Sequential and batch-sequential Bayesian sampling strategies for the identification of an excursion set

Clément Chevalier (IRSN and University of Bern)

supervised by David Ginsbourger (University of Bern) and Yann Richet (IRSN)

Mascot-SAMO 2013, Ph.D. student day, Nice, July 1st 2013
Outline

1. Motivations

2. SUR strategies for inversion, state of the art and contributions

3. SUR strategies for robust inversion
1 Motivations

2 SUR strategies for inversion, state of the art and contributions

3 SUR strategies for robust inversion
Let us consider an expensive-to-evaluate black box simulator $f$:

$$f : \mathbb{X} \subset \mathbb{R}^d \mapsto \mathbb{R}$$

$$x \mapsto f(x)$$
Let us consider an expensive-to-evaluate black box simulator $f$:

$$f : \mathbb{R}^d \mapsto \mathbb{R}$$

$$x \mapsto f(x)$$

- No analytical expression for $f$. 

Let us consider an expensive-to-evaluate black box simulator $f$:

$$f : \mathbb{X} \subset \mathbb{R}^d \mapsto \mathbb{R}$$

- **No analytical expression for $f$.**
- **Evaluating $f(x)$ at one point $x \in \mathbb{X}$ takes a lot of time.**
Let us consider an expensive-to-evaluate black box simulator $f$:

$$ f : \mathbb{R}^d \rightarrow \mathbb{R} $$

- No analytical expression for $f$.
- Evaluating $f(x)$ at one point $x \in \mathbb{X}$ takes a lot of time.
- Very small budget to evaluate $f$. 
Let $T \in \mathbb{R}$ be a fixed threshold.
Let $T \in \mathbb{R}$ be a fixed threshold.

**Problem 1: Inversion**
Identify the excursion set: $\Gamma^\text{inversion} := \{ \mathbf{x} \in \mathbb{X} : f(\mathbf{x}) \geq T \}$
Let $T \in \mathbb{R}$ be a fixed threshold.

**Problem 1: Inversion**
Identify the excursion set: $\Gamma_{\text{inversion}}^* := \{x \in X : f(x) \geq T\}$

**Problem 2: Robust inversion**
Among the $d$ inputs, $d_c$ are controlled and $d_u$ are not. We want to identify the set:

$$\Gamma_{\text{rob.inv}}^* := \{x_c \in X_c : \forall x_u \in X_u, f(x_c, x_u) \leq T\}$$
Let $T \in \mathbb{R}$ be a fixed threshold.

**Problem 1: Inversion**
Identify the excursion set: $\Gamma_{\text{inversion}}^* := \{x \in X : f(x) \geq T\}$

**Problem 2: Robust inversion**
Among the $d$ inputs, $d_c$ are controlled and $d_u$ are not. We want to identify the set:

$$\Gamma_{\text{rob.inv}}^* := \{x_c \in X_c : \forall x_u \in X_u, f(x_c, x_u) \leq T\}$$

- Which strategy will guide our sequential evaluations of $f$?
Let $T \in \mathbb{R}$ be a fixed threshold.

**Problem 1: Inversion**  
Identify the excursion set: $\Gamma_{\text{inversion}}^* := \{x \in X : f(x) \geq T\}$

**Problem 2: Robust inversion**  
Among the $d$ inputs, $d_c$ are controlled and $d_u$ are not. We want to identify the set:

$$\Gamma_{\text{rob.inv}}^* := \{x_c \in X_c : \forall x_u \in X_u, f(x_c, x_u) \leq T\}$$

- Which strategy will guide our sequential evaluations of $f$?

Our strategy will be based on a *Kriging* metamodel.
Motivations
SUR strategies for inversion, state of the art and contributions
SUR strategies for robust inversion

Context
Prior: $f$ is considered as a realization of a Gaussian process $\xi$. 
Prior: $f$ is considered as a realization of a Gaussian process $\xi$.

When we have $n$ observations $A_n = (\xi(x_1), \ldots, \xi(x_n))$, the posterior distribution of $\xi$ is still Gaussian.
Prior: $f$ is considered as a realization of a Gaussian process $\xi$.

When we have $n$ observations $A_n = (\xi(x_1), \ldots, \xi(x_n))$, the posterior distribution of $\xi$ is still Gaussian.

- Kriging mean: $m_n(x)$
- Kriging variance: $s_n^2(x)$
- $\mathcal{L}(\xi(x)|A_n) = \mathcal{N}(m_n(x), s_n^2(x))$
Short description of a SUR strategy:
Short description of a SUR strategy:

- Defining a measure of uncertainty for the problem at hand
SUR strategies for inversion

Short description of a SUR strategy:

- Defining a measure of uncertainty for the problem at hand

- Optimal 1-step lookahead criterion: Expectation of the future uncertainty if an observation $x_{n+1}$ is added.
Short description of a SUR strategy:

- Defining a measure of uncertainty for the problem at hand

- Optimal 1-step lookahead criterion: Expectation of the future uncertainty if an observation $x_{n+1}$ is added.

- SUR strategy: sampling sequentially at the location where the criterion is minimized.
Starting point:

Sequential design of computer experiments for the estimation of a probability of failure.

introduces a definition for the “uncertainty”.
Starting point:


introduces a definition for the “uncertainty”.

Set of interest:

\[ \Gamma := \{ \mathbf{x} \in \mathbb{X} : \xi(\mathbf{x}) \geq T \} \]

Uncertainty:

\[ H_n := \text{Var}_n(\mathbb{P}_X(\Gamma)) \]
In Bect et. al., the uncertainty $H_n$ is judged intractable so that a different definition of the uncertainty is used:

$$H_n \leq \tilde{H}_n := \int_X p_n(1 - p_n)d\mathbb{P}_X$$

where

$$p_n(x) := P(x \in \Gamma|A_n) = P(\xi(x) \geq T|A_n)$$

The function $p_n(\cdot)$ is called excursion (or coverage) probability function.
The optimal one-step lookahead SUR criteria that need to be minimized to find the next evaluation location are:

\[ J_n(x) := E_n(H_{n+1} | X_{n+1} = x) \]

\[ \tilde{J}_n(x) := E_n(\tilde{H}_{n+1} | X_{n+1} = x) \]
The optimal one-step lookahead SUR criteria that need to be minimized to find the next evaluation location are:

\[ J_n(x) := \mathbb{E}_n(H_{n+1} | X_{n+1} = x) \]
\[ \tilde{J}_n(x) := \mathbb{E}_n(\tilde{H}_{n+1} | X_{n+1} = x) \]

Main issue with SUR strategies:

- Computer intensive
- Designed to sample one point at a time, whereas we often have \( r > 1 \) CPUs to evaluate \( f \) in parallel.
**SUR strategies for inversion**

The optimal one-step lookahead SUR criteria that need to be minimized to find the next evaluation location are:

\[
J_n(x) := \mathbb{E}_n(H_{n+1} | X_{n+1} = x)
\]

\[
\tilde{J}_n(x) := \mathbb{E}_n(\tilde{H}_{n+1} | X_{n+1} = x)
\]

**Main issue with SUR strategies:**
- **Computer intensive**
- Designed to sample one point at a time, whereas we often have \( r > 1 \) CPUs to evaluate \( f \) in parallel.

**A first contribution:** definition of a **multi-points** generalization of the SUR criteria and “fast” formulas to compute both the one-point and multi-points version of \( J_n \) and \( \tilde{J}_n \).
SUR strategies for inversion

Proposition (1)

\[
\tilde{J}_n(x^{(r)}) = \int_X \Phi_2 \left( \begin{pmatrix} a(x) \\ -a(x) \end{pmatrix}, \begin{pmatrix} c(x) & 1-c(x) \\ 1-c(x) & c(x) \end{pmatrix} \right) dP_X(x),
\]

\[
J_n(x^{(r)}) = \gamma_n - \int_{X \times X} \Phi_2 \left( \begin{pmatrix} a(z_1) \\ -a(z_2) \end{pmatrix}, \begin{pmatrix} c(z_1) & d(z_1, z_2) \\ d(z_1, z_2) & c(z_2) \end{pmatrix} \right) dP_X(dz_1) dP_X(dz_2)
\]

- \( \Phi_2(\cdot, M) \) is the c.d.f. of the centered bivariate Gaussian with covariance matrix \( M \)
- \( a(x) := (m_n(x) - T)/s_{n+r}(x) \),
- \( b(x) := \frac{1}{s_{n+r}(x)} \Sigma^{-1}(k_n(x, x_{n+1}), \ldots, k_n(x, x_{n+r})^\top \)
- \( c(x) := s_n^2(x)/s_{n+r}^2(x) \)
- \( d(z_1, z_2) := b(z_1)^\top \Sigma b(z_2) \)
- \( \Sigma \) is conditional covariance matrix of \((\xi(x_{n+1}), \ldots, \xi(x_{n+r}))^\top \).
- \( \gamma_n \) does not depend on \((x_{n+1}, \ldots, x_{n+r})\).
Details and proofs can be found in:

Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set.

Accepted with minor revisions to Technometrics, preprint available on HAL.
Details and proofs can be found in:

  *Accepted with minor revisions to Technometrics, preprint available on HAL.*

The proof widely rely on the “kriging update” formulas:

- X. Emery
  The Kriging update equations and their application to the selection of neighbouring data
  *Computational Geosciences, 13, 269–280, 2009.*

- C. C, D. Ginsbourger, X. Emery
  Corrected Kriging update formulae for batch-sequential data assimilation.
  *Proceedings of the IAMG2013 conference, preprint available on HAL.*
Details and proofs can be found in:


The proof widely rely on the “kriging update” formulas:


A key step in the proof is to write the updated kriging mean function, $m_{n+r}(\cdot)$, as a function of the unknown response $\xi(x^{(r)})$. 
Applications
Nuclear Safety

Figure: 3 iterations of a parallel SUR algorithm (criterion $\tilde{J}_n$ with $r = 4$)
Applications
Nuclear Safety

Figure: Evolution of $\int_X p_n(1 - p_n) dP_X$ during the inversion.
**Applications**

Nuclear Safety

---

**Figure:** Evolution of $\int_X p_n(1 - p_n) dP_X$ during the inversion.

---

C. C, V. Picheny and D. Ginsbourger.

The KrigInv package: An efficient and user-friendly R implementation of Kriging-based inversion algorithms.

*Computational Statistics & Data Analysis, 2013*
**Motivation:** the presented SUR strategies are meant to reduce the variance of the excursion’s volume (or a bound of this volume).
Motivation: the presented SUR strategies are meant to reduce the variance of the excursion’s volume (or a bound of this volume).

If we are interested in:

$$\Gamma^*_{\text{inversion}} := \{ \mathbf{x} \in \mathbb{X} : f(\mathbf{x}) \geq T \},$$
**Motivation:** the presented SUR strategies are meant to reduce the variance of the excursion’s volume (or a bound of this volume).

If we are interested in:

\[ \Gamma^*_{\text{inversion}} := \{ \mathbf{x} \in \mathbb{X} : f(\mathbf{x}) \geq T \}, \]

we would like to define a “variance” for the random excursion set:

\[ \Gamma := \{ \mathbf{x} \in \mathbb{X} : \xi(\mathbf{x}) \geq T \} \]
Figure: Real (unknown) excursion set of a two dimensional function, with a threshold $T = 80$. 
**Figure:** Real (unknown) excursion set of a two dimensional function, with a threshold $T = 80$. 
**Figure:** Coverage probability calculated using Kriging, on a two dimensional function, with a threshold $T = 80$. 
Figure: Random set realizations, obtained with GP conditional simulations.
There exist different approaches to define the expectation and the variance of a random set.

I. Molchanov
Theory of Random Sets.  
*Springer*, 2005.
There exist different approaches to define the expectation and the variance of a random set.

I. Molchanov

Theory of Random Sets.  
*Springer, 2005.*

One of them is the Vorob’ev approach.
Vorob’ev expectation:

\[ Q_{n,\alpha_n} := \{ x \in D : p_n(x) \geq \alpha_n \}, \]
Vorob’ev expectation:

\[ Q_{n,\alpha_n} := \{ x \in D : p_n(x) \geq \alpha_n \}, \]

where \( \alpha_n \) satisfies the equation:

\[ \mathbb{P}_X(Q_{n,\alpha_n}) = \int_X p_n d\mathbb{P}_X := \nu_n \]
Vorob’ev expectation:

\[ Q_{n,\alpha_n} := \{ x \in D : p_n(x) \geq \alpha_n \}, \]

where \( \alpha_n \) satisfies the equation:

\[ P_X(Q_{n,\alpha_n}) = \int_X p_n \, dP_X := \nu_n \]

Vorob’ev deviation, our “variance”

It is proven that, \( \forall \) closed set \( S \) with \( P_X(S) = \nu_n \),

\[ E_n(P_X(Q_{n,\alpha_n} \triangle \Gamma)) \leq E_n(P_X(S \triangle \Gamma)). \]
Vorob’ev expectation:

\[ Q_{n,\alpha_n} := \{ \mathbf{x} \in D : p_n(\mathbf{x}) \geq \alpha_n \}, \]

where \( \alpha_n \) satisfies the equation:

\[
\mathbb{P}_{\mathbf{x}}(Q_{n,\alpha_n}) = \int_{\mathbf{x}} p_n d\mathbb{P}_{\mathbf{x}} := \nu_n
\]

Vorob’ev deviation, our “variance”

It is proven that, \( \forall \) closed set \( S \) with \( \mathbb{P}_{\mathbf{x}}(S) = \nu_n \),

\[
\mathbb{E}_n(\mathbb{P}_{\mathbf{x}}(Q_{n,\alpha_n} \triangle \Gamma)) \leq \mathbb{E}_n(\mathbb{P}_{\mathbf{x}}(S \triangle \Gamma)).
\]

We can thus define our variance of a random set:

\[ Var_n(\Gamma) := \mathbb{E}_n(\mathbb{P}_{\mathbf{x}}(Q_{n,\alpha_n} \triangle \Gamma)) \]
**SUR strategies for inversion**

**Random sets**

*Figure:* Random set realizations, obtained with GP conditional simulations.
Figure: Random set realizations, obtained with GP conditional simulations.
Figure: Symmetrical differences between the realizations and the Vorob’ev Expectation
Construction of a SUR strategy based on the uncertainty:

\[ H_n := \text{Var}_n(\Gamma) \]
Construction of a SUR strategy based on the uncertainty:

\[ H_n := \text{Var}_n(\Gamma) \]

- \( H_n \) can be computed quite easily.
- An efficient implementation of the one-step lookahead optimal SUR criteria relies on closed-form formulas derived, again, from the kriging update formulas.
Construction of a SUR strategy based on the uncertainty:

\[ H_n := Var_n(\Gamma) \]

- \( H_n \) can be computed quite easily.
- An efficient implementation of the one-step lookahead optimal SUR criteria relies on closed-form formulas derived, again, from the kriging update formulas.

C. C, D. Ginsbourger, J. Bect and I. Molchanov
Estimating and quantifying uncertainties on level sets using the Vorob’ev expectation and deviance with Gaussian process models.

*moda 10, Advances in Model-Oriented Design and Analysis, Physica-Verlag HD, 2013.*
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 0.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 1.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 2.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 3.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $Var_n(\Gamma)$. Iteration 4.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $Var_n(\Gamma)$. Iteration 5.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $Var_n(\Gamma)$. Iteration 6.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 7.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 8.
Sequential sampling strategy

Figure: SUR sampling strategy reducing $\text{Var}_n(\Gamma)$. Iteration 9.
Sequential sampling strategy

Excursion set

Excursion probability
Sequential sampling strategy

Vorob'ev criterion

sur criterion
Outline

1. Motivations
2. SUR strategies for inversion, state of the art and contributions
3. SUR strategies for robust inversion
In robust inversion, we aim at identifying the set:

$$\Gamma_{\text{optinv}}^* := \{x_c \in \mathbb{X}_c : \forall x_u \in \mathbb{X}_u, f(x_c, x_u) \leq T\}$$
In robust inversion, we aim at identifying the set:

$$\Gamma_{\text{optinv}}^* := \{x_c \in X_c : \forall x_u \in X_u, f(x_c, x_u) \leq T\}$$

This set is a subset of $X_c$. 
In robust inversion, we aim at identifying the set:

\[ \Gamma_{\text{optinv}}^* := \{ \mathbf{x}_c \in \mathbb{X}_c : \forall \mathbf{x}_u \in \mathbb{X}_u, f(\mathbf{x}_c, \mathbf{x}_u) \leq T \} \]

This set is a subset of \( \mathbb{X}_c \).

An uncertainty measure can be defined in the same spirit than in inversion.
SUR strategies for robust inversion

Excursion probability:

\[ \widetilde{\rho}_n(x_c) := P \left( \max_{x_u \in X_u} \xi(x_c, x_u) \leq T | A_n \right) \]
SUR strategies for robust inversion

Excursion probability:

\[ \tilde{\rho}_n(x_c) := P \left( \max_{x_u \in X_u} \xi(x_c, x_u) \leq T | A_n \right) \]

\[ H_n := \int_{X_c} \tilde{\rho}_n(x_c)(1 - \tilde{\rho}_n(x_c))dP_{X_c}(x_c) \]

is a possible measure to quantify uncertainties on \( \Gamma_{optinv}^* \).
Excursion probability:

$$\tilde{\rho}_n(x_c) := P \left( \max_{x_u \in X_u} \xi(x_c, x_u) \leq T | A_n \right)$$

$$H_n := \int_{X_c} \tilde{\rho}_n(x_c)(1 - \tilde{\rho}_n(x_c))d\mathbb{P}_{X_c}(x_c)$$

is a possible measure to quantify uncertainties on $\Gamma_{\text{optinv}}^*$. 

Main issue

The associated SUR criterion is too expensive to compute!
**Goal**: Find a “fast” approximation of $\tilde{\rho}_n(x_c)$. 
**Goal**: Find a “fast” approximation of $\tilde{p}_n(x_c)$.

**First idea**

$$
\tilde{p}_n(x_c) := P \left( \max_{x_u \in \{x_u^{(1)}, \ldots, x_u^{(q)}\}} \xi(x_c, x_u) \leq T \middle| A_n \right)
$$

where $x_u^{(1)}, \ldots, x_u^{(q)}$ are chosen so that $\tilde{p}_n(x_c)$ is as close as possible to $\tilde{p}_n(x_c)$. 
The SUR criteria can be efficiently computed using a generalization of the results obtained in inversion:

\[
J_n(x^{(r)}) := \mathbb{E}_n \left( \int_{x_c} \hat{p}_{n+r}(x_c)(1 - \hat{p}_{n+r}(x_c)) d\mathbb{P}_{x_c}(x_c) \right)
\]

\[
= \int_{x_c} \left( \hat{p}_n(x_c) - \Phi_2 q \left( \begin{pmatrix} T - m_n^{(q)} \\ B^\top \sum_n^{(r)} B \end{pmatrix}, \begin{pmatrix} \sum_n^{(q)} & B^\top \sum_n^{(r)} B \\ \sum_n^{(q)} & \sum_n^{(q)} \end{pmatrix} \right) \right) d\mathbb{P}_{x_c}(x_c)
\]
SUR strategies for robust inversion

Figure: \( \hat{p}_n(x_c) \) function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{\rho}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
**Goal**: Find a “fast” approximation of $\tilde{\rho}_n(x_c)$. 
**Goal**: Find a “fast” approximation of $\tilde{p}_n(x_c)$.

**Second idea: GP realization updates**

$$\tilde{p}_n(x_c) := \frac{1}{M} \# \{i : \max_{1 \leq j \leq q} z_{i,n}^j \leq T\}$$

where $z_{i,n}^1, \ldots, z_{i,n}^q$ is a realization (conditioned on $n$ obs.) of $\xi(x_c, \cdot)$ in $q$ locations.
Idea: from the kriging update formulas, it is possible to calculate how a GP realization is modified by new observations.
Idea: from the kriging update formulas, it is possible to calculate how a GP realization is modified by new observations.
SUR strategies for robust inversion

number of simulations not exceeding $T$

response at location $x_{n+1}$
SUR strategies for robust inversion

\[ \text{pn.hat} (1 - \text{pn.hat}) \]

response at location \( x_{n+1} \)
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: \( \hat{p}_n(x_c) \) function at current iteration.
Figure: $\hat{p}_n(x_c)$ function at current iteration.
Figure: $\hat{p}_n(x_c)$ function at current iteration.
SUR strategies for robust inversion

Figure: $\hat{p}_n(x_c)$ function at current iteration.
Future work:

- Other approximations of the exceedance probability relying on e.g. the work of Adler and Taylor.
Future work:

- Other approximations of the exceedance probability relying on e.g. the work of Adler and Taylor.

- Convergence results.
Future work:

- Other approximations of the exceedance probability relying on e.g. the work of Adler and Taylor.
- Convergence results.
- Choice of the integration points to compute the integrals: Sequential Monte-Carlo methods.
Thank you for your attention!
Future work

Thank you for your attention! any questions?

Acknowledgement: Clément Chevalier acknowledges support from the IRSN and the ReDICE consortium.