Bayesian inverse problems with Laplacian noise

Remo Kretschmann

Faculty of Mathematics, University of Duisburg-Essen

Graz, 28 September 2017
Classical setting

- $X$, $Y$ separable Hilbert spaces,
- $F: X \to Y$.

Given observed data $y \in Y$ find unknown $u \in X$, where

$$y = F(u) + \eta$$

with observational noise $\eta \in Y$. 
Bayesian approach

- $X, Y$ separable Hilbert spaces,
- $F: X \rightarrow Y$,
- probability measures $\mu_0$ on $(X, \mathcal{B}(X))$, $Q_0$ on $(Y, \mathcal{B}(Y))$,
- prior $u \sim \mu_0$, noise $\eta \sim Q_0$, $\eta$ independent of $u$ and

$$y = F(u) + \eta.$$ 

Given observed data $y \in Y$ find posterior distribution $\mu^y$, the conditional distribution of $u|y$.

- Extract information out of $\mu^y$ in the form of estimators.
Bayesian inverse problems in function spaces:

- [Dashti, Law, Stuart, Voss 2013]: Nonlinear inverse problems with Gaussian prior and noise that satisfies certain conditions. In this case, the MAP estimator can be described as the minimiser of the Onsager-Machlup functional.

- [Dashti, Stuart 2015]: Inverse heat equation with Gaussian noise and different priors (i.a. Gaussian).
Bayesian inverse problems in function spaces:

- [Dashti, Law, Stuart, Voss 2013]: Nonlinear inverse problems with Gaussian prior and noise that satisfies certain conditions. In this case, the MAP estimator can be described as the minimiser of the Onsager-Machlup functional.
- [Dashti, Stuart 2015]: Inverse heat equation with Gaussian noise and different priors (i.a. Gaussian).

Questions:

- What happens if the prior is Gaussian but the noise is non-Gaussian?
- Does Laplacian noise lead to an $\ell^1$-discrepancy term?
Motivation

- We study the **inverse heat equation** with **Laplacian noise** in combination with a **Gaussian prior**.

- **Problem**: Laplacian noise violates the conditions of [Dashti et al 2013].

- Existence of a solution?

- Connection: MAP estimator – optimisation problem?

- Does the MAP estimator converge towards the true solution, as the variance of the noise tends to zero?

- Behaviour of the CM estimator?
An example

- $X = Y = \mathbb{R}$, $F(u) = \frac{1}{2} u$, so that

$$y = \frac{1}{2} u + \eta,$$

- centred Gaussian prior $u$ with variance $2$,
- centred Laplacian noise $\eta$ with variance $\frac{1}{2}$. 
An example

- $X = Y = \mathbb{R}, \ F(u) = \frac{1}{2} u$, so that
  
  $\ y = \frac{1}{2} u + \eta,$

- centred Gaussian prior $u$ with variance 2,
- centred Laplacian noise $\eta$ with variance $\frac{1}{2}$.
- Joint probability density of $(u, y)$,
  
  $\ p_{(u,y)}(u, y) = p_u(u)p_{y|u}(y) = p_u(u)p_{\eta}(y - \frac{1}{2} u),$

- conditional probability density of $u$ given $y$,
  
  $\ p_{u|y}(u) = \frac{p_u(u)p_{y|u}(y)}{p_y(y)} = \frac{p_u(u)p_{y|u}(y)}{\int_{\mathbb{R}} p_u(u)p_{y|u}(y)du}.$
Prior and noise distribution

Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
Conditional distribution of $y$ given $u$

$$p_{y|u}(y) = p_\eta(y - \frac{1}{2}u)$$
Joint distribution of \((u, y)\)

\[ p(u, y)(u, y) = p_u(u)p_{y|u}(y) \]
Posterior distribution

\[ p_{u|y}(u) = \frac{1}{\int_{\mathbb{R}} p_{u}(u)p_{y|u}(y)du} p_{u}(u)p_{y|u}(y) \]

Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
Inverse heat equation Bayesian inversion MAP and CM estimators Numerical results
Outline

1. Inverse heat equation and Laplacian measures
2. Bayesian inversion
3. Maximum a posteriori and conditional mean estimators
4. Numerical results
The heat conduction equation

- $D \subset \mathbb{R}^d$ bounded domain, $\partial D \in C^k$ for some $k \geq 1$,
- $A := -\Delta$ defined on $\mathcal{D}(A) = H^2(D) \cap H^1_0(D)$.

For every $u \in L^2(D)$ there is a unique solution

$$v \in C([0, \infty), L^2(D)) \cap C^1((0, \infty), \mathcal{D}(A))$$

of the heat equation on $D$ with Dirichlet boundary conditions,

$$\begin{cases}
\frac{dv}{dt}(t) = -Av(t) & \text{for } t > 0, \\
v(0) = u,
\end{cases}$$

given by

$$v(t) = \exp(-At)u \quad \text{for all } t \geq 0.$$
The inverse problem (outline)

- Fix $t = 1$, i.e. $F(u) = v(1) = e^{-A}u$.

Given temperature measurement $y$ at time $t = 1$, find initial temperature $u \in L^2(D)$ at time $t = 0$, where

$$y = e^{-A}u + \eta.$$
Given temperature measurement $y$ at time $t = 1$, find conditional distribution of the **posterior** $u|y$, where

$$y = e^{-A}u + \eta.$$  

We assume that

- $-A$ is a Laplace-like operator,
- the noise $\eta$ has a **centred Laplacian distribution** with covariance operator $A^{s-\beta}$, and
- the prior $u$ has a **centred Gaussian distribution** with covariance operator $A^{-\tau}$. 

**Inverse heat equation**  
**Bayesian inversion**  
**MAP and CM estimators**  
**Numerical results**
Laplace-like operators

We assume that the operator $A$ in $L^2(D)$ satisfies the following properties:

1. The eigenvectors $\{\varphi_k\}_{k \in \mathbb{N}}$ of $A$ form an orthonormal basis of $L^2(D)$.
2. The respective eigenvalues $a_1 \geq a_2 \geq \cdots > 0$ of $A$ satisfy
   \[
   \frac{1}{C_A} k^\frac{2}{d} \leq a_k \leq C_A k^\frac{2}{d} \quad \text{for all } k \in \mathbb{N}
   \]
   and a constant $C_A > 1$.
3. $A$ is densely defined and surjective.
4. $A$ is self-adjoint.
A induces a **Hilbert scale** \( \{ \mathcal{H}^s \}_{s \in \mathbb{R}} \), where

\[
\mathcal{H}^s := A^{-s}(L^2(D)) = \left\{ u \in L^2(D) : \sum_{k=1}^{\infty} \alpha_{k}^{2s} |(u, \varphi_k)_{L^2}|^2 < \infty \right\}
\]

for all \( s \geq 0 \), equipped with

\[
\| u \|_{\mathcal{H}^s} := \| A^{s/2} u \|_{L^2} \quad \text{and} \quad (u, v)_{\mathcal{H}^s} := (A^{s/2} u, A^{s/2} v)_{L^2}.
\]

Now we set \( X := L^2(D) = \mathcal{H}^0 \) and \( Y := \mathcal{H}^s \) with \( s \geq 0 \), i.e.,

\[
u \in L^2(D) \quad \text{and} \quad \eta, y \in \mathcal{H}^s.
\]
Standard Laplacian measure on $\mathbb{R}$
For $a \in \mathbb{R}$ and $\lambda > 0$ define probability measure $\mathcal{L}_{a,\lambda}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by
\[
\mathcal{L}_{a,\lambda}(B) = \frac{1}{\sqrt{2\lambda}} \int_B e^{-\frac{\sqrt{2}|x-a|}{\sqrt{\lambda}}} \, dx \quad \text{for all } B \in \mathcal{B}(\mathbb{R}).
\]
Then $\mathcal{L}_{a,\lambda}$ has mean $a$ and variance $\lambda$, i.e.,
\[
\int_{\mathbb{R}} x \mathcal{L}_{a,\lambda}(dx) = a, \quad \int_{\mathbb{R}} (x-a)^2 \mathcal{L}_{a,\lambda}(dx) = \lambda.
\]
Laplacian measure on a Hilbert space

- Gaussian measure $\mu$ on Hilbert space $H$ defined by the property, that for every $h \in H$, the pushforward $\mu \circ p_h^{-1}$ under the projection $p_h = (\cdot, h)_H$ is a Gaussian measure on $\mathbb{R}$.

- Gaussian measure on separable Hilbert space can be constructed as infinite dimensional product measure of Gaussian measures on $\mathbb{R}$ [Da Prato 2001].
Laplacian measure on a Hilbert space

- Gaussian measure $\mu$ on Hilbert space $H$ defined by the property, that for every $h \in H$, the pushforward $\mu \circ p_h^{-1}$ under the projection $p_h = (\cdot, h)_H$ is a Gaussian measure on $\mathbb{R}$.

- Gaussian measure on separable Hilbert space can be constructed as infinite dimensional product measure of Gaussian measures on $\mathbb{R}$ [Da Prato 2001].

- We will construct a Laplacian measure on a separable Hilbert space $H$ as **infinite dimensional product measure** of Laplacian measures on $\mathbb{R}$.

- For given $a \in H$ and $Q \in \mathcal{L}(H)$ positive definite trace class operator we want to define a Laplacian measure $\mathcal{L}_{a,Q}$ with with mean $a$ and covariance operator $Q$. 

Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
Infinite-dimensional product measure

- $H$ separable real Hilbert space.
- For every compact self-adjoint operator $Q$ on $H$ there is an orthonormal basis $\{e_k\}_{k \in \mathbb{N}}$ of $H$ consisting of eigenvectors of $Q$.
- Identify $H$ with $\ell^2$ by means of $x \mapsto \{(x, e_k)_H\}_{k \in \mathbb{N}}$.

For any $a \in H$ and any positive definite trace class operator $Q \in \mathcal{L}(H)$ define **Laplacian measure** $\mathcal{L}_{a,Q}$ on $(\ell^2, \mathcal{B}(\ell^2))$ as the product measure

$$
\mathcal{L}_{a,Q} = \bigotimes_{k=1}^{\infty} \mathcal{L}_{a_k,\lambda_k},
$$

with $a_k := (a, e_k)_H$ and $\lambda_k := (Qe_k, e_k)_H$ for all $k \in \mathbb{N}$.
Infinite-dimensional product measure

- $H$ separable real Hilbert space.
- For every compact self-adjoint operator $Q$ on $H$ there is an orthonormal basis $\{e_k\}_{k \in \mathbb{N}}$ of $H$ consisting of eigenvectors of $Q$.
- Identify $H$ with $\ell^2$ by means of $x \mapsto \{(x, e_k)_H\}_{k \in \mathbb{N}}$.

For any $a \in H$ and any positive definite trace class operator $Q \in \mathcal{L}(H)$ define **Laplacian measure** $\mathcal{L}_{a,Q}$ on $(\ell^2, \mathcal{B}(\ell^2))$ as the product measure

$$\mathcal{L}_{a,Q} = \bigotimes_{k=1}^{\infty} \mathcal{L}_{a_k, \lambda_k},$$

with $a_k := (a, e_k)_H$ and $\lambda_k := (Qe_k, e_k)_H$ for all $k \in \mathbb{N}$.

Caution: This definition depends on the choice of the basis $\{e_k\}_{k \in \mathbb{N}}$. 

Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
Basic properties

- This way, $\mathcal{L}_{a,Q}$ has mean $a$ and covariance operator $Q$, i.e.,

$$\int_H (x, y)_H \mathcal{L}_{a,Q}(dx) = (a, y)_H \quad \text{for all } y \in H,$$

$$\int_H (x - a, y)_H (x - a, z)_H \mathcal{L}_{a,Q}(dx) = (Qy, z)_H \quad \text{for all } y, z \in H.$$

- Equivalently, we can define a random variable $x \sim \mathcal{L}_{a,Q}$ by

$$x := a + \sum_{k=1}^{\infty} \sqrt{\lambda_k} x_k e_k,$$

where $\{x_k\}_{k \in \mathbb{N}}$ are i.i.d. with $x_k \sim \mathcal{L}_{0,1}$ and $\lambda_k := (Q e_k, e_k)_H$.

- In case $a = 0$ we write $\mathcal{L}_Q := \mathcal{L}_{0,Q}$. 

Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
The Bayesian inverse problem

Given \( y \in Y \), find conditional distribution of \( u|y \) on \( X \), where

- **noise** \( \eta \sim \mathcal{L}_{A^{s-\beta}} \) with **Laplacian measure** \( \mathcal{L}_{A^{s-\beta}} \) on \( Y := \mathcal{H}^s \)
  using basis \( e_k := a_k^{-\frac{s}{2}} \varphi_k \) and \( 0 \leq s < \beta - \frac{d}{2} \),

- **prior** \( u \sim \mathcal{N}_{A^{-\tau}} \) independent from \( \eta \) with **Gaussian measure** \( \mathcal{N}_{A^{-\tau}} \) on \( X := \mathcal{L}^2(D) = \mathcal{H}^0 \) and \( \tau > \frac{d}{2} \),

- \( y = e^{-A} u + \eta \).

**Idea:** Use Bayes’ Theorem to obtain posterior distribution.
Outline

1. Inverse heat equation and Laplacian measures

2. Bayesian inversion

3. Maximum a posteriori and conditional mean estimators

4. Numerical results
Bayes’ Theorem

- \((X, A), (Y, B)\) measurable spaces,
- \(\nu, \nu_0\) probability measures on \(X \times Y\), such that \(\nu \ll \nu_0\), i.e., \(\nu\) is absolutely continuous with respect to \(\nu_0\).
- Then \(\nu\) has a density \(f = \frac{d\nu}{d\nu_0}\) with respect to \(\nu_0\), i.e., \(\nu = f \nu_0\).

Theorem (Bayes)

Assume that the conditional random variable \(x|y\) exists under \(\nu_0\) with probability distribution \(\nu_0^y\) on \(X\). Then the conditional random variable \(x|y\) exists under \(\nu\) with probability distribution \(\nu^y\) on \(X\), and \(\nu^y \ll \nu_0^y\). If additionally, \(Z(y) := \int_X \frac{d\nu}{d\nu_0}(x, y) \nu_0^y(dx) > 0\), then

\[
\frac{d\nu^y}{d\nu_0^y}(x) = \frac{1}{Z(y)} \frac{d\nu}{d\nu_0}(x, y).
\]
In our case, \((u, \eta) \sim \nu_0\) and \((u, y) \sim \nu\) on \(X \times Y = L^2(D) \times \mathcal{H}^s\).

In order for \(\nu \ll \nu_0\) to hold, we require \(\mathcal{L}_{e^{-A}u, A^{s-\beta}} \ll \mathcal{L}_{A^{s-\beta}}\) for all \(u \in X\).

Then by Bayes’ Theorem, the **posterior measure** \(\mu^y\) of \(u|y\) is absolutely continuous with respect to the prior measure \(\mathcal{N}_{A^{-\tau}}\) with the density

\[
\frac{d\mu^y}{d\mathcal{N}_{A^{-\tau}}}(u) = \frac{1}{Z(y)} \exp(-\Phi(u, y))\quad \nu_0\text{-a.e.},
\]

\[
\Phi(u, y) = \sqrt{2} \sum_{k=1}^{\infty} \beta \left( |y_k - e^{-\alpha_k}u_k| - |y_k| \right),
\]

where \(y_k := (y, \varphi_k)_X\), \(u_k := (u, \varphi_k)_X\).
Admissible shifts

- $H$ separable Hilbert space, $Q \in \mathcal{L}(H)$ positive definite trace class operator.

**Theorem**

1. If $a \notin Q^\frac{1}{2}(H)$ then $\mathcal{L}_{a,Q}$ and $\mathcal{L}_Q$ are singular.

2. If $a \in Q^\frac{1}{2}(H)$ then $\mathcal{L}_{a,Q}$ and $\mathcal{L}_Q$ are equivalent ($\mathcal{L}_{a,Q} \ll \mathcal{L}_Q$ and $\mathcal{L}_Q \ll \mathcal{L}_{a,Q}$) and

$$
\frac{d\mathcal{L}_{a,Q}}{d\mathcal{L}_Q}(y) = \exp \left( -\sqrt{2} \sum_{k=1}^{\infty} \frac{|y_k - a_k| - |y_k|}{\sqrt{\lambda_k}} \right) \mathcal{L}_Q\text{-a.e.,}
$$

where $y_k := (y, e_k)_H$, $a_k := (a, e_k)_H$ and $\lambda_k = (Q e_k, e_k)_H$.

**Idea of proof:** Apply Kakutani’s Theorem.
Admissible shifts (2)

- In our case \((H = Y = \mathcal{H}_s, a = e^{-A}u, Q = A^{s-\beta})\),

\[
Q^{\frac{1}{2}}(H) = A^{\frac{s-\beta}{2}}(\mathcal{H}_s) = \mathcal{H}^\beta
\]

and \(e^{-A}u \in \mathcal{H}^\beta\) is true for all \(u \in X = L^2(D)\).

- So \(\mathcal{L}_{e^{-A}u, A^{s-\beta}}\) and \(\mathcal{L}_{A^{s-\beta}}\) are equivalent for all \(u \in X\), and

\[
\frac{d\mathcal{L}_{e^{-A}u, A^{s-\beta}}}{d\mathcal{L}_{A^{s-\beta}}}(y) = \exp(-\Phi(u, y)) \quad \mathcal{L}_{A^{s-\beta}}\text{-a.e.}
\]

with the potential \(\Phi: X \times Y \rightarrow \mathbb{R}\),

\[
\Phi(u, y) := \sqrt{2} \sum_{k=1}^{\infty} a_k^{\frac{\beta}{2}} \left( |y_k - e^{-\alpha_k} u_k| - |y_k| \right),
\]

where \(y_k := (y, \varphi_k)_X\), \(u_k := (u, \varphi_k)_X\).
1. Inverse heat equation and Laplacian measures

2. Bayesian inversion

3. Maximum a posteriori and conditional mean estimators

4. Numerical results
Let $\mu$ be a probability measure on a separable Hilbert space $X$. Any point $\hat{u} \in X$ satisfying

$$\lim_{\varepsilon \to 0} \frac{\mu(B_\varepsilon(\hat{u}))}{\sup_{u \in X} \mu(B_\varepsilon(u))} = 1$$

is called a maximum a posteriori (MAP) estimator for $\mu$. 
$I: E \to \mathbb{R}$ is called **Onsager-Machlup functional** for $\mu$, if

$$
\lim_{\epsilon \to 0} \frac{\mu(B_\epsilon(u))}{\mu(B_\epsilon(v))} = \exp(I(v) - I(u))
$$

for all $u, v \in E$, where $E \subseteq X$ denotes the space of all admissible shifts that yield an equivalent measure.
Onsager-Machlup functional

$I: \mathcal{E} \to \mathbb{R}$ is called **Onsager-Machlup functional** for $\mu$, if

$$
\lim_{\varepsilon \to 0} \frac{\mu(B_\varepsilon(u))}{\mu(B_\varepsilon(v))} = \exp(I(v) - I(u))
$$

for all $u, v \in \mathcal{E}$, where $\mathcal{E} \subseteq X$ denotes the space of all admissible shifts that yield an equivalent measure.

- For a centered Gaussian measure $\mathcal{N}_Q$ on $X$, $\mathcal{E} = Q^{1/2}(X)$ and

$$
I(u) = \frac{1}{2} \| Q^{-1/2} u \|^2_X \quad \text{for all } u \in \mathcal{E}
$$
Characterisation of MAP estimators

- \( \mu_0 \) centred Gaussian measure on \( X \), \( \mu^y \) posterior measure on \( X \) with \( \frac{d\mu^y}{d\mu_0}(u) = \exp(-\Phi(u)) \) \( \mu_0 \)-a.e., \( \Phi: X \to \mathbb{R} \),

- \( X \) separable Banach space, \( E \subseteq X \) space of admissible shifts for \( \mu_0, \mu^y \) that yield an equivalent measure.

**Theorem [Dashti et al 2013]**

Assume that

1. \( \Phi \) is bounded from below,
2. \( \Phi \) is locally bounded from above,
3. \( \Phi \) is locally Lipschitz continuous.

Then \( u \in E \) is a MAP estimator for \( \mu^y \) if and only if it minimises the Onsager-Machlup functional \( I \) for \( \mu^y \).
For $\mu_0 = \mathcal{N}_{\mathcal{A}^{-\tau}}$ and $\mu^\gamma$, the space of admissible shifts is given by

$$E = \mathcal{A}^{-\frac{\tau}{2}}(L^2(D)) = \mathcal{H}_\tau.$$

In our case, Onsager-Machlup functional $I: \mathcal{H}_\tau \to \mathbb{R}$ for $\mu^\gamma$,

$$I(u) := \Phi(u) + \frac{1}{2} \|u\|_{\mathcal{H}_\tau}^2$$

$$= \sqrt{2} \sum_{k=1}^{\infty} \alpha_k \frac{\beta}{2} (|y_k - e^{-\alpha_k} u_k| - |y_k|) + \frac{1}{2} \sum_{k=1}^{\infty} \alpha_k \tau |u_k|^2,$$

where $u_k := (u, \varphi_k)_{L^2}$ and $y_k := (y, \varphi_k)_{L^2}$. 

Problem: For Laplacian noise, $\Phi$ is not bounded from below.

Upside: $\Phi$ is globally Lipschitz continuous.
For $\mu_0 = \mathcal{N}_{A^{-\tau}}$ and $\mu^y$, the space of admissible shifts is given by

$$E = A^{-\frac{\tau}{2}}(L^2(D)) = \mathcal{H}^\tau.$$

In our case, Onsager-Machlup functional $I: \mathcal{H}^\tau \to \mathbb{R}$ for $\mu^y$,

$$I(u) := \Phi(u) + \frac{1}{2} \|u\|_{\mathcal{H}^\tau}^2$$

$$= \sqrt{2} \sum_{k=1}^{\infty} \alpha_k \frac{\beta}{2} (|y_k - e^{-\alpha_k} u_k| - |y_k|) + \frac{1}{2} \sum_{k=1}^{\infty} \alpha_k \tau |u_k|^2,$$

where $u_k := (u, \varphi_k)_{L^2}$ and $y_k := (y, \varphi_k)_{L^2}$.

**Problem:** For Laplacian noise, $\Phi$ is not bounded from below.

**Upside:** $\Phi$ is globally Lipschitz continuous.
Characterisation of MAP estimators (2)

- \( \mu_0 \) centred Gaussian measure on \( X \), \( \mu^y \) posterior measure on \( X \) with \( \frac{d\mu^y}{d\mu_0}(u) = \exp(-\Phi(u)) \) \( \mu_0 \)-a.e., \( \Phi: X \to \mathbb{R} \),
- \( X \) separable Hilbert space, \( E \subseteq X \) space of admissible shifts for \( \mu_0, \mu^y \) that yield an equivalent measure.

Theorem

Assume that \( \Phi \) is globally Lipschitz continuous. Then \( u \in E \) is a MAP estimator for \( \mu^y \) if and only if it minimises the Onsager-Machlup functional \( I \) for \( \mu^y \).
Characterisation of MAP estimators (3)

Idea of proof:
- Show that $\{u_\varepsilon\}_{\varepsilon > 0}$,

$$u_\varepsilon := \arg\max_{u \in X} \mu^y(B_\varepsilon(u)),$$

contains a subsequence $\{u_{\varepsilon_n}\}_{n \in \mathbb{N}}$ that converges in $X$ and its limit $u_0 \in E$ is both a MAP estimator for $\mu^y$ and a minimiser of $I$.

- Show that every MAP estimator $\hat{u} \in X$ also minimises $I$.

- Show that every minimiser $\tilde{u} \in E$ of $I$ also is a MAP estimator.
Consistency of the MAP estimator

- Does the MAP estimator converge towards the true solution, as the variance of the noise tends to zero?
- How to choose the variance of the prior appropriately?
Scaled distributions

- Noise distribution $\mathcal{L}_{b^2 A^{s-\beta}}$, $b > 0$
- Prior distribution $\mathcal{N}_{r^2 A^{-\tau}}$, $r > 0$.
- Associated Onsager-Machlup functional $I: \mathcal{H}^\tau \to \mathbb{R}$ for $y \in \mathcal{H}^s$,

$$I(u) = \frac{1}{b} \Phi(u) + \frac{1}{2r^2} \|u\|^2_{\mathcal{H}^\tau}.$$
Scaled distributions

- Noise distribution $\mathcal{L}_{b^2A^{s-\beta}}$, $b > 0$
- Prior distribution $\mathcal{N}_{r^2A^{-\tau}}$, $r > 0$.
- Associated Onsager-Machlup functional $I: \mathcal{H}^\tau \to \mathbb{R}$ for $y \in \mathcal{H}^s$,

$$I(u) = \frac{1}{b} \Phi(u) + \frac{1}{2r^2} \|u\|^2_{H^\tau}. $$

- By the previous theorem, the minimiser $\tilde{u}(y)$ of $I$ is a MAP estimator for $\mu^y$.
- Its components are

$$ (\tilde{u}(y), \varphi_k)_{L^2} = \max \{-R_k, \min \{e^{a_k}(y, \varphi_k)_{L^2}, R_k\} \}, $$

where

$$ R_k := \sqrt{2} \frac{r^2a_k^{-\tau}}{ba_k^{-2}} \exp(-a_k). $$

Inverse heat equation  Bayesian inversion  MAP and CM estimators  Numerical results
Frequentist consistency

- True solution $u^\dagger \in L^2(D)$ (fixed, no prior),
- positive sequences $\{b_n\}_{n \in \mathbb{N}}, \{r_n\}_{n \in \mathbb{N}}$ with $b_n \to 0$,
- Laplacian noise $\eta^n \in H^s$ with $\eta^n \sim \mathcal{L}_{b_n^2 A^{s-\beta}}$ and

$$y^n = e^{-A} u^\dagger + \eta^n.$$

- Let $u^n$ denote the respective minimisers of $I_n: H^{\tau} \to \mathbb{R}$,

$$I_n(u) := \frac{1}{b_n} \Phi(u, y^n) + \frac{1}{2r_n^2} \|u\|^2_{H^{\tau}}.$$
### Convergence in mean square

**Theorem**

If a \( w \in \mathcal{H}^{2-\beta} \cap L^2(D) \) with \( \| w \|_{\mathcal{H}^{2-\beta}} \leq \rho \) exists, such that

\[
    u^\dagger = e^{-A}w,
\]

and if \( C > 0 \) and \( N \in \mathbb{N} \) exist, such that

\[
    \rho \frac{1}{2} b_n^2 \leq r_n \leq C \frac{1}{2} b_n^2 \quad \text{for all } n \geq N,
\]

then

\[
    \mathbb{E} \left[ \| u^n - u^\dagger \|_{L^2}^2 \right] \leq 2C \text{Tr } A^{-\tau} b_n \quad \text{for all } n \geq N.
\]
Lemma

Let $\bar{u}(y)$ be a minimiser of $I: \mathcal{H}^\tau \rightarrow \mathbb{R}$, $I(u) = \frac{1}{b} \Phi(u) + \frac{1}{2r^2} \|u\|_{\mathcal{H}^\tau}^2$, and $y \sim \mathcal{L}_{e^{-A} u^\dagger, b^2 A^{s-\beta}}$. Then

$$\mathbb{E} \left[ \left\| (\bar{u} - u^\dagger, \varphi_k)_{L^2} \right\|^2 \right]$$

$$= \frac{1}{S_k^2} f \left( S_k \left| R_k + \left\| (u^\dagger, \varphi_k)_{L^2} \right\| \right) + \frac{1}{S_k^2} f \left( S_k \left| R_k - \left\| (u^\dagger, \varphi_k)_{L^2} \right\| \right)$$

$$+ \chi_{[0, \infty)} \left( \left\| (u^\dagger, \varphi_k)_{L^2} \right\| - R_k \right) \frac{1}{S_k^2} g \left( S_k \left| R_k - \left\| (u^\dagger, \varphi_k)_{L^2} \right\| \right),$$

where $f(t) := 1 - e^{-t} - te^{-t}$, $g(t) := t^2 - 2f(t)$,

$R_k := \sqrt{2r^2 \frac{b}{b} \alpha_k^{\frac{\beta}{2} - \tau}} e^{-\alpha_k}$ and $S_k := \sqrt{2} \frac{1}{b} \alpha_k^{\frac{\beta}{2}} e^{-\alpha_k}$.
Convergence in mean square (3)

Idea of proof:

- \( f(t) = 1 - e^{-t} - te^{-t} \leq 1 - e^{-t} \leq t \) for all \( t \geq 0 \).
- If \( |(u^\dagger, \varphi_k)_{L^2}| \leq R_k \), then

\[
\mathbb{E} \left[ |(\tilde{u} - u^\dagger, \varphi_k)_{L^2}|^2 \right] = \frac{1}{S_k^2} f \left( S_k \left( R_k + |(u^\dagger, \varphi_k)_{L^2}| \right) \right) \\
+ \frac{1}{S_k^2} f \left( S_k \left( R_k - |(u^\dagger, \varphi_k)_{L^2}| \right) \right) \leq \frac{2R_k}{S_k} = 2r^2 a_k^{-\tau}.
\]

- Indeed, the source condition and the choice of \( r_n \) ensure that

\[
|(u^\dagger, \varphi_k)_{L^2}| = a_k^{\frac{\beta}{2} - \tau} \left| (e^{-Aw}, a_k^{\frac{\beta}{2} - \tau} \varphi_k)_{H^2} \right| \\
\leq a_k^{\frac{\beta}{2} - \tau} e^{-a_k \rho} \leq \sqrt{2} a_k^{\frac{\beta}{2} - \tau} e^{-a_k} \frac{r_n^2}{b_n} = R_k.
\]
Convergence in mean square (4)

- By the choice of $r_n$,

$$\mathbb{E} \left[ \left| \left( u^n - u^\dagger, \varphi_k \right)_{L^2} \right|^2 \right] \leq 2r_n^2 a_k^{-\tau} \leq 2Cb_n a_k^{-\tau}$$

for all $n \geq N$.

- Summing up yields

$$\mathbb{E} \left[ \left\| u^n - u^\dagger \right\|^2_{L^2} \right] \leq 2Cb_n \text{Tr } A^{-\tau}$$

for all $n \geq N$. 
The conditional mean (CM) estimator $\hat{u}_{CM}$ is defined by

$$
\hat{u}_{CM}(y) := \mathbb{E}_{\mu^y} u = \int_{L^2(D)} u \, \mu^y(du) \quad \text{for all } y \in \mathcal{H}^s.
$$

Its components are

$$(\hat{u}_{CM}(y), \varphi_k)_{L^2} = R_k \frac{\text{erfcx}(\gamma_-) - \text{erfcx}(\gamma_+)}{\text{erfcx}(\gamma_-) + \text{erfcx}(\gamma_+)},$$

where $\text{erfcx}(x) := \exp(x^2) \text{erfc}(x)$, $\text{erfc}(x) := \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2} \, dt$,

$$R_k := \sqrt{\frac{r^2}{b}} \alpha_k^{\frac{b\tau}{2} - \tau} e^{-\alpha_k}, \quad \gamma_- := \frac{1}{\sqrt{2r}} \alpha_k^{\frac{\tau}{2}} \left( R_k - e^{\alpha_k(y, \varphi_k)_{L^2}} \right),$$

$$\gamma_+ := \frac{1}{\sqrt{2r}} \alpha_k^{\frac{\tau}{2}} \left( R_k + e^{\alpha_k(y, \varphi_k)_{L^2}} \right).$$
Outline

1. Inverse heat equation and Laplacian measures
2. Bayesian inversion
3. Maximum a posteriori and conditional mean estimators
4. Numerical results
Setting

- Set \( d = 1, \ D = [0, 1], \ X = Y = L^2(D) \).
- The operator
  \[
  A := -\Delta = -\frac{\partial^2}{\partial x^2}
  \]
  in \( L^2(D) \) has the eigenfunctions \( \{\phi_k\}_{k \in \mathbb{N}} \), given by
  \[
  \phi_k(x) = \sqrt{2} \sin(\pi k x) \quad \text{for all } x \in [0, 1],
  \]
  and the eigenvalues \( a_k = \pi^2 k^2 \) for all \( k \in \mathbb{N} \).
- We assume, that \( u \sim \mathcal{N}_{r^2 A^{-\tau}}, \ \eta \sim \mathcal{L}_{b^2 A^{-\beta}} \) and
  \[
  y = e^{-tA}u + \eta
  \]
  with \( r, b, t > 0 \).
- In the following, set \( \beta = 0.65, \ \tau = 0.55, \ t = 0.002 \).
Instead of exact data \( y \) we measure

\[
y^N := P_N y = \sum_{k=1}^{N} (y, \varphi_k)_{L^2} \varphi_k
\]

for some \( N \in \mathbb{N} \).

Then both

\[
\hat{u}_{\text{MAP}}(y^N) \to \hat{u}_{\text{MAP}}(y) \quad \text{and} \quad \hat{u}_{\text{CM}}(y^N) \to \hat{u}_{\text{CM}}(y)
\]

in \( L^2(D) \) as \( N \to \infty \).

In the following, set \( N = 180 \).
First look (scenario 1)

\[ \beta = 0.65, \quad \tau = 0.55, \quad t = 0.002, \quad N = 180, \quad b = 0.132, \]

source condition \( u^\dagger = e^{-tA}w \) with \( \|w\|_{H^{2\tau-\beta}} = \rho = 6.70, \)

\[ r = \frac{1}{\sqrt{2}} \rho^{\frac{1}{2}} b^{\frac{1}{2}} = 0.791 \] chosen a priori.

Inverse heat equation

Bayesian inversion

MAP and CM estimators

Numerical results
$\beta = 0.65, \tau = 0.55, t = 0.002, N = 180, b = 0.146, $ 

no source condition, $r = 2.18\sqrt{b} = 1.39.$
Consistency (scenario 1)

Mean squared error $\mathbb{E} \left[ \| \hat{u}(y^N) - u^\dagger \|_{L^2}^2 \right]$, 1000 noise samples, source condition, $r = \frac{1}{\sqrt{2}} \rho^\frac{1}{2} b^\frac{1}{2}$ chosen a priori.
Consistency (scenario 2)

Mean squared error $\text{IE} \left[ \| \hat{u}(y^N) - u^\dagger \|^2_{L^2} \right]$, 1000 noise samples, no source condition, $r = 2.18 \sqrt{b}$. 

Inverse heat equation Bayesian inversion MAP and CM estimators Numerical results
Conclusion

Bayesian inverse heat equation with Laplacian noise:

- The posterior distribution exists.
- Every minimiser of the Onsager-Machlup functional is a MAP estimator.
- The MAP estimator is consistent in a frequentist sense.

Outlook:

- Direct posterior sampling
- Variational characterisation of MAP estimators for different prior and noise