# Bayesian Inverse Problems in PDEs

Andrew Stuart[1]

[1] Mathematics Institute and
Centre for Scientific Computing
University of Warwick

DSPDE10, Barcelona
June 4th 2010

Funded by EPSRC, ERC, ONR

# Outline

# Outline

# Data Assimilation in Oceanography

S.L. Cotter, M. Dashti, J.C. Robinson and AMS:
Inverse Problems 2009.

- **Stokes operator** is $A = -\triangle$ on divergence free space.
- Consider the **Navier-Stokes equation**:

$$\frac{dv}{dt} + \nu Av + B(v, v) = f, \quad v(0) = u$$

- **Find** $u(x)$ (or $u(x)$ and $f(x, t)$).
- **Given** noisy *Lagrangian observations* $y$:

$$y_{j,k} = z_j(t_k) + \eta_{j,k}$$
$$\frac{dz_j}{dt} = v(z_j, t), \quad z_j(0) = z_{j,0}$$

# Crossing Energy Barriers in Molecular Dynamics

F. Pinski and AMS, J. Chem. Phys. 2010

- **Find** $u(t)$ solving

$$\frac{du}{dt} = -\nabla V(u) + \sqrt{\frac{2}{\beta}}\frac{dB}{dt}.$$

- **Given** $u(0) = u^{-}$ and $u(T) = u^{+}$.

# Groundwater Flow

M. Dashti, S. Harris and AMS, 2010

- Let $\mathcal{H} = L^2(\Omega, \mathbb{R}^3)$.
- Consider Darcy's Law

$$\nabla \cdot \Big(\exp(u)\nabla p\Big) = 0, \quad x \in \Omega$$

$$p = \phi, \quad x \in \partial\Omega.$$

- **Find** $u(x)$.
- **Given** noisy *observations* $y$ of the pressure:

$$y_j = p(x_j) + \eta_j$$

# Outline

# Idea 1

Everyone is familiar with ideas like:

WELL-POSED $\oplus$ SMALL PERTURBATION TO EQUATION
$\Rightarrow$ SMALL PERTURBATION TO SOLUTION

STABILITY $\oplus$ CONSISTENCY $\Rightarrow$ CONVERGENCE

Inverse problems are ill-posed. But we will develop a (probabilistic) well-posedness for inverse problems leading to a variety of useful results such as:

WELL-POSED $\oplus$ APPROXIMATION OF FORWARD PROBLEM
$\Rightarrow$ APPROXIMATION OF INVERSE PROBLEM

# Idea 2

Minimizing $I(u)$ is closely related to:

Exploring the probability density function $\exp\big(-I(u)\big)$

We will thus connect Bayesian probability and optimal control.

# Outline

# Least Squares

- The equation $y = \mathcal{G}(u)$ for $u \in X$ is ill-posed: (no solution, many solutions, sensitive dependence).
- Replace by least squares problem

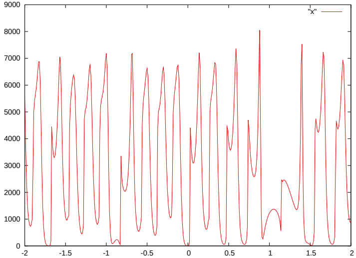
$$\min_{u \in X} \Phi(u; y), \quad \Phi(u; y) = \frac{1}{2}\|y - \mathcal{G}(u)\|_Y^2.$$

- Can have $\Phi(u_n; y) \to 0$ but $u_n$ does not converge in $X$.
- Tikhonov Regularization:

$$\min_{u \in E} I(u), \quad I(u) = \Phi(u; y) + \frac{1}{2}\|u\|_E^2.$$

# Problems With Least Squares

- Complex objective functions. Lorenz equations:




- How should $\|\cdot\|_Y, \|\cdot\|_E$ be chosen?

# Outline

# Bayesian Approach to Inverse Problems

- Jointly varying random variable $(u, y) \in X \times Y$.
- u prior $\mu_0(du) = \mathbb{P}(du)$ on $u$ :

$$\mu_0 = \mathcal{N}(0, \mathcal{C}_0), \quad \mu_0(X) = 1.$$

- y|u data $y \in Y$

$$y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathcal{N}(0, \Gamma).$$

- u|y posterior $\mu^y(du) = \mathbb{P}(du|y)$ on $u$ :

$$\frac{d\mu^y}{d\mu_0}(u) \propto \mathbb{P}(y|u) \propto \exp\Big(-\Phi(u; y)\Big).$$

# Increasing Number of Lagrangian Trajectories

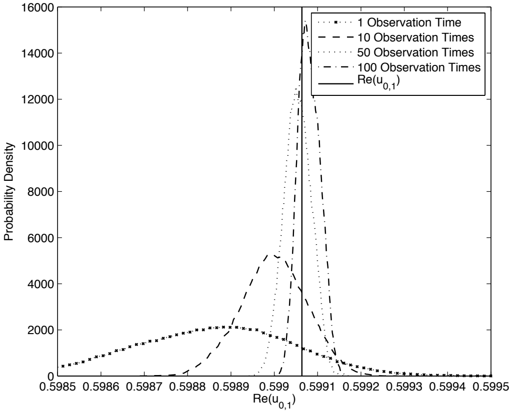



Figure: Posterior on Fourier mode for increasing number of Lagrangian trajectories.

# Potential

## Conditions on the Potential

- there is, $\forall \epsilon > 0, r > 0$ an $M = M(\epsilon, r) > 0$ such that

$$\Phi(u; y) \geq -\epsilon \|u\|_X^2 + M \quad \forall u \in X, \|y\|_Y < r;$$

- $\forall r > 0$ there is $K(r) > 0$ such that,

$$|\Phi(u_1; y) - \Phi(u_2; y)| \leq K(r) \|u_1 - u_2\|_X, \quad \forall \|u_i\|_X, \|y\|_Y < r;$$

- $\forall \epsilon > 0, r > 0$, there is $K = K(\epsilon, r) > 0$ such that,

$$|\Phi(u; y_1) - \Phi(u; y_2)| \leq K e^{\epsilon \|u\|_X^2} \|y_1 - y_2\|_Y, \ \forall u \in X, \|y_i\|_Y < r.$$

# Well-Posed Inverse Problem

S.L. Cotter, M. Dashti, J.C. Robinson and AMS:
Inverse Problems 2009.

**Theorem**

*Assume that* POTENTIAL CONDITIONS *hold and that* $\mu_0(X) = 1$. *There is* $C = C(r) > 0$ *such that, for all* $y_1, y_2$ *with* $\max\{\|y_1\|_Y, \|y_2\|_Y\} \le r$,

$$d_{Hell}(\mu^{y_1}, \mu^{y_2}) \le C\|y_1 - y_2\|_Y.$$

The metric $d_{Hell}$ is a useful one because:

$$\|\mathbb{E}^{\mu} f - \mathbb{E}^{\nu} f\| \le 2\Big(\mathbb{E}^{\mu}\|f\|^2 + \mathbb{E}^{\nu}\|f\|^2\Big)^{\frac{1}{2}} d_{Hell}(\mu, \nu).$$

# Outline

# Probability Maximizers and Optimal Control

M. Dashti and AMS 2010.

$$I(u) := \frac{1}{2}\|u\|_E^2 + \Phi(u;y).$$

Let $B^\delta(z)$ be a ball of radius $\delta$ in $X$ centred at $z \in E$.

**Theorem**

*Assume that* POTENTIAL CONDITIONS *hold and that* $\mu_0(X) = 1$. *Then*

$$\lim_{\delta \to 0} \frac{\mu^y\big(B^\delta(z_1)\big)}{\mu^y\big(B^\delta(z_2)\big)} = \exp\big(I(z_2) - I(z_1)\big).$$

Thus probability maximizers are minimizers of the regularized least squares functional $I$.

# Existence of Optimal Control Solutions

S.L. Cotter, M. Dashti, J.C. Robinson and AMS: Inverse Problems 2009.

The minimization is well-defined:

**Theorem**

*Assume that* POTENTIAL CONDITIONS *hold and that* $\mu_0(X) = 1$. *Then there exists* $\overline{u} \in E$ *such that*

$$I(\overline{u}) = \bar{I} := \inf\{I(u) : u \in E\}.$$

*Furthermore, if* $\{u_n\}$ *is a minimizing sequence satisfying* $I(u_n) \to I(\overline{u})$ *then there is a subsequence* $\{u_{n'}\}$ *that converges strongly to* $\overline{u}$ *in* $E$.

# Outline

# Forward Approximation Gives Inverse Approximation

S.L. Cotter, M. Dashti and AMS: SINUM 2010.

Consider two measures $\mu$ and $\mu^N$ :

$$\frac{d\mu}{d\mu_0}(u) \propto \exp\Big(-\Phi(u)\Big), \quad \frac{d\mu^N}{d\mu_0}(u) \propto \exp\Big(-\Phi^N(u)\Big).$$

## Theorem

*Assume that:* $\Phi$ *and* $\Phi^N$ *satisfy* POTENTIAL CONDITIONS, *that* $\mu_0(X) = 1$ *and that, for any* $\epsilon > 0$ *there is* $K = K(\epsilon) > 0$ :

$$|\Phi(u) - \Phi^N(u)| \le K \exp\big(\epsilon \|u\|_X^2\big)\psi(N) \tag{1}$$

*where* $\psi(N) \to 0$ *as* $N \to \infty$. *Then there is a constant* $C$, *independent of* $N$, *and such that*

$$d_{Hell}(\mu, \mu^N) \le C\psi(N). \tag{2}$$

# Data Assimilation in Oceanography

Let $\{\phi_k\}_{k\in\mathbb{K}}$ denote an orthonormal basis in $H$ made from eigenfunctions of the Stokes operator $A$. Let $P^N$ denote orthogonal projection, in $H$, into $\{\phi_k\}_{|k|<N}$.
Approximate forward problem by solving

$$\frac{dv}{dt} + \nu P^N A v = P^N f, \quad v(0) = P^N u$$

and assume particle trajectories are integrated exactly.

**Theorem**

*There is $\gamma > 0$, depending on covariance operator of the prior measure, for which*

$$d_{Hell}(\mu, \mu^N) \le cN^{-\gamma}.$$

# Increasing Approximation Dimension

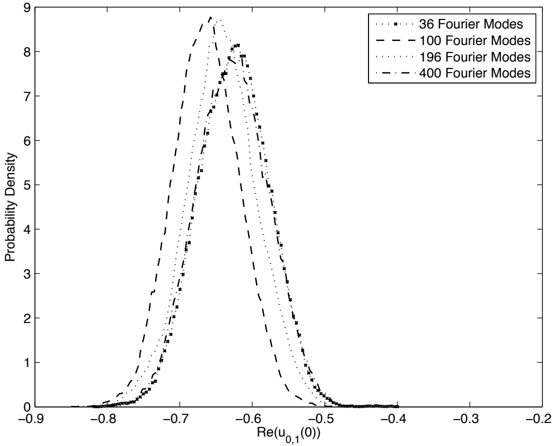



Figure: Posterior for increasing number of Fourier modes.

# Outline

# What We Have Shown

We have shown that:

- **Applications:** Many inverse problems in differential equations can be formulated in the framework of Bayesian statistics on function space.
- **Common Structure:** These problems share a common mathematical structure leading to *well-posed* inverse problems for measures and a link to *optimal control*.
- **Approximation:** This well-posedness leads to a transfer of approximation properties from the forward problem to the inverse problem, in the Hellinger metric.
- **Algorithms:** MCMC methods can be defined on function space: algorithms robust to discretization.

# References


- A.M. Stuart. ”Inverse Problems: A Bayesian Perspective.” Acta Numerica 19(2010).
- S.L.Cotter, M. Dashti, J.C.Robinson, A.M.Stuart. ”Bayesian inverse problems for functions and applications to fluid mechanics.” Inverse Problems, 25(2009), doi:10.1088/0266-5611/25/11/115008.
- S.L. Cotter, M. Dashti and A.M.Stuart. ”Approximation of Bayesian Inverse Problems”. SINUM, 48(2010), 322–345.
- F. Pinski and A.M. Stuart “Transition paths in molecules at finite temperature”. J. Chem. Phys. 132(2010), 184104.

# References (Continued)


- A. Beskos and A.M. Stuart. ”MCMC Methods for Sampling Function Space”. Invited Lectures, ICIAM 2007, Published by the European Mathematical Society, 2009.
- A. Beskos and A.M. Stuart. “Computational complexity of Metropolis-Hastings methods in high dimensions”. Plenary Lectures MCQMC2008, Springer, 2009.
- A. Beskos, G.O. Roberts and A.M. Stuart. ”Optimal scalings for local Metropolis-Hastings chains on non-product targets in high dimensions.” Ann. Appl. Prob. **19**(2009), 863–898.

# References (Continued)


- M. Hairer, A.M.Stuart and J. Voss. "Sampling the posterior: an approach to non-Gaussian data assimilation." PhysicaD, **230**(2007), 50–64.
- M. Hairer, A.M.Stuart, P. Wiberg and J. Voss. "Analysis of SPDEs Arising in Path Sampling. Part 1: The Gaussian Case." Comm. Math. Sci. 3(2005), 587–603
- M. Hairer, A.M.Stuart and J. Voss. "Analysis of SPDEs Arising in Path Sampling. Part 2: The Nonlinear Case." Ann. Appl. Prob. 17(2007), 1657–1706.
- For all papers see:

  *http://www.maths.warwick.ac.uk/~masdr/sample.html*

# The Heat Equation

- Let $\mathcal{H} = L^2(D)$, $A = -\triangle$ with $D(A) = H^2(D) \bigcap H_0^1(D)$.
- Consider the heat equation:

$$\frac{dv}{dt} + Av = 0, \quad v(0) = u$$

- Find $u$.
- Given noisy observations

$$y = v(1) + \eta.$$

# Data Assimilation in Atmospheric Sciences

- Let $H = \{u \in L^2(\mathbb{T}^2, \mathbb{R}^2) : \nabla \cdot u = 0, \int_{\mathbb{T}^2} u dx = 0\}$. Then Stokes operator is $A = -\triangle$ with $D(A) = H^2(\mathbb{T}^2, \mathbb{R}^2) \bigcap H$.
- Consider the Navier-Stokes equation:

$$\frac{dv}{dt} + \nu A v + B(v, v) = f, \quad v(0) = u$$

- Find $u$ (or $u$ and $f$).
- Given noisy *Eulerian observations*

$$y_{j,k} = v(x_j, t_k) + \eta_{j,k}$$

# Oil Recovery

- Consider the equations for transport in a porous medium:

$$
\begin{aligned}
-\nabla \cdot v &= h, \\
v &= -\lambda(S) \exp(u) \nabla p, \\
\frac{\partial S}{\partial t} &= -v \cdot \nabla f(S) + \nu \triangle S
\end{aligned}
$$

- Find log permeability $u(x)$.
- Given noisy observations

$$
y_j = 1 - \frac{\int_{\partial\Omega_{\text{out}}} f\big(S(x, t_j)\big) v_n d\ell}{\int_{\partial\Omega_{\text{out}}} v_n d\ell} + \eta_j.
$$

# The Heat Equation

- Because the observation operator $\mathcal{G}$ is linear, the posterior measure $\mu^y$ is Gaussian $\mathcal{N}(m, \mathcal{C})$.
- Assume $A, \mathcal{C}$ and $\Gamma$ all commute for simplicity.
- If we define $K = I + \exp(-2AT)\Gamma^{-1}\mathcal{C}_0$ then

$$m = \exp(-AT)\Gamma^{-1}\mathcal{C}_0 K^{-1} y$$
$$\mathcal{C} = \mathcal{C}_0 K^{-1}.$$

- Formally, in the limit $\Gamma \to 0$, we recover

$$m \to \exp(AT)y$$
$$\mathcal{C} \to 0.$$

# Heat Equation

Let $\{\phi_k\}_{k\in\mathbb{K}}$ denote an orthonormal basis for $\mathcal{H}$ made from eigenvalues of $A$. Let $P^N$ denote orthogonal projection, in $\mathcal{H}$, into $\{\phi_k\}_{|k|<N}$.
Approximate forward problem by solving

$$\frac{dv}{dt} + P^N A v = 0, \quad v(0) = P^N u.$$

## Theorem

$$d_{Hell}(\mu, \mu^N) \leq c_1 \exp(-c_2 N^2).$$

*Consequently the mean and covariance operator of $\mu^y$ and $\mu^N$ are $\mathcal{O}(\exp(-c_2 N^2))$ close in the $\mathcal{H}$ and $\mathcal{H}-$operator norms respectively.*

# Data Assimilation in Atmospheric Sciences

Let $\{\phi_k\}_{k\in\mathbb{K}}$ denote an orthonormal basis for Stokes operator $A$ in $H$. Let $P^N$ denote orthogonal projection, in $H$, into $\{\phi_k\}_{|k|<N}$. Approximate forward problem by solving

$$\frac{dv}{dt} + \nu P^N Av + P^N B(v, v) = P^N f, \quad v(0) = P^N u.$$

**Theorem**

*We have*

$$d_{Hell}(\mu, \mu^N) \to 0$$

*as $N \to \infty$.*

Constant in forward error does not exhibit the desired control.
Weakening of the preceding theory.