Stochastic inversion under functional uncertainties

Reda El Amri

MASCOT-NUM annual conference. 18-20 March 2019, IFPEN, Rueil-Malmaison.

In collaboration with Céline Helbert†, Olivier Lepreux¶, Miguel Munoz Zuniga¶, Clémentine Prieur§ and Delphine Sinoquet¶

† École Centrale de Lyon, Institut Camille Jordan, France
¶ IFPEN, France
§ Université Grenoble Alpes, Laboratoire Jean Kuntzmann, Grenoble, France
Contents

1. Context

2. Data-driven stochastic inversion via functional quantization

3. Stochastic inversion via meta-modelling in the joint space

4. Application

5. Conclusion and outlook
IFPEN test case: Selective Catalytic Reduction

Scalar outputs and thresholds \( \begin{cases} \text{NO}_x^{\text{out}} \leq 80 \ mg.km^{-1} \\ \text{NH}_3^{\text{out}} \leq 30 \ ppm \end{cases} \)
Data-driven stochastic inversion via functional quantization
Stochastic inversion via meta-modelling in the joint space
Application
Conclusion and outlook

Context

Stochastic inversion under functional uncertainties

Reda El Amri

Stochastic inversion under functional uncertainties
Context
Data-driven stochastic inversion via functional quantization
Stochastic inversion via meta-modelling in the joint space
Application
Conclusion and outlook

Input: NEDC Driving Cycle

Output: NO\textsubscript{eq}

score DeNOx = 51.55 (mg/km)

Input: Random Driving Cycle

Output: NO\textsubscript{eq}

score DeNOx = 183 (mg/km)
The framework

\[ f : X \times V \rightarrow \mathbb{R} \]
\[ (x, v) \mapsto f(x, v) \]

where \( X \subset \mathbb{R}^p \) is a compact and \( V \) functional space. For a fixed \( c \in \mathbb{R} \), some objects of interest can be

\[ \Gamma^* := \{ x \in X \text{ s.t. } f(x, V) \leq c \text{ almost surely} \}, \]
\[ \Gamma^* := \{ x \in X \text{ s.t. } g(x) = \mathbb{E}[f(x, V)] \leq c \} := g^{-1}(C), \text{ where } C = (-\infty, c]. \]

Objective

- Estimate \( \Gamma^* \), the excursion set of g below c.
The framework

- each evaluation \( f(x, v) \) is expensive-to-compute,
- \( V \) is only known through \( N \) realizations \( \Xi = \{ v_1, v_2, \ldots, v_N \} \)
  - redundant information
  - rough curves
- \( g \) is evaluated at \( X_n = (x_1, \ldots, x_n) \) with values \( g_{X_n} = (g(x_1), \ldots, g(x_n)) \), where \( g(x_i) = \mathbb{E}[f(x_i, V)] \).
The framework

- each evaluation $f(x, v)$ is expensive-to-compute,
- $V$ is only known through $N$ realizations $\Xi = \{v_1, v_2, \ldots, v_N\}$
  - redundant information
  - rough curves
- $g$ is evaluated at $X_n = (x_1, \ldots, x_n)$ with values $g_{X_n} = (g(x_1), \ldots, g(x_n))$, where $g(x_i) = \mathbb{E}[f(x, V)]$.

Methodology I

- build a meta-model for the response $g$ and choose $x_{n+1} \in X$ by SUR strategy,
- estimate the expectation $\mathbb{E}[f(x_{n+1}, V)]$ (requires $l \leq N$ evaluations of $f$).

Methodology II

- build a meta-model in the joint space,
- choose $(x_{n+1}, v_{n+1}) \in X \times V$ (only one evaluation of $f$ is required).
Gaussian process interpolation and SUR strategy

$g$ is seen as a realization of a Gaussian Process $(Z_x)_{x \in X}$ with prior mean $m$ and covariance kernel $k$,

$\Gamma^*$ is a realization of $\Gamma := \{ x \in X : Z_x|Z_{X_n} = g_{X_n} \leq c \}$.

Stepwise Uncertainty Reduction

find the best next evaluation point $x_{n+1}$ that optimally reduces the expected uncertainty $\mathcal{H}_{n+1}$ on the future estimate, i.e.,

$$x_{n+1} = \arg\min_{x \in X} \mathbb{E}_{n,x}[\mathcal{H}^{\text{uncert}}_{n+1}(x)].$$
Problem formulation

Let $V$ be a stochastic process defined in the space $\mathcal{H} = L^2(\Omega, \mathcal{F}, P; L^2([0, T]))$ such that

$$||V||_{L^2} = \left(\mathbb{E}[||V||^2]\right)^{1/2} = \left(\mathbb{E}\left[\int_0^T V^2 dt\right]\right)^{1/2},$$

- **Aim:** Replace $V$ by a r.v. taking finite number of values close to $V$ in some sense.

- Given a (finite) 'grid' $\Theta_I = \{v_1, \ldots, v_l\} \subset V$, we define a (Borel) Nearest neighbor projection
  - let $\left(C_{v_i}(\Theta_I)\right)_{1 \leq i \leq l}$ be a Voronoi partition of $V$ generated by $\Theta_I$,
  - $\pi_{\Theta_I} : \mathcal{V} \rightarrow \Theta_I$ the induced $\Theta_I$ Neighbour projection

$$V \mapsto \sum_{i=1}^{l} v_i \mathbf{1}_{C_{v_i}(\Theta_I)}(V).$$
We define the Voronoi Quantization of the random process $\mathbf{V}$ as

$$\hat{\mathbf{V}}_l = \sum_{i=1}^{l} v_i \mathbf{1}_{C_{v_i}(\Theta_l)}(\mathbf{V}),$$

Distribution of $\hat{\mathbf{V}}_l$: weights associated to each $v_i$

$$\mathbb{P}(\hat{\mathbf{V}}_l = v_i) = \mathbb{P}(\mathbf{V} \in C_{v_i}(\Theta_l)), \quad i = 1 \ldots, l$$

We define the $l$-optimal quantization error at level $l$ as

$$e_l(\mathbf{V}) = \inf \left\{ \left( \mathbb{E} \| \mathbf{V} - \hat{\mathbf{V}}_l \| ^2 \right)^{1/2}, \hat{\mathbf{V}}_l : \Omega \to \mathcal{V}, \text{card}(\hat{\mathbf{V}}_l(\Omega)) \leq l \right\}.$$
Expectation estimation

Let \( h : \mathcal{V} = L^2([0, T]) \rightarrow \mathbb{R} \) be a continuous function, and let \( \hat{\mathcal{V}}_l \) be a \textit{l}-quantization. It is natural to approximate \( \mathbb{E}[h(\mathcal{V})] \) by \( \mathbb{E}[h(\hat{\mathcal{V}}_l)] \). This quantity \( \mathbb{E}[h(\hat{\mathcal{V}}_l)] \) is simply the finite weighted sum:

\[
\mathbb{E}[h(\hat{\mathcal{V}}_l)] = \sum_{i=1}^{l} h(\hat{v}_i) P(\hat{\mathcal{V}}_l = \hat{v}_i).
\]

If \( h \) is Lipschitz continuous

\[
\left| \mathbb{E}[h(\mathcal{V})] - \mathbb{E}[h(\hat{\mathcal{V}}_l)] \right| \leq \mathbb{E}|h(\mathcal{V}) - h(\hat{\mathcal{V}}_l)|
\leq [h]_{\text{Lip}} \mathbb{E}\|\mathcal{V} - \hat{\mathcal{V}}_l\|
\leq [h]_{\text{Lip}} (\mathbb{E}\|\mathcal{V} - \hat{\mathcal{V}}_l\|^2)^{1/2}.
\]
Numerical computation of quantizers

- **Vectorial case**: let $U : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^m, \mathcal{B}or(\mathbb{R}^m), ||.||)$ be a random vector such that $\mathbb{E}||U||^2 < \infty$
  - Lloyd’s algorithm, Competitive Learning Vector Quantization algorithm (Gradient based alg.)

**Figure**: Quantizers for $\mathcal{N}(0, l_2)$ of size $l = 150$ in $(\mathbb{R}^2, ||.||)$. 
Numerical computation of quantizers

- **Vectorial case:** let \( U : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^m, Bor(\mathbb{R}^m), ||.||) \) be a random vector such that \( \mathbb{E}||U||^2 < \infty \)
  - Llyod’s algorithm, Competitive Learning Vector Quantization algorithm (Gradient based alg.)
  - Greedy vector quantization \(^2\)

Numerical computation of quantizers

**Vectorial case:** let $\mathbf{U} : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^m, \mathcal{B}or(\mathbb{R}^m), ||.||) \to (\mathbb{R}^m, \mathcal{B}or(\mathbb{R}^m), ||.||)$ be a random vector such that $\mathbb{E}||\mathbf{U}||^2 < \infty$

- Llyod’s algorithm, Competitive Learning Vector Quantization algorithm (Gradient based alg.)
- Greedy vector quantization

**Functional case:**

- Optimal quantization for Gaussian processes (e.g. Brownian motion) \(^3\) Available online at quantize.maths-fi.com
- Centroidal Voronoi Tessellation method (probabilistic approach of Lloyd’s alg.) \(^4\)


\(^3\) Harald Luschgy, Gilles Pagès, and Benedikt Wilbertz (2010). “Asymptotically optimal quantization schemes for Gaussian processes on Hilbert spaces”. In: *ESAIM: Probability and Statistics* 14

\(^4\) M Miranda and P Bocchini (2013). “Functional quantization of stationary Gaussian and non-Gaussian random processes”. In: *Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures*
The Karhunen-Loève expansion

Let \( V \in H \) be a random process with zero mean and continuous covariance function \( C(t, s) \). Then

\[
V(t) \simeq \sum_{i=1}^{m_{KL}} u_i \psi_i(t),
\]

where \( \{\psi_i\}_{i=1}^{m_{KL}} \) are orthogonal and normalized eigenfunctions of the integral operator corresponding to \( C \). The \( \{u_i\}_{i=1}^{m_{KL}} \) denotes a set of orthogonal random variables with zero mean and variance \( \lambda_i \), where \( \lambda_i \) is the eigenvalue corresponding to the eigenfunction \( \psi_i \).

**Computational method for functional PCA** \( C(s, t) \) is estimated from \( \Xi = \{v_i\}_{i=1}^{N} \) by:

\[
C^N(t, s) = \frac{1}{N} \sum_{j=1}^{N} v_j(s)v_j(t).
\]

---

Greedy Functional Quantization: the first proposal

\[ \hat{D}_1 = \{ \hat{u}_1 \} \] where \( \hat{u}_1 \) is a solution of \( e_1(U) \)

\[ \forall l \geq 2, \quad \hat{D}_l = \hat{D}_{l-1} \cup \{ \hat{u}_l \} \]

where \( \hat{u}_l \in \arg \min_{u \in G} \left( \mathbb{E} \| \mathbf{U} - \mathbf{U}_l \|^2 \right)^{1/2} \),

where \( \mathbf{U} \sim \mathbb{U}_G \) and \( \mathbf{U}_l \) is the \( l \)-quantization induced by \( \{ \hat{u}_1, \ldots, \hat{u}_{l-1} \} \cup \{ u \} \).
Greedy Functional Quantization: the second proposal

Initialization: $\tilde{D}_1 = \{\tilde{u}_1\}$ where $\tilde{u}_1$ is randomly chosen

$\forall l \geq 2$, $\tilde{D}_l = \tilde{D}_{l-1} \cup \{\tilde{u}_l\}$

where $\tilde{u}_l \in \arg \max_{u \in G} \phi_{\text{Maximin}}(\tilde{D}_{l-1} \cup \{u\})$. 
Application 1 We consider a functional $f$ defined as

$$
\begin{align*}
  f : (x, v) \mapsto \max_t v_t \cdot |0.1 \cos(x_1 \max_t v_t) \sin(x_2) (x_1 + x_2 \min_t v_t)^2| \cdot \int_0^T (30 + v_t)^{\frac{x_1 \cdot x_2}{20}} dt ,
\end{align*}
$$

- $x = (x_1, x_2) = (2.95, 3.97)$
- $m_{KL} = 2$
- $\text{card}(\Xi) = 200$ (Brownian motion)
Algorithm 1 Data-driven stochastic inversion via functional quantization \textsuperscript{6}

1: Create an initial DoE of \( n \) points in the control space \( X \)
2: \textbf{while} Stopping criterion not met (SUR) \textbf{do}
3: \hspace{1em} \( x_{n+1} \leftarrow \) Sampling criterion \( J_n \)
4: \hspace{1em} Set \( l = 1 \).
5: \hspace{1em} \textbf{while} Stopping criterion not met (Expectation Estimation) \textbf{do}
6: \hspace{2em} Approximate the expectation by \( E[f(x_{n+1}, \hat{V}_l)] \)
7: \hspace{2em} Set \( l = l + 1 \)
8: \hspace{1em} \textbf{end while}
9: Update DoE
10: Set \( n = n + 1 \)
11: \textbf{end while}
12: \textbf{end}

\textsuperscript{6}Reda El Amri et al. (Feb. 2018). “Data-driven stochastic inversion via functional quantization”. URL: https://hal.inria.fr/hal-01704189
1575 calls to the simulator
Dimension reduction:

\[ f : \text{control variables} \rightarrow f(x, v) \]

\[ V : \text{functional variable} \xrightarrow{\text{dimension Reduction}} \text{U : scalar variables} \]
Dimension reduction:

\[ f : X \times \mathbb{R}^{m_{KL}} \to \mathbb{R} \]

\[ (x, u) \mapsto f(x, u) \]

The new framework:

Objective

- Estimate \( \{ x \in X \ s.t. \ g(x) = \mathbb{E}[f(x, U)] \leq c \} \)
GP model for $f$

By $Y_{(x,u)}$ we denote Gaussian Process with prior mean $m_Y(x,u)$ and covariance function $\text{cov}_Y(x,u; x', u')$

\[
\mathbb{E}[Y_{(x,u)}] = m_Y(x,u) \\
\text{COV}(Y_{(x,u)}, Y_{(x',u')}) = \text{cov}_Y(x,u; x', u')
\]
GP model for $f$

By $Y_{(x,u)}$ we denote Gaussian Process with prior mean $m_Y(x,u)$ and covariance function $\text{cov}_Y(x,u;x',u')$

$$
\mathbb{E}[Y_{(x,u)}] = m_Y(x,u)
$$

$$
\text{COV}(Y_{(x,u)}, Y_{(x',u')}) = \text{cov}_Y(x,u;x',u')
$$

Projected process to model $g$

$$
Z^n_{(x)} = \mathbb{E}_{U}[Y^n_{(x,u)}] = \int_{\mathbb{R}^{m_{KL}}} Y^n_{(x,u)} d\rho(u),
$$

where $d\rho(u)$ probability measure on $u$ and $Y^n$ is the GP conditioned on the $n$ observations.

$Z^n_{(x)}$ remains Gaussian, therefore, it is entirely characterized by its two first moments

---

Step I: choose the control variable site

(Stepwise Uncertainty Reduction) $x_{n+1} = \arg\min_{x \in X} \mathbb{E}_{n,x}[H_{n+1}^{uncert}(x)]$. 
Step I: choose the control variable site

(Stepwise Uncertainty Reduction) \( x_{n+1} = \arg \min_{x \in X} \mathbb{E}_{n,x}[\mathcal{H}_{n+1}^{\text{uncert}}(x)] \).

Step II: choose the uncertain variable site

Choose the point in the uncertain space that reduces the most the uncertainty of the estimated \( \mathbb{E}_{U}[g(x, U)] \) at the point \( x_{n+1} \).

\[
\mathbf{u}_{n+1} = \arg \min_{\tilde{u} \in \mathbb{R}^{m_{KL}}} \text{VAR}(Z^{n+1}_{(x_{n+1})})
\]

\[
\text{VAR}(Z^{n+1}_{(x_{n+1})}) = \int_{\mathbb{R}^{m_{KL}}} \int_{\mathbb{R}^{m_{KL}}} \text{cov}_{Y}(x_{n+1},u;x_{n+1},u')d\rho(u)d\rho(u') - \int_{\mathbb{R}^{m_{KL}}} \int_{\mathbb{R}^{m_{KL}}} \text{cov}_{Y}(x_{n+1},u;X^{n+1},U^{n+1})d\rho(u)d\rho(u')
\]

where \( X^{n+1} = (X^{n},x_{n+1}) \) and \( U^{n+1} = (U^{n},\tilde{u}) \) and \( X^{n}, U^{n} \) are observed data points.
Algorithm 2 Stochastic inversion via joint space modelling

1: Create an initial DoE of $n$ points in the joint space $\mathbb{X}$
2: Simulator responses
3: while $n \leq$ budget do
4:   $x_{n+1} \leftarrow$ Sampling criterion $J_n$
5:   $u_{n+1} \leftarrow \text{argmin}_{u \in \mathbb{R}^{m_{KL}}} \text{VAR}(Z_{(x_{n+1})}^{n+1})$
6:   Simulator response at point $(x_{n+1}, v_{n+1})$
7:   Update DoE
8:   Set $n = n + 1$
9: end while
10: end
IFPEN test case: control strategy for an automotive NO\textsubscript{x} depollution system

- NH\textsubscript{3} function of \( x_1 \) and \( x_2 \),
  \( \mathbb{X} = [0, 0.6] \times [0, 0.6] \),
- continuous function, expensive to evaluate (kinetic models),
- 100 random driving cycles,
IFPEN test case: control strategy for an automotive NO\textsubscript{x} depollution system

- NH\textsubscript{3} function of \(x_1\) and \(x_2\), \(X = [0, 0.6] \times [0, 0.6]\),
- continuous function, expensive to evaluate (kinetic models),
- 100 random driving cycles,
- \(n = 5(2 + m_{KL}) = 110\) observations (black triangles),
- constant mean function \(m\),
- Matérn covariance kernel (\(\rho = 3/2\)).
IFPEN test case: control strategy for an automotive NO\textsubscript{x} depollution system

- NH\textsubscript{3} function of \(x_1\) and \(x_2\), \(\mathbb{X} = [0, 0.6] \times [0, 0.6]\),
- continuous function, expensive to evaluate (kinetic models),
- 100 random driving cycles,

- \(n = 5(2 + m_{KL}) = 110\) observations (black triangles),
- constant mean function \(m\),
- Matérn covariance kernel \((\rho = 3/2)\),
**IFPEN test case:** control strategy for an automotive NO\textsubscript{x} depollution system

- NH\textsubscript{3} function of \( x_1 \) and \( x_2 \), \( \mathbb{X} = [0, 0.6] \times [0, 0.6] \),
- continuous function, expensive to evaluate (kinetic models),
- 100 random driving cycles,
- \( n = 5(2 + m_{KL}) = 110 \) observations (black triangles),
- constant mean function \( m \),
- Matérn covariance kernel \( (\rho = 3/2) \),

After 510 calls to the simulator.

green set (Algo I), red boundary (algo II)
Conclusion

- Robustness measure: Expectation,
- dealing with functional uncertainties
  - greedy expectation estimation via **functional quantization**, 
  - sequential strategy to wisely choose the points in **joint space**.
Conclusion

- Robustness measure: Expectation,
- dealing with functional uncertainties
  - greedy expectation estimation via functional quantization,
  - sequential strategy to wisely choose the points in joint space.

Outlook

- Perform a convergence analysis of the greedy functional quantization algorithm.
- extend the proposed methods to the case of correlated responses,
- other functionals of the distribution may be of greater importance.
Conclusion

- Robustness measure: Expectation,
- dealing with functional uncertainties
  - greedy expectation estimation via functional quantization,
  - sequential strategy to wisely choose the points in joint space.

Outlook

- Perform a convergence analysis of the greedy functional quantization algorithm,
- extend the proposed methods to the case of correlated responses,
- other functionals of the distribution may be of greater importance.

Thanks for your attention!