

Quand l'exception fait la règle

-

La simulation d'évènements rares dans les systèmes complexes

Jérôme Morio – ONERA/DCPS
22/05/2014



retour sur innovation

Plan de la présentation

- Problématique de recherche
- Synthèse sur les méthodes d'estimation d'évènements rares
- Développements récents réalisés à l'ONERA
- Conclusion - Enjeux scientifiques

Contexte

- **Une des missions de l'Onera : évaluer les performances de systèmes complexes**

➔ **Nécessité : quantifier les risques**

Collision
débris-satellite



Effets collatéraux
d'une retombée
d'engin



Cas extrêmes
de turbulence
de sillage



➔ **Focalisation sur les méthodes de simulation et d'estimation d'évènements rares**

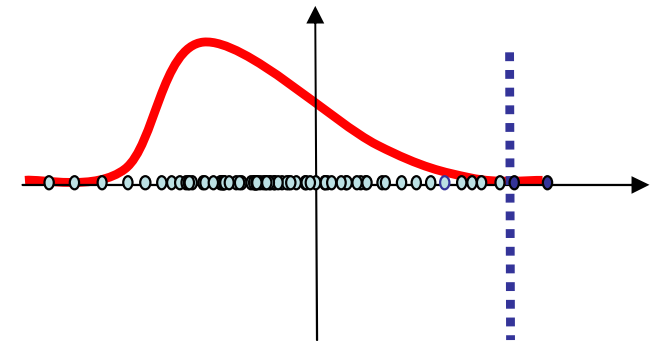
Caractérisation d'un évènement rare

■ Les évènements rares considérés:

- Evènement de forte gravité, critique pour la sécurité ou la performance du système



- Probabilité faible $\ll 1/N$ (N, le budget de simulations) **difficilement estimable** par simulation directe (Monte-Carlo)

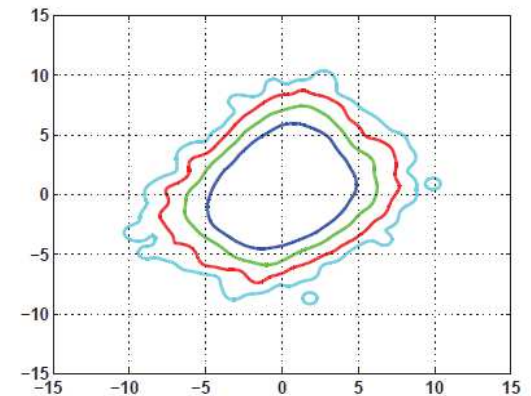
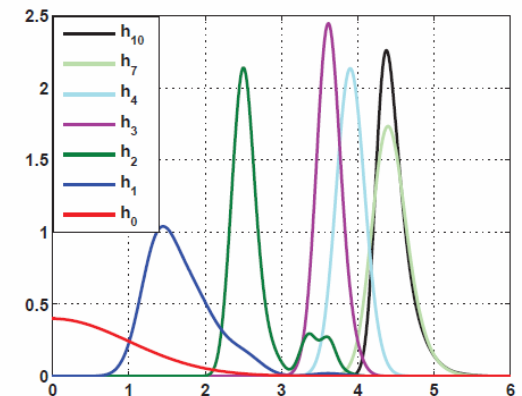


- Nécessité d'une faible erreur sur cette probabilité. Conséquences graves d'une erreur sur l'estimation



Problématique de recherche

- Analyser l'applicabilité des méthodes de simulations d'évènements rares existantes.
- Proposer des améliorations de celles-ci pour traiter des cas de complexité accrue
- Répondre à des problèmes nouveaux induits par des applications aérospatiales

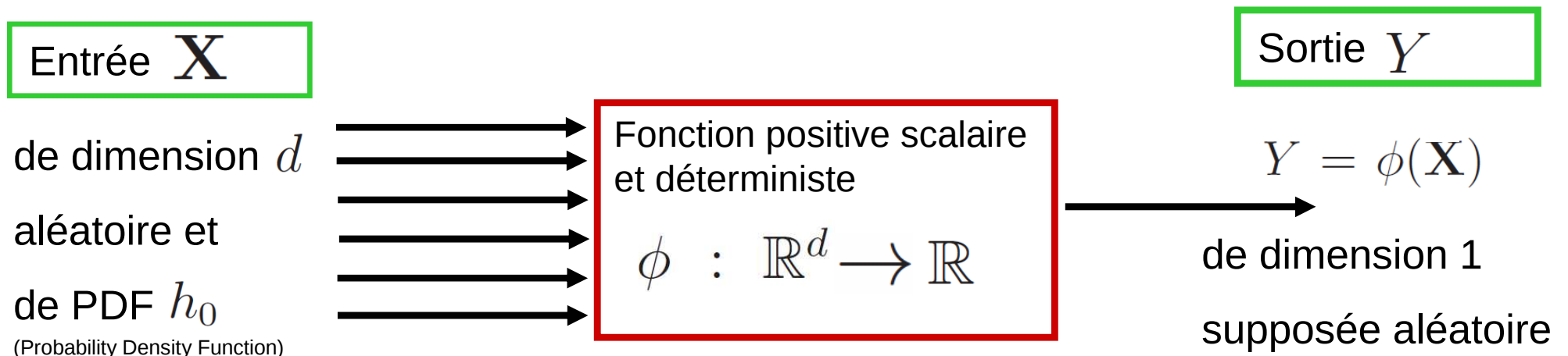


Plan de la présentation

- Problématique de recherche
- Synthèse sur les méthodes d'estimation d'évènements rares
- Quelques développements récents réalisés à l'Onera
- Conclusion - Enjeux scientifiques

Système considéré pour la synthèse

- Système statique décrit par une fonction boîte noire entrée-sortie



- **Estimation de :**
 - Probabilités de type $\mathbb{P} = P(\phi(\mathbf{X}) > S)$, S un seuil scalaire
 - α -quantiles de la sortie Y

Exemple de système

■ Estimation de la probabilité de collision entre un débris et un satellite

- **Entrée:** positions et vitesses du débris spatial et du satellite
- **Boîte noire:** code de propagation de trajectoires
- **Sortie:** distance entre le débris et le satellite
- **Probabilité** de collision entre le débris et le satellite



■ Évaluation des cas extrêmes de turbulence de sillage

- **Entrée:** caractéristiques de l'avion, conditions météorologiques, etc
- **Boîte noire:** code de prédiction de turbulences
- **Sortie:** circulation des turbulences
- **Probabilité** de dépassement de seuil de la circulation



Exemple simple

Entrée $\mathbf{X}$

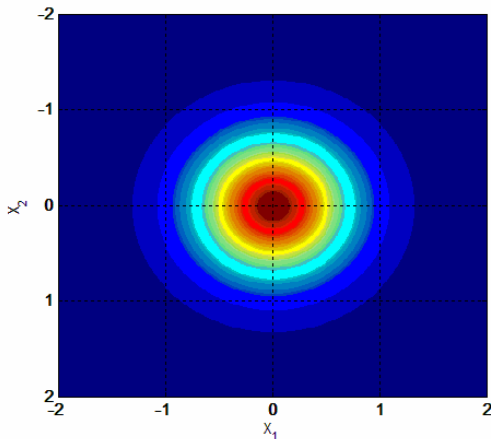
Sortie Y

de dimension 2
et de loi gaussienne
standard

Code de
calcul

ϕ

$Y = \phi(\bar{\mathbf{X}})$
de dimension 1
supposée aléatoire



Rencontres du réseau

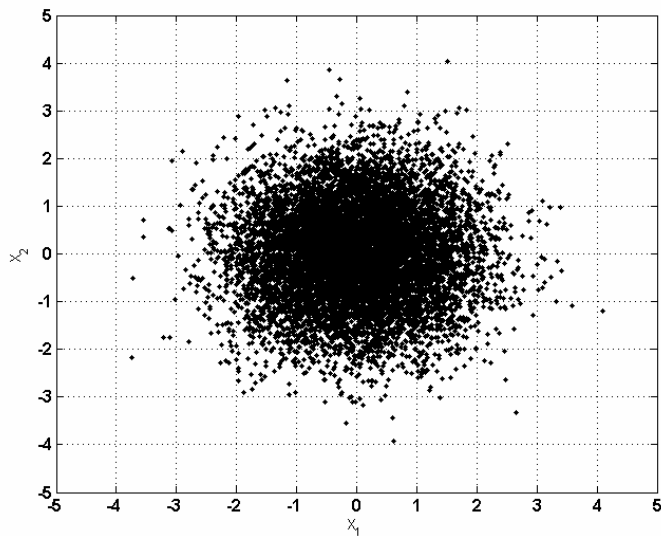
Comment estimer $\mathbb{P} = P(\phi(\mathbf{X}) > S) \ ?$

Méthodes Monte-Carlo

Estimation de la probabilité en simulant les incertitudes sur $\mathbf{X}$ et en les propageant à travers ϕ

[Silverman 86, Sobol 95, Niederreiter 00, Robert 10...]

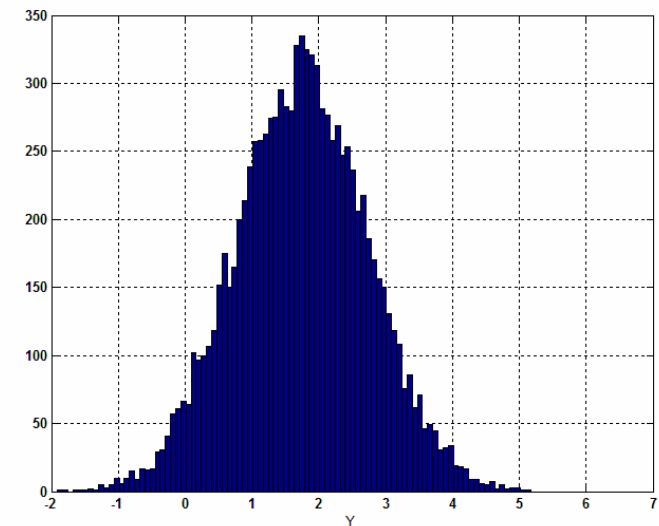
Entrée $\mathbf{X}$



ϕ



Sortie Y



Histogramme de

$\phi(\mathbf{X}_1), \dots, \phi(\mathbf{X}_N)$

$\mathbf{X}_1, \dots, \mathbf{X}_N$ générés selon h_0

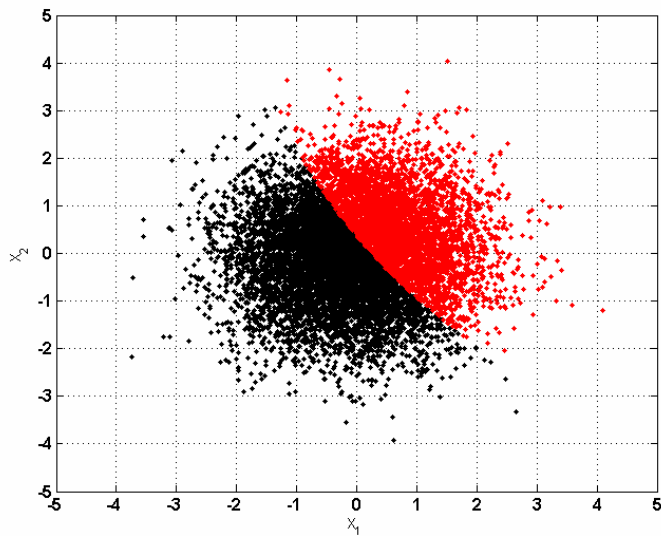
10000 échantillons

Méthodes Monte-Carlo

Estimation de la probabilité en simulant les incertitudes sur $\mathbf{X}$ et en les propageant à travers ϕ

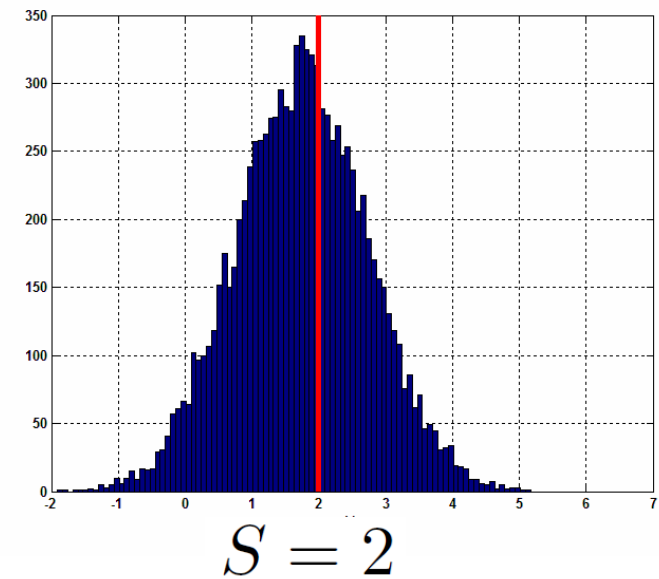
[Silverman 86, Sobol 95, Niederreiter 00, Robert 10...]

Entrée $\mathbf{X}$



ϕ

Sortie Y



$$\hat{\mathbb{P}}^{CMC} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\phi(\mathbf{x}_i) > S}$$

$$\hat{\mathbb{P}}^{CMC}(Y > 2) = 0.40 \text{ (1\%)}$$

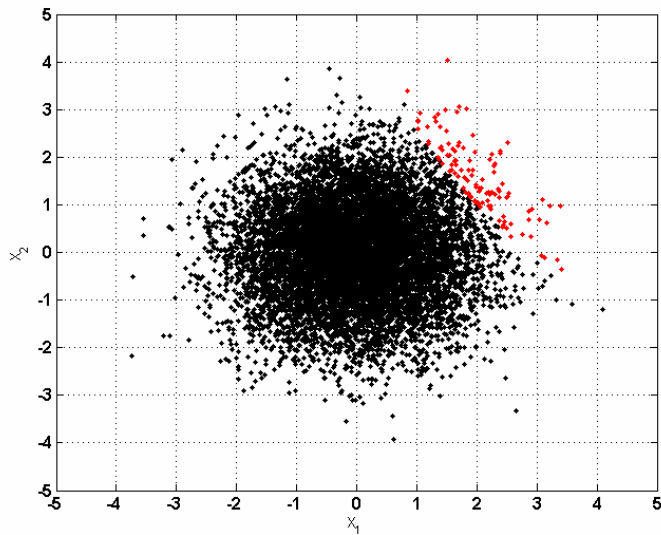
N= 10000 échantillons

Méthodes Monte-Carlo

Estimation de la probabilité en simulant les incertitudes sur $\mathbf{X}$ et en les propageant à travers ϕ

[Silverman 86, Sobol 95, Niederreiter 00, Robert 10...]

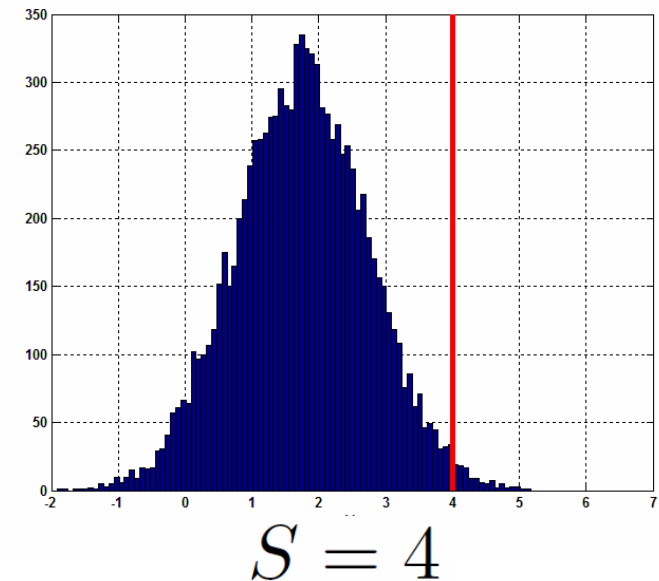
Entrée $\mathbf{X}$



ϕ



Sortie Y



$$\hat{\mathbb{P}}^{CMC}(Y > 4) = 1.2 \cdot 10^{-2} \text{ (10\%)}$$

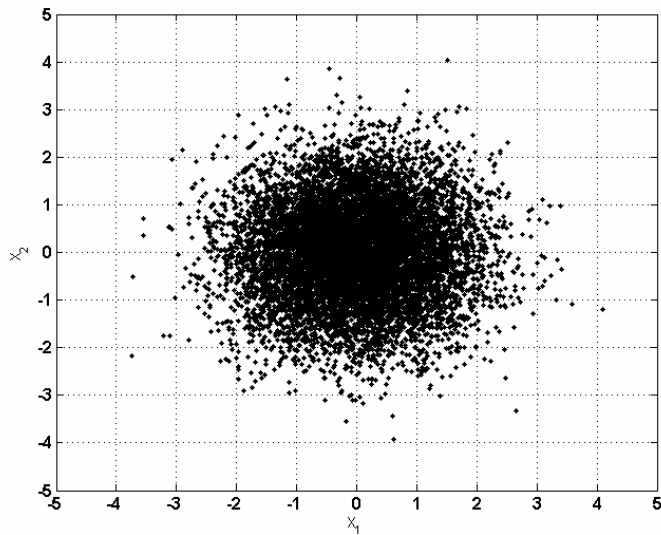
N= 10000 échantillons

Méthodes Monte-Carlo

Estimation de la probabilité en simulant les incertitudes sur $\mathbf{X}$ et en les propageant à travers ϕ

[Silverman 86, Sobol 95, Niederreiter 00, Robert 10...]

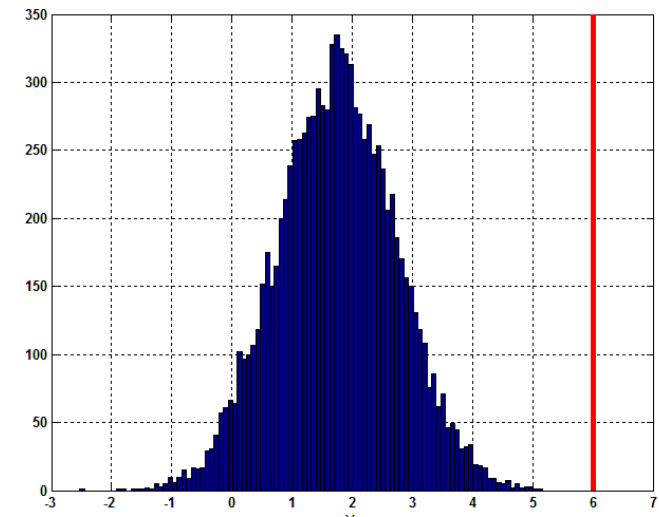
Entrée $\mathbf{X}$



ϕ



Sortie Y



N= 10000 échantillons

$S = 6$

$$\hat{\mathbb{P}}^{CMC}(Y > \textcircled{6}) = 1.2 \cdot 10^{-5} \textcircled{(273\%)}$$

Méthodes Monte-Carlo

- Facile à implémenter en pratique
- Pas de condition de régularité sur le code de calcul
- Convergence indépendante de la dimension
- Budget de simulations important
- La précision varie en $N^{-1/2}$ (10^6 échantillons nécessaires pour estimer une probabilité de l'ordre de 10^{-4} avec 10% d'erreur)

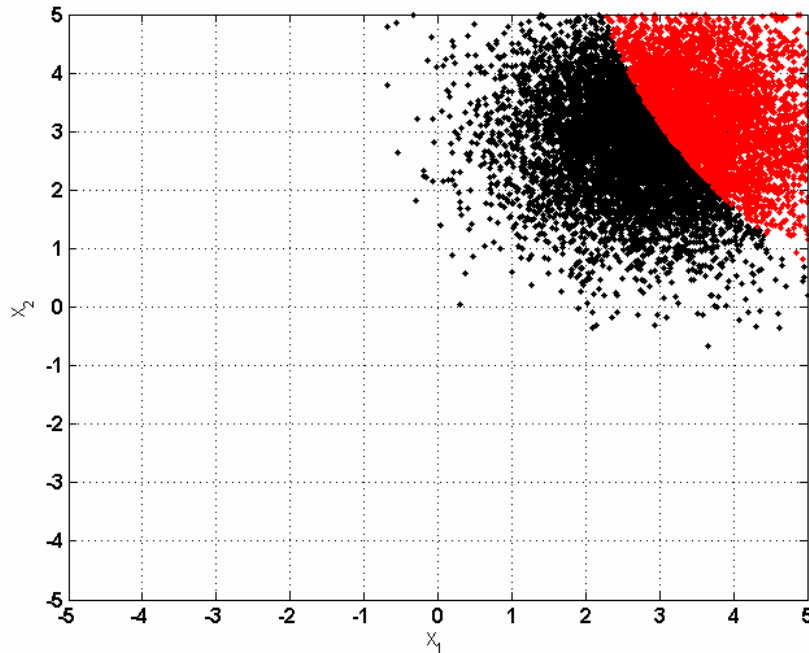
Importance Sampling

Générer les échantillons en entrée selon une densité
auxiliaire h favorisant les évènements rares $\phi(\mathbf{X}) > S$

[Hammersley 64, Bucklew 04, Richard 07, Kroese 12...]

Entrée $\mathbf{X}$

$\mathbf{X}_1, \dots, \mathbf{X}_N$
générés selon h



$S = 6$

$$\hat{\mathbb{P}}^{IS} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\phi(\mathbf{x}_i) > S} \frac{h_0(\mathbf{X}_i)}{h(\mathbf{X}_i)}$$

$$\hat{\mathbb{P}}^{IS}(Y > 6) = 1.0 \cdot 10^{-5} \text{ (3\%)}$$

N= 10000 échantillons

$$\hat{\mathbb{P}}^{CMC}(Y > 6) = 1.2 \cdot 10^{-5} \text{ (273\%)}$$

Importance Sampling

- Facile à implémenter en pratique
- Très forte réduction potentielle de la variance de l'estimation
- Trouver une densité auxiliaire efficace assez complexe dans un cadre réaliste (paramétrique ou non paramétrique ?)
- Dimension du problème

Importance splitting/subset sampling/SMC

(sequential Monte-Carlo)

- On cherche à estimer $\mathbb{P} = P(\mathbf{X} \in \mathbf{A})$
avec $\mathbf{A} = \{\mathbf{X} \in \mathbb{R}^d | \phi(\mathbf{X}) > S\}$
- Définition de sous-espaces emboîtés dans l'espace d'entrée :

$$\mathbf{A}_0 = \mathbb{R}^d \supset \mathbf{A}_1 \supset \dots \supset \mathbf{A}_{n-1} \supset \mathbf{A}_n = \mathbf{A}$$

- Relation de Bayes :

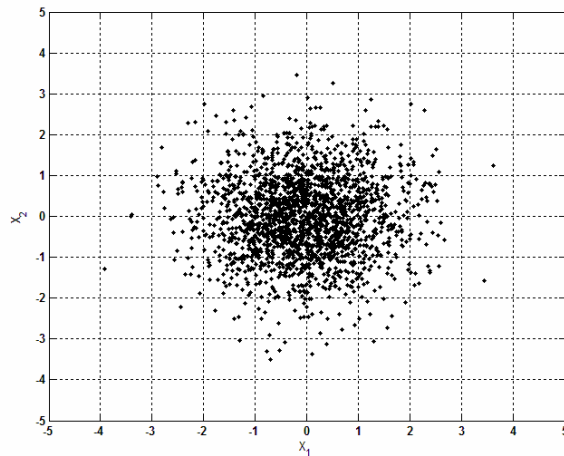
$$P(\mathbf{X} \in \mathbf{A}) = \prod_{k=1}^n P(\mathbf{X} \in \mathbf{A}_k | \mathbf{X} \in \mathbf{A}_{k-1})$$

➤ Principe du splitting (analogie avec les algorithmes génétiques)

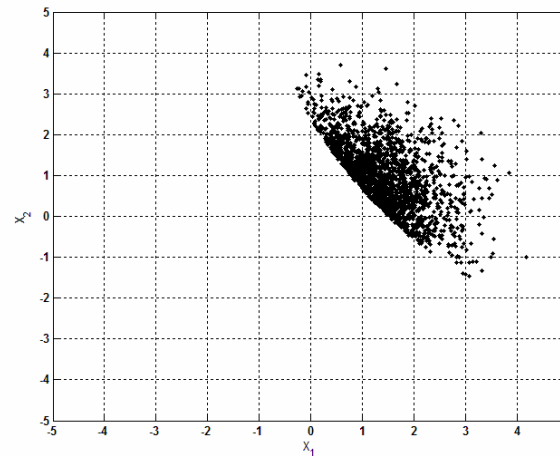
Génération d'une population initiale d'échantillons puis alternance de :

- **sélection** des échantillons de la population les plus prometteurs pour atteindre l'évènement rare
- **mutation** des échantillons pour explorer l'espace d'entrée sans changement de distribution

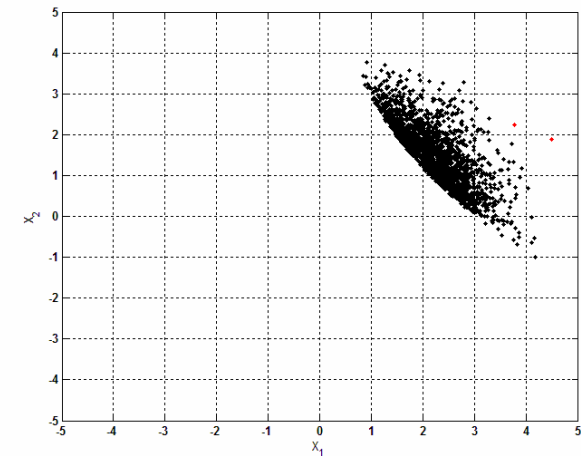
Importance splitting/subset sampling/SMC (sequential Monte-Carlo)



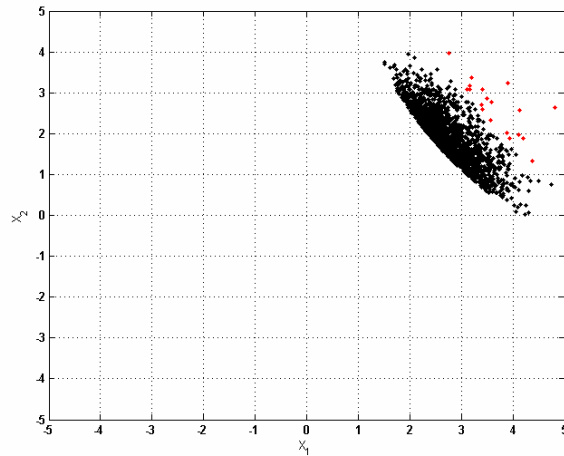
Echantillons initiaux



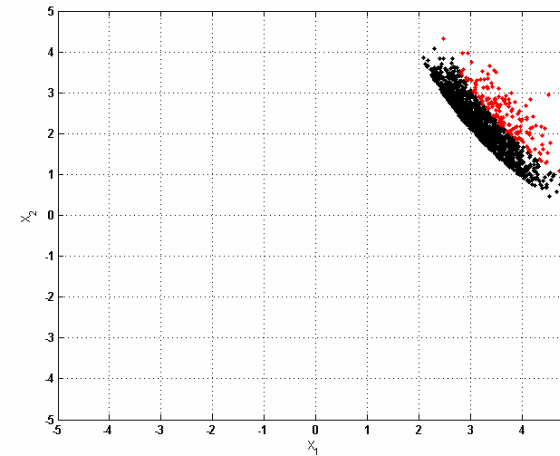
Itération 1



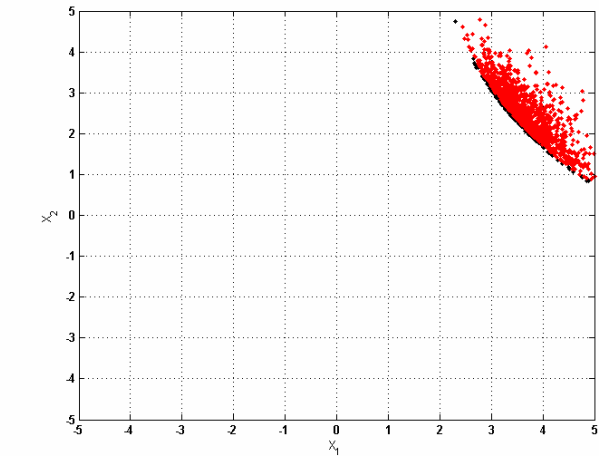
Itération 2



Itération 3



Itération 4



Itération 5

$$\hat{\mathbb{P}}^{ISp}(Y > 6) = 1.1 \cdot 10^{-5} \text{ (15\%)}$$

Importance splitting/subset sampling/SMC

(sequential Monte-Carlo)

- Très rentable pour de très faibles probabilités et quantiles (à partir de 10^{-4})
- Efficace en grande dimension
- Difficultés d'ajustement des paramètres de l'algorithme
- Peut nécessiter un nombre d'échantillons très important

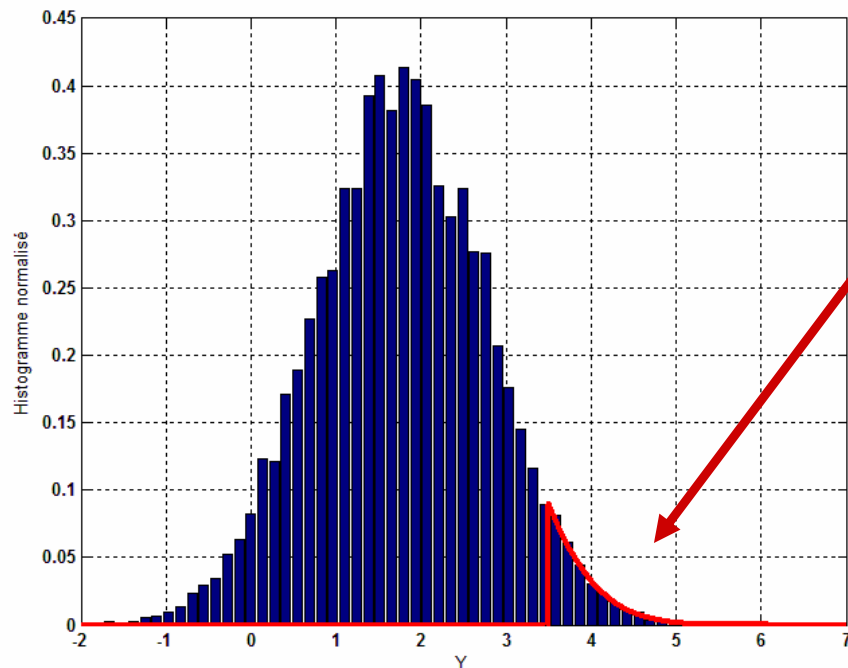
Théorie des valeurs extrêmes

A partir d'un jeu d'échantillons de la sortie du code,
estimation paramétrique de la densité en sortie du code

➤ Loi de Pareto généralisée

$$P(Y \leq t | Y > \mu) \rightarrow H(t)$$

$$H(t) = 1 - \left(1 + \varepsilon \frac{t - \mu}{\sigma}\right)^{-\frac{1}{\varepsilon}}$$



$$S = 6$$

avec
 $(\mu, \sigma, \varepsilon)$ estimés à partir des échantillons

N= 10000 échantillons

$$\hat{\mathbb{P}}^{TVE}(Y > 6) = 1.7 \cdot 10^{-5} \text{ (40\%)}$$

Théorie des valeurs extrêmes

- Peu d'alternatives à cette méthode si un rééchantillonnage n'est pas possible
- Peu d'échantillons nécessaires pour obtenir des résultats
- Difficultés d'estimation des paramètres
- Moins efficace que les techniques de simulation quand le rééchantillonnage est possible

- **Techniques utilisant des métamodèles :**

couplage métamodèles et techniques de simulation, voire optimisation

[Echard 11, Bect 12, Dubourg 13...]

- **FORM/SORM:**

estimation par approximation paramétrique de la probabilité de défaillance

[Bjerager 91, Yan-Gang 99, Low 13...]

- **Line sampling:**

approximation par échantillonnage suivant des directions de l'espace d'entrée

[Koutsourelakis 04, Zio 13...]

-
.....

Contribution à une synthèse des méthodes d'évènements rares

- **Catalogue de méthodes plus ou moins efficaces selon les situations**
- **Réglages souvent complexes et cas-dépendants**
- **Dans un cas concret pratique, quelles méthodes choisir?**

Comparaison des différents algorithmes et synthèse de leur applicabilité

- J. Morio, R. Pastel et F. Le Gland, Estimation de probabilités et de quantiles rares pour la caractérisation d'une zone de retombée d'un engin, Journal de la Société Française de Statistique
- J. Morio, M. Balesdent, D. Jacquemart and C. Vergé, A survey of rare event estimation methods for static input-output models, in minor revision, Simulation Practice modeling and Theory

Synthèse

- Peut-on utiliser le code pour obtenir des **échantillons** de la sortie Y ?
- La fonction entrée-sortie est-elle **analytiquement connue** ?
- Quelle est la **dimension** du problème ?
- Quel est le **budget de simulations** disponible ?
- La région de l'entrée du code de calcul qui aboutit l'évènement rare est-elle **multimodale** ?
- La région de l'entrée du code de calcul qui aboutit à l'évènement rare est-elle **approximativement connue** ?

Synthèse

Tableau de synthèse

	Impossibility of resampling	Density of Y known	ϕ known analytically	ϕ and Y unknown	Region $Y > S$ partially known	Region $Y > S$ disjoint - info not available	$d < 10$	$d > 10$	$N > 1000$	$N < 1000$	ϕ non linear
AST				✓				✓	✓		✓
Anti.Var.		✓	✓								
CE				✓	✓	×		✓	✓		
Cont.Var.		✓	✓								
EVT	✓										
Exp.Tw		✓	✓								
FORM /SORM				✓	✓	×				✓	×
LDT		✓	✓								
LS				✓		×	✓			✓	
Surrogates	✓				✓			×		✓	×
NAIS				✓	✓		✓	×	✓		
SS					✓						

Tableau de synthèse - ✓(resp. ×): la méthode présente des avantages (resp. inconvénients) pour la caractéristique considérée

- Souvent difficile de choisir *a priori* la méthode la plus adaptée à une situation donnée
- Éviter l'utilisation à contre-emploi de certains algorithmes (famille de densité unimodale à optimiser par CE pour densité optimale multimodale...)
- Étudier avec précision les caractéristiques du code de calcul (paramètres influents,...), les PDF des entrées...
- Caractériser l'estimation de la probabilité de l'évènement rare par son erreur relative (bornes, estimation Monte-Carlo...)

Conclusion

- Les communautés de recherche les plus actives concernent les métamodèles, les techniques de splitting et les valeurs extrêmes
- Expertise nécessaire pour les réglages des algorithmes. L'automatisation de ces réglages est un axe de recherche
- Les méthodes Monte-Carlo souvent considérées comme un mètre étalon
- Rédaction en cours d'un livre de synthèse pour Elsevier (sortie en 2015)

Estimation of rare event probabilities in complex aerospace (and other) systems - a practical approach

Plan de la présentation

- Problématique de recherche
- Synthèse sur les méthodes d'estimation d'évènements rares
- **Développements récents réalisés à l'ONERA**
 - Importance sampling
 - Incertitudes sur les paramètres de modèle
- Conclusion - Enjeux scientifiques

Importance Sampling

- Fonction positive scalaire : $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$
- Entrées : $\mathbf{X}$ de dimension d , aléatoires de densité h_0
- Probabilité visée : $\mathbb{P} = P(\phi(\mathbf{X}) > S)$ avec S un seuil
- Estimateur Monte-Carlo : $\mathbf{X}_1, \dots, \mathbf{X}_N$ i.i.d. suivant h_0

$$\hat{\mathbb{P}}^{CMC} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\phi(\mathbf{X}_i) > S}$$

- Estimateur Importance Sampling : $\mathbf{X}_1, \dots, \mathbf{X}_N$ i.i.d. suivant h

$$\hat{\mathbb{P}}^{IS} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\phi(\mathbf{X}_i) > S} \frac{h_0(\mathbf{X}_i)}{h(\mathbf{X}_i)}$$

Comment déterminer une densité auxiliaire efficace ?

Importance Sampling

- Densité auxiliaire optimale, *i.e.* minimisant la variance :

$$h_{opt}(\mathbf{X}) = \frac{\mathbf{1}_{\phi(\mathbf{X}) > sh_0}(\mathbf{X})}{\mathbb{P}}$$

- Par exemple, approcher itérativement la densité auxiliaire optimale :

- de manière paramétrique (Cross-entropy)

- [Rubinstein 97, 04]

- Minimiser la distance de Kullback-Leibler avec la densité auxiliaire optimale

- de manière non paramétrique [Zhang 96, Neddermeyer 10,12]

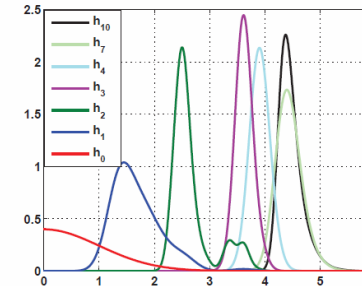
- Estimateur à noyau de densité = NAIS (non parametric adaptive importance sampling)

Contributions de recherche sur NAIS

■ Améliorer l'applicabilité de NAIS

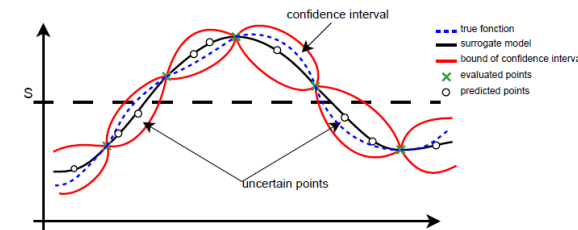
J. Morio, Extreme quantile estimation with nonparametric adaptive importance sampling, Simulation Modelling Practice and Theory

J. Morio, Nonparametric Adaptive Importance Sampling for the probability estimation of a launcher impact position, Reliability Engineering & System Safety



■ Couplage avec un métamodèle (krigeage)

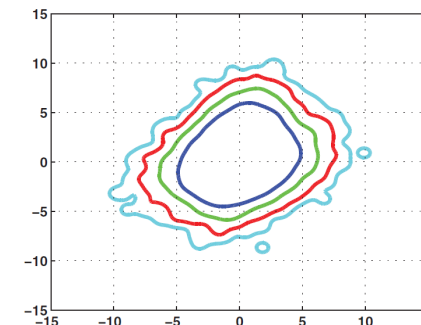
M. Balesdent, J. Morio and J. Marzat, Kriging-based adaptive importance sampling algorithms for rare event estimation. Structural Safety



■ Extension de NAIS dans le cas où la sortie du code de calcul est multidimensionnelle (Minimum Volume Set)

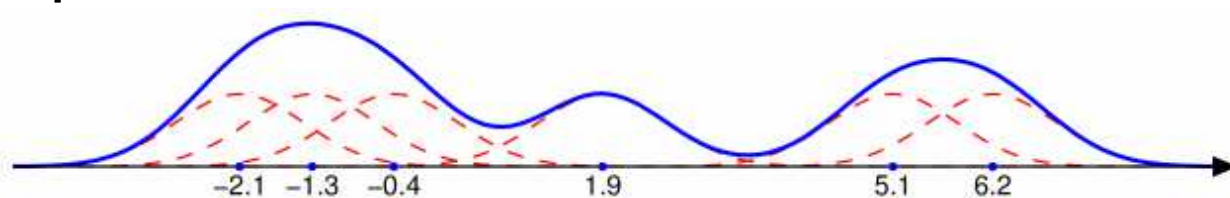
Thèse de R. Pastel

J. Morio and R. Pastel, Plug-in estimation of d-dimensional density minimum volume set of a rare event in a complex system, Proceedings of the IMechE, Part O, Journal of Risk and Reliability



Importance Sampling NAIS

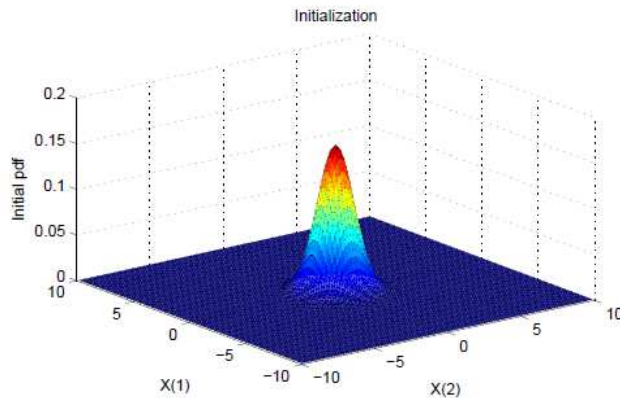
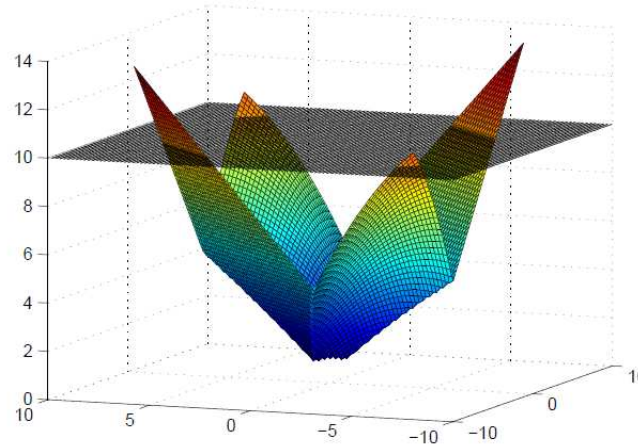
- **Améliorer l'applicabilité de NAIS** [Zhang 96, Neddermeyer 10,12]
 - Difficilement applicable en pratique (densité initiale doit déjà générer des échantillons tels que $\phi(\mathbf{X}) > S$)
 - Proposition d'amélioration de NAIS à l'aide de seuils itératifs, permettant d'initialiser l'algorithme par h_0 , écartant ainsi l'hypothèse précédente
 - Avantages et défauts inhérents à l'estimation par noyaux toujours présents



Cas à deux dimensions avec densité optimale de tirage multimodale

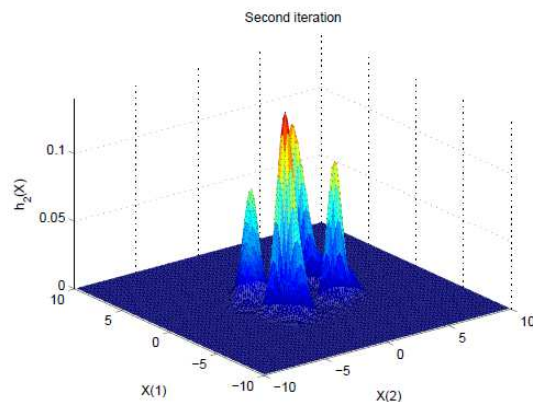
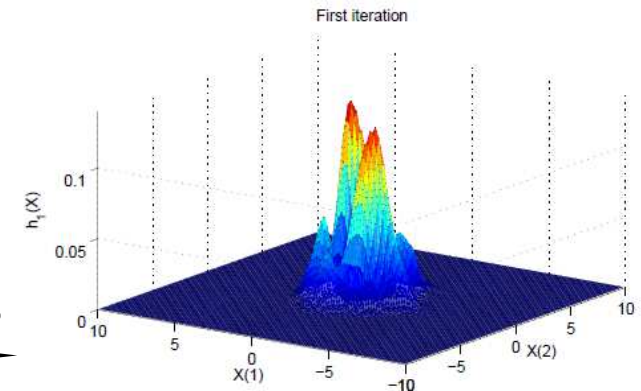
Fonction ϕ

X suit une loi
gaussienne
centrée réduite



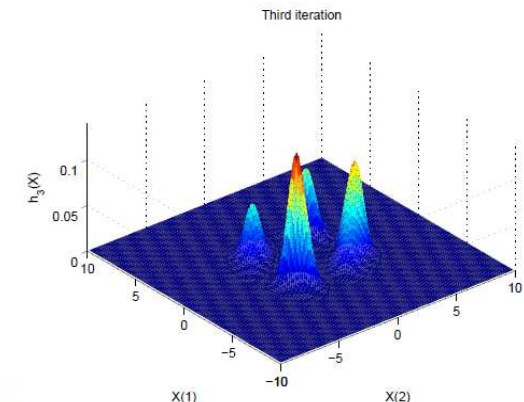
Initialisation NAIS

1ère itération NAIS



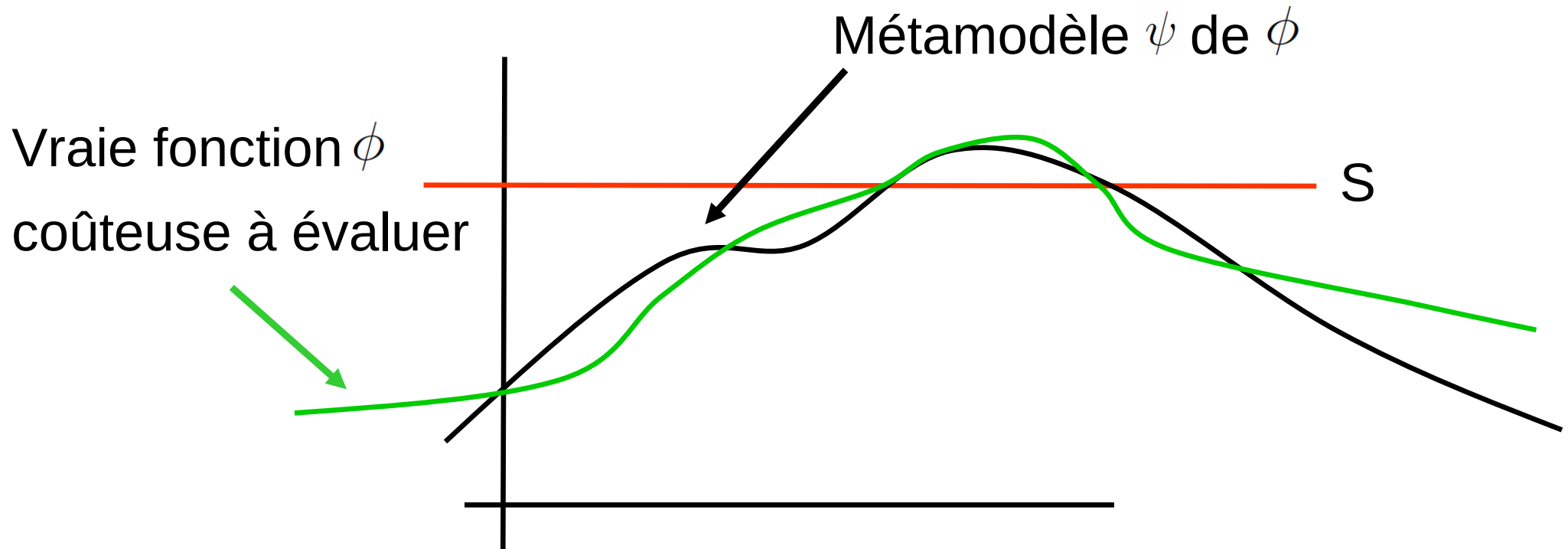
2ème itération NAIS

3ème itération NAIS



Couplage importance sampling - krigage

- Couplage avec un métamodèle (krigeage)

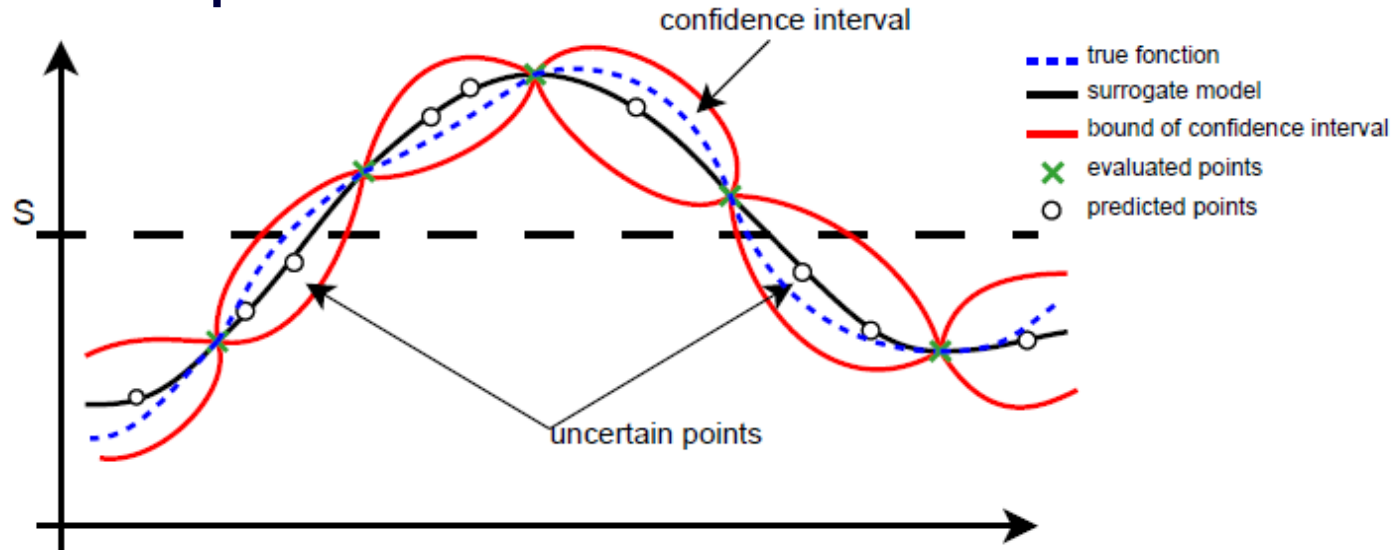


Si $\mathbf{X}$ est une variable aléatoire, $P(\phi(\mathbf{X}) > S) = P(\psi(\mathbf{X}) > S)$

Le métamodèle ψ n'a pas besoin d'être exactement représentatif de la fonction ϕ (seulement dans les zones d'intersection avec le seuil S)

Couplage importance sampling - krigage

- Métamodèle choisi, **le krigage** [Matheron 63] :
 - - Modèle de substitution **probabiliste**
 - - **Interpolant**
 - - Fournissant en tout point une estimation de la **variance de l'erreur de prédiction**



- Proposition d'une politique d'échantillonnage de la fonction ϕ dans le cas d'un estimateur importance sampling adaptatif

Couplage importance sampling - krigeage

- A chaque itération de l'algorithme :
 - Proposition d'échantillons par Importance Sampling (NAIS)
 - On construit un métamodèle adaptatif de la fonction ϕ **uniquement dans les zones où il y a « ambiguïté »**
 - Détermination des points à évaluer sur ϕ par **optimisation globale** (CMA-ES) [Hansen 08] d'un critère représentant **l'amélioration globale espérée** du nuage de points incertains **(obtenu grâce au krigeage)** jusqu'à ce que tous les points soient déterminés avec un degré de confiance donné

Couplage importance sampling - krigage

Résultats :

- Diminution d'un facteur 4 à 20 du nombre d'échantillons réellement évalués
- Estimation de l'erreur introduite par le métamodèle
- Couplage avec krigage aussi réalisé pour l'importance sampling paramétrique (Cross-Entropy)

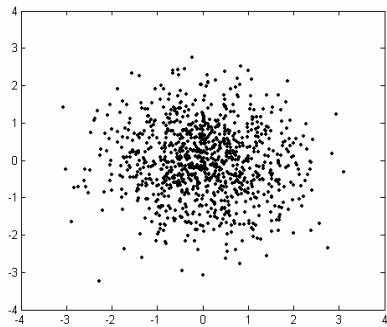
Perspectives :

- Technique à mettre en place sur des cas plus complexes
- Comparer avec les autres approches proposées dans le domaine (SUR [Bect 12], AK-IS [Dubourg 13]...)
- Améliorer la caractérisation heuristique de l'erreur introduite par le métamodèle

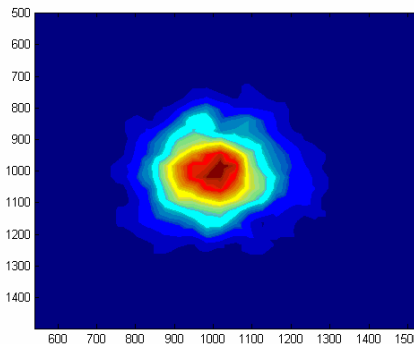
Minimum Volume Set (MVS)

- **Extension de NAIS dans le cas où la sortie du code de calcul est multidimensionnelle (Minimum Volume Set)**
 - Supposons que la sortie Y est de dimension r et de densité g
 - Définition d'une ligne de niveau $\mathbf{L}(t) = \{y \in \mathbb{R}^r : g(y) \geq t\}$
[Polonik 97]
 - Mapping avec les probabilités :
Une ligne de niveau t est associé à une aire de volume minimal de probabilité α (réciproque fausse) [Nunez-Garcia 03]

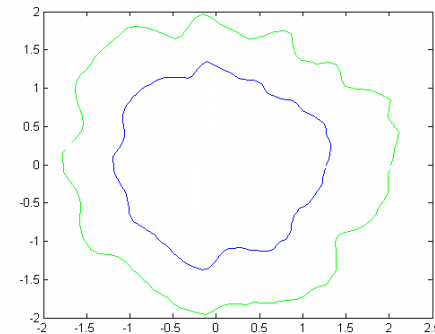
Exemple en dimension 2 :



1000 simulations



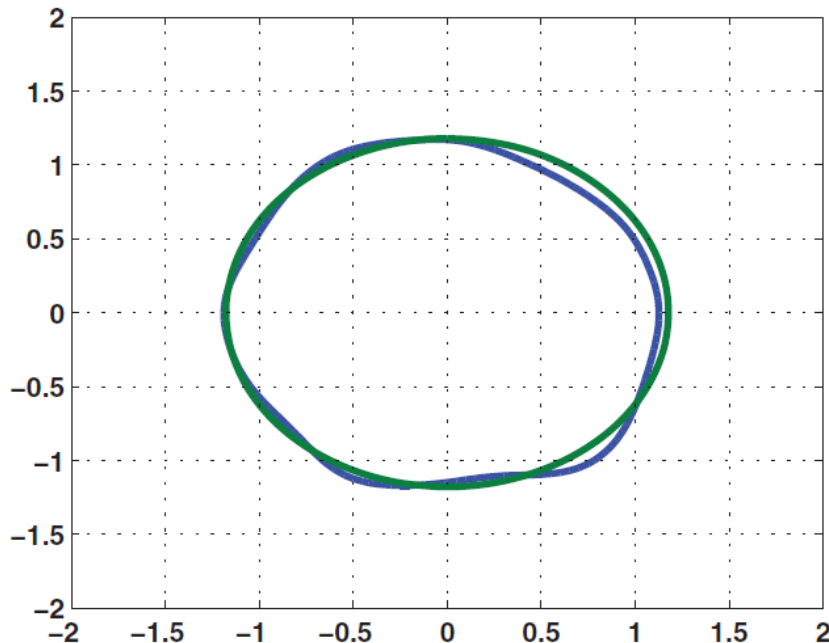
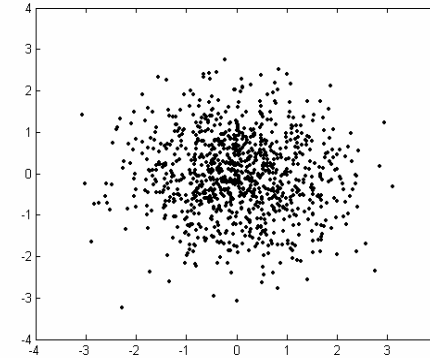
Densité de probabilité



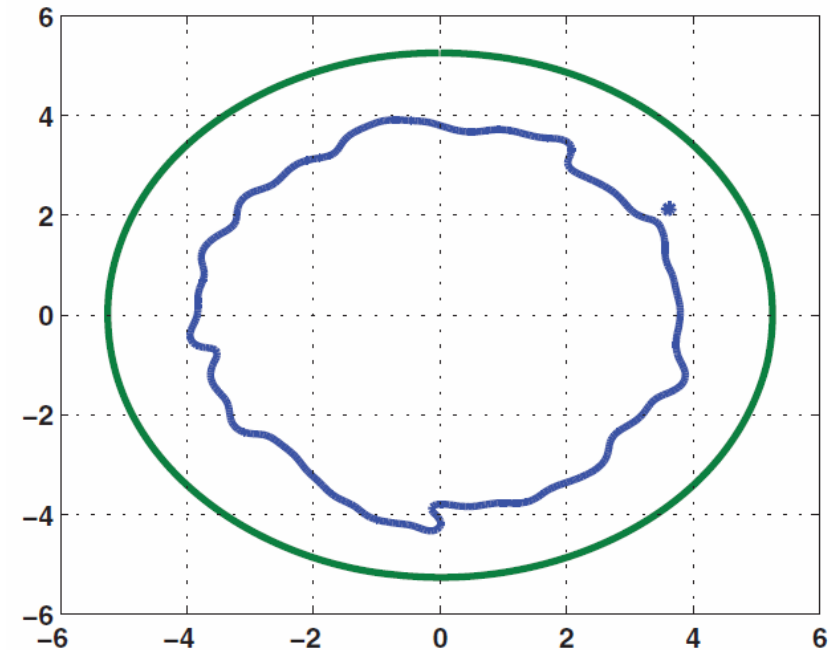
MVS à 50% et 80%

Minimum Volume Set

Cas simple gaussien
en dimension 2
avec méthodes Monte-Carlo
(10000 échantillons)



$\alpha = 0.5$

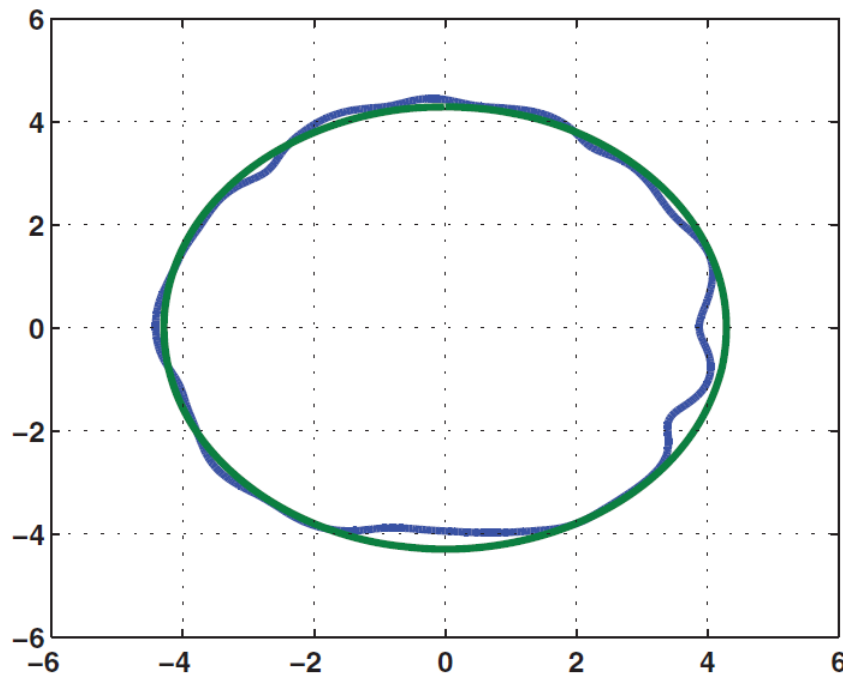


$\alpha = 0.99999$

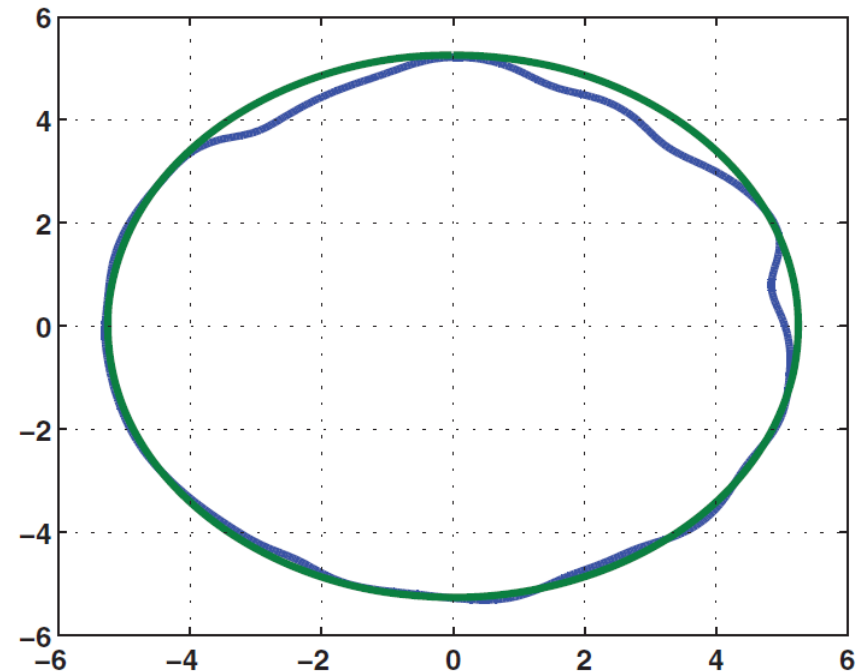
Courbe bleue : estimation Monte Carlo - courbe verte : théorie

Minimum Volume Set

Cas simple gaussien en dimension 2 avec NAIS (10000 échantillons)



$\alpha = 0.999$



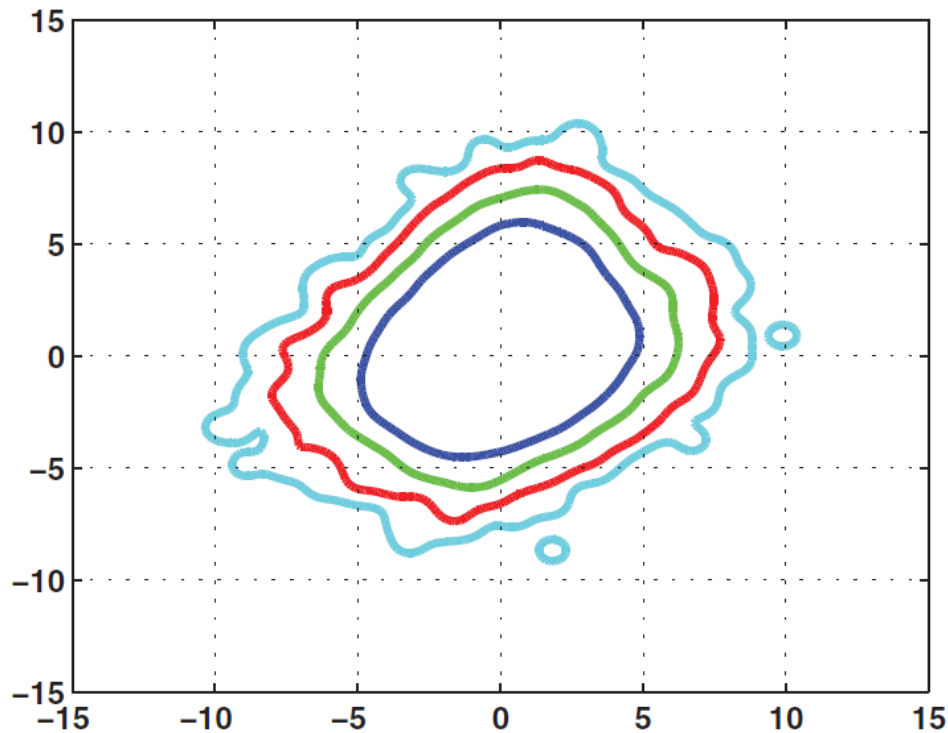
$\alpha = 0.99999$

Courbe bleue : estimation NAIS - courbe verte : théorie

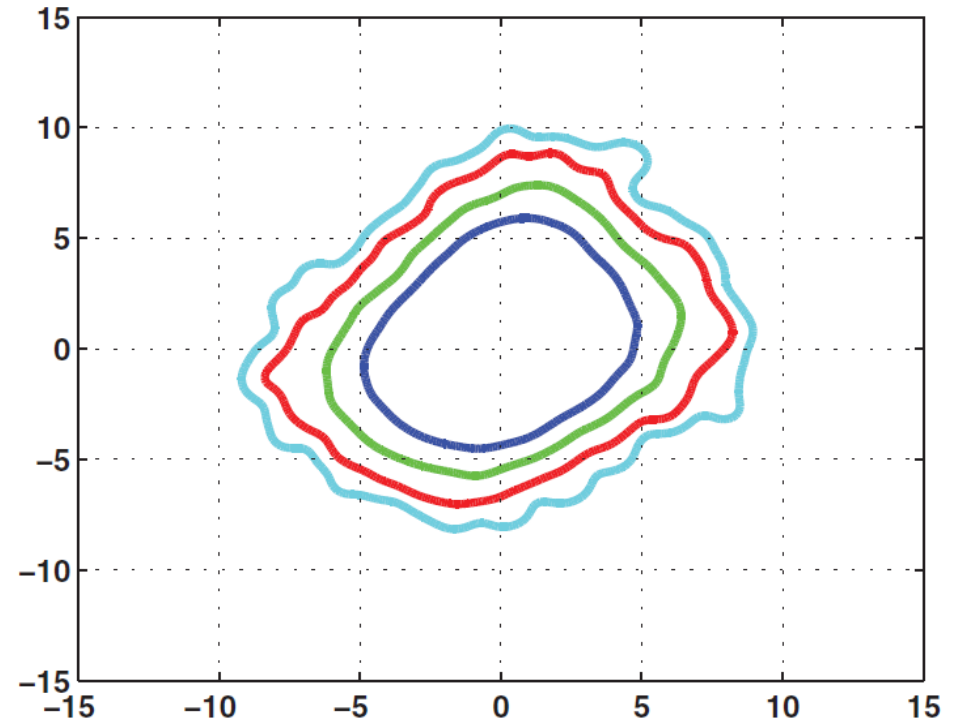
Minimum Volume Set

■ Cas d'une retombée d'engin

NAIS (10^4 échantillons)



Monte-Carlo (10^6 échantillons)



Niveau : 0.95, 0.99, 0.999, 0.99999

Perspectives :

■ Calcul de la variance d'un MVS

Plan de la présentation

- Problématique de recherche
- Synthèse sur les méthodes d'estimation d'évènements rares
- **Développements récents réalisés à l'ONERA**
 - Importance sampling
 - **Incertitudes sur les paramètres de modèle**
- Conclusion - Enjeux scientifiques

Incertitudes sur les paramètres de modèle

- Système statique décrit par une fonction boîte noire entrée-sortie

Fonction positive scalaire : $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$

Entrées : $\mathbf{X}$ de dimension d , aléatoires de densité h_0

Densité h_0 paramétrée par : $\Theta = (\Theta_1, \dots, \Theta_m)$

Probabilité visée : $\mathbb{P}_\theta = P(\phi(\mathbf{X}) > S | \Theta = \theta)$

Contributions de recherche à l'analyse des incertitudes sur les paramètres de modèle

- **Estimer à moindre coût $\mathbb{P}_\theta$ pour différentes valeurs θ de Θ**

J. Morio and M. Balesdent, Parametric and non parametric importance sampling algorithms for rare event probability estimation with uncertain input distributions, submitted to Proceedings of the IMechE, Part O, Journal of Risk and Reliability

- **Déterminer les extrema de $\mathbb{P}_\theta$**

M. Balesdent, J. Morio and L. Brevault, Rare event probability estimation using Kriging-based adaptive importance sampling in presence of epistemic uncertainty on input probability distribution parameters, conditionnally accepted in Methodology and Computing in Applied Probability

- **Analyser l'influence de Θ sur $\mathbb{P}_\theta$**

J. Morio, Influence of input PDF parameters of a model on a failure probability estimation, Simulation Modelling Practice and Theory

- **Estimer la loi de Θ sachant $\phi(\mathbf{X}) > S$**

Thèse de C. Vergé

C. Vergé, J. Morio and P Del Moral, An island particle Markov chain Monte Carlo algorithm for safety analysis, submitted to Structural Safety

Incertitudes sur les paramètres de modèle

- **Estimer à moindre coût $\mathbb{P}_\theta$ pour différentes valeurs θ de Θ**

Approche proposée avec NAIS et Cross-Entropy (CE) :

- réutiliser les échantillons générés pour estimer $\mathbb{P}_\theta$
- améliorer la densité initiale de NAIS et CE pour l'estimation de $P(\phi(\mathbf{X}) > S | \Theta = \theta')$
- réduction du nombre d'échantillons

- **Déterminer les extrema de $\mathbb{P}_\theta$**

Si les composantes de Θ varient dans un intervalle, développement d'une approche couplée Importance Sampling + krigeage + optimisation

Incertitudes sur les paramètres de modèle

- Analyser l'influence de Θ sur $\mathbb{P}_\theta$

Un estimateur possible de $\mathbb{P}_\theta$:

$$\hat{\mathbb{P}}_\theta = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\phi(\mathbf{X}_i) > S} \frac{h_{0, \Theta=\theta}(\mathbf{X}_i)}{h_{\langle \theta \rangle}(\mathbf{X}_i)}$$

avec $h_{\langle \theta \rangle}$ une densité efficace de tirage pour $\Theta = \langle \theta \rangle$

On obtient alors une relation du type :

$$\hat{\mathbb{P}}_\theta = \Gamma(\Theta) = \Gamma(\Theta_1, \dots, \Theta_m)$$

Analyse de sensibilité (indices de Sobol par exemple) permet de caractériser l'influence des Θ_i sur $\mathbb{P}_\theta$

Incertitudes sur les paramètres de modèle

Code simplifié d'évaluation de turbulences de sillage:

- densité de l'entrée paramétrée par $\Theta = (\Theta_1, \dots, \Theta_6)$ de dimension 6
- influence des Θ_i sur $\mathbb{P}_\theta$?

Paramètre	Indice de Sobol total
Θ_1	0.07
Θ_2	0.12
Θ_3	0.32
Θ_4	0.44
Θ_5	0.18
Θ_6	0.58

Perspectives :

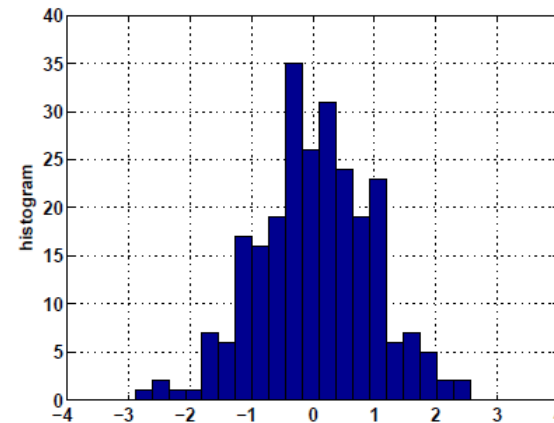
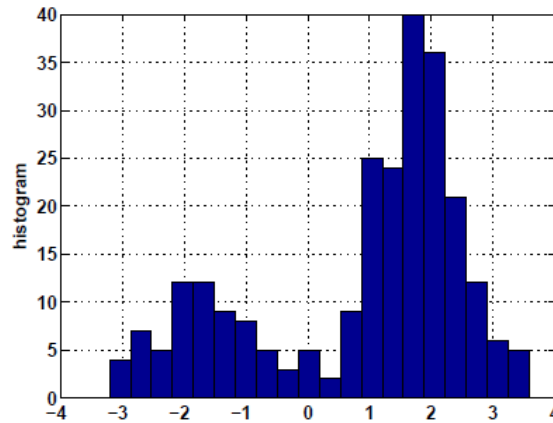
- Approche heuristique (validation numérique) mais peu coûteuse
- Comparaison avec d'autres méthodes [Lemaître 13]

Incertitudes sur les paramètres de modèle

- **Estimer la loi de Θ sachant $\phi(\mathbf{X}) > S$**
 - Déterminer les valeurs θ de Θ les plus vraisemblables avec valeur élevée de $\mathbb{P}_\theta$ (événement critique)
 - Algorithme SMC² (Sequential Monte-Carlo square) [Chopin 13] :
SMC sur Θ = algorithme génétique avec alternance d'étapes de sélection et de mutation basées sur une probabilité de dépassement de seuil, estimée, elle aussi par SMC
 - En sortie de l'algorithme SMC², estimation de la loi de Θ sachant $\phi(\mathbf{X}) > S$

Incertitudes sur les paramètres de modèle

- **Code simplifié d'une retombée d'engin :**
 - 4 entrées aléatoires gaussiennes de moyenne Θ
 - Θ suit une loi gaussienne centrée réduite de dimension 4.



Composantes de Θ : Θ_1 Θ_3

- **Histogrammes des lois de Θ_1 et de Θ_3 sachant $\phi(\mathbf{X}) > S$**

Perspectives :

- Travail sur l'interprétation de ces lois
- Réduire le temps de calcul
- Lien avec l'analyse de sensibilité
- Caractériser la convergence de l'algorithme

Plan de la présentation

- Problématique de recherche
- Synthèse sur les méthodes d'estimation d'évènements rares
- Développements récents réalisés à l'ONERA
- Conclusion - Enjeux scientifiques

Conclusion

- Établir un lien entre méthodes et problématiques industrielles
- Améliorer l'applicabilité des algorithmes (comparaison, automatisation, évaluation, réglages)
- Cas des systèmes dynamiques (modélisation par chaînes de Markov) – trajectoires d'aéronefs, largages de missiles...
- Opportunités de collaborations scientifiques ?

- Traitement des systèmes très multidimensionnels et coûteux en temps de calcul
- Parallélisation de code et/ou méthodes d'estimation d'évènements rares
- Lien entre théorie des valeurs extrêmes et simulation
- Fiabilisation des méthodes d'évènements rares

Merci de votre attention.

Questions?