TUBAF

The University of Resources.
Since 1765.

Sensitivity of uncertainty propagation, risk, and Bayesian inversion

Joint work with Oliver Ernst and Alois Pichler (TU Chemnitz)

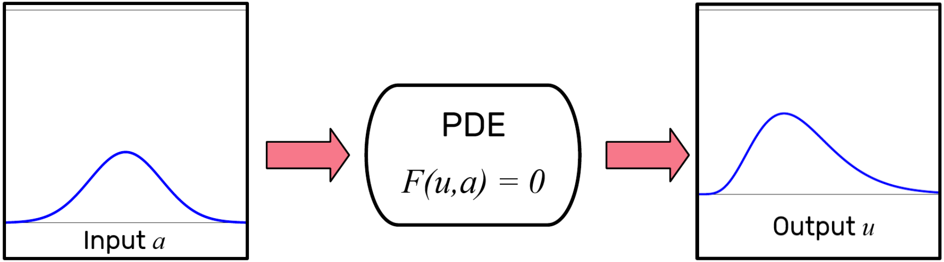
Björn Sprungk

Workshop “Uncertainty Quantification for High-Dimensional Problems”, CWI Amsterdam

November 13th, 2024

Uncertainty Propagation

UQ scheme

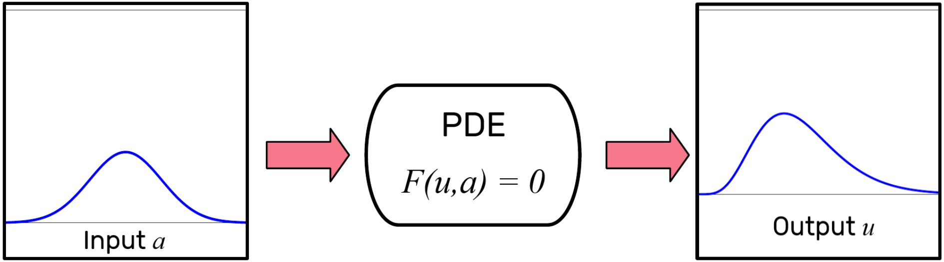


But: Distribution μ of a often obtained by estimation or subjective knowledge

Question

Can we control the effect of perturbations of μ on output distribution ν?

UQ scheme

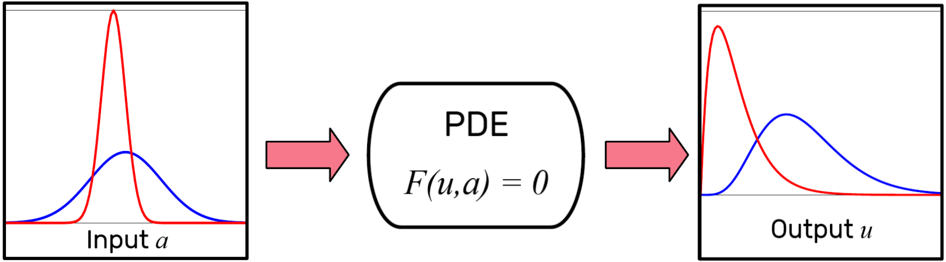


But: Distribution μ of a often obtained by estimation or subjective knowledge

Question

Can we control the effect of perturbations of μ on output distribution ν?

UQ scheme

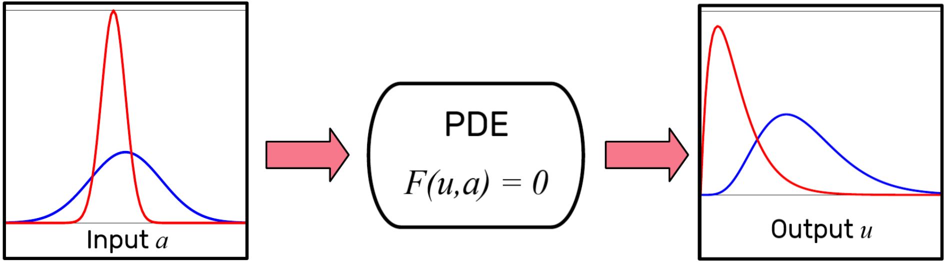


But: Distribution μ of *a* often obtained by estimation or subjective knowledge

Question

Can we control the effect of perturbations of μ on output distribution ν?

UQ scheme



But: Distribution μ of *a* often obtained by estimation or subjective knowledge

Question

Can we control the effect of perturbations of μ on output distribution ν?

Setting

- Consider general PDE with solution $u \in \mathscr{U}$ and uncertain coefficient(s) $a \in \mathscr{A}$ given by

$$\mathscr{F}(u, a) = 0$$

  with associated solution operator $S \colon \mathscr{A} \to \mathscr{U}$ mapping coefficient a to unique solution u

- **Running example**: Elliptic diffusion equation on compact $D \subseteq \mathbb{R}^2$,

$$-\nabla \cdot (\mathrm{e}^{a} \nabla u) = f, \qquad u\big|_{\partial D} \equiv 0$$

  with $\mathscr{U} = H_0^1(D)$ and $\mathscr{A} = L^\infty(D)$ (weak form)

- **UQ approach:**
  1. Describe uncertainty about a by probability measure μ on $\mathscr{A}$
  2. Compute pushforward distribution $S_*\mu$ on $\mathscr{U}$ of random solution $u = S(a)$, $a \sim \mu$

Setting

- Consider general PDE with solution $u \in \mathscr{U}$ and uncertain coefficient(s) $a \in \mathscr{A}$ given by

$$\mathscr{F}(u, a) = 0$$

  with associated solution operator $S \colon \mathscr{A} \to \mathscr{U}$ mapping coefficient a to unique solution u

- **Running example**: Elliptic diffusion equation on compact $D \subseteq \mathbb{R}^2$,

$$-\nabla \cdot (e^{a} \nabla u) = f, \qquad u\big|_{\partial D} \equiv 0$$

  with $\mathscr{U} = H_0^1(D)$ and $\mathscr{A} = L^\infty(D)$ (weak form)

- **UQ approach:**
  1. Describe uncertainty about a by probability measure μ on $\mathscr{A}$
  2. Compute pushforward distribution $S_*\mu$ on $\mathscr{U}$ of random solution $u = S(a)$, $a \sim \mu$

Gaussian random fields as coefficients

- Often we assume $a \sim \mathrm{N}(m, c)$ with mean and covariance function m and c

$$\mathrm{E}\left[a(x)\right] = m(x), \qquad \mathrm{Cov}[a(x), a(y)] = c(x, y), \qquad a(x) \sim \mathrm{N}\left(m(x), c(x, x)\right)$$

- Common parametrized class: Matérn covariance functions

$$c_{\sigma^2,\rho,k+\frac{1}{2}}(x, y) := \sigma^2 \, \mathrm{e}^{-\frac{\sqrt{2k+1}}{\rho}|x-y|} \, P_k\left(\frac{\sqrt{2k+1}}{\rho}|x - y|\right)$$

with variance $\sigma^2 > 0$, correlation length $\rho > 0$, smoothness $k + \frac{1}{2}$, $k \in \mathbb{N}_0$

Gaussian random fields as coefficients

- Often we assume $a \sim \mathrm{N}(m, c)$ with mean and covariance function m and c

$$\mathrm{E}\left[a(x)\right] = m(x), \qquad \mathrm{Cov}[a(x), a(y)] = c(x, y), \qquad a(x) \sim \mathrm{N}\left(m(x), c(x, x)\right)$$

- Common parametrized class: Matérn covariance functions

$$c_{\sigma^2, \rho, k+\frac{1}{2}}(x, y) := \sigma^2 \, \mathrm{e}^{-\frac{\sqrt{2k+1}}{\rho}|x-y|} \, P_k\left(\frac{\sqrt{2k+1}}{\rho}|x - y|\right)$$

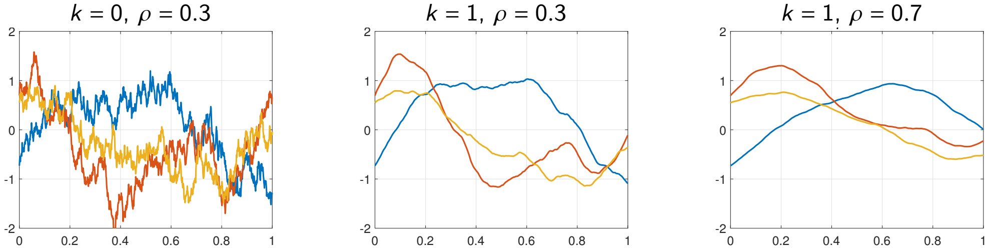
with variance $\sigma^2 > 0$, correlation length $\rho > 0$, smoothness $k + \frac{1}{2}$, $k \in \mathbb{N}_0$

Gaussian random fields as coefficients

- Often we assume $a \sim \mathrm{N}(m, c)$ with mean and covariance function m and c
- Common parametrized class: Matérn covariance functions

$$c_{\sigma^2,\rho,k+\frac{1}{2}}(x,y) := \sigma^2 \, e^{-\frac{\sqrt{2k+1}}{\rho}|x-y|} \, P_k\left(\frac{\sqrt{2k+1}}{\rho}|x-y|\right)$$

with variance $\sigma^2 > 0$, correlation length $\rho > 0$, smoothness $k + \frac{1}{2}$, $k \in \mathbb{N}_0$



Gaussian random fields as coefficients

- Often we assume $a \sim \mathrm{N}(m, c)$ with mean and covariance function m and c
- Common parametrized class: Matérn covariance functions

$$c_{\sigma^2,\rho,k+\frac{1}{2}}(x,y) := \sigma^2 \, \mathrm{e}^{-\frac{\sqrt{2k+1}}{\rho}|x-y|} \, P_k\left(\frac{\sqrt{2k+1}}{\rho}|x-y|\right)$$

with variance $\sigma^2 > 0$, correlation length $\rho > 0$, smoothness $k + \frac{1}{2}$, $k \in \mathbb{N}_0$

- **In practice:** Use estimates $\widehat{\sigma}^2$, $\widehat{k}$, $\widehat{\rho}$ given data $a(x_j), j = 1, \ldots, n$

Motivational Question

How does estimation error, e.g., for σ^2, affect the output of the UQ analysis?

Gaussian random fields as coefficients

- Often we assume $a \sim \mathrm{N}(m, c)$ with mean and covariance function m and c
- Common parametrized class: Matérn covariance functions

$$c_{\sigma^2,\rho,k+\frac{1}{2}}(x,y) := \sigma^2 \, \mathrm{e}^{-\frac{\sqrt{2k+1}}{\rho}|x-y|} \, P_k\left(\frac{\sqrt{2k+1}}{\rho}|x-y|\right)$$

with variance $\sigma^2 > 0$, correlation length $\rho > 0$, smoothness $k + \frac{1}{2}$, $k \in \mathbb{N}_0$

- **In practice:** Use estimates $\widehat{\sigma}^2$, $\widehat{k}$, $\widehat{\rho}$ given data $a(x_j), j = 1, \ldots, n$

Motivational Question

How does estimation error, e.g., for σ^2, affect the output of the UQ analysis?

Smoothness of forward map

- Often forward map S is (locally) Lipschitz: with monotonically increasing $\mathrm{Lip}_S : \mathbb{R}_+ \to \mathbb{R}_+$

$$\|S(a) - S(\widehat{a})\|_{\mathscr{U}} \le \mathrm{Lip}_S(r) \; \|a - \widehat{a}\|_{\mathscr{A}} \qquad \forall \|a\|_{\mathscr{A}}, \|\widehat{a}\|_{\mathscr{A}} \le r$$

- **Running example:** For elliptic problem $-\nabla \cdot (\mathrm{e}^a \nabla u) = f$ we have if $\|a\|_{L^\infty(D)}, \|\widehat{a}\|_{L^\infty(D)} \le r$

$$\|u - \widehat{u}\|_{H_0^1(D)} = \|S(a) - S(\widehat{a})\|_{H_0^1(D)} \le c_f \, \mathrm{e}^{3r} \; \|a - \widehat{a}\|_{L^\infty(D)}$$

- Does Lipschitz continuity of S yield Lipschitz continuity of

$$\mu \mapsto S_* \mu \; ?$$

Smoothness of forward map

- Often forward map S is (locally) Lipschitz: with monotonically increasing $\mathrm{Lip}_S : \mathbb{R}_+ \to \mathbb{R}_+$

$$\|S(a) - S(\widehat{a})\|_{\mathscr{U}} \leq \mathrm{Lip}_S(r) \;\; \|a - \widehat{a}\|_{\mathscr{A}} \qquad \forall \|a\|_{\mathscr{A}}, \|\widehat{a}\|_{\mathscr{A}} \leq r$$

- **Running example:** For elliptic problem $-\nabla \cdot (e^a \nabla u) = f$ we have if $\|a\|_{L^\infty(D)}, \|\widehat{a}\|_{L^\infty(D)} \leq r$

$$\|u - \widehat{u}\|_{H_0^1(D)} = \|S(a) - S(\widehat{a})\|_{H_0^1(D)} \;\leq\; c_f \, e^{3r} \; \|a - \widehat{a}\|_{L^\infty(D)}$$

- Does Lipschitz continuity of S yield Lipschitz continuity of

$$\mu \;\mapsto\; S_* \mu \;?$$

Smoothness of forward map

- Often forward map S is (locally) Lipschitz: with monotonically increasing $\mathrm{Lip}_S : \mathbb{R}_+ \to \mathbb{R}_+$

$$\|S(a) - S(\widehat{a})\|_{\mathscr{U}} \leq \mathrm{Lip}_S(r) \;\; \|a - \widehat{a}\|_{\mathscr{A}} \qquad \forall \|a\|_{\mathscr{A}}, \|\widehat{a}\|_{\mathscr{A}} \leq r$$

- **Running example:** For elliptic problem $-\nabla \cdot (\mathrm{e}^a \nabla u) = f$ we have if $\|a\|_{L^\infty(D)}, \|\widehat{a}\|_{L^\infty(D)} \leq r$

$$\|u - \widehat{u}\|_{H_0^1(D)} = \|S(a) - S(\widehat{a})\|_{H_0^1(D)} \;\leq\; c_f \, \mathrm{e}^{3r} \;\; \|a - \widehat{a}\|_{L^\infty(D)}$$

- Does Lipschitz continuity of S yield Lipschitz continuity of

$$\mu \;\mapsto\; S_* \mu \;?$$

Suitable Metric

- Common choice: total variation distance

$$d_{\mathrm{TV}}(\mu, \widehat{\mu}) = \sup_{A \subseteq \mathscr{A}} |\mu(A) - \widehat{\mu}(A)| = \frac{1}{2} \sup_{f \colon \|f\|_{L^\infty} \le 1} \left| \mathrm{E}_{X \sim \mu}[f(X)] - \mathrm{E}_{X \sim \widehat{\mu}}[f(X)] \right|$$

- Then, we have

$$d_{\mathrm{TV}}(S_* \mu, S_* \widehat{\mu}) \le d_{\mathrm{TV}}(\mu, \widehat{\mu}), \qquad \text{for any measurable } S \colon \mathscr{A} \to \mathscr{U}$$

- **But:** In infinite dimensions, i.e., for Gaussian random fields $a \sim \mathrm{N}(m, c)$ we have, e.g.,

$$d_{\mathrm{TV}}\big(\mathrm{N}(m, c), \mathrm{N}(m, \sigma^2 c)\big) = 1 \qquad \text{if } \sigma \neq 1$$

Suitable Metric

- Common choice: total variation distance

$$d_{\mathrm{TV}}(\mu, \widehat{\mu}) = \sup_{A \subseteq \mathscr{A}} |\mu(A) - \widehat{\mu}(A)| = \frac{1}{2} \sup_{f \colon \|f\|_{L^\infty} \leq 1} \left| \mathrm{E}_{X \sim \mu}[f(X)] - \mathrm{E}_{X \sim \widehat{\mu}}[f(X)] \right|$$

- Then, we have

$$d_{\mathrm{TV}}(S_*\mu, S_*\widehat{\mu}) \leq d_{\mathrm{TV}}(\mu, \widehat{\mu}), \qquad \text{for any measurable } S \colon \mathscr{A} \to \mathscr{U}$$

- **But:** In infinite dimensions, i.e., for Gaussian random fields $a \sim \mathrm{N}(m, c)$ we have, e.g.,

$$d_{\mathrm{TV}}\big(\mathrm{N}(m, c), \mathrm{N}(m, \sigma^2 c)\big) = 1 \qquad \text{if } \sigma \neq 1$$

Suitable Metric

- Common choice: total variation distance

$$d_{\mathrm{TV}}(\mu, \widehat{\mu}) = \sup_{A \subseteq \mathscr{A}} |\mu(A) - \widehat{\mu}(A)| = \frac{1}{2} \sup_{f \colon \|f\|_{L^\infty} \leq 1} \left| \mathrm{E}_{X \sim \mu}[f(X)] - \mathrm{E}_{X \sim \widehat{\mu}}[f(X)] \right|$$

- Then, we have

$$d_{\mathrm{TV}}(S_* \mu, S_* \widehat{\mu}) \leq d_{\mathrm{TV}}(\mu, \widehat{\mu}), \qquad \text{for any measurable } S \colon \mathscr{A} \to \mathscr{U}$$

- **But:** In infinite dimensions, i.e., for Gaussian random fields $a \sim \mathrm{N}(m, c)$ we have, e.g.,

$$d_{\mathrm{TV}}\big(\mathrm{N}(m, c), \mathrm{N}(m, \sigma^2 c)\big) = 1 \qquad \text{if } \sigma \neq 1$$

Wasserstein distance

- Instead, we consider the p-Wasserstein distance, $p \geq 1$,

$$\mathrm{W}_p(\mu, \widehat{\mu}) := \inf_{X \sim \mu,\ \widehat{X} \sim \widehat{\mu}} \mathrm{E}\left[\|X - \widehat{X}\|^p\right]^{1/p}$$

- We have $\mathrm{W}_p(\mu, \widehat{\mu}) \leq \mathrm{W}_{p'}(\mu, \widehat{\mu})$ for $p \leq p'$ and

$$\mathrm{W}_1(\mu, \widehat{\mu}) = \sup_{f:\ \mathrm{Lip}_f \leq 1} \left|\mathrm{E}_{X \sim \mu}\left[f(X)\right] - \mathrm{E}_{X \sim \widehat{\mu}}\left[f(X)\right]\right|$$

- Also reasonable for measures which are singular w. r. t. each other, e. g., W_2-distance of Gaussian random field measures explicitly known [Gelbrich, 1990][1]

[1] M. Gelbrich. On a Formula for the L2 Wasserstein Metric between Measures on Euclidean and Hilbert Spaces. *Mathematische Nachrichten* **147(1)**:185-203, 1990.

Wasserstein distance

- Instead, we consider the p-Wasserstein distance, $p \geq 1$,

$$\mathrm{W}_p(\mu, \widehat{\mu}) \;:=\; \inf_{X \sim \mu,\ \widehat{X} \sim \widehat{\mu}} \mathrm{E}\left[\|X - \widehat{X}\|^p\right]^{1/p}$$

- We have $\mathrm{W}_p(\mu, \widehat{\mu}) \;\leq\; \mathrm{W}_{p'}(\mu, \widehat{\mu})$ for $p \leq p'$ and

$$\mathrm{W}_1(\mu, \widehat{\mu}) = \sup_{f\,:\ \mathrm{Lip}_f \leq 1} \left| \mathrm{E}_{X \sim \mu}\left[f(X)\right] - \mathrm{E}_{X \sim \widehat{\mu}}\left[f(X)\right] \right|$$

- Also reasonable for measures which are singular w. r. t. each other, e. g., W_2-distance of Gaussian random field measures explicitly known [Gelbrich, 1990][1]

[1] M. Gelbrich. On a Formula for the L2 Wasserstein Metric between Measures on Euclidean and Hilbert Spaces. *Mathematische Nachrichten* **147(1)**:185-203, 1990.

Wasserstein distance

- Instead, we consider the p-Wasserstein distance, $p \geq 1$,

$$\mathrm{W}_p(\mu, \widehat{\mu}) := \inf_{X \sim \mu,\ \widehat{X} \sim \widehat{\mu}} \mathrm{E}\left[\|X - \widehat{X}\|^p\right]^{1/p}$$

- We have $\mathrm{W}_p(\mu, \widehat{\mu}) \leq \mathrm{W}_{p'}(\mu, \widehat{\mu})$ for $p \leq p'$ and

$$\mathrm{W}_1(\mu, \widehat{\mu}) = \sup_{f:\ \mathrm{Lip}_f \leq 1} \left|\mathrm{E}_{X\sim\mu}\left[f(X)\right] - \mathrm{E}_{X\sim\widehat{\mu}}\left[f(X)\right]\right|$$

- Also reasonable for measures which are singular w. r. t. each other, e. g., W_2-distance of Gaussian random field measures explicitly known [Gelbrich, 1990][1]

[1] M. Gelbrich. On a Formula for the L2 Wasserstein Metric between Measures on Euclidean and Hilbert Spaces. *Mathematische Nachrichten* **147(1)**:185-203, 1990.

Theorem (Wasserstein sensitivity)

Let $S\colon \mathscr{A} \to \mathscr{U}$ be globally Lipschitz,

$$\|S(a) - S(\widehat{a})\|_{\mathscr{U}} \le \mathrm{Lip}_S \, \|a - \widehat{a}\|_{\mathscr{A}} \qquad \forall \|a\|_{\mathscr{A}}, \|\widehat{a}\|_{\mathscr{A}} < \infty.$$

Then for any $\mu, \widehat{\mu}$ we have

$$\mathrm{W}_p\left(S_*\mu, S_*\widehat{\mu}\right) \le \mathrm{Lip}_S \mathrm{W}_p\left(\mu, \widehat{\mu}\right), \qquad p \ge 1.$$

If S is not globally Lipschitz, then $\mu \mapsto S_*\mu$ as well: for $\mu = \mathrm{N}(0,1)$, $\mu_n = \mathrm{N}(n,1)$, $S(x) := e^x$, $x \in \mathbb{R}$,

$$\mathrm{W}_p(S_*\mu, S_*\mu_n)/\mathrm{W}_2(\mu, \mu_n) \xrightarrow{n \to +\infty} +\infty, \qquad p \ge 1$$

Theorem (Wasserstein sensitivity)

Let $S\colon \mathscr{A} \to \mathscr{U}$ be locally Lipschitz,

$$\|S(a) - S(\widehat{a})\|_{\mathscr{U}} \le \mathrm{Lip}_S(r)\ \|a - \widehat{a}\|_{\mathscr{A}} \qquad \forall \|a\|_{\mathscr{A}}, \|\widehat{a}\|_{\mathscr{A}} \le r.$$

Then for any $\mu, \widehat{\mu}$ with

$$\mathrm{E}_{\mu}\left[\mathrm{Lip}_S^{2p}(\|a\|_{\mathscr{A}})\right],\ \mathrm{E}_{\widehat{\mu}}\left[\mathrm{Lip}_S^{2p}(\|a\|_{\mathscr{A}})\right] \le C < \infty$$

we have

$$\mathrm{W}_p\left(S_*\mu, S_*\widehat{\mu}\right) \ \le\ 2C^{1/2p}\ \mathrm{W}_{2p}\left(\mu, \widehat{\mu}\right), \qquad p \ge 1.$$

If S is not globally Lipschitz, then $\mu \mapsto S_*\mu$ as well: for $\mu = \mathrm{N}(0,1)$, $\mu_n = \mathrm{N}(n,1)$, $S(x) := \mathrm{e}^x$, $x \in \mathbb{R}$,

$$\mathrm{W}_p(S_*\mu, S_*\mu_n)/\mathrm{W}_2(\mu,\mu_n) \xrightarrow{n\to+\infty} +\infty, \qquad p \ge 1$$

Theorem (Wasserstein sensitivity)

Let $S\colon \mathscr{A} \to \mathscr{U}$ be locally Lipschitz,

$$\|S(a) - S(\widehat{a})\|_{\mathscr{U}} \le \mathrm{Lip}_S(r)\ \|a - \widehat{a}\|_{\mathscr{A}} \qquad \forall \|a\|_{\mathscr{A}}, \|\widehat{a}\|_{\mathscr{A}} \le r.$$

Then for any $\mu, \widehat{\mu}$ with

$$\mathrm{E}_\mu\left[\mathrm{Lip}_S^{2p}(\|a\|_{\mathscr{A}})\right],\ \mathrm{E}_{\widehat{\mu}}\left[\mathrm{Lip}_S^{2p}(\|a\|_{\mathscr{A}})\right] \le C < \infty$$

we have

$$\mathrm{W}_p\left(S_*\mu, S_*\widehat{\mu}\right) \ \le\ 2C^{1/2p}\ \mathrm{W}_{2p}\left(\mu, \widehat{\mu}\right), \qquad p \ge 1.$$

If S is not globally Lipschitz, then $\mu \mapsto S_*\mu$ as well: for $\mu = \mathrm{N}(0,1)$, $\mu_n = \mathrm{N}(n,1)$, $S(x) := \mathrm{e}^x$, $x \in \mathbb{R}$,

$$\mathrm{W}_p(S_*\mu, S_*\mu_n)/\mathrm{W}_2(\mu, \mu_n) \xrightarrow{n\to+\infty} +\infty, \qquad p \ge 1$$

Special case: Gaussian random fields

- Consider **subclass $\mathscr{G}$ of Gaussian measures on** $C(D)$ **with Matérn covariance** $c_{\sigma^2,\rho,k+\frac{1}{2}}$

$$\mathscr{G} = \left\{ \mathrm{N}\left(m, c_{\sigma^2,\rho,k+\frac{1}{2}}\right) : \|m\|_{C(D)} \leq M, \quad \sigma \leq \sigma_{\max}, \quad \rho \geq \rho_{\min}, \quad k \in \{0, \ldots, k_{\max}\} \right\}$$

- Fernique's theorem + Dudley's entropy bound yield $\sup_{\mu \in \mathscr{G}} \mathrm{E}_{a\sim\mu}\left[\exp\left(\beta \|a\|_{C(D)}\right)\right] < \infty, \beta > 0$

Theorem

If $S\colon C(D) \to \mathscr{U}$ is locally Lipschitz with $\mathrm{Lip}_S(r) \in \mathscr{O}(e^{\beta r})$ for a $\beta > 0$, then, there exists a constant $C = C(\mathscr{G}) < \infty$ such that

$$\mathrm{W}_p\left(S_*\mu, S_*\widehat{\mu}\right) \leq C\ \mathrm{W}_{2p}\left(\mu, \widehat{\mu}\right) \qquad \forall \mu, \widehat{\mu} \in \mathscr{G}.$$

Special case: Gaussian random fields

- Consider **subclass $\mathscr{G}$ of Gaussian measures on $C(D)$ with Matérn covariance** $c_{\sigma^2,\rho,k+\frac{1}{2}}$

$$\mathscr{G} = \left\{ \mathrm{N}\left(m, c_{\sigma^2,\rho,k+\frac{1}{2}}\right) : \|m\|_{C(D)} \leq M, \quad \sigma \leq \sigma_{\max}, \quad \rho \geq \rho_{\min}, \quad k \in \{0, \ldots, k_{\max}\} \right\}$$

- Fernique's theorem + Dudley's entropy bound yield $\sup_{\mu \in \mathscr{G}} \mathrm{E}_{a \sim \mu}\left[\exp\left(\beta \, \|a\|_{C(D)}\right)\right] < \infty$, $\beta > 0$

Theorem

If $S\colon C(D) \to \mathscr{U}$ is locally Lipschitz with $\mathrm{Lip}_S(r) \in \mathscr{O}(\mathrm{e}^{\beta r})$ for a $\beta > 0$, then, there exists a constant $C = C(\mathscr{G}) < \infty$ such that

$$\mathrm{W}_p\left(S_*\mu, S_*\widehat{\mu}\right) \leq C \; \mathrm{W}_{2p}\left(\mu, \widehat{\mu}\right) \qquad \forall \mu, \widehat{\mu} \in \mathscr{G}.$$

Related work

- [Amir Sagev, 2020][2]: Upper and lower bounds for

$$\mathrm{W}_p\left(S_*\mu, \widetilde{S}_*\mu\right)$$

  for approximations $\widetilde{S}$ of forward map $S\colon \mathscr{A} \to \mathscr{U}$, e.g., numerical approximations

- [Owhadi et al., 2013][3]: Robust UQ in terms of upper and lower bounds on failure probabilities

$$\mathrm{P}_{a\sim\mu}(q(u) > \mathrm{tol})$$

  by varying (μ, q) or $(\mu, q \circ S)$ within an admissible set of input distributions and forward maps

[2] A. Sagiv. The Wasserstein distances between pushed-forward measures with applications to uncertainty quantification. *Commun. Math. Sci.* **18(3)**:707–724, 2020.

[3] H. Owhadi, C. Sovel, T. Sullivan, M. McKerns, M. Ortiz. Optimal Uncertainty Quantification. *SIAM Review* **55(2)**:271–345, 2013.

Risk

Risk Functionals

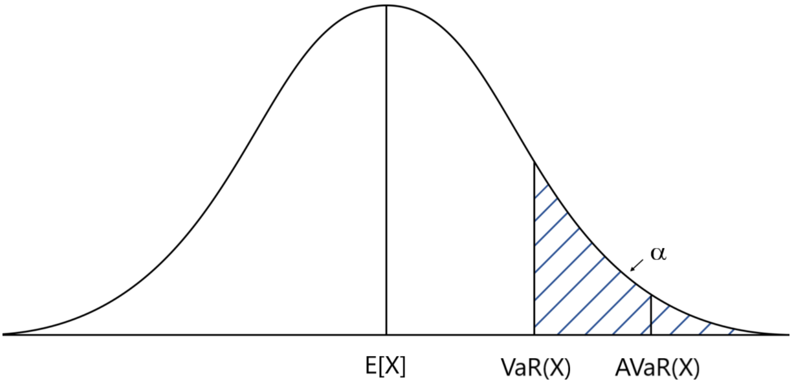
- Consider now scalar quantity of interest $q\colon \mathscr{U} \to \mathbb{R}$ of solution u of (random) PDE
- Risk functionals R assign real numbers $\mathrm{R}(X) \in \mathbb{R}$ to (real-valued) random variables X which quantify the risk associated with their random outcomes:

Expectation:	$\mathrm{R}(X) = \mathrm{E}[X]$	
Value-at-Risk (VaR):	$\mathrm{R}(X) := F_X^{-1}(1-\alpha),$	$\alpha \in (0,1)$
Average Value-at-Risk (AVaR):	$\mathrm{R}(X) = \dfrac{1}{\alpha}\displaystyle\int_{1-\alpha}^{1} F_X^{-1}(t)\mathrm{d}t,$	$\alpha \in (0,1)$
Spectral risk functional:	$\mathrm{R}(X) = \displaystyle\int_0^1 w(t)\, F_X^{-1}(t)\mathrm{d}t,$	$w \in L^1(\mathbb{R}_+)$

Risk Functionals

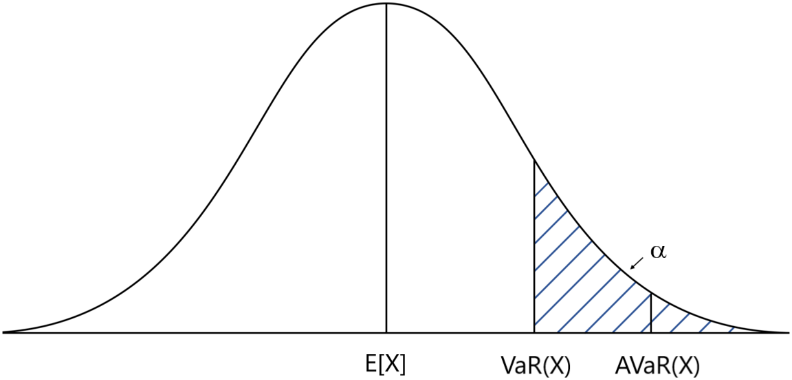
- Consider now scalar quantity of interest $q\colon \mathscr{U} \to \mathbb{R}$ of solution u of (random) PDE
- Risk functionals R assign real numbers $\mathrm{R}(X) \in \mathbb{R}$ to (real-valued) random variables X which quantify the risk associated with their random outcomes:

Expectation:	$\mathrm{R}(X) = \mathrm{E}\left[X\right]$	
Value-at-Risk (VaR):	$\mathrm{R}(X) := F_X^{-1}(1-\alpha),$	$\alpha \in (0,1)$
Average Value-at-Risk (AVaR):	$\mathrm{R}(X) = \dfrac{1}{\alpha}\displaystyle\int_{1-\alpha}^{1} F_X^{-1}(t)\mathrm{d}t,$	$\alpha \in (0,1)$
Spectral risk functional:	$\mathrm{R}(X) = \displaystyle\int_0^1 w(t)\; F_X^{-1}(t)\mathrm{d}t,$	$w \in L^1(\mathbb{R}_+)$



Question

Can we control the effect of perturbations on μ on the risk value $R(q(u))$, $u = S(a), a \sim \mu$?



Question

Can we control the effect of perturbations on μ on the risk value $\mathrm{R}(q(u))$, $\;u = S(a), a \sim \mu$?

Sensitivity of risk

For spectral risk functionals and larger class of coherent risk functionals [Artzner et al., 1997] we have dual representation

$$\mathrm{R}(X) = \sup_{Z \in \mathscr{Z}} \mathrm{E}\,[Z\, X]\,, \qquad \mathscr{Z} \subseteq \{Z \colon Z \geq 0 \text{ a. s. and } \mathrm{E}\,[Z] = 1\},$$

i.e., Z basically represent probability density functions

Theorem

For Hölder-continuous quantity $q \colon \mathscr{U} \to \mathbb{R}$, i.e., $|q(u) - q(\widehat{u})| \leq C_q \|u - \widehat{u}\|_{\mathscr{U}}^{\beta}$, $\beta > 0$, we have for any coherent risk functional R that

$$|\mathrm{R}(q(u)) - \mathrm{R}(q(\widehat{u}))| \leq C_{\mathrm{R},p,q}\, \mathrm{W}_p\,(\nu, \widehat{\nu})^{\beta}, \qquad p \geq 1$$

where $u \sim \nu$ and $\widehat{u} \sim \widehat{\nu}$.

Sensitivity of risk

For spectral risk functionals and larger class of coherent risk functionals [Artzner et al., 1997] we have dual representation

$$\mathrm{R}(X) = \sup_{Z \in \mathscr{Z}} \mathrm{E}\,[Z\ X]\,, \qquad \mathscr{Z} \subseteq \{Z \colon Z \geq 0 \text{ a. s. and } \mathrm{E}\,[Z] = 1\},$$

i.e., Z basically represent probability density functions

Theorem

For Hölder-continuous quantity $q \colon \mathscr{U} \to \mathbb{R}$, i.e., $|q(u) - q(\widehat{u})| \leq C_q \|u - \widehat{u}\|_{\mathscr{U}}^{\beta}$, $\beta > 0$, we have for any coherent risk functional R that

$$|\mathrm{R}(q(u)) - \mathrm{R}(q(\widehat{u}))| \leq C_{\mathrm{R},p,q}\ \mathrm{W}_p\,(\nu, \widehat{\nu})^{\beta}, \qquad p \geq 1$$

where $u \sim \nu$ and $\widehat{u} \sim \widehat{\nu}$.

Corollary

For Hölder-continuous $q\colon \mathscr{U} \to \mathbb{R}$ and locally Lipschitz $S\colon \mathscr{A} \to \mathscr{U}$ we have for any spectral risk measures R and suitable measures $\mu, \widehat{\mu}$ on $\mathscr{A}$

$$|\mathrm{R}(q(u)) - \mathrm{R}(q(\widehat{u}))| \leq C_{q,w,p}\, \mathrm{W}_{2p}\,(\mu, \widehat{\mu})^{\beta}\,, \qquad p \geq 1,$$

where $u = S(a)$, $a \sim \mu$, and $\widehat{u} = S(\widehat{a})$, $\widehat{a} \sim \widehat{\mu}$.

Example: For elliptic problem $-\nabla \cdot (e^{a} \nabla u) = f$ with lognormal diffusion coefficients we have for $a \sim \mathrm{N}(m, c_{\sigma^2,\rho,k+\frac{1}{2}})$, $\widehat{a} \sim \mathrm{N}(m, c_{\widehat{\sigma}^2,\rho,k+\frac{1}{2}})$

$$|\mathrm{AVaR}(q(u)) - \mathrm{AVaR}(q(\widehat{u}))| \leq C_{\sigma_{\max}}\, |\sigma - \widehat{\sigma}|^{\beta} \qquad \forall \sigma, \widehat{\sigma} \leq \sigma_{\max}$$

for Hölder-continuous $q\colon H_0^1(D) \to \mathbb{R}$

Corollary

For Hölder-continuous $q\colon \mathscr{U} \to \mathbb{R}$ and locally Lipschitz $S\colon \mathscr{A} \to \mathscr{U}$ we have for any spectral risk measures R and suitable measures $\mu, \widehat{\mu}$ on $\mathscr{A}$

$$|\mathrm{R}(q(u)) - \mathrm{R}(q(\widehat{u}))| \leq C_{q,w,p}\, \mathrm{W}_{2p}\,(\mu, \widehat{\mu})^{\beta}\,, \qquad p \geq 1,$$

where $u = S(a)$, $a \sim \mu$, and $\widehat{u} = S(\widehat{a})$, $\widehat{a} \sim \widehat{\mu}$.

Example: For elliptic problem $-\nabla \cdot (\mathrm{e}^{a} \nabla u) = f$ with lognormal diffusion coefficients we have for $a \sim \mathrm{N}(m, c_{\sigma^2,\rho,k+\frac{1}{2}})$, $\widehat{a} \sim \mathrm{N}(m, c_{\widehat{\sigma}^2,\rho,k+\frac{1}{2}})$

$$|\mathrm{AVaR}(q(u)) - \mathrm{AVaR}(q(\widehat{u}))| \leq C_{\sigma_{\max}}\, |\sigma - \widehat{\sigma}|^{\beta} \qquad \forall \sigma, \widehat{\sigma} \leq \sigma_{\max}$$

for Hölder-continuous $q\colon H_0^1(D) \to \mathbb{R}$

Bayesian Inference

Bayesian approach to inverse problem

$$y = G(a) + \varepsilon, \qquad G: \mathscr{A} \to \mathbb{R}^k, \qquad \varepsilon \sim \mathrm{N}(0, \Sigma),$$

- Condition prior measure $a \sim \mu$ given data $y = G(a) + \varepsilon$ yields posterior measure

Bayes' rule

$$\mu_\Phi(\mathrm{d}a) \;\propto\; \mathrm{e}^{-\Phi(a)}\,\mu(\mathrm{d}a), \qquad \Phi(a) := \frac{1}{2}\|y - G(a)\|^2_{\Sigma^{-1}}.$$

- BIP well-posed, i.e., local Lipschitz dependence of μ_Φ on data $y \in \mathbb{R}^k$ [Stuart, 2010], [Hosseini, 2017], [Sullivan, 2017], [Latz, 2020],...

- **Question:** How sensitively depends posterior μ_Φ on (subjective) choice of prior μ ?

Bayesian approach to inverse problem

$$y = G(a) + \varepsilon, \qquad G\colon \mathscr{A} \to \mathbb{R}^k, \qquad \varepsilon \sim \mathrm{N}(0, \Sigma),$$

- Condition prior measure $a \sim \mu$ given data $y = G(a) + \varepsilon$ yields posterior measure

Bayes' rule

$$\mu_\Phi(\mathrm{d}a) \;\propto\; \mathrm{e}^{-\Phi(a)}\,\mu(\mathrm{d}a), \qquad \Phi(a) := \frac{1}{2}\|y - G(a)\|^2_{\Sigma^{-1}}.$$

- BIP well-posed, i.e., local Lipschitz dependence of μ_Φ on data $y \in \mathbb{R}^k$ [Stuart, 2010], [Hosseini, 2017], [Sullivan, 2017], [Latz, 2020],...

- **Question:** How sensitively depends posterior μ_Φ on (subjective) choice of prior μ ?

Bayesian approach to inverse problem

$$y = G(a) + \varepsilon, \qquad G\colon \mathscr{A} \to \mathbb{R}^k, \qquad \varepsilon \sim \mathrm{N}(0, \Sigma),$$

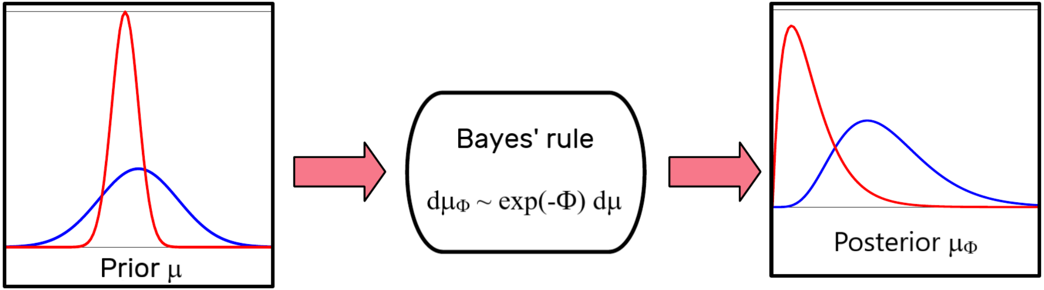
- Condition prior measure $a \sim \mu$ given data $y = G(a) + \varepsilon$ yields posterior measure

Bayes' rule

$$\mu_\Phi(\mathrm{d}a) \;\propto\; \mathrm{e}^{-\Phi(a)}\,\mu(\mathrm{d}a), \qquad \Phi(a) := \frac{1}{2}\|y - G(a)\|^2_{\Sigma^{-1}}.$$

- BIP well-posed, i.e., local Lipschitz dependence of μ_Φ on data $y \in \mathbb{R}^k$ [Stuart, 2010], [Hosseini, 2017], [Sullivan, 2017], [Latz, 2020],...

- **Question:** How sensitively depends posterior μ_Φ on (subjective) choice of prior μ ?

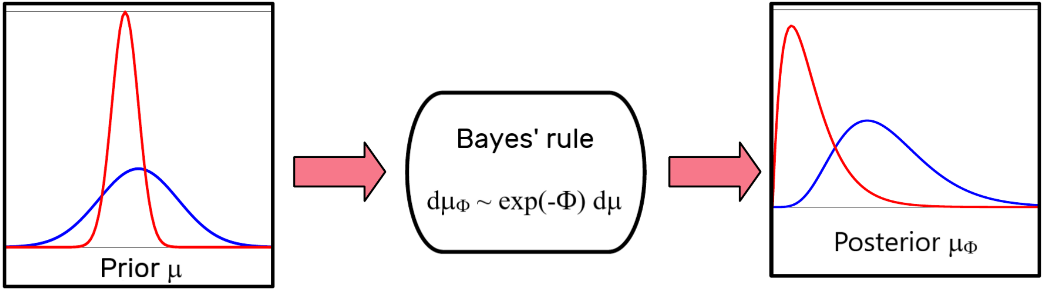


Theorem (informal)

For d being TV, Hellinger, or 1-Wasserstein distance or KL divergence we have under suitable assumptions a locally Lipschitz continuity:

$$d(\mu_\Phi, \widehat{\mu}_\Phi) \;\le\; C_\Phi(r)\; d(\mu, \widehat{\mu}), \qquad \text{if} \;\; d(\mu, \widehat{\mu}) \le r$$

But: $C_\Phi(r) \longrightarrow \infty$ as data y more informative, e.g., noise $\varepsilon \longrightarrow 0$

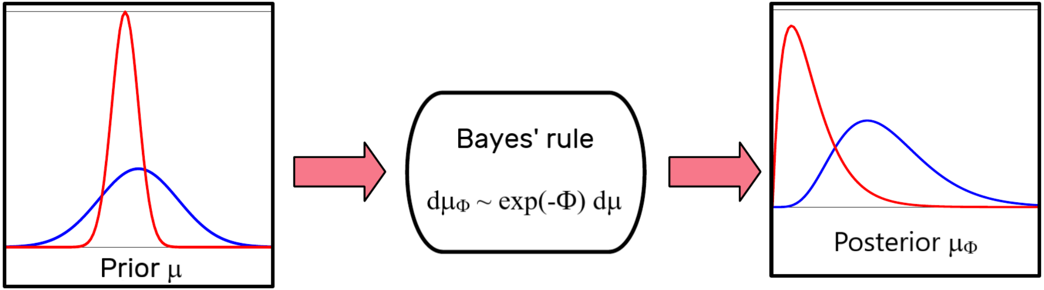


Theorem (informal)

For d being TV, Hellinger, or 1-Wasserstein distance or KL divergence we have under suitable assumptions a locally Lipschitz continuity:

$$d(\mu_\Phi, \widehat{\mu}_\Phi) \;\leq\; C_\Phi(r)\; d(\mu, \widehat{\mu}), \qquad \text{if} \;\; d(\mu, \widehat{\mu}) \leq r$$

But: $C_\Phi(r) \longrightarrow \infty$ as data y more informative, e.g., noise $\varepsilon \longrightarrow 0$



Theorem (informal)

For d being TV, Hellinger, or 1-Wasserstein distance or KL divergence we have under suitable assumptions a locally Lipschitz continuity:

$$d(\mu_\Phi, \widehat{\mu}_\Phi) \;\leq\; C_\Phi(r)\; d(\mu, \widehat{\mu}), \qquad \text{if} \;\; d(\mu, \widehat{\mu}) \leq r$$

But: $C_\Phi(r) \to \infty$ as data y more informative, e.g., noise $\varepsilon \to 0$

- For the TV-distance, we have

$$d_{\mathrm{TV}}(\mu_\Phi, \widehat{\mu}_\Phi) \le \frac{2}{Z}\, d_{\mathrm{TV}}(\mu, \widehat{\mu}), \qquad Z := \int \mathrm{e}^{-\Phi} \mathrm{d}\mu$$

- If $\mathscr{A}$ is bounded and $\mathrm{e}^{-\Phi}\colon \mathscr{A} \to \mathbb{R}_+$ globally Lipschitz, then

$$\mathrm{W}_1(\mu_\Phi, \widehat{\mu}_\Phi) \le \left(\frac{2 + 2c_{\mathscr{A}}\mathrm{Lip}_{\mathrm{e}^{-\Phi}}}{Z}\right)^2 \mathrm{W}_1(\mu, \widehat{\mu}) \qquad \forall \widehat{\mu}\colon \mathrm{W}_1(\mu, \widehat{\mu}) \le \frac{Z}{2\mathrm{Lip}_{\mathrm{e}^{-\Phi}}}$$

- For continuous $\Phi\colon \mathscr{A} \to \mathbb{R}_+$ we also have continuity in p-Wasserstein distance, i.e.,

$$\lim_{n\to\infty} \mathrm{W}_p\left(\mu, \widehat{\mu}^{(n)}\right) = 0 \quad \Longrightarrow \quad \lim_{n\to\infty} \mathrm{W}_p\left(\mu_\Phi, \widehat{\mu}_\Phi^{(n)}\right) = 0, \qquad p \ge 1.$$

- For the TV-distance, we have

$$d_{\mathsf{TV}}(\mu_\Phi, \widehat{\mu}_\Phi) \leq \frac{2}{Z} d_{\mathsf{TV}}(\mu, \widehat{\mu}), \qquad Z := \int e^{-\Phi} d\mu$$

- If $\mathscr{A}$ is bounded and $e^{-\Phi}\colon \mathscr{A} \to \mathbb{R}_+$ globally Lipschitz, then

$$W_1(\mu_\Phi, \widehat{\mu}_\Phi) \leq \left(\frac{2 + 2c_{\mathscr{A}} \mathrm{Lip}_{e^{-\Phi}}}{Z}\right)^2 W_1(\mu, \widehat{\mu}) \qquad \forall \widehat{\mu}\colon W_1(\mu, \widehat{\mu}) \leq \frac{Z}{2\mathrm{Lip}_{e^{-\Phi}}}$$

- For continuous $\Phi\colon \mathscr{A} \to \mathbb{R}_+$ we also have continuity in p-Wasserstein distance, i.e.,

$$\lim_{n\to\infty} W_p\left(\mu, \widehat{\mu}^{(n)}\right) = 0 \quad \Rightarrow \quad \lim_{n\to\infty} W_p\left(\mu_\Phi, \widehat{\mu}_\Phi^{(n)}\right) = 0, \qquad p \geq 1.$$

- For the TV-distance, we have

$$d_{\mathrm{TV}}(\mu_\Phi, \widehat{\mu}_\Phi) \le \frac{2}{Z} d_{\mathrm{TV}}(\mu, \widehat{\mu}), \qquad Z := \int \mathrm{e}^{-\Phi} \mathrm{d}\mu$$

- If $\mathscr{A}$ is bounded and $\mathrm{e}^{-\Phi} \colon \mathscr{A} \to \mathbb{R}_+$ globally Lipschitz, then

$$\mathrm{W}_1(\mu_\Phi, \widehat{\mu}_\Phi) \le \left(\frac{2 + 2c_{\mathscr{A}} \mathrm{Lip}_{\mathrm{e}^{-\Phi}}}{Z} \right)^2 \mathrm{W}_1(\mu, \widehat{\mu}) \qquad \forall \widehat{\mu} \colon \mathrm{W}_1(\mu, \widehat{\mu}) \le \frac{Z}{2\mathrm{Lip}_{\mathrm{e}^{-\Phi}}}$$

- For continuous $\Phi \colon \mathscr{A} \to \mathbb{R}_+$ we also have continuity in p-Wasserstein distance, i.e.,

$$\lim_{n\to\infty} \mathrm{W}_p\left(\mu, \widehat{\mu}^{(n)}\right) = 0 \quad \Longrightarrow \quad \lim_{n\to\infty} \mathrm{W}_p\left(\mu_\Phi, \widehat{\mu}_\Phi^{(n)}\right) = 0, \qquad p \ge 1.$$

Related work

- Robust Bayesian statistics, e.g., [Berger, 1994]: Consider span of resulting posterior predictions for ϵ-contamination classes for prior

$$\Gamma_\epsilon(\mu) := \{(1-\epsilon)\mu + \epsilon\tilde{\mu} : \ \tilde{\mu} \in \mathscr{P}\}, \qquad \epsilon > 0$$

- [Diaconis & Freedman, 1986]: Fréchet derivative $\partial T_\Phi(\mu)$ of $T_\Phi(\mu) := \mu_\Phi$ w.r.t. TV topology

$$\|\partial T_\Phi(\mu)\| \sim \frac{1}{Z}, \qquad Z := \int e^{-\Phi} d\mu$$

- [Gustafson & Wasserman, 1995]: Sensitivity measure based on Gâteaux derivative w.r.t perturbations based with explicit growth for increasing number of i.i.d. data

- Bayesian brittleness [Owhadi et al., 2015]: Span of posterior expectations for $f : \mathscr{A} \to \mathbb{R}$ for all priors μ which satisfy finitely many generalized moment conditions coincides with range(f)

Related work

- Robust Bayesian statistics, e.g., [Berger, 1994]: Consider span of resulting posterior predictions for ϵ-contamination classes for prior

$$\Gamma_\epsilon(\mu) := \{(1-\epsilon)\mu + \epsilon\tilde{\mu} : \ \tilde{\mu} \in \mathscr{P}\}, \qquad \epsilon > 0$$

- [Diaconis & Freedman, 1986]: Fréchet derivative $\partial T_\Phi(\mu)$ of $T_\Phi(\mu) := \mu_\Phi$ w.r.t. TV topology

$$\|\partial T_\Phi(\mu)\| \sim \frac{1}{Z}, \qquad Z := \int e^{-\Phi} d\mu$$

- [Gustafson & Wasserman, 1995]: Sensitivity measure based on Gâteaux derivative w.r.t perturbations based with explicit growth for increasing number of i.i.d. data
- Bayesian brittleness [Owhadi et al., 2015]: Span of posterior expectations for $f : \mathscr{A} \to \mathbb{R}$ for all priors μ which satisfy finitely many generalized moment conditions coincides with range(f)

Related work

- Robust Bayesian statistics, e.g., [Berger, 1994]: Consider span of resulting posterior predictions for ϵ-contamination classes for prior

$$\Gamma_\epsilon(\mu) := \{(1-\epsilon)\mu + \epsilon\tilde{\mu} : \ \tilde{\mu} \in \mathscr{P}\}, \qquad \epsilon > 0$$

- [Diaconis & Freedman, 1986]: Fréchet derivative $\partial T_\Phi(\mu)$ of $T_\Phi(\mu) := \mu_\Phi$ w.r.t. TV topology

$$\|\partial T_\Phi(\mu)\| \sim \frac{1}{Z}, \qquad Z := \int e^{-\Phi} d\mu$$

- [Gustafson & Wasserman, 1995]: Sensitivity measure based on Gâteaux derivative w.r.t perturbations based with explicit growth for increasing number of i.i.d. data

- Bayesian brittleness [Owhadi et al., 2015]: Span of posterior expectations for $f : \mathscr{A} \to \mathbb{R}$ for all priors μ which satisfy finitely many generalized moment conditions coincides with range(f)

Related work

- Robust Bayesian statistics, e.g., [Berger, 1994]: Consider span of resulting posterior predictions for ϵ-contamination classes for prior

$$\Gamma_\epsilon(\mu) := \{(1-\epsilon)\mu + \epsilon\tilde{\mu} : \; \tilde{\mu} \in \mathscr{P}\}, \qquad \epsilon > 0$$

- [Diaconis & Freedman, 1986]: Fréchet derivative $\partial T_\Phi(\mu)$ of $T_\Phi(\mu) := \mu_\Phi$ w.r.t. TV topology

$$\|\partial T_\Phi(\mu)\| \sim \frac{1}{Z}, \qquad Z := \int \mathrm{e}^{-\Phi} \mathrm{d}\mu$$

- [Gustafson & Wasserman, 1995]: Sensitivity measure based on Gâteaux derivative w.r.t perturbations based with explicit growth for increasing number of i.i.d. data
- Bayesian brittleness [Owhadi et al., 2015]: Span of posterior expectations for $f : \mathscr{A} \to \mathbb{R}$ for all priors μ which satisfy finitely many generalized moment conditions coincides with range(f)

References

1. O. Ernst, A. Pichler, BS. Wasserstein Sensitivity of Risk and Uncertainty Propagation. *SIAM/ASA Journal on Uncertainty Quantification* **10(3)**:915–948, 2022.

2. BS. On the local Lipschitz stability of Bayesian inverse problems. *Inverse Problems* **36**:055015, 2020.