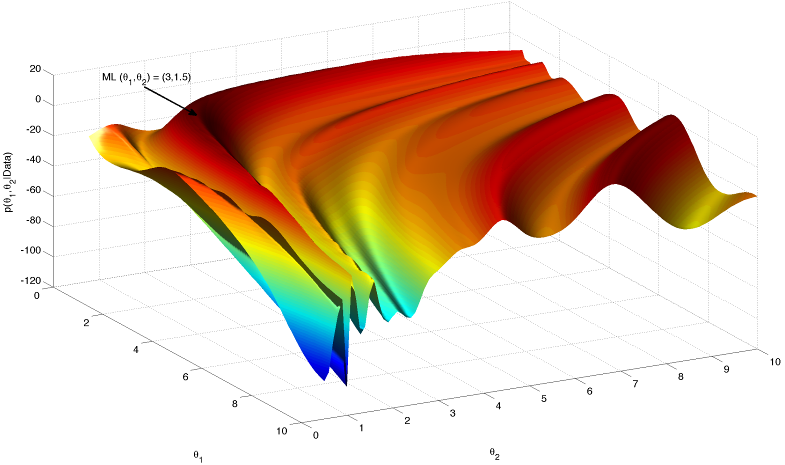
- Why statistical inference for mechanistic models is hugely challenging
- One approach is to attack the problem exploiting intrinsic geometry of mechanistic dynamics
- Mechanistic dynamics described via PDE and ODE systems with non-analytic solutions
- How to account for uncertainty induced by implicit definition of dynamics via PDE and ODE representation
- Oftentimes under fine spatial mesh refinement computing a likelihood exactly may be infeasible
- Forward and Inverse inference can still progress exploiting pseudo-marginal constructions in general form of Russian Roulette

## Simple Dynamics

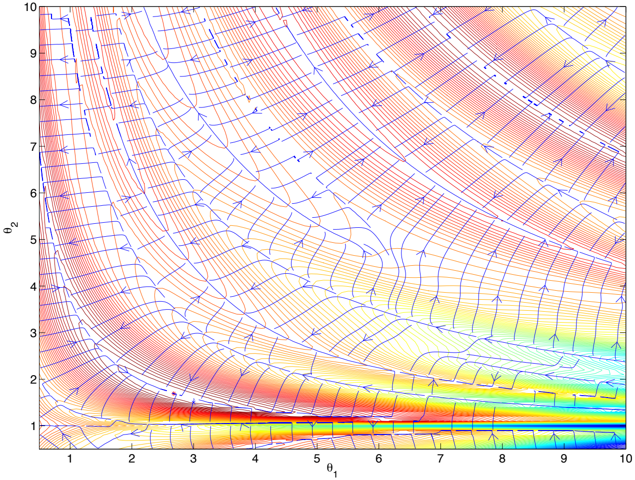
$$\frac{dx}{dt} = \theta_1 yz$$
$$\frac{dy}{dt} = -xz$$
$$\frac{dz}{dt} = -\theta_2 xy$$



# Simple Dynamics



# Simple Dynamics



- Simple nonlinear model - protein autoregulation.

$$\frac{dx}{dt} = \frac{72}{36 + y} - \alpha$$
$$\frac{dy}{dt} = \beta x - 1$$

- Simple nonlinear model - protein autoregulation.

$$\frac{dx}{dt} = \frac{72}{36+y} - \alpha$$
$$\frac{dy}{dt} = \beta x - 1$$

- 120 equally spaced measurements of system from $t = 0 \ldots 60 seconds$ with Normal errors having known variance 0.5, $\alpha = 3, \beta = 1$.

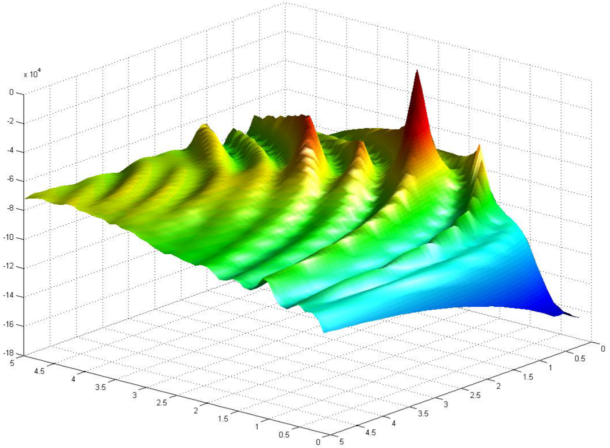
- Simple nonlinear model - protein autoregulation.

$$\frac{dx}{dt} = \frac{72}{36+y} - \alpha$$
$$\frac{dy}{dt} = \beta x - 1$$

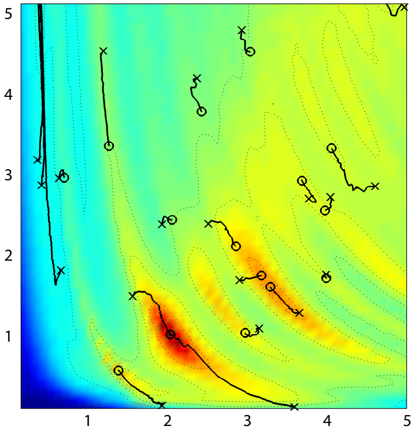
- 120 equally spaced measurements of system from $t = 0...60 seconds$ with Normal errors having known variance 0.5, $\alpha = 3, \beta = 1$.

- Induces a data density posing many challenges for simulation based inference

## Systems Identification - Posterior Inference



# Mixing of Markov Chains



# Geometric Concepts in MCMC

## Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

## Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold

## Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- For nonlinear DE's $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$ observed under Normal additive error with covariance , then FR metric tensor

## Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- For nonlinear DE's $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$ observed under Normal additive error with covariance , then FR metric tensor

$$\mathbf{G}(\boldsymbol{\theta}) = \partial_{\boldsymbol{\theta}}\mathbf{x}\mathbf{C}^{-1}\partial_{\boldsymbol{\theta}}\mathbf{x}^T$$

## Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- For nonlinear DE's $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$ observed under Normal additive error with covariance , then FR metric tensor

$$\mathbf{G}(\boldsymbol{\theta}) = \partial_{\boldsymbol{\theta}}\mathbf{x}\mathbf{C}^{-1}\partial_{\boldsymbol{\theta}}\mathbf{x}^T$$

- Sensitivity equations define metric can be exploited to locally transform coordinate system for sampling

## Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- For nonlinear DE's $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$ observed under Normal additive error with covariance , then FR metric tensor

$$\mathbf{G}(\boldsymbol{\theta}) = \partial_{\boldsymbol{\theta}}\mathbf{x}\mathbf{C}^{-1}\partial_{\boldsymbol{\theta}}\mathbf{x}^T$$

- Sensitivity equations define metric can be exploited to locally transform coordinate system for sampling
- Discretised Langevin diffusion (MALA) on manifold defines efficient proposal mechanism

$$\boldsymbol{\theta}'_d = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2\sum_{i,j}^{D}\boldsymbol{G}(\boldsymbol{\theta})^{-1}_{ij}\Gamma^d_{ij} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

# Geometric Concepts in MCMC

- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- For nonlinear DE's $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$ observed under Normal additive error with covariance , then FR metric tensor

$$\mathbf{G}(\boldsymbol{\theta}) = \partial_{\boldsymbol{\theta}}\mathbf{x}\mathbf{C}^{-1}\partial_{\boldsymbol{\theta}}\mathbf{x}^T$$

- Sensitivity equations define metric can be exploited to locally transform coordinate system for sampling
- Discretised Langevin diffusion (MALA) on manifold defines efficient proposal mechanism

$$\boldsymbol{\theta}'_d = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2 \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})^{-1}_{ij}\Gamma^d_{ij} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

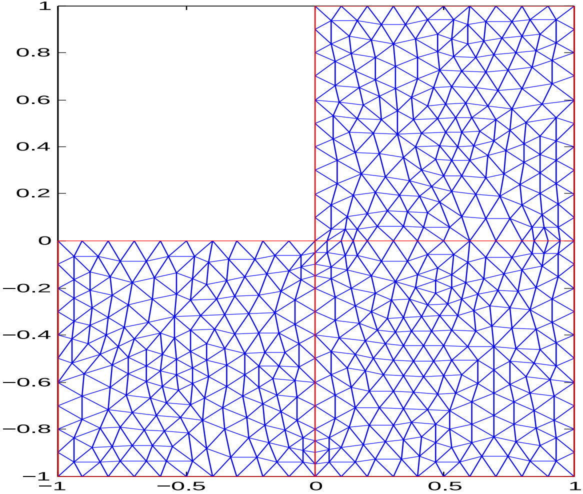
- Non-Euclidean geometry can exploit geodesics equations to devise sampling schemes (RMHMC)

# Infinite and finite dimensional mismatch

- Uncertainty induced by discrete mesh

## Infinite and finite dimensional mismatch

- Uncertainty induced by discrete mesh



## Infinite and finite dimensional mismatch

- Models relate $\dot{\mathbf{x}}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ to states $\mathbf{x}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ by vector field $f_{\boldsymbol{\theta}}(t, \cdot) : \mathbb{R}^P \to \mathbb{R}^P$ indexed by parameter vector, $\boldsymbol{\theta} \in \Theta$

## Infinite and finite dimensional mismatch

- Models relate $\dot{\mathbf{x}}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ to states $\mathbf{x}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ by vector field $f_{\boldsymbol{\theta}}(t, \cdot) : \mathbb{R}^P \to \mathbb{R}^P$ indexed by parameter vector, $\boldsymbol{\theta} \in \Theta$
- Observations, $\mathbf{y}(t) \in \mathbb{R}^R$, via transformation, $\mathcal{G} : \mathbb{R}^P \to \mathbb{R}^R$, of the *true or exact* solution, $\mathbf{x}(t, \theta)$, of system at $T$ time points.

## Infinite and finite dimensional mismatch

- Models relate $\dot{\mathbf{x}}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ to states $\mathbf{x}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ by vector field $f_{\boldsymbol{\theta}}(t, \cdot) : \mathbb{R}^P \to \mathbb{R}^P$ indexed by parameter vector, $\boldsymbol{\theta} \in \Theta$
- Observations, $\mathbf{y}(t) \in \mathbb{R}^R$, via transformation, $\mathcal{G} : \mathbb{R}^P \to \mathbb{R}^R$, of the *true or exact* solution, $\mathbf{x}(t, \theta)$, of system at $T$ time points.

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\big(\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})\big) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0}, \Sigma^{(r)}),$$

where $\mathcal{G}^{(r)}$ is the $r$th output of the observation model.

# Infinite and finite dimensional mismatch

- Models relate $\dot{\mathbf{x}}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ to states $\mathbf{x}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ by vector field $f_{\boldsymbol{\theta}}(t, \cdot) : \mathbb{R}^P \to \mathbb{R}^P$ indexed by parameter vector, $\boldsymbol{\theta} \in \Theta$
- Observations, $\mathbf{y}(t) \in \mathbb{R}^R$, via transformation, $\mathcal{G} : \mathbb{R}^P \to \mathbb{R}^R$, of the *true or exact* solution, $\mathbf{x}(t, \theta)$, of system at $T$ time points.

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\left(\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})\right) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0}, \Sigma^{(r)}),$$

  where $\mathcal{G}^{(r)}$ is the $r$th output of the observation model.

- If unique analytic, $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, at each observation time (and spatial position in the case of PDEs), given data $\mathbf{y}(\mathbf{t})$ posterior takes the form

## Infinite and finite dimensional mismatch

- Models relate $\dot{\mathbf{x}}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ to states $\mathbf{x}(t, \boldsymbol{\theta}) \in \mathbb{R}^P$ by vector field $f_{\boldsymbol{\theta}}(t, \cdot) : \mathbb{R}^P \rightarrow \mathbb{R}^P$ indexed by parameter vector, $\boldsymbol{\theta} \in \Theta$

- Observations, $\mathbf{y}(t) \in \mathbb{R}^R$, via transformation, $\mathcal{G} : \mathbb{R}^P \rightarrow \mathbb{R}^R$, of the *true or exact* solution, $\mathbf{x}(t, \theta)$, of system at $T$ time points.

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\left(\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})\right) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0}, \Sigma^{(r)}),$$

  where $\mathcal{G}^{(r)}$ is the $r$th output of the observation model.

- If unique analytic, $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, at each observation time (and spatial position in the case of PDEs), given data $\mathbf{y}(\mathbf{t})$ posterior takes the form

$$p(\boldsymbol{\theta}, \mathbf{x}_0 | \mathbf{y}(\mathbf{t})) \propto p\left(\mathbf{y}(\mathbf{t}) | \mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\right) \times \pi\left(\boldsymbol{\theta}, \mathbf{x}_0\right).$$

## Infinite and finite dimensional mismatch

- Due to non-analytic $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, likelihood $p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big)$ unavailable

## Infinite and finite dimensional mismatch

- Due to non-analytic $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, likelihood $p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big)$ unavailable
- Replace $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$ with finite-dimensional representation $\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})$ - Finite spatial and temporal mesh, i.e. numerically integrated solution

## Infinite and finite dimensional mismatch

- Due to non-analytic $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, likelihood $p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big)$ unavailable
- Replace $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$ with finite-dimensional representation $\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})$ - Finite spatial and temporal mesh, i.e. numerically integrated solution
- Disregard the additional uncertainty induced in the discrete solution

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\big(\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})\big) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0}, \Sigma^{(r)}),$$

## Infinite and finite dimensional mismatch

- Due to non-analytic $\mathbf{x}(\mathbf{t},\boldsymbol{\theta})$, likelihood $p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{x}_0\big)$ unavailable
- Replace $\mathbf{x}(\mathbf{t},\boldsymbol{\theta})$ with finite-dimensional representation $\mathbf{x}^N(\mathbf{t},\boldsymbol{\theta})$ - Finite spatial and temporal mesh, i.e. numerically integrated solution
- Disregard the additional uncertainty induced in the discrete solution

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\big(\mathbf{x}^N(\mathbf{t},\boldsymbol{\theta})\big) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0},\Sigma^{(r)}),$$

Yielding posterior measure

$$p^N\big(\boldsymbol{\theta},\mathbf{x}_0|\mathbf{y}(\mathbf{t})\big) \propto p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}^N(\mathbf{t},\boldsymbol{\theta}),\mathbf{x}_0\big) \times \pi\big(\boldsymbol{\theta},\mathbf{x}_0\big).$$

## Infinite and finite dimensional mismatch

- Due to non-analytic $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, likelihood $p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big)$ unavailable
- Replace $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$ with finite-dimensional representation $\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})$ - Finite spatial and temporal mesh, i.e. numerically integrated solution
- Disregard the additional uncertainty induced in the discrete solution

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\big(\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})\big) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0}, \Sigma^{(r)}),$$

Yielding posterior measure

$$p^N\big(\boldsymbol{\theta}, \mathbf{x}_0|\mathbf{y}(\mathbf{t})\big) \propto p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big) \times \pi\big(\boldsymbol{\theta}, \mathbf{x}_0\big).$$

- Seemingly innocuous implicit assumption may lead to serious statistical bias and misleading inferences as unaccounted errors accumulate in the discretisation

## Infinite and finite dimensional mismatch

- Due to non-analytic $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$, likelihood $p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big)$ unavailable
- Replace $\mathbf{x}(\mathbf{t}, \boldsymbol{\theta})$ with finite-dimensional representation $\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})$ - Finite spatial and temporal mesh, i.e. numerically integrated solution
- Disregard the additional uncertainty induced in the discrete solution

$$y^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\big(\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta})\big) + \epsilon^{(r)}(\mathbf{t}), \qquad \epsilon^{(r)}(\mathbf{t}) \sim \mathcal{N}_T(\mathbf{0}, \Sigma^{(r)}),$$

Yielding posterior measure

$$p^N\big(\boldsymbol{\theta}, \mathbf{x}_0|\mathbf{y}(\mathbf{t})\big) \propto p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}^N(\mathbf{t}, \boldsymbol{\theta}), \mathbf{x}_0\big) \times \pi\big(\boldsymbol{\theta}, \mathbf{x}_0\big).$$

- Seemingly innocuous implicit assumption may lead to serious statistical bias and misleading inferences as unaccounted errors accumulate in the discretisation

$$\begin{aligned} y^{(r)}(\mathbf{t}) &= \mathcal{G}^{(r)}\big(\mathbf{x}(\mathbf{t})\big) + \epsilon^{(r)}(\mathbf{t}) \\ &= \mathcal{G}^{(r)}\big(\mathbf{x}^N(\mathbf{t})\big) + \delta^{(r)}(\mathbf{t}) + \epsilon^{(r)}(\mathbf{t}) \end{aligned}$$

where $\delta^{(r)}(\mathbf{t}) = \mathcal{G}^{(r)}\big(\mathbf{x}(\mathbf{t})\big) - \mathcal{G}^{(r)}\big(\mathbf{x}^N(\mathbf{t})\big)$.

# Full Bayesian posterior measure for model uncertainty

- Define a probability measure over functions

$$p\big(\mathbf{x}(\mathbf{t},\theta),\mathbf{f}|\theta,\mathbf{x}_0,\Psi\big)$$

## Full Bayesian posterior measure for model uncertainty

- Define a probability measure over functions

$$p\big(\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f}|\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)$$

where $\Psi$ are parameters of error model for $\mathbf{x}(\mathbf{t},\boldsymbol{\theta})$, and $\mathbf{f}$ evaluations of vector field at discretised approximate solution

## Full Bayesian posterior measure for model uncertainty

- Define a probability measure over functions

$$p\big(\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f}|\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)$$

  where $\Psi$ are parameters of error model for $\mathbf{x}(\mathbf{t},\boldsymbol{\theta})$, and $\mathbf{f}$ evaluations of vector field at discretised approximate solution

- The full posterior distribution therefore follows as

$$p\big(\boldsymbol{\theta},\mathbf{x}_0,\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f},\Psi|\mathbf{y}(\mathbf{t})\big) \propto \underbrace{p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t},\boldsymbol{\theta})\big)}_{\text{Likelihood}} \times \underbrace{p\big(\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f}|\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)}_{\text{Probabilistic Solution}} \times \underbrace{\pi\big(\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)}_{\text{Prior}}$$

## Full Bayesian posterior measure for model uncertainty

- Define a probability measure over functions

$$p\big(\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f}|\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)$$

  where $\Psi$ are parameters of error model for $\mathbf{x}(\mathbf{t},\boldsymbol{\theta})$, and $\mathbf{f}$ evaluations of vector field at discretised approximate solution

- The full posterior distribution therefore follows as

$$p(\boldsymbol{\theta},\mathbf{x}_0,\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f},\Psi|\mathbf{y}(\mathbf{t})) \propto \underbrace{p\big(\mathbf{y}(\mathbf{t})|\mathbf{x}(\mathbf{t},\boldsymbol{\theta})\big)}_{\text{Likelihood}} \times \underbrace{p\big(\mathbf{x}(\mathbf{t},\boldsymbol{\theta}),\mathbf{f}|\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)}_{\text{Probabilistic Solution}} \times \underbrace{\pi\big(\boldsymbol{\theta},\mathbf{x}_0,\Psi\big)}_{\text{Prior}}$$

- Uncertainty in the probabilistic solution $\mathbf{x}(\mathbf{t},\boldsymbol{\theta})$ is made explicit taking into account the mismatch between state solution and a finite approximation

## Full Bayesian posterior measure for model uncertainty

- Approximate solutions at $N$ knots, $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$

## Full Bayesian posterior measure for model uncertainty

- Approximate solutions at $N$ knots, $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$
- From these obtain values of $N$ approximate vector field evaluations as,

$$\mathbf{f}_{1:N} = \left[f_\theta(s_1, \tilde{x}(s_1)), \cdots, f_\theta(s_N, \tilde{x}(s_N))\right]$$

## Full Bayesian posterior measure for model uncertainty

- Approximate solutions at $N$ knots, $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$
- From these obtain values of $N$ approximate vector field evaluations as,

$$\mathbf{f}_{1:N} = \left[f_\theta(s_1, \tilde{x}(s_1)), \cdots, f_\theta(s_N, \tilde{x}(s_N))\right]$$

- Error between $\mathbf{f}_{1:N}$ and $\dot{x}(\mathbf{s}) = [\dot{x}(s_1), \cdots, \dot{x}(s_N)]$ define probabilistically

## Full Bayesian posterior measure for model uncertainty

- Approximate solutions at $N$ knots, $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$
- From these obtain values of $N$ approximate vector field evaluations as,

$$\mathbf{f}_{1:N} = \left[f_\theta(s_1, \tilde{x}(s_1)), \cdots, f_\theta(s_N, \tilde{x}(s_N))\right]$$

- Error between $\mathbf{f}_{1:N}$ and $\dot{x}(\mathbf{s}) = [\dot{x}(s_1), \cdots, \dot{x}(s_N)]$ define probabilistically
- Infinite dimension function space, no Lebesgue measure

# Full Bayesian posterior measure for model uncertainty

- Approximate solutions at $N$ knots, $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$
- From these obtain values of $N$ approximate vector field evaluations as,

$$\mathbf{f}_{1:N} = \left[f_\theta(s_1, \tilde{x}(s_1)), \cdots, f_\theta(s_N, \tilde{x}(s_N))\right]$$

- Error between $\mathbf{f}_{1:N}$ and $\dot{x}(\mathbf{s}) = [\dot{x}(s_1), \cdots, \dot{x}(s_N)]$ define probabilistically
- Infinite dimension function space, no Lebesgue measure
- Radon-Nikodym derivative of posterior measure with respect to GP prior

$$\frac{d\mu^f}{d\mu_0^f}(\dot{x}(\mathbf{s})) \propto \exp\left(-\frac{1}{2}||\dot{x}(\mathbf{s}) - \mathbf{f}_{1:N}||^2_{\mathbf{\Lambda}_N}\right)$$

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined
- Eigenfunctions of covariance operator form basis for derivative space

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined
- Eigenfunctions of covariance operator form basis for derivative space
- $m_0^f(t_1) = \ell(t_1), \quad \mathcal{C}_0^f(t_1, t_2) = \alpha^{-1} \int_{\mathbb{R}} \mathrm{R}_\lambda(t_1, z)\mathrm{R}_\lambda(t_2, z)\mathrm{d}z = \mathrm{RR}(t_1, t_2).$

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined
- Eigenfunctions of covariance operator form basis for derivative space
- $m_0^f(t_1) = \ell(t_1), \quad \mathcal{C}_0^f(t_1, t_2) = \alpha^{-1} \int_{\mathbb{R}} \mathrm{R}_\lambda(t_1, z)\mathrm{R}_\lambda(t_2, z)\mathrm{d}z = \mathrm{RR}(t_1, t_2).$
- Posterior measure for $\dot{x}(\mathbf{t})$ denoted by $\mu_n^f = \mathcal{N}\left(m_n^f(\mathbf{t}), \mathcal{C}_n^f(\mathbf{t}, \mathbf{t})\right)$

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined
- Eigenfunctions of covariance operator form basis for derivative space
- $m_0^f(t_1) = \ell(t_1), \quad \mathcal{C}_0^f(t_1, t_2) = \alpha^{-1} \int_{\mathbb{R}} \mathrm{R}_\lambda(t_1, z)\mathrm{R}_\lambda(t_2, z)\mathrm{d}z = \mathrm{RR}(t_1, t_2).$
- Posterior measure for $\dot{x}(\mathbf{t})$ denoted by $\mu_n^f = \mathcal{N}\left(m_n^f(\mathbf{t}), \mathcal{C}_n^f(\mathbf{t},\mathbf{t})\right)$

$$m_n^f(\mathbf{t}) = m_0^f(\mathbf{t}) + \mathrm{RR}(\mathbf{t},\mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s},\mathbf{s})\right)^{-1}\left(\mathbf{f}_{1:n} - m_0^f(\mathbf{s})\right)$$
$$\mathcal{C}_n^f(\mathbf{t},\mathbf{t}) = \mathrm{RR}(\mathbf{t},\mathbf{t}) - \mathrm{RR}(\mathbf{t},\mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s},\mathbf{s})\right)^{-1}\mathrm{RR}(\mathbf{s},\mathbf{t})$$

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined
- Eigenfunctions of covariance operator form basis for derivative space
- $m_0^f(t_1) = \ell(t_1), \quad \mathcal{C}_0^f(t_1, t_2) = \alpha^{-1} \int_{\mathbb{R}} \mathrm{R}_\lambda(t_1, z)\mathrm{R}_\lambda(t_2, z)\mathrm{d}z = \mathrm{RR}(t_1, t_2).$
- Posterior measure for $\dot{x}(\mathbf{t})$ denoted by $\mu_n^f = \mathcal{N}\big(m_n^f(\mathbf{t}), \mathcal{C}_n^f(\mathbf{t}, \mathbf{t})\big)$

$$m_n^f(\mathbf{t}) = m_0^f(\mathbf{t}) + \mathrm{RR}(\mathbf{t}, \mathbf{s})\big(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\big)^{-1}\big(\mathbf{f}_{1:n} - m_0^f(\mathbf{s})\big)$$
$$\mathcal{C}_n^f(\mathbf{t}, \mathbf{t}) = \mathrm{RR}(\mathbf{t}, \mathbf{t}) - \mathrm{RR}(\mathbf{t}, \mathbf{s})\big(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\big)^{-1}\mathrm{RR}(\mathbf{s}, \mathbf{t})$$

- Linearity of integral operator provides Gaussian prior measure for $x(t)$

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined

- Eigenfunctions of covariance operator form basis for derivative space

- $m_0^f(t_1) = \ell(t_1), \quad \mathcal{C}_0^f(t_1, t_2) = \alpha^{-1} \int_{\mathbb{R}} \mathrm{R}_\lambda(t_1, z)\mathrm{R}_\lambda(t_2, z)\mathrm{d}z = \mathrm{RR}(t_1, t_2).$

- Posterior measure for $\dot{x}(\mathbf{t})$ denoted by $\mu_n^f = \mathcal{N}\left(m_n^f(\mathbf{t}), \mathcal{C}_n^f(\mathbf{t}, \mathbf{t})\right)$

$$m_n^f(\mathbf{t}) = m_0^f(\mathbf{t}) + \mathrm{RR}(\mathbf{t}, \mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\right)^{-1}\left(\mathbf{f}_{1:n} - m_0^f(\mathbf{s})\right)$$

$$\mathcal{C}_n^f(\mathbf{t}, \mathbf{t}) = \mathrm{RR}(\mathbf{t}, \mathbf{t}) - \mathrm{RR}(\mathbf{t}, \mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\right)^{-1}\mathrm{RR}(\mathbf{s}, \mathbf{t})$$

- Linearity of integral operator provides Gaussian prior measure for $x(t)$

- Posterior measure follows as $p\left(x(\mathbf{t}) \middle| \mathbf{f}_{1:N}, \boldsymbol{\theta}, \mathbf{x}_0, \Psi\right) = \mathcal{N}_T\left(m_N(\mathbf{t}), \mathcal{C}_N(\mathbf{t}, \mathbf{t})\right)$

## Full Bayesian posterior measure for model uncertainty

- Gaussian measure, $\mu_0^f = \mathcal{N}(m_0^f, \mathcal{C}_0^f)$, on a Hilbert space, $\mathcal{H}$, mean function $m_0^f$, covariance operator $\mathcal{C}_0^f$ well defined
- Eigenfunctions of covariance operator form basis for derivative space
- $m_0^f(t_1) = \ell(t_1), \quad \mathcal{C}_0^f(t_1, t_2) = \alpha^{-1} \int_{\mathbb{R}} \mathrm{R}_\lambda(t_1, z)\mathrm{R}_\lambda(t_2, z)\mathrm{d}z = \mathrm{RR}(t_1, t_2).$
- Posterior measure for $\dot{x}(\mathbf{t})$ denoted by $\mu_n^f = \mathcal{N}\left(m_n^f(\mathbf{t}), \mathcal{C}_n^f(\mathbf{t}, \mathbf{t})\right)$

$$m_n^f(\mathbf{t}) = m_0^f(\mathbf{t}) + \mathrm{RR}(\mathbf{t}, \mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\right)^{-1}\left(\mathbf{f}_{1:n} - m_0^f(\mathbf{s})\right)$$
$$\mathcal{C}_n^f(\mathbf{t}, \mathbf{t}) = \mathrm{RR}(\mathbf{t}, \mathbf{t}) - \mathrm{RR}(\mathbf{t}, \mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\right)^{-1}\mathrm{RR}(\mathbf{s}, \mathbf{t})$$

- Linearity of integral operator provides Gaussian prior measure for $x(t)$
- Posterior measure follows as $p\left(x(\mathbf{t}) \middle| \mathbf{f}_{1:N}, \boldsymbol{\theta}, \mathbf{x}_0, \Psi\right) = \mathcal{N}_T\left(m_N(\mathbf{t}), \mathcal{C}_N(\mathbf{t}, \mathbf{t})\right)$

$$m_n(\mathbf{t}) = m_0(\mathbf{t}) + \mathrm{QR}(\mathbf{t}, \mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\right)^{-1}\left(\mathbf{f}_{1:n} - m_0^f(\mathbf{s})\right)$$
$$\mathcal{C}_n(\mathbf{t}, \mathbf{t}) = \mathrm{QQ}(\mathbf{t}, \mathbf{t}) - \mathrm{QR}(\mathbf{t}, \mathbf{s})\left(\Lambda_n + \mathrm{RR}(\mathbf{s}, \mathbf{s})\right)^{-1}\mathrm{RQ}(\mathbf{s}, \mathbf{t}),$$

## Probabilistic Integration

- Posterior measure of functional solutions in $\mathcal{H}$ is Gaussian Process

## Probabilistic Integration

- Posterior measure of functional solutions in $\mathcal{H}$ is Gaussian Process
- Approximate solution at knots $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$ probability measure for solution values at any $t$ follows in Gaussian conditional

## Probabilistic Integration

- Posterior measure of functional solutions in $\mathcal{H}$ is Gaussian Process
- Approximate solution at knots $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$ probability measure for solution values at any $t$ follows in Gaussian conditional
- Provides calibrated probability measure of uncertainty due to finite solution and infinite model mismatch

## Probabilistic Integration

- Posterior measure of functional solutions in $\mathcal{H}$ is Gaussian Process
- Approximate solution at knots $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$ probability measure for solution values at any $t$ follows in Gaussian conditional
- Provides calibrated probability measure of uncertainty due to finite solution and infinite model mismatch

$$p\big(\boldsymbol{\theta}, \mathbf{x}_0, \mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{f}, \Psi | \mathbf{y}(\mathbf{t})\big) \propto \underbrace{p\big(\mathbf{y}(\mathbf{t}) | \mathbf{x}(\mathbf{t}, \boldsymbol{\theta})\big)}_{\text{Likelihood}} \times \underbrace{p\big(\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{f} | \boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)}_{\text{Probabilistic Solution}} \times \underbrace{\pi\big(\boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)}_{\text{Prior}}$$

## Probabilistic Integration

- Posterior measure of functional solutions in $\mathcal{H}$ is Gaussian Process
- Approximate solution at knots $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$ probability measure for solution values at any $t$ follows in Gaussian conditional
- Provides calibrated probability measure of uncertainty due to finite solution and infinite model mismatch

$$p(\boldsymbol{\theta}, \mathbf{x}_0, \mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{f}, \Psi | \mathbf{y}(\mathbf{t})) \propto \underbrace{p\big(\mathbf{y}(\mathbf{t}) | \mathbf{x}(\mathbf{t}, \boldsymbol{\theta})\big)}_{\text{Likelihood}} \times \underbrace{p\big(\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{f} | \boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)}_{\text{Probabilistic Solution}} \times \underbrace{\pi\big(\boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)}_{\text{Prior}}$$

- Important in quantifying uncertainty (or uncertainty reduction) in moving from coarse to fine meshing

## Probabilistic Integration

- Posterior measure of functional solutions in $\mathcal{H}$ is Gaussian Process
- Approximate solution at knots $\tilde{x}(s_1), \cdots, \tilde{x}(s_N)$ probability measure for solution values at any $t$ follows in Gaussian conditional
- Provides calibrated probability measure of uncertainty due to finite solution and infinite model mismatch

$$p(\boldsymbol{\theta}, \mathbf{x}_0, \mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{f}, \Psi | \mathbf{y}(\mathbf{t})) \propto \underbrace{p\big(\mathbf{y}(\mathbf{t}) | \mathbf{x}(\mathbf{t}, \boldsymbol{\theta})\big)}_{\text{Likelihood}} \times \underbrace{p\big(\mathbf{x}(\mathbf{t}, \boldsymbol{\theta}), \mathbf{f} | \boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)}_{\text{Probabilistic Solution}} \times \underbrace{\pi\big(\boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)}_{\text{Prior}}$$

- Important in quantifying uncertainty (or uncertainty reduction) in moving from coarse to fine meshing
- Suggests probabilistic construction (integration), sampling of solutions

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t}u = -u\frac{\partial}{\partial x}u - \frac{\partial^2}{\partial x^2}u - \frac{\partial^4}{\partial x^4}u,$$

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t}u = -u\frac{\partial}{\partial x}u - \frac{\partial^2}{\partial x^2}u - \frac{\partial^4}{\partial x^4}u,$$

- Domain $x \in [0, 32\pi]$, $t \in [0, 150]$

# Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t} u = -u \frac{\partial}{\partial x} u - \frac{\partial^2}{\partial x^2} u - \frac{\partial^4}{\partial x^4} u,$$

- Domain $x \in [0, 32\pi]$, $t \in [0, 150]$
- Initial function $u(0, x) = \cos(x/16)\{1 + \sin(x/16)\}$

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t}u = -u\frac{\partial}{\partial x}u - \frac{\partial^2}{\partial x^2}u - \frac{\partial^4}{\partial x^4}u,$$

- Domain $x \in [0, 32\pi]$, $t \in [0, 150]$
- Initial function $u(0, x) = \cos(x/16)\left\{1 + \sin(x/16)\right\}$
- Discretize spatial domain, obtaining a high-dimensional (128 dimensions) system of stiff ODEs

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t}u = -u\frac{\partial}{\partial x}u - \frac{\partial^2}{\partial x^2}u - \frac{\partial^4}{\partial x^4}u,$$

- Domain $x \in [0, 32\pi]$, $t \in [0, 150]$
- Initial function $u(0, x) = \cos(x/16)\left\{1 + \sin(x/16)\right\}$
- Discretize spatial domain, obtaining a high-dimensional (128 dimensions) system of stiff ODEs
- Use the integrating factor method to transform the system to one of purely nonlinear ODEs

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t}u = -u\frac{\partial}{\partial x}u - \frac{\partial^2}{\partial x^2}u - \frac{\partial^4}{\partial x^4}u,$$

- Domain $x \in [0, 32\pi]$, $t \in [0, 150]$
- Initial function $u(0, x) = \cos(x/16)\{1 + \sin(x/16)\}$
- Discretize spatial domain, obtaining a high-dimensional (128 dimensions) system of stiff ODEs
- Use the integrating factor method to transform the system to one of purely nonlinear ODEs
- Probabilistic IVP solutions sampled using 2K uniform solver knots

## Kuramoto-Sivashinsky model of reaction-diffusion

- Kuramoto-Sivashinsky model of reaction-diffusion systems

$$\frac{\partial}{\partial t}u = -u\frac{\partial}{\partial x}u - \frac{\partial^2}{\partial x^2}u - \frac{\partial^4}{\partial x^4}u,$$

- Domain $x \in [0, 32\pi]$, $t \in [0, 150]$
- Initial function $u(0, x) = \cos(x/16)\left\{1 + \sin(x/16)\right\}$
- Discretize spatial domain, obtaining a high-dimensional (128 dimensions) system of stiff ODEs
- Use the integrating factor method to transform the system to one of purely nonlinear ODEs
- Probabilistic IVP solutions sampled using 2K uniform solver knots
- Fifteen solution samples illustrate uncertainty over domain propagates through system resulting in noticeably distinct dynamics, not captured by deterministic numerical solvers.

## Kuramoto-Sivashinsky model of reaction-diffusion

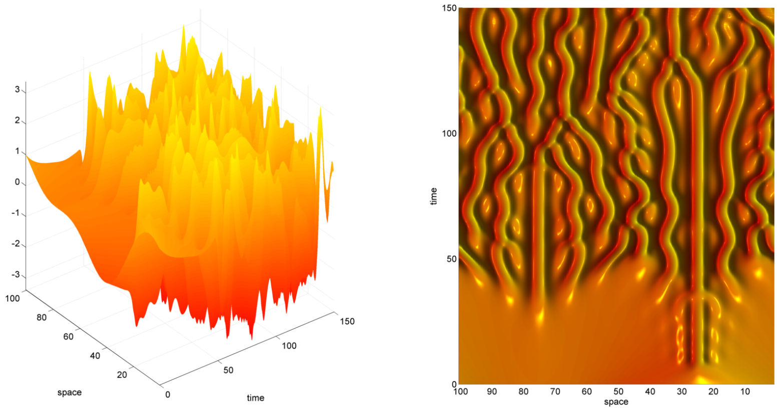



Figure: Side view and top view of a probabilistic solution realization of the Kuramoto-Sivashinsky PDE with initial function $u(0,x) = \cos(x/16)\left\{1 + \sin(x/16)\right\}$ and domain $x \in [0, 32\pi]$, $t \in [0, 150]$.

# Kuramoto-Sivashinsky model of reaction-diffusion

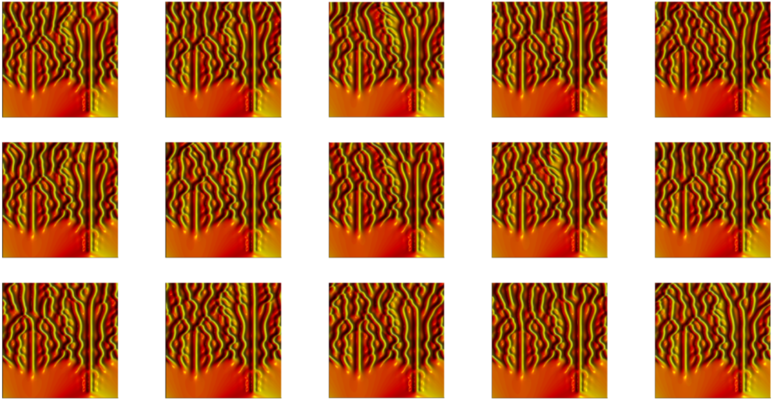


Figure: Fifteen realizations of the probabilistic solution of the Kuramoto-Sivashinsky PDE using a fixed initial function. The solution is known to exhibit temporal chaos. Deterministic numerical solutions only capture one type of behaviour given a fixed initial function, which can lead to bias when used in conjunction with data-based inference.

## Full Bayesian Uncertainty Quantification

- Consider now joint parameter and solution inference

## Full Bayesian Uncertainty Quantification

- Consider now joint parameter and solution inference
- Draw $K$ samples from the posterior distribution $p\big(\boldsymbol{\theta}, \mathbf{x}(\mathbf{t}), \mathbf{f}_{1:N} | \mathbf{y}(\mathbf{t}), \mathbf{x}_0, \Psi\big)$

## Full Bayesian Uncertainty Quantification

- Consider now joint parameter and solution inference
- Draw $K$ samples from the posterior distribution $p\big(\boldsymbol{\theta}, \mathbf{x}(\mathbf{t}), \mathbf{f}_{1:N} | \mathbf{y}(\mathbf{t}), \mathbf{x}_0, \Psi\big)$

  Initialize $\boldsymbol{\theta}$;
  **for** $k = 1 : K$ **do**
  Propose $\boldsymbol{\theta}^\star \sim q(\boldsymbol{\theta}^\star | \boldsymbol{\theta})$;
  Conditionally simulate a solution realisation $\mathbf{x}^\star(\mathbf{t})$ from $p\big(\mathbf{x}(\mathbf{t}), \mathbf{f}_{1:N} \big| \boldsymbol{\theta}, \mathbf{x}_0, \Psi\big)$
  Compute:

$$\rho(\boldsymbol{\theta}, \mathbf{x}(\mathbf{t}) \to \boldsymbol{\theta}^\star, \mathbf{x}^\star(\mathbf{t})) = \frac{q(\boldsymbol{\theta}|\boldsymbol{\theta}^\star)}{q(\boldsymbol{\theta}^\star|\boldsymbol{\theta})} \frac{p(\boldsymbol{\theta}^\star)}{p(\boldsymbol{\theta})} \frac{p\big(\mathbf{y}(\mathbf{t}) | \mathcal{G}\big(\mathbf{x}^\star(\mathbf{t})\big), \Sigma\big)}{p\big(\mathbf{y}(\mathbf{t}) | \mathcal{G}\big(\mathbf{x}(\mathbf{t})\big), \Sigma\big)};$$

  **if** $\min[1, \rho(\boldsymbol{\theta} \to \boldsymbol{\theta}^\star)] > \mathrm{U}[0, 1]$ **then**
  Update $\boldsymbol{\theta}, \mathbf{x}(\mathbf{t}) = \boldsymbol{\theta}^\star, \mathbf{x}^\star(\mathbf{t})$;
  **end if**
  Return $\boldsymbol{\theta}, \mathbf{x}(\mathbf{t})$.
  **end for**

# Inference for model of cellular signal transduction

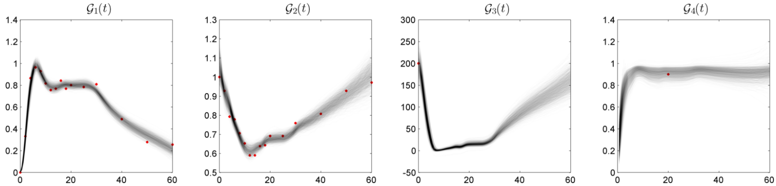



Figure: Experimental data and sample paths of the observation processes obtained by transforming a sample from marginal posterior state distribution by observation function

# Inference for model of cellular signal transduction

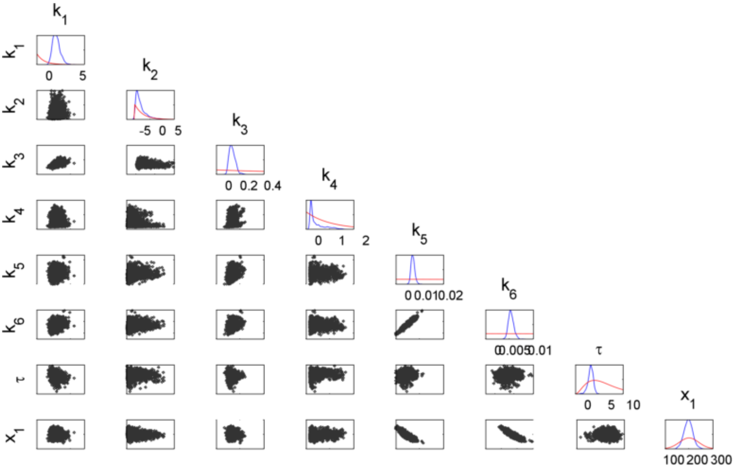
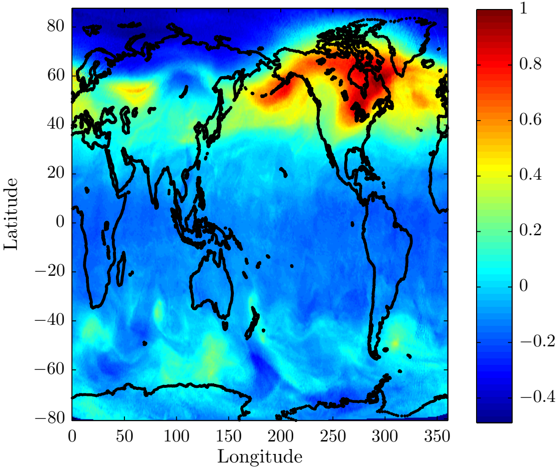



Figure: Marginal parameter posterior based on sample of size 100K generated by a parallel tempering algorithm utilizing seven chains, with the first 10K samples removed. Prior densities are shown in red.

## Intractable Likelihoods under Mesh Refinement

What can we do?

# Large Scale GMRF Ozone Column Model



## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements

## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements
- Data and spatial extent has precluded full Bayesian analysis to date

## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements
- Data and spatial extent has precluded full Bayesian analysis to date
- Matern covariance function triangulated over 196,002 vertices on sphere

## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements
- Data and spatial extent has precluded full Bayesian analysis to date
- Matern covariance function triangulated over 196,002 vertices on sphere

$$p(\mathbf{x}|\theta) = \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}\left[\log|\mathbf{C}_\theta| + \mathbf{x}^\mathsf{T}\mathbf{C}_\theta^{-1}\mathbf{x}\right]\right\}$$

## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements
- Data and spatial extent has precluded full Bayesian analysis to date
- Matern covariance function triangulated over 196,002 vertices on sphere

$$
\begin{aligned}
p(\mathbf{x}|\theta) &= \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}\left[\log|\mathbf{C}_\theta| + \mathbf{x}^\mathsf{T}\mathbf{C}_\theta^{-1}\mathbf{x}\right]\right\} \\
&= \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{\frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{E}\{\mathbf{z}^\mathsf{T}(\mathbf{I}-\mathbf{C}_\theta)^n\mathbf{z}\} - \frac{1}{2}\sum_{m=0}^{\infty}\mathbf{x}^\mathsf{T}(\mathbf{I}-\mathbf{C}_\theta)^m\mathbf{x}\right\}
\end{aligned}
$$

## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements
- Data and spatial extent has precluded full Bayesian analysis to date
- Matern covariance function triangulated over 196,002 vertices on sphere

$$
\begin{aligned}
p(\mathbf{x}|\theta) &= \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}\left[\log|\mathbf{C}_\theta| + \mathbf{x}^\mathrm{T}\mathbf{C}_\theta^{-1}\mathbf{x}\right]\right\} \\
&= \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{\frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{E}\{\mathbf{z}^\mathrm{T}(\mathbf{I}-\mathbf{C}_\theta)^n\mathbf{z}\} - \frac{1}{2}\sum_{m=0}^{\infty}\mathbf{x}^\mathrm{T}(\mathbf{I}-\mathbf{C}_\theta)^m\mathbf{x}\right\}
\end{aligned}
$$

$$
\begin{aligned}
p(\mathbf{x} \mid \theta) &= \frac{1}{(2\pi)^{N/2}} \exp\left\{\frac{1}{2}\log|\mathbf{Q}_\theta| - \frac{1}{2}\mathbf{x}^T\mathbf{Q}_\theta\mathbf{x}\right\} \\
&= \frac{1}{(2\pi)^{N/2}} \exp\left\{\frac{1}{2}\mathbb{E}_\mathbf{z}\{\mathbf{z}^\mathbf{T}\log(\mathbf{Q}_\theta)\mathbf{z}\} - \mathbf{\frac{1}{2}}\mathbf{x}^\mathbf{T}\mathbf{Q}_\theta\mathbf{x}\right\}
\end{aligned}
$$

## Large Scale GMRF Ozone Column Model

- Cressie, (2008) comprised of 173,405 ozone measurements
- Data and spatial extent has precluded full Bayesian analysis to date
- Matern covariance function triangulated over 196,002 vertices on sphere

$$
\begin{aligned}
p(\mathbf{x}|\theta) &= \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}\left[\log|\mathbf{C}_\theta| + \mathbf{x}^\mathsf{T}\mathbf{C}_\theta^{-1}\mathbf{x}\right]\right\} \\
&= \frac{1}{(2\pi)^{\frac{N}{2}}} \exp\left\{\frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{E}\{\mathbf{z}^\mathsf{T}(\mathbf{I}-\mathbf{C}_\theta)^n\mathbf{z}\} - \frac{1}{2}\sum_{m=0}^{\infty}\mathbf{x}^\mathsf{T}(\mathbf{I}-\mathbf{C}_\theta)^m\mathbf{x}\right\}
\end{aligned}
$$

$$
\begin{aligned}
p(\mathbf{x} \mid \theta) &= \frac{1}{(2\pi)^{N/2}} \exp\left\{\frac{1}{2}\log|\mathbf{Q}_\theta| - \frac{1}{2}\mathbf{x}^T\mathbf{Q}_\theta\mathbf{x}\right\} \\
&= \frac{1}{(2\pi)^{N/2}} \exp\left\{\frac{1}{2}\mathbb{E}_\mathbf{z}\{\mathbf{z}^\mathsf{T}\log(\mathbf{Q}_\theta)\mathbf{z}\} - \frac{\mathbf{1}}{\mathbf{2}}\mathbf{x}^\mathsf{T}\mathbf{Q}_\theta\mathbf{x}\right\}
\end{aligned}
$$

- Exploit Pseudo-Marginal construction - Andrieu & Roberts, 2009 - Russian Roulette unbiased truncation of infinite series - MCMC based inference can proceed..... in principle ;-)

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC
- Probabilistic solution of DE's

# Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC
- Probabilistic solution of DE's
- Addressing intractable nature of likelihoods under mesh refinement

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC
- Probabilistic solution of DE's
- Addressing intractable nature of likelihoods under mesh refinement
- EQUIP is awesome

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC
- Probabilistic solution of DE's
- Addressing intractable nature of likelihoods under mesh refinement
- EQUIP is awesome
- EQUIP is awesome

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC
- Probabilistic solution of DE's
- Addressing intractable nature of likelihoods under mesh refinement
- EQUIP is awesome
- EQUIP is awesome
- EQUIP is awesome

## Conclusions and Discussion

- EQUIP presents some of the most exciting open research problems at the leading edge of computational statistics
- Geometric approaches to MCMC
- Probabilistic solution of DE's
- Addressing intractable nature of likelihoods under mesh refinement
- EQUIP is awesome
- EQUIP is awesome
- EQUIP is awesome