The Bayesian approach to inverse problems

Masoumeh Dashti

Department of Mathematics
University of Sussex

UNQW01: Key UQ methodologies and motivating applications

Isaac Newton Institute, January 11^{th} 2018

Inverse problems

Encountered when indirect measurements, y, of quantity of interest, u, is available

$$y \approx \mathcal{G}(u)$$

with $\mathcal{G}$ well-posed.

Inverse problems

Encountered when indirect measurements, y, of quantity of interest, u, is available

$$y \approx \mathcal{G}(u)$$

with $\mathcal{G}$ well-posed.

- $\mathcal{G}^{-1}$ might not be well-posed
- measurements are typically inaccurate, hence y might lie outside of range of $\mathcal{G}$

Example. Inverse heat equation

$$\begin{aligned} &u_t - u_{xx} = 0, \quad x \in (0, \pi), \quad t > 0 \\ &u(0, t) = u(\pi, t) = 0, \quad t \geq 0. \end{aligned}$$

Find $u_0(x) := u(x, 0)$, given measurements $y(x) \approx u(x, 1)$.

Example. Inverse heat equation

$$u_t - u_{xx} = 0, \quad x \in (0, \pi), \quad t > 0$$
$$u(0, t) = u(\pi, t) = 0, \quad t \geq 0.$$

Find $u_0(x) := u(x, 0)$, given measurements $y(x) \approx u(x, 1)$.

We have

$$u(x, t) = \sum_{n=1}^{\infty} c_n \, \mathrm{e}^{-n^2 t} \sin nx$$

for $u_0(x) = \sum_{n=1}^{\infty} c_n \sin nx$.

Example. Inverse heat equation

$$u_t - u_{xx} = 0, \quad x \in (0,\pi), \quad t > 0$$
$$u(0,t) = u(\pi,t) = 0, \quad t \geq 0.$$

Find $u_0(x) := u(x,0)$, given measurements $y(x) \approx u(x,1)$.

We have

$$u(x,t) = \sum_{n=1}^{\infty} c_n \, \mathrm{e}^{-n^2 t} \sin nx$$

for $u_0(x) = \sum_{n=1}^{\infty} c_n \sin nx$.

The above inverse problem is ill-posed: For

$$y(x) - u(x,1) = \alpha \sin 100x,$$

$$v_0(x) - u_0(x) = \alpha \, \mathrm{e}^{10^4} \sin 100x$$

Example. Inverse problem for an elliptic PDE:

$$-\nabla \cdot \left(e^{u(x)} \nabla p(x)\right) = f, \quad x \in D \subset \mathbb{R}^d$$
$$p(x) = 0, \quad x \in \partial D.$$

Find u, given

$$y_j = p(x_j) + \eta_j, \quad j = 1, \ldots, J$$

$$\mathcal{G}(u) = (p(x_1), \cdots, p(x_J))^T$$

- $\mathcal{G}$ captures both the forward model and the observation mechanism

Classical approach

aiming to find a reasonable estimate of the quantity of interest

- Least-squares solution

$$u^* = \arg\min_{u \in X} \|y - \mathcal{G}(u)\|_Y$$

- Regularised solution

$$u^* = \arg\min_{u \in X} \|y - \mathcal{G}(u)\|_Y^2 + \alpha \mathcal{J}(u)$$

(typically $\mathcal{J}(u) = \|u\|_E^p$ for $E \subseteq X$ and $p \geq 1$)

Bayesian approach

Let

$$y = \mathcal{G}(u) + \eta$$

with $u \in X$, $y \in \mathbb{R}^J$ (X separable Banach spaces), and suppose

- prior $u \sim \mu_0$
- statistics of noise is known: $\eta \sim \rho_\eta$

solution: posterior $\mathrm{Prob}(u \mid \text{data}) \propto \mathrm{Prob}(\text{data} \mid u)\, \mathrm{Prob}(u)$

When $\dim X = \infty$ implementation of Bayesian approach requires *high-dimensional discretization*

Do these scale well with respect to refining discretization?

When $\dim X = \infty$ implementation of Bayesian approach requires *high-dimensional discretization*

Do these scale well with respect to refining discretization?

- Discretization invariance theory

  M. Lassas and S. Siltanen. *Inverse Problems* (2004)

  M. Lassas, E. Saksman, and S. Siltanen. *Inverse Probl. Imaging* (2009)

When $\dim X = \infty$ implementation of Bayesian approach requires *high-dimensional discretization*

Do these scale well with respect to refining discretization?

- Discretization invariance theory

  M. Lassas and S. Siltanen. *Inverse Problems* (2004)

  M. Lassas, E. Saksman, and S. Siltanen. *Inverse Probl. Imaging* (2009)

- Establishing well-posedness of posterior on X itself

  A. M. Stuart. *Acta Numerica* (2010)

Outline

1. Well-posedness
2. Approximations and uncertainty quantification
3. Posterior consistency

Let

$$y = \mathcal{G}(u) + \eta$$

with $u \in X$, $y \in \mathbb{R}^J$ (X separable Banach spaces), and suppose

- prior $u \sim \mu_0$
- statistics of noise is known: $\eta \sim \rho_\eta$

solution: posterior $\mathrm{Prob}(\mathrm{u} \mid \mathsf{data}) \propto \mathrm{Prob}(\mathsf{data} \mid u)\, \mathrm{Prob}(u)$

Let

$$y = \mathcal{G}(u) + \eta$$

with $u \in X$, $y \in \mathbb{R}^J$ (X separable Banach spaces), and suppose

- prior $u \sim \mu_0$
- statistics of noise is known: $\eta \sim \rho_\eta$

solution: posterior $\mathrm{Prob}(\mathrm{u} \,|\, \text{data}) \propto \mathrm{Prob}(\text{data} \,|\, u)\, \mathrm{Prob}(u)$

Noting that for $y|u \sim \rho_\eta(y - \mathcal{G}(u))$,

posterior μ^y (*when well-defined*) satisfies

$$\mu^y(\mathrm{d}u) \propto \rho_\eta(y - \mathcal{G}(u))\, \mu_0(\mathrm{d}u).$$

Well-definedness

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(u) \propto \rho_\eta(y - \mathcal{G}(u)) =: \mathrm{e}^{-\Phi(u,y)}$$

ρ_η density of measure $\mathbb{Q}$ measure on $Y = \mathbb{R}^J$,
X separable Banach space

Well-definedness

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(u) \propto \rho_\eta(y - \mathcal{G}(u)) =: \mathrm{e}^{-\Phi(u,y)}$$

ρ_η density of measure $\mathbb{Q}$ measure on $Y = \mathbb{R}^J$,
X separable Banach space

Theorem. (A. Stuart 2010) Assume that

- $\mu_0(X) = 1$
- $\Phi : X \times \mathbb{R}^J \to \mathbb{R}$ is $\mu_0 \otimes \mathbb{Q}$ measurable
- $Z := \int_X \exp(-\Phi(u, y))\, \mathrm{d}\mu_0(u) < \infty$.

Then $\mu^y \ll \mu_0$ and

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(u) = \frac{1}{Z} \exp(-\Phi(u, y))$$

Stability in data

If $y, y' \in \mathbb{R}^J$ are close, are μ^y and $\mu^{y'}$ 'close'?

To measure the distance between probability measures we use

$$\textit{Hellinger metric:}\quad d^2_{\mathrm{Hell}}(\mu, \mu') = \frac{1}{2}\int_X \Big(\sqrt{\frac{d\mu}{d\nu}} - \sqrt{\frac{d\mu'}{d\nu}}\,\Big)^2 d\nu$$

Stability in data

If $y, y' \in \mathbb{R}^J$ are close, are μ^y and $\mu^{y'}$ 'close'?

To measure the distance between probability measures we use

$$\textit{Hellinger metric:}\;\; d^2_{\text{Hell}}(\mu, \mu') = \frac{1}{2}\int_X \Big(\sqrt{\frac{d\mu}{d\nu}} - \sqrt{\frac{d\mu'}{d\nu}}\,\Big)^2 d\nu$$

Theorem. (A. Stuart 2010) *Suppose also that for* $\max(|y_1|, |y_2|) < r$

$$|\Phi(u, y_1) - \Phi(u, y_2)| \;\leq\; M(r, \|u\|_X)\,|y_1 - y_2|$$

$$\textit{with}\quad \int_X M(r, \|u\|_X)\,\mathrm{d}\mu_0 < \infty$$

Then

$$d_{\text{Hell}}(\mu^{y_1}, \mu^{y_2}) \leq C|y_1 - y_2|$$

- With Hellinger metric

$$|\mathbb{E}^{\mu} f(u) - \mathbb{E}^{\mu'} f(u)| \leq C\, d_{\mathrm{Hell}}(\mu, \mu')$$

The prior

- $(X, \|\cdot\|)$ separable Banach space with basis $\{\psi_j\}_{j\in\mathbb{N}}$
- *Define μ_0 through Karhunen-Loève expansion of its draws:*

$$u(x) = \sum_{j\in\mathbb{N}} \alpha_j \, \xi_j \, \psi_j(x)$$

$\{\psi_j\}$ orthonormal basis in L^2,

ξ_j i.i.d random variables,

$\{\alpha_j\}$ decreasing sequence determining smoothness of u

The prior

- $(X, \|\cdot\|)$ separable Banach space with basis $\{\psi_j\}_{j\in\mathbb{N}}$
- *Define μ_0 through Karhunen-Loève expansion of its draws:*

$$u(x) = \sum_{j\in\mathbb{N}} \alpha_j \, \xi_j \, \psi_j(x)$$

$\{\psi_j\}$ orthonormal basis in L^2,

ξ_j i.i.d random variables,

$\{\alpha_j\}$ decreasing sequence determining smoothness of u

Gaussian	Besov (Lassas, Saksman, Siltanen 09)
$\xi_j \sim c \exp(-\frac{1}{2}\|x\|^2)$	$\xi_j \sim c_q \exp(-\|x\|^q), \quad q \geq 1$
$\{\psi_j\}$ an orthonormal basis	$\{\psi_j\}$ orthonormal wavelet basis

$q = 1$ especially interesting

Besov priors:

- based on wavelet expansions
- appropriate when we expect u to be
  smooth with a few local irregularities
- promotes sparsity

Besov priors:

- based on wavelet expansions
- appropriate when we expect u to be
  smooth with a few local irregularities
- promotes sparsity

Idea rooted in statistical literature

I. Johnstone. *Statistical Decision Theory and Related Topics V* (1994)

E. Candes and D. Donoho. *Tech report, Stanford Dept of Statistics* (2000)

D. Donoho and M. Elad. *PNAS* (2003), ...

also central in compressed sensing

I. Daubechies et. al. *Comm. Pure Appl. Math* 57(11) (2004)

Example. Inverse problem for an elliptic PDE:

$$-\nabla \cdot \left(e^{u(x)}\,\nabla p(x)\right) = f, \quad x \in D \subset \mathbb{R}^d$$
$$p(x) = 0, \quad x \in \partial D.$$

Find u, given $y_j = p(x_j) + \eta_j$, $j = 1, \ldots, J$, $\eta_j \sim \mathcal{N}(0, \gamma_j)$

$\mathcal{G}(u) = (p(x_1), \cdots, p(x_J))^T \qquad \Phi(u, y) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$

Example. Inverse problem for an elliptic PDE:

$$-\nabla \cdot \left(\mathrm{e}^{u(x)} \nabla p(x)\right) = f, \quad x \in D \subset \mathbb{R}^d$$
$$p(x) = 0, \quad x \in \partial D.$$

Find u, given $y_j = p(x_j) + \eta_j,\ j = 1, \ldots, J, \quad \eta_j \sim \mathcal{N}(0, \gamma_j)$

$\mathcal{G}(u) = (p(x_1), \cdots, p(x_J))^T \qquad \Phi(u, y) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$

Let $X = L^\infty$ and choose μ_0 such that $\mu_0(X) = 1$:

e.g.:

- $\mu_0 \sim \mathcal{N}(0, (-\Delta)^{-s})$, $s > d/2$ gives $\mu_0(C^t) = 1$, $t < s - d/2$

Example. Inverse problem for an elliptic PDE:

$$-\nabla \cdot \left(e^{u(x)} \nabla p(x)\right) = f, \quad x \in D \subset \mathbb{R}^d$$
$$p(x) = 0, \quad x \in \partial D.$$

Find u, given $y_j = p(x_j) + \eta_j$, $j = 1, \ldots, J$, $\eta_j \sim \mathcal{N}(0, \gamma_j)$

$\mathcal{G}(u) = (p(x_1), \cdots, p(x_J))^T \qquad \Phi(u, y) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$

Let $X = L^\infty$ and choose μ_0 such that $\mu_0(X) = 1$:

e.g.:

- $\mu_0 \sim \mathcal{N}(0, (-\Delta)^{-s})$, $s > d/2$ gives $\mu_0(C^t) = 1$, $t < s - d/2$
- $\mu_0 \sim B_1^s$-Besov measure with $s > d$ defined by KL expansion $u = \sum_j j^{-\frac{s}{d}+\frac{1}{2}} \xi_j \psi_j$ gives $\mu_0(C^t) = 1$, $t < s - d$

M. D., S. Harris, A. Stuart. *Inverse Probl. Imaging* (2012)

We have

$$|\Phi(u_1, y) - \Phi(u_2, y)| \leq L_{u_1, u_2} \, \|u_1 - u_2\|_{L^\infty},$$

hence Φ is measurable wrt $\mu_0 \times \mathbb{Q}$ and μ with

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0} \propto \exp(-\Phi(u, y))$$

is well-posed.

Approximations and uncertainty quantification

Consider μ and μ^N:

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0}(u) \propto \exp(-\Phi(u)), \qquad \frac{\mathrm{d}\mu^N}{\mathrm{d}\mu_0}(u) \propto \exp(-\Phi^N(u))$$

Consider μ and μ^N:

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0}(u) \propto \exp(-\Phi(u)), \qquad \frac{\mathrm{d}\mu^N}{\mathrm{d}\mu_0}(u) \propto \exp(-\Phi^N(u))$$

Theorem. (A. Stuart 2010) *Assume that*

$$|\Phi(u) - \Phi^N(u)| \;\leq\; M(\|u\|_X)\,\psi(N)$$

where $\displaystyle\int_X M(r, \|u\|_X)\,\mathrm{d}\mu_0 < \infty$ *and* $\psi(N) \to 0$ *as* $N \to \infty$

Then

$$d_{\mathrm{Hell}}(\mu, \mu^N) \leq C\,\psi(N).$$

Finite dimensional approximation

- $(X, \|\cdot\|)$ Hilbert space with orthonormal basis $\{\psi_j\}_{j\in\mathbb{N}}$
- $P^N : X \to \mathbb{R}^n$ orthogonal projection onto $W = \mathrm{span}\big\{\{\psi_\mathrm{j}\}_{\mathrm{j}=1}^{\mathrm{N}}\big\}$
- $\Phi^N(u) = \Phi(P^N u)$ (with $|\Phi - \Phi^N| \leq M\psi(N)$)

Finite dimensional approximation

- $(X, \|\cdot\|)$ Hilbert space with orthonormal basis $\{\psi_j\}_{j\in\mathbb{N}}$
- $P^N : X \to \mathbb{R}^n$ orthogonal projection onto $W = \mathrm{span}\{\{\psi_j\}_{j=1}^N\}$
- $\Phi^N(u) = \Phi(P^N u)$ $\quad$ (with $|\Phi - \Phi^N| \le M\psi(N)$)
- Constructing μ_0 from KL expansion using $\{\psi_j\}$ we get

$$\mu_0 = \mu_0^N \otimes \mu_0^\perp$$

- Then $\mu^N = \nu^N \otimes \mu_0^\perp$ with

$$\frac{\mathrm{d}\nu^N}{\mathrm{d}\mu_0^N}(u) \propto \exp\left(-\Phi^N(u)\right), \quad u \in W$$

Finite dimensional approximation

- $(X, \|\cdot\|)$ Hilbert space with orthonormal basis $\{\psi_j\}_{j\in\mathbb{N}}$
- $P^N : X \to \mathbb{R}^n$ orthogonal projection onto $W = \mathrm{span}\{\{\psi_j\}_{j=1}^N\}$
- $\Phi^N(u) = \Phi(P^N u)$ (with $|\Phi - \Phi^N| \leq M\psi(N)$)
- Constructing μ_0 from KL expansion using $\{\psi_j\}$ we get

$$\mu_0 = \mu_0^N \otimes \mu_0^\perp$$

- Then $\mu^N = \nu^N \otimes \mu_0^\perp$ with

$$\frac{\mathrm{d}\nu^N}{\mathrm{d}\mu_0^N}(u) \propto \exp\left(-\Phi^N(u)\right), \quad u \in W$$

$$\|\mathbb{E}^\mu F(u) - \mathbb{E}^{\nu^N} F(u^N, 0)\|_S \leq C\,\psi(N), \qquad (F : X \to S)$$

M. D., A. Stuart. *SIAM J. Num. Anal.* (2011)

Posterior measure: updated prior (using data)

- improved estimation of our degree of info./uncertainty in input
  results in *improved quantification of uncertainty*
- Weak error estimates we just found are part of this analysis

For e.g., for our elliptic problem

$$-\nabla \cdot \left(\mathrm{e}^{u(x)} \nabla p(x)\right) = f, \quad x \in D \subset \mathbb{R}^d$$
$$p(x) = 0, \quad x \in \partial D.$$

with

- observations $y_j = p(x_j) + \eta_j,\ j = 1, \dots, J, \quad \eta_j \sim \mathcal{N}(0, \gamma_j)$
- $\mu_0 \sim B_1^s$-Besov measure with $s > d$ constructed using the wavelet basis $\{\psi_j\}_{j \in \mathbb{N}}$ and with $W = \mathrm{span}\{\{\psi_j\}_{j=1}^N\}$.

For e.g., for our elliptic problem

$$-\nabla \cdot \left(\mathrm{e}^{u(x)} \nabla p(x)\right) = f, \quad x \in D \subset \mathbb{R}^d$$
$$p(x) = 0, \quad x \in \partial D.$$

with

- observations $y_j = p(x_j) + \eta_j,\ j = 1, \ldots, J, \quad \eta_j \sim \mathcal{N}(0, \gamma_j)$
- $\mu_0 \sim B_1^s$-Besov measure with $s > d$ constructed using the wavelet basis $\{\psi_j\}_{j\in\mathbb{N}}$ and with $W = \mathrm{span}\big\{\{\psi_\mathrm{j}\}_{\mathrm{j}=1}^{\mathrm{N}}\big\}$.

We have

$$\|\mathbb{E}^{\mu} p - \mathbb{E}^{\nu^N} p^N\|_{H^1} \ \le\ C\, N^{-t/d}$$
$$\|\mathbb{E}^{\mu}(p - \bar{p}) \otimes (p - \bar{p})\ -\ \mathbb{E}^{\nu^N}(p^N - \bar{p}^N) \otimes (p^N - \bar{p}^N)\|_S \ \le\ C\, N^{-t/d}$$

M. D., S. Harris, A. Stuart. *Inverse Probl. Imaging* (2012)

For μ_0 Gaussian:

M.D., A. Stuart *SIAM J. Num. Anal.(2011)*

Posterior consistency

Posterior consistency

We look for an estimate of the underlying true $u^\dagger$ from:

$$y_j = \mathcal{G}(u^\dagger) + \eta_j, \quad j = 1, \ldots, n$$

with $y_j \in \mathbb{R}^K$,

$\mathcal{G} : X \to \mathbb{R}^K$, and $\eta_j \sim \mathcal{N}(0, \Gamma)$, i.i.d.

Posterior consistency

We look for an estimate of the underlying true $u^\dagger$ from:

$$y_j = \mathcal{G}(u^\dagger) + \eta_j, \quad j = 1, \ldots, n$$

with $y_j \in \mathbb{R}^K$,

$$\mathcal{G} : X \to \mathbb{R}^K, \quad \text{and} \quad \eta_j \sim \mathcal{N}(0, \Gamma), \quad \text{i.i.d.}$$

Given $\mu_0 \sim \mathcal{N}(0, \mathcal{C}_0)$ we have

$$\frac{\mathrm{d}\mu^{y_1,\ldots,y_n}}{\mathrm{d}\mu_0}(u) \propto \exp\left(-\frac{1}{2}\sum_{j=1}^{n} \left|y_j - \mathcal{G}(u)\right|_\Gamma^2\right).$$

Posterior consistency

We look for an estimate of the underlying true $u^\dagger$ from:

$$y_j = \mathcal{G}(u^\dagger) + \eta_j, \quad j = 1, \ldots, n$$

with $y_j \in \mathbb{R}^K$,

$$\mathcal{G} : X \to \mathbb{R}^K, \quad \text{and} \quad \eta_j \sim \mathcal{N}(0, \Gamma), \quad \text{i.i.d.}$$

Given $\mu_0 \sim \mathcal{N}(0, \mathcal{C}_0)$ we have

$$\frac{\mathrm{d}\mu^{y_1,\ldots,y_n}}{\mathrm{d}\mu_0}(u) \propto \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\left|y_j - \mathcal{G}(u)\right|_\Gamma^2\right).$$

As n increases, does $\mu^{y_1,\ldots,y_n}$ gives ‘better’ estimates of $u^\dagger$?

It is reasonable to look at *most likely functions* under the posterior

(*weak consistency*)

MAP estimators

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0}(u) \propto \exp(-\Phi(u)), \quad \mu_0 \sim \mathcal{N}(0, \mathcal{C}_0)$$

The finite-dimensional case is straight-forward:
If $X = \mathbb{R}^d$, MAP estimators are maximisers of the density function

$$\rho_\mu = c \exp\left(-\Phi(u) - \frac{1}{2}|\mathcal{C}_0^{-1/2} u|^2\right)$$

which is maximised at minimisers of

$$I(u) := \Phi(u) + \frac{1}{2}|\mathcal{C}_0^{-1/2} u|^2$$

For infinite-dimensional space X:

let $B^\delta(z)$ be a ball of radius δ and centre z in X.

- Fix δ and find z^δ such that $B^\delta(z^\delta)$ has maximal probability
- Look at the 'limit' of $\{z^\delta\}_\delta$ as δ shrinks to zero

For infinite-dimensional space X:

let $B^\delta(z)$ be a ball of radius δ and centre z in X.

- Fix δ and find z^δ such that $B^\delta(z^\delta)$ has maximal probability
- Look at the 'limit' of $\{z^\delta\}_\delta$ as δ shrinks to zero

Definition. (M.D., K.Law, A.Stuart, J.Voss 2013) *Let*

$$z^\delta = \operatorname{argmax}_{z \in X} \mu(B^\delta(z)).$$

Any $\tilde{z} \in X$ satisfying

$$\lim_{\delta \to 0} \frac{\mu(B^\delta(\tilde{z}))}{\mu(B^\delta(z^\delta))} = 1,$$

is a MAP estimator.

'weak' MAP estimators: T. Helin, M. Burger *Inverse Problems* (2016)

Theorem. (M.D., K.Law, A.Stuart, J.Voss 2013)

Let μ satisfy

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0} \propto \exp(-\Phi(u)),$$

with

$\mu_0 \sim \mathcal{N}(0, \mathcal{C}_0)$ *and* $\Phi : X \to \mathbb{R}$ *bounded below and locally Lipschitz.*

Then MAP estimators are characterised by the minimisers of

$$I(u) = \begin{cases} \Phi(u) + \frac{1}{2}\|u\|_E^2 & \text{if } u \in E, \\ +\infty & \text{otherwise,} \end{cases}$$

where E is the space of admissible shifts of μ_0.

Also true for Besov priors: A. Agapiou, M.D., T. Helin, M. Burger (2017)

Posterier weak consistency – large data

$$y_j = \mathcal{G}(u^\dagger) + \eta_j, \quad j = 1, \dots, n$$

$$\frac{\mathrm{d}\mu^{y_1,\dots,y_n}}{\mathrm{d}\mu_0}(u) \propto \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\left|y_j - \mathcal{G}(u)\right|_\Gamma^2\right); \quad \mu_0 \sim \mathcal{N}(0, \mathcal{C}_0)$$

MAP estimators: $u_n := \underset{E}{\operatorname{argmin}} \|u\|_E^2 + \sum_{j=1}^{n} |y_j - \mathcal{G}(u)|_{\mathcal{C}_1}^2.$

Theorem. (M.D., K.Law, A.Stuart, J.Voss 2013)

Assume that

$u^\dagger \in E$.

Then $\exists u^* \in E$ *and a subseq of* $\{u_n\}_{n\in\mathbb{N}}$ *such that*

$$u_n \xrightarrow{in\ E} u^*, \quad a.s. \quad \left(hence\ u_n \xrightarrow{in\ X} u^*\right)$$

For any such u^* *we have* $\mathcal{G}(u^*) = \mathcal{G}(u^\dagger)$.

- Posterior consistency with contraction rates

  linear and Gaussian well-understood

  B. Knapik et al, *The Annals of Statistics* (2011)

  A. Agapiou et al, *Stoch. Process. Appl.* (2013)

  linear and non-Gaussian being developed

  K. Ray, *Electron. J. Stat* (2013),

  B. Knapik & J.-B. Salomond (2017)

  nonlinear only some partial results:

  Y. Pokern et al, *Stochastic Process. Appl.* (2013)

  S. Vollmer, *Inverse Problems* (2013)

  J. van Waajj, H. van Zanten et al, *Electron. J. Stat* (2016)