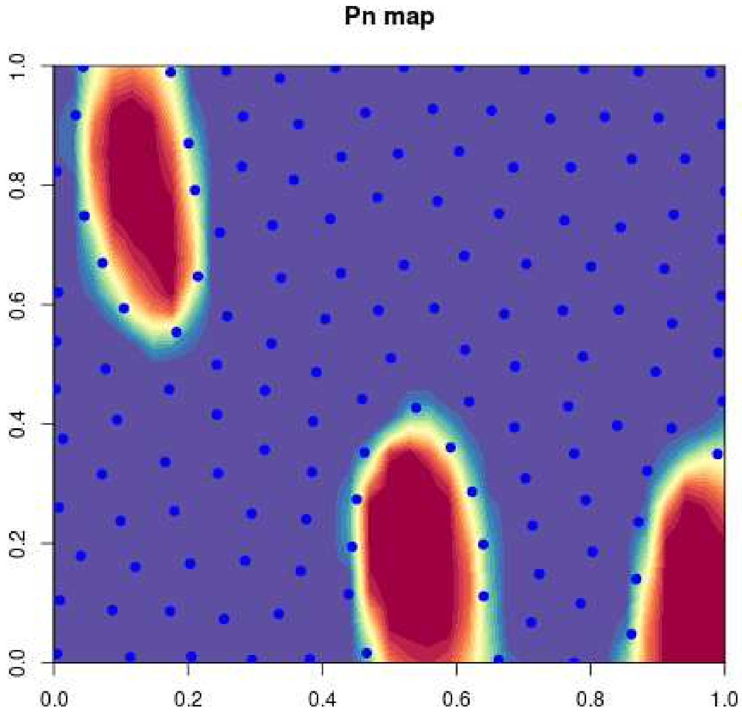
Learning and space-filling design space constrained from hidden failure of computer experiments

Morgane Menz, Miguel Munoz Zuniga, Delphine Sinoquet

IHP, 10-11 December 2024

LEARNING AND SPACE-FILLING OBJECTIVES

- Accelerate: learning of the hidden constraint and generation of a space-filling design of experiments in the hidden constrained space
  - Sequential strategy: propose point dedicated **alternatively** to learning and space-filling
  - Coupled strategy: propose point dedicated **simultaneously** to learning and space-filling



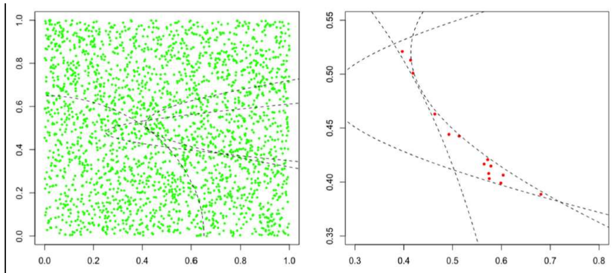
SPACE-FILLING IN CONSTRAINED SPACE

CoMinED [Huang et al.. 2021] : **Sequential** construction technique of **space-filling designs** with **continuous constraints**

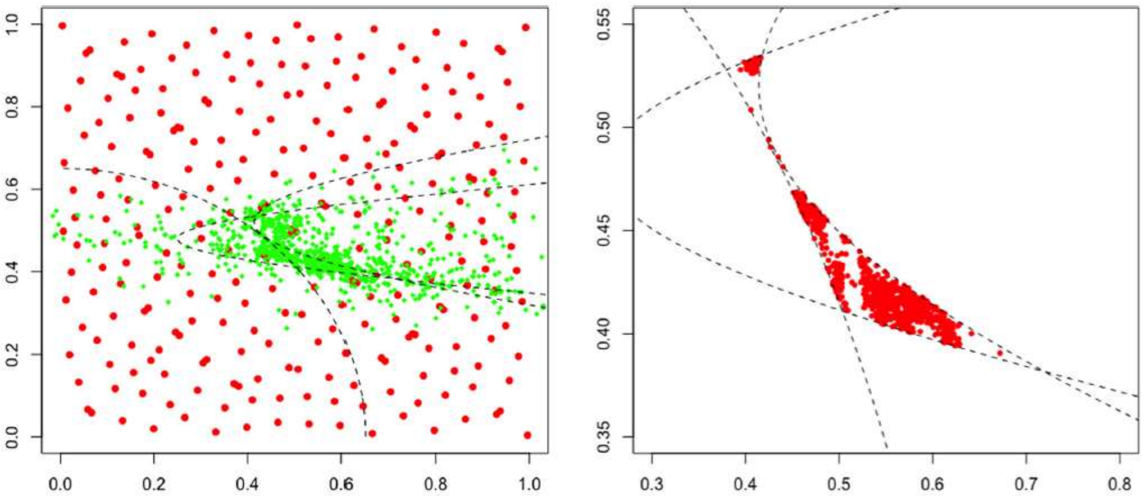
- **Candidate Generation: generate a large set of uniformly distributed candidates in X.**
- Design Construction: choose points from the set of candidates by a desired criterion.

$$\{x \in \Omega, g(x) \leq 0\}$$

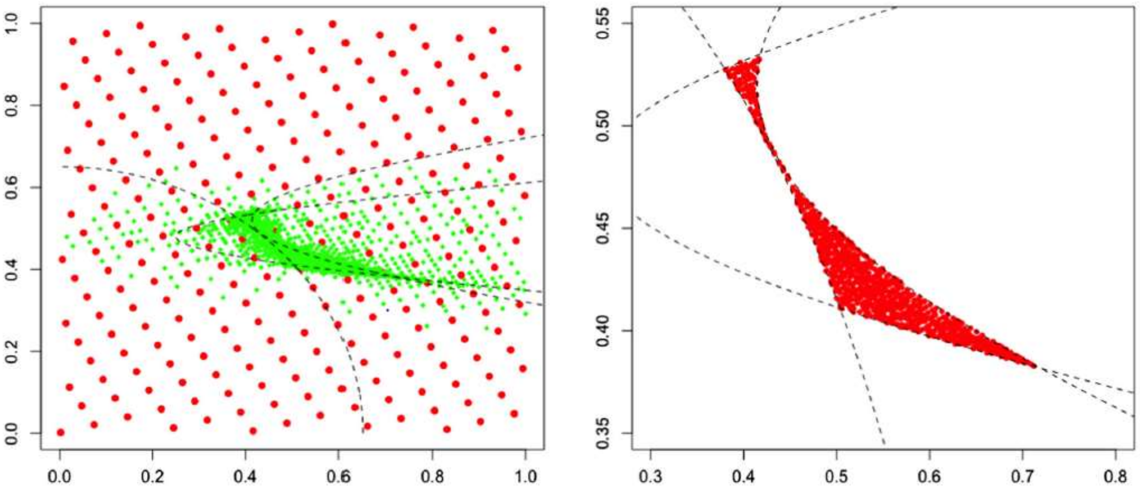
GRAPHICAL MOTIVATION [Huang et al.. 2021]



Rejection/Acceptation



Red left:
265 Sobol' points
Green left:
Adaptive Sequentially Constrained Monte Carlo
Golchi and Loeppky, 2015
Red Right:
feasible points

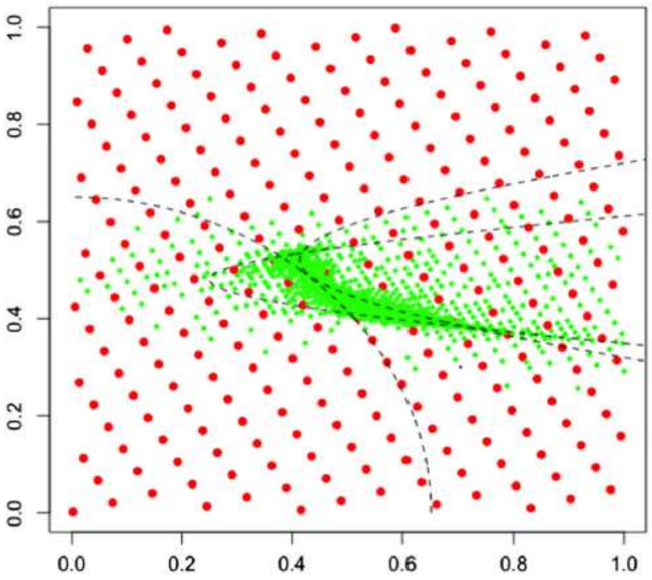


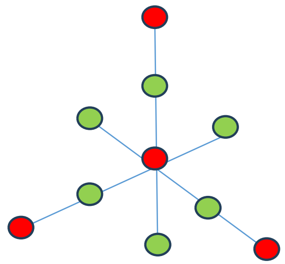
Red left:
lattice grid of points
Green left:
CoMinED population: Q-NN, middle and reflection
Red Right:
feasible points

CONSTRAINED MINIMUM ENERGY DESIGNS ALGORITHM

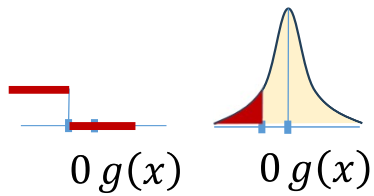
Augmentation of the set of candidate samples $\mathcal{C}^t$

Q-NN, middle point and reflection





CONSTRAINED MINIMUM ENERGY DESIGNS ALGORITHM



Constrained greedy maximin

$$\max_x \mathbb{1}_{g(x)\leq 0}\mathbb{1}_{g(x_i)\leq 0}\min_i \|x_i - x\|_s$$

$$\mathbb{1}_{g(x)\leq 0} \sim \Phi(-\tau g(x))$$

Inputs:

- sequence of rigidity parameters $(\tau_t)_{t=1,\dots,T}$.
- continuous constraints $g(x)$ -> g(x)<0
- parameters for candidate set augmentation : Q

- Initialisation of the initial candidate set $\mathcal{C}^1$
- For t= 1. … . T
  - Construction of a constrained DoE $\mathcal{D}_c$: **one-point-at-a-time Greedy-algorithm** with τ_t on candidate samples $\mathcal{C}^t$

$$x_{m+1}^t = \underset{x\in C^t\setminus\{x_l\}_{l=1}^m}{\operatorname{argmax}} \Phi(-\tau_t g(x))\left[\min_{i=1,\dots,m} \Phi(-\tau_t g(x_i))\|x_i - x\|_s^{2p}\right]$$

  - Augmentation of the set of candidate samples $\mathcal{C}^t$

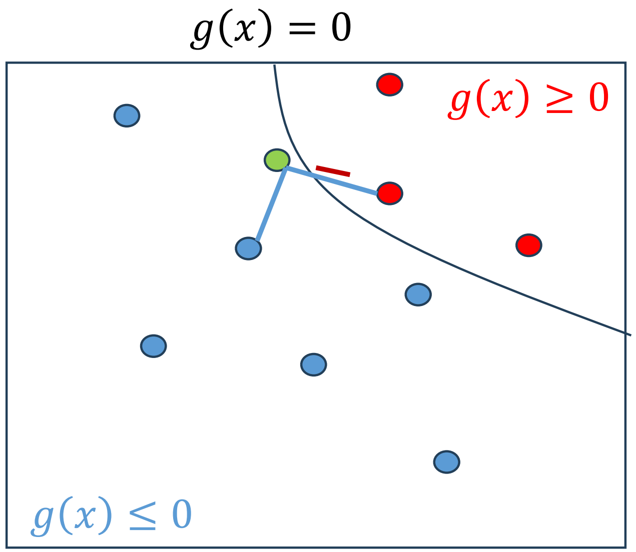
$$\|u\|_s = \left(\frac{1}{p}\sum_{l=1}^{p}|u_l|^s\right)^{\frac{1}{s}}$$ $\mathcal{D}_c = \{x_l\}_{l=1}^n$ Φ: Gaussian cumulative prob. function

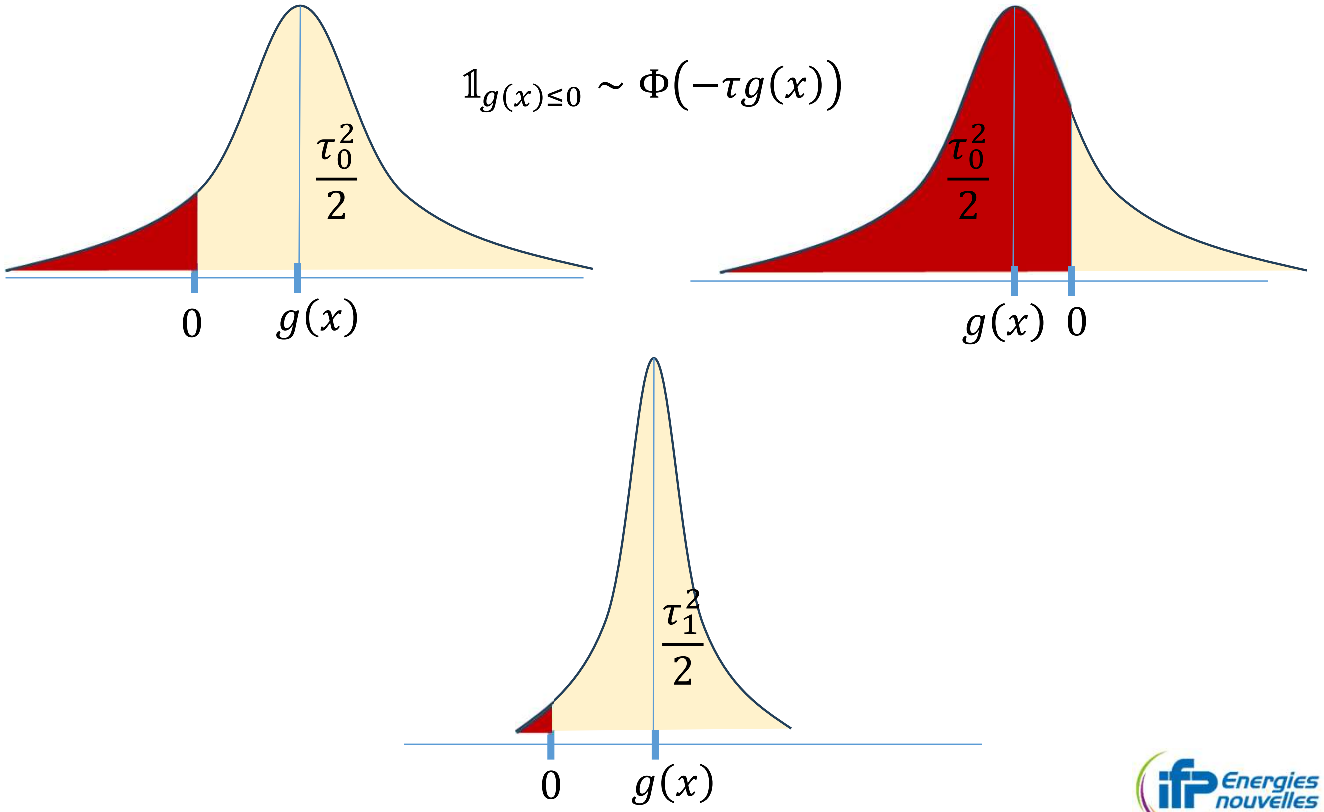
COMINED CRITERION FROM WITHIN

$0 < \tau_0 < \cdots < \tau_T$

$$x_{m+1}^t = \underset{x \in C^t \backslash \{x_l\}_{l=1}^m}{\text{argmax}} \; \Phi(-\tau_t g(x)) \left[\min_{i=1,\ldots,m} \Phi(-\tau_t g(x_i)) \|x_i - x\|_s^{2p} \right]$$

$\mathbb{1}_{g(x) \leq 0} \sim \Phi\big(-\tau g(x)\big)$

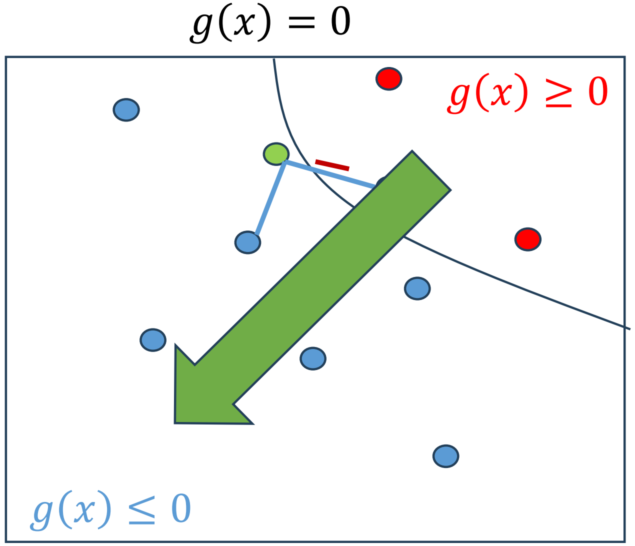


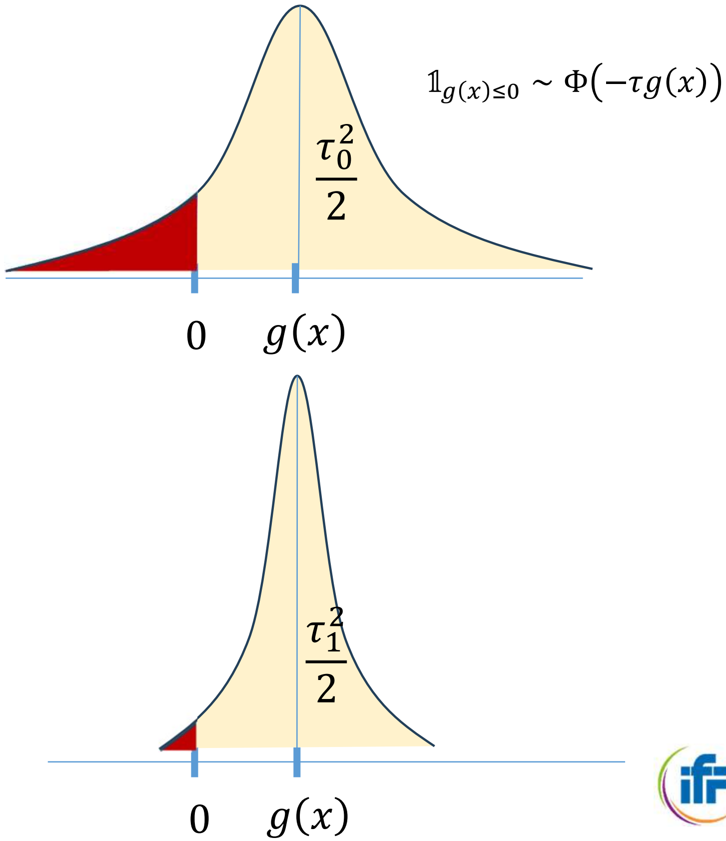


COMINED CRITERION FROM WITHIN

$0 < \tau_0 < \cdots < \tau_T$

$$x_{m+1}^t = \underset{x \in C^t \setminus \{x_l\}_{l=1}^m}{\operatorname{argmax}} \Phi(-\tau_t g(x)) \left[\min_{i=1,\ldots,m} \Phi(-\tau_t g(x_i)) \|x_i - x\|_s^{2p} \right]$$





APPLICATION WITH HIDDEN CONSTRAINTS

In **hidden** constraints context: one **binary** constraint $y(x) \in \{0,1\}$

- $y(x)$ result of **expensive simulations**
- $y(x)$ **non-continuous constraint**: untractable with usual constrained space-filling construction techniques

Consider a conditioned **Gaussian Process Classifier [Bachoc et al. 2020]** model $Y_n(x) = Y(x)|x_n, y_n$ of the hidden constraint $y(x)$. It predicts: $p_n(x) = \mathbb{P}[Y_n(x) = 1]$

CoMinED can be applied with :

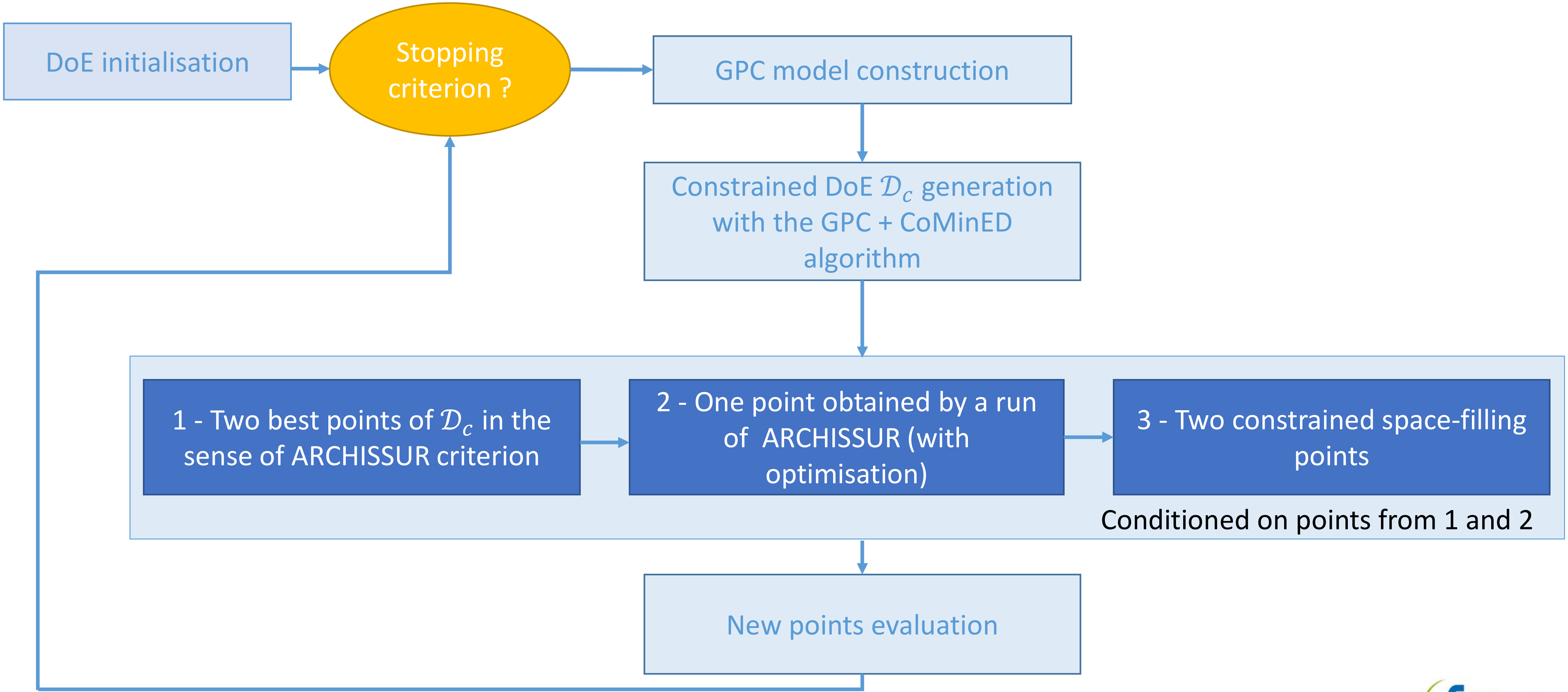
$$g(x) = \alpha^* - p_n(x)$$

$$x_{m+1}^t = \underset{x \in \mathrm{C}^t \backslash \{x_l\}_{l=1}^m}{\operatorname{argmax}} \Phi(-\tau_t g(x)) \left[\min_{i=1,\ldots,m} \Phi(-\tau_t g(x_i)) \|x_i - x\|_s^{2p} \right]$$

α^* estimated within the Vorob'ev expectation of $\Gamma = \{x \in \Omega | Y_n(x) = 1\}$ estimation

$$\hat{\boldsymbol{\Gamma}}_{\boldsymbol{n}} = \{\boldsymbol{x} \in \boldsymbol{\Omega}, \boldsymbol{p_n}(\boldsymbol{x}) \geq \boldsymbol{\alpha}^*\}$$

PROPOSED ADAPTIVE SPACE-FILLING STRATEGY



STOPPING CRITERION OF THE ADAPTIVE ALGORITHM

1st scenario: **reaching a total number of simulations or "convergence"**

At iteration i stop if:

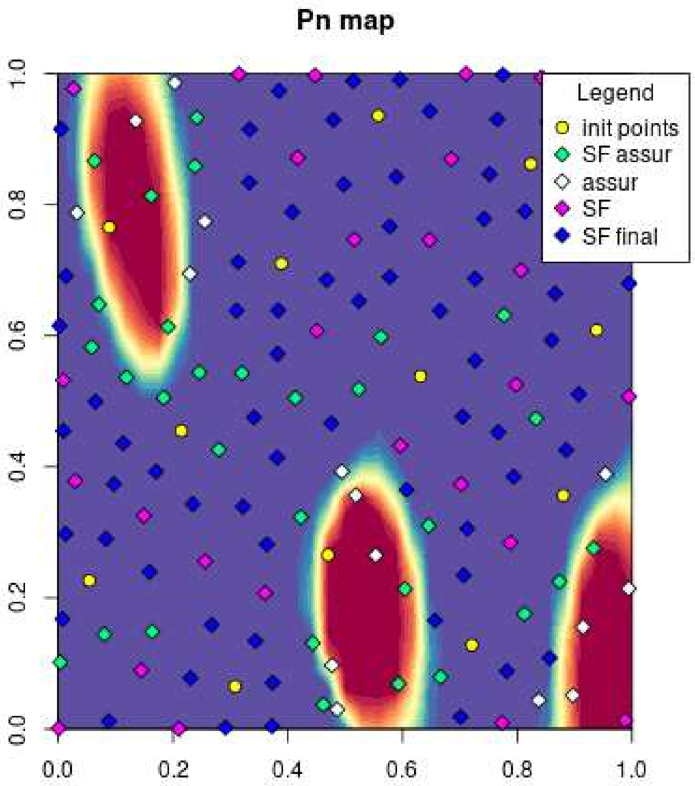
$$n_{sim}^{i} > n_{sim}^{adapt} \quad or \quad \frac{|Var_j(\Gamma) - Var_{j-1}(\Gamma)|}{\mathbb{E}_j[\mu(\Gamma)]} < \epsilon \quad \forall j \in \{i - l. \dots . i\}$$

$$n_{sim}^{adapt} = \lfloor C n_{sim} \rfloor$$

With the remaining simulation budget: Constrained space-filling with current GPC model

2nd scenario: **reaching a number of successful simulations**

ILLUSTRATION ON BRANIN EXAMPLE



Inputs:

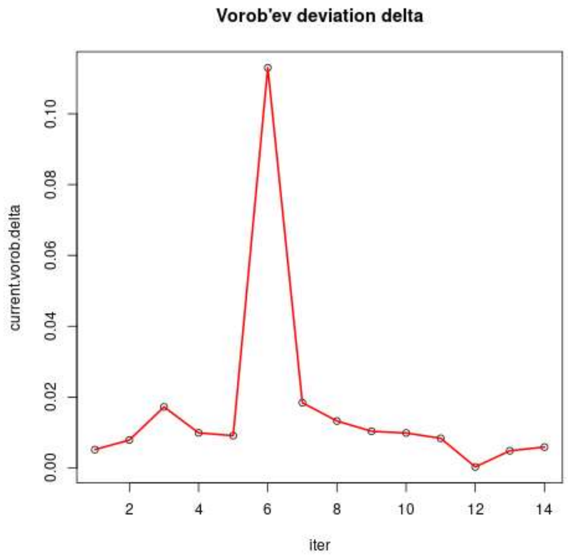
$n_{init} = 12$

$n_{sim}^{max} = 12 + 30 * 5$

$n_{sim}^{adapt} = 24 * 5$

$l = 4$

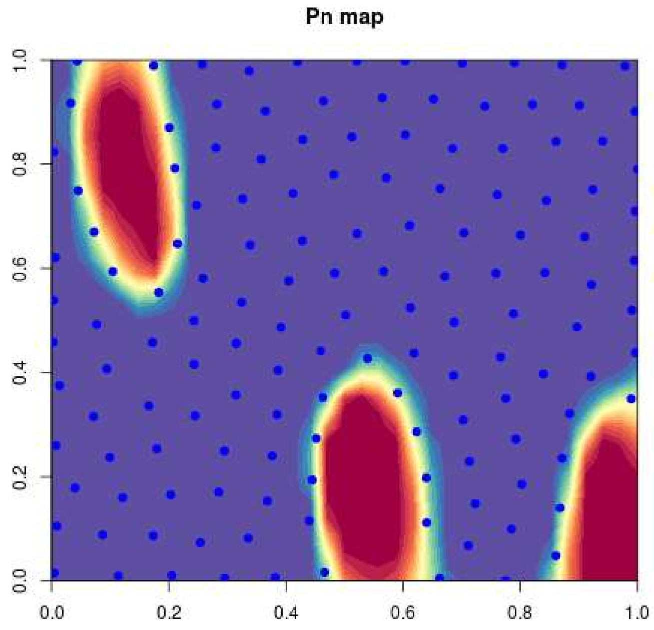
$\epsilon = 10^{-2}$



COMPARAISON OF RESULTS

- Comparison with **optimal** constrained DoE
  → In practice : obtained with simulated annealing [Auffray. 2012] adapted to non-connex domains [internal IFPEN report] on the final DoE of the adaptive procedure

- Comparison with a **crude rejection** method

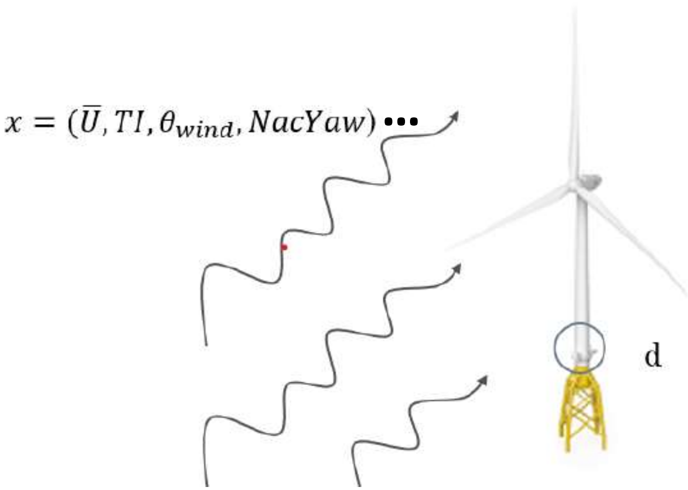


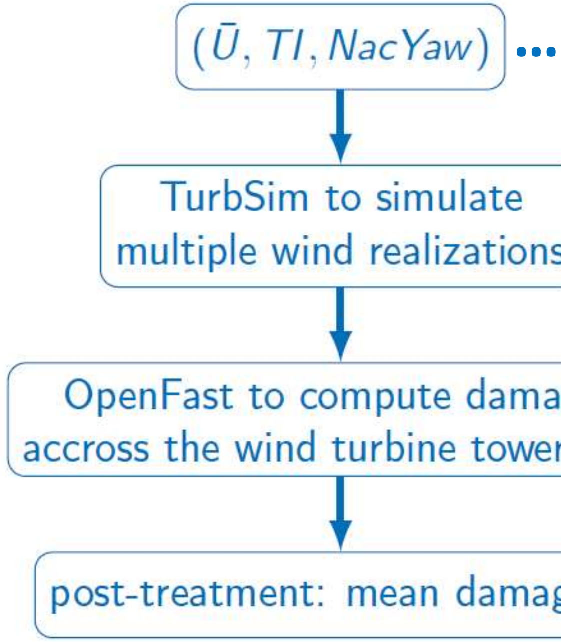
DAMAGE PREDICTION OF A WIND TURBINE

Wind loads are described by 13 parameters:

- wind speed $\bar{U}$,
- turbulence intensity TI,
- misalignment angle $NacYaw$.

…





 |

COMPARAISON OF RESULTS – WIND TURBINE

Initial doe size: 100
total simulation points: 1000
dimension: 13

$$Q_{\alpha^*} = \{x \in \Omega : p_n(x) \geq \alpha^*\}$$

	MINMAX	MAXMIN	EFFICIENCY
LHS MAXIMIN + REJECTION	2.12	1.25	0.42
Initial GPC + COMINED	1.68	0.85	**0.93**
ASUR SF	1.66	0.98	0.75
ASUR SF + COMINED (Greedy) (80% + 20%)	**1.63**	**0.94**	**0.79**

Robustesse ?
More conservative ?
Points on frontier more
Spread on the feasable area

$$\text{Efficiency} = \frac{n_{success}}{n_{sim}}$$

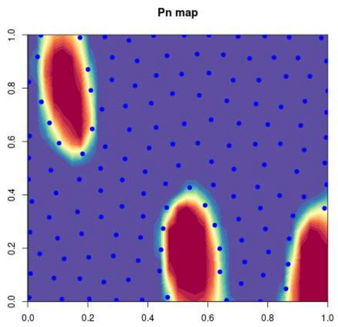
COMPARAISON OF RESULTS – WIND TURBINE

Initial doe size: 100
total simulation points: 1000
dimension: 13

$$Q_\alpha = \{x \in \Omega : p_n(x) \geq \alpha\}, \alpha \in (0, 1]$$

Q2 ->	Proba 0.2	Proba 0.4	Proba 0.5	Proba 0.6	Proba 0.8	Proba0.95
REJECTION (0.067)	NA	NA	NA	NA	NA	NA
Initial GPC + COMINED	**0.57**	**0.61**	0.61	0.63	0.64	0.76
ASUR SF	0.08	0.45	0.77	**0.86**	**0.92**	0.97
ASUR SF + COMINED (80% + 20%)	-0.002	**0.61**	**0.81**	0.85	0.91	**0.98**

$$\text{Efficiency} = \frac{n_{success}}{n_{sim}}$$



Work in progress

A PROCESS OF SUR-IFICATION

Stepwise Uncertainty Reduction (SUR) and assimilated one-step-lookahead strategies

criteria

- Variance
- Expected Improvement
- Integrated Bernouilli Variance
- Vorob'ev deviation
- **Bichon**
- Ranjan

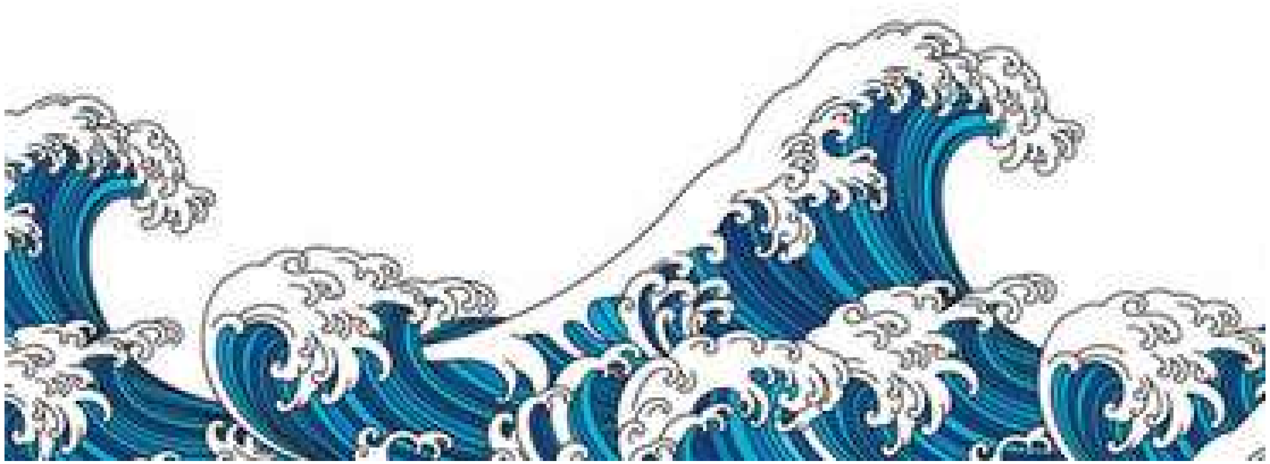
…

A PROCESS OF SUR-IFICATION

Stepwise Uncertainty Reduction (SUR) and assimilated one-step-lookahead strategies

Proofs of convergence
of the SUR based estimator

SUR-ification of criteria



criteria

A SAMPLE FROM THE PROCESS OF SUR-IFICATION

Stepwise Uncertainty Reduction (SUR) and assimilated one-step-lookahead strategies

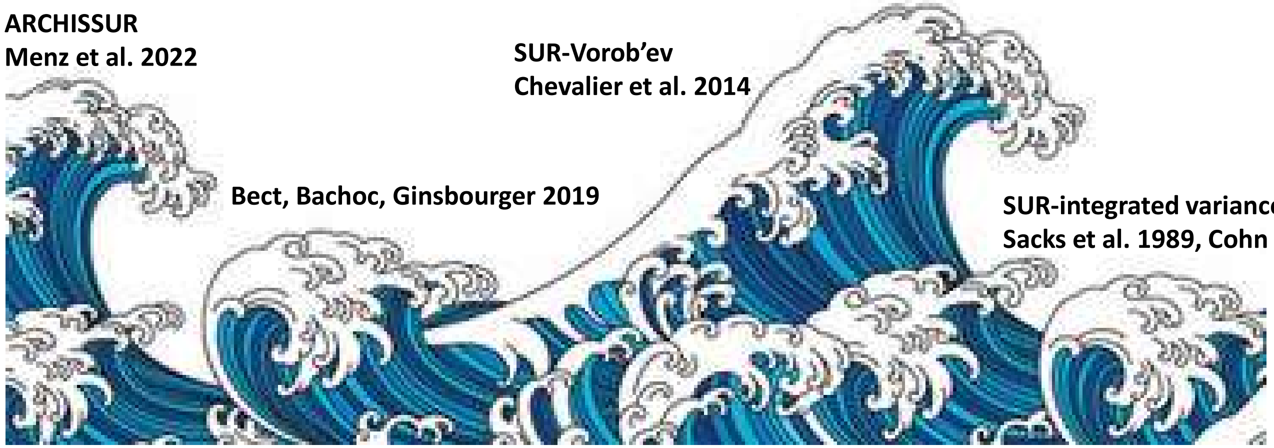
Proofs of convergence
of the SUR based estimator

SUR-ification of criteria

SUR-EER
Roy et al. 2001

ARCHISSUR
Menz et al. 2022

SUR-Vorob'ev
Chevalier et al. 2014



Bect, Bachoc, Ginsbourger 2019

SUR-integrated variance
Sacks et al. 1989, Cohn et al. 1996

SUR-Bichon
Duhamel et al. 2024

criteria

SUR-Integrated Bernouilli
Bect et al. 2012

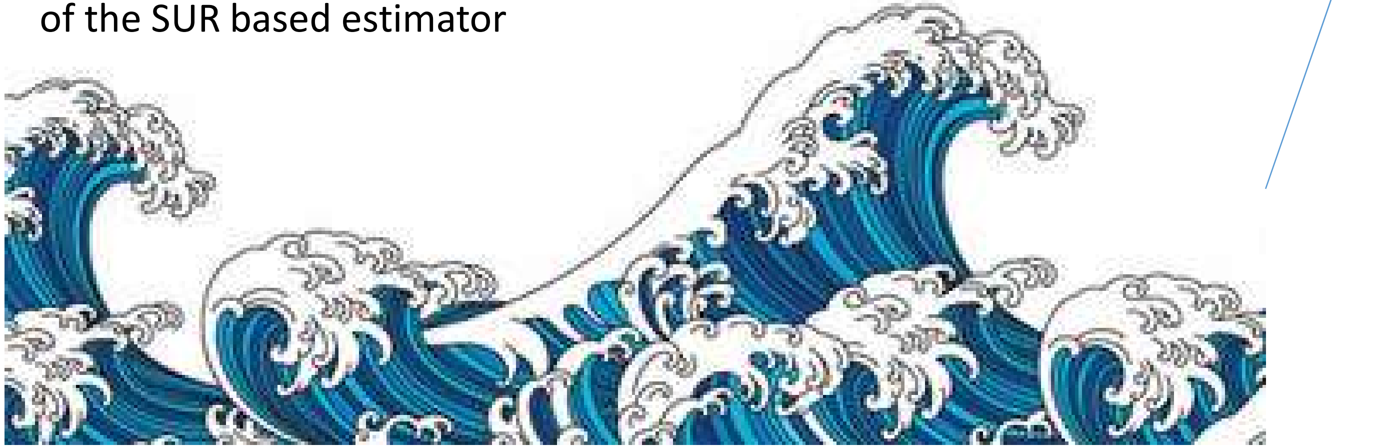
Bect Vazquez 2010
Bull 11
Srinivas et al 2012

PROCESS OF SUR-IFICATION

Stepwise Uncertainty Reduction - SUR

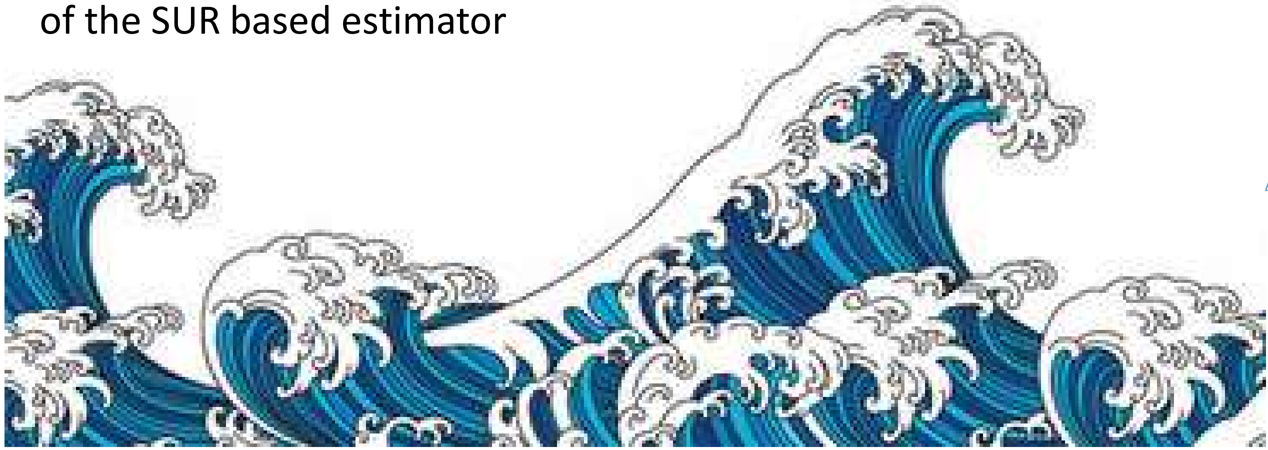
Proofs of convergence
of the SUR based estimator

SUR-ification of criteria



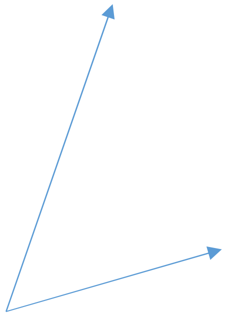
PROCESS OF SUR-IFICATION

Stepwise Uncertainty Reduction - SUR



Proofs of convergence
of the SUR based estimator

SUR-ification of criteria

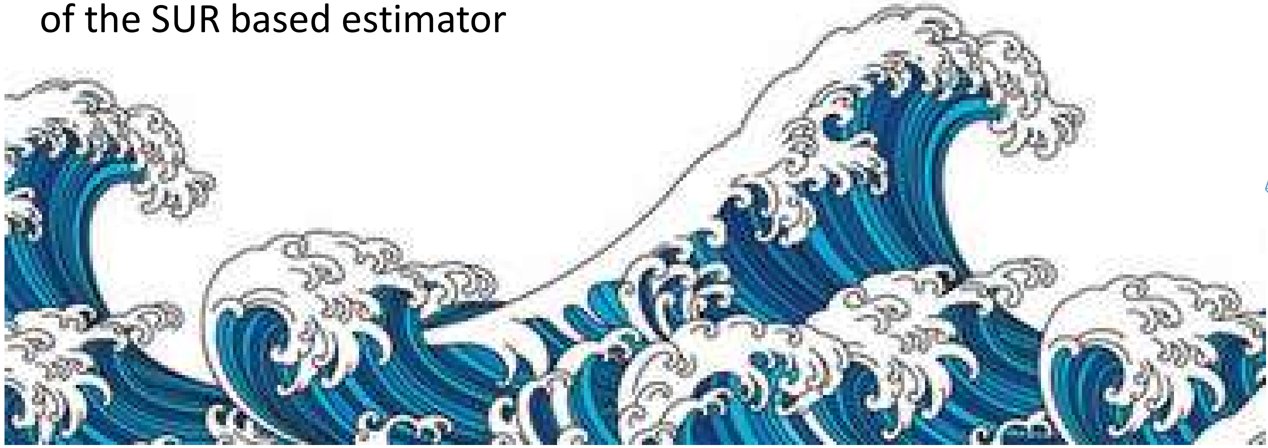


PROCESS OF SUR-IFICATION

Stepwise Uncertainty Reduction - SUR

Proofs of convergence
of the SUR based estimator

SUR-ification of criteria



1. VANILLA ONE-STEP-LOOKAHEAD

$$g_{t,m+1}(x) = \alpha_m^* - p_{t,m+1}(x)$$

$$\alpha_m^* \sim 1/2$$

$$\underset{x_{m+1}\in C^t \backslash \{x_l\}_{l=1}^m}{argmax} \mathbb{E}_{Y_{m+1}}[U_{t,m+1}]$$

$$U_{t,m+1} = \min_{i=1,\dots,m} \frac{1}{2p} \log\Big(\Phi\big(-\tau_t g_{t,m+1}(x_{m+1})\big)\Big) + \frac{1}{2p} \log\Big(\Phi\big(-\tau_t g_{t,m}(x_i)\big)\Big) + \log(\|x_i - x_{m+1}\|_s)$$

$$\underset{x_{m+1}}{argmax}\, \mathbb{E}_{Y_{m+1}}\big[\Phi\big(-\tau_t g_{t,m+1}(x_{m+1})\big)\big] \min_{i=1,\dots,m} \Phi\big(-\tau_t g_{t,m}(x_i)\big)\|x_i - x_{m+1}\|_s^{2p}$$

$$\mathbb{E}_{Y_{m+1}}\big[\Phi\big(-\tau_t g_{t,m+1}(x_{m+1})\big)\big] = p_{t,m}(x_{m+1})\Phi\big(-\tau_t(\alpha_m^* - 1)\big) + \big(1 - p_{t,m}(x_{m+1})\big)\Phi(-\tau_t \alpha_m^*)$$

2. A STEP BACK

$$g_{t,m}(x) = \alpha_m^* - p_{t,m}(x)$$

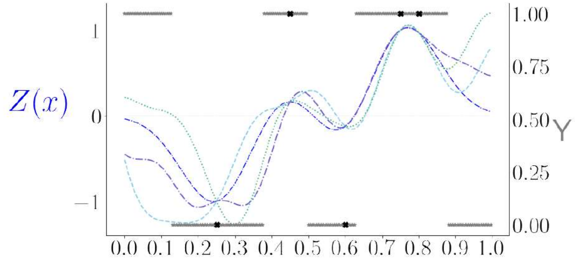
$$\underset{x_{m+1}}{argmax} \left[\Phi\left(-\tau_t g_{t,m}(x_{m+1})\right) \min_{i=1,\dots,m} \Phi\left(-\tau_t g_{t,m}(x_i)\right) \|x_i - x_{m+1}\|_s \right]$$

$$\prod_{i=1}^{N} \Phi\left(\tau_t\left(\mu_{t,m}^i(x_{m+1}) - \sigma_{t,m}(x_{m+1})\Phi^{-1}(\alpha_m^*)\right)\right)$$

$$\prod_{i=1}^{N} \Phi\left(\tau_t\left(\mu_{t,m}^i(x_{m+1}) + k\,\sigma_{t,m}(x_{m+1})\right)\right)$$

In [Bachoc et al., 2020], the probability of non-failure is modeled using the sign of a latent GP $Z(x) \sim GP(m_n(x, z_n), k_n(x))$:

$$p_n(x) = \mathbb{P}[\mathbb{1}_{Z(x)>0} = 1 | x, \mathcal{X}, \mathcal{Y}] = \int_{\mathbb{R}^n} \phi_{\mathcal{Y}}^{Z_n}(z_n) \bar{\Phi}\left(\frac{-m_n(x, z_n)}{\sqrt{k_n(x)}}\right) \mathrm{d}z_n$$



$$\underset{x_{m+1}}{argmax} \left[\prod_{i=1}^{N} \Phi\left(\tau_t\left(\mu_{t,m}^i(x_{m+1}) + k\,\sigma_{t,m}(x_{m+1})\right)\right) \min_{i=1,\dots,m} \Phi\left(-\tau_t g_{t,m}(x_i)\right) \|x_i - x_{m+1}\|_s^{2p} \right]$$

One-step-lookahead

$$\underset{x_{m+1}}{argmax}\, \mathbb{E}_{Y_{m+1}} \left[\prod_{i=1}^{N} \Phi\left(\tau_t \mu_{t,m+1}^i(x_{m+1})\right) \right] \min_{i=1,\dots,m} \Phi\left(-\tau_t g_{t,m}(x_i)\right) \|x_i - x_{m+1}\|_s^{2p}$$

3. ONE-STEP-LOOKAHEAD MIXED

$$\underset{x_{m+1}}{argmax}\, \mathbb{E}_{Y_{m+1}}\left[\Phi\left(-\tau_t g_{t,m+1}(x_{m+1})\right)\right] \min_{i=1,\dots,m} \Phi\left(-\tau_t g_{t,m}(x_i)\right) \|x_i - x_{m+1}\|_s^{2p}$$

$$\underset{x_{m+1}}{argmax}\, \mathbb{E}_{Y_{m+1}}\left[\left(1 - \mathrm{Var}_{\mathrm{t,m+1}}^{\Gamma}\right) \times \Phi\left(-\tau_t g_{t,m+1}(x_{m+1})\right)\right] \min_{i=1,\dots,m} \Phi\left(-\tau_t g_{t,m}(x_i)\right) \|x_i - x_{m+1}\|_s^{2p}$$

Good order of magnitude ?

PUBLICATIONS

- *Estimation of simulation failure set with active learning based on Gaussian process classifiers and random set theory.*
  Submitted to Structural Safety, Oct. 2024.
  Menz, M., Munoz Zuniga, M., Sinoquet, D.

- *Learning and space-filling the hidden design space associated to failure of computer experiments*
  To be submitted beginning 2024*.*
  Menz, M., Munoz Zuniga, M., Sinoquet, D.

REFERENCES I

- Charles Audet and John E Dennis Jr. Mesh adaptive direct search algorithms for constrained optimization. SIAM Journal on optimization, 17(1):188–217, 2006.
- Charles Audet, Sébastien Le Digabel, Viviane Rochon Montplaisir, and Christophe Tribes. Nomad version 4: Nonlinear optimization with the mads algorithm.
- Yves Auffray, Pierre Barbillon, and Jean-Michel Marin. Maximin design on non hypercube domains and kernel interpolation. Statistics and Computing, 22:703–712, 2012.
- François Bachoc, Céline Helbert, and Victor Picheny. Gaussian process optimization with failures: classification and convergence proof. Journal of Global Optimization, 78(3):483–506, November 2020.
- Julien Bect, David Ginsbourger, Ling Li, Victor Picheny, and Emmanuel Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. Statistics and Computing, 22(3):773–793, May 2012.
- Clément Chevalier. Fast uncertainty reduction strategies relying on Gaussian process models. PhD Thesis, 2013.
- Bin Dai, Shilin Ding, and Grace Wahba. Multivariate Bernoulli distribution. Bernoulli, 19(4):1465 – 1483, 2013.
- Sébastien Le Digabel and Stefan M Wild. A taxonomy of constraints in simulation-based optimization.
- Mohamed Reda El Amri, Céline Helbert, Olivier Lepreux, Miguel Munoz Zuniga, Clémentine Prieur, and Delphine Sinoquet. Data-driven stochastic inversion via functional quantization. Statistics and Computing, 30(3):525–541, 2020.
- Hoël Langouët. Optimisation sans dérivées sous contraintes : deux applications industrielles en ingénierie de réservoir et en calibration des moteurs. Theses, Université Nice Sophia Antipolis, June 2011.
- Morgane Menz, Miguel Munoz-Zuniga, and Delphine Sinoquet. Learning hidden constraints using a Stepwise Uncertainty Reduction strategy based on Gaussian Process Classifiers. working paper or preprint, November 2022.
- Ilya Molchanov. Theory of random sets, volume 19. Springer, 2005.
- Michael JD Powell. The newuoa software for unconstrained optimization without derivatives. Large-scale nonlinear optimization, pages 255–297, 2006.
- Oleg Yu. Vorobyev and Natalia A. Lukyanova. A Mean Probability Event for a Set of Events. Journal of Siberian Federal University. Mathematics and Physics., page 128–136, 2013.
- Guang Zhao, Edward R Dougherty, Byung-Jun Yoon, Francis J Alexander, and Xiaoning Qian. Efficient Active Learning for Gaussian Process Classification by Error Reduction. page 13, 2021.
- Huang. Chaofan. V. Roshan Joseph. et Douglas M. Ray. « Constrained Minimum Energy Designs ». Statistics and Computing 31. no 6 (novembre 2021): 80.
- Sacks, J., Schiller, S. B., and Welch, W. J. (1989). Designs for computer experiments. Technometrics, pages 41–47

REFERENCES II

- Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. Journal of Artificial Intelligence Research, 4, 129–145.
- D. Geman and B. Jedynak. An active testing model for tracking roads in satellite images. IEEE Trans. Pattern Analysis and Machine Intelligence, 18(1):1–14, 1996. 30
- F. Fleuret and D. Geman. Graded learning for object detection. In Proceedings of the workshop on Statistical and Computational Theories of Vision of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR/SCTV), 1999. 2
- Bect, J., Ginsbourger, D., Li, L., Picheny, V., and Vazquez, E. (2012). Sequential design of computer experiments for the estimation of a probability of failure. Statistics and Computing, 22 (3):773–793.
- Chevalier, C. (2013). Fast Uncertainty Reduction Strategies Relying on Gaussian Process Models. PhD thesis, University of Bern.
- Villemonteix, J., Vazquez, E., and Walter, E. (2009). An informational approach to the global optimization of expensive-to-evaluate functions. Journal of Global Optimization, 44(4):509–534.
- R. Bettinger. Inversion dun systeme par krigeage. Application `a la synthese de catalyseurs `a haut d´ebit. PhD thesis, PhD thesis, universit´e de Nice-Sophia Antipolis, 2009. 30
- D. Ginsbourger and R. Le Riche. Towards gaussian process-based optimization with finite time horizon. In mODa 9–Advances in Model-Oriented Design and Analysis, pages 89–96. Springer, 2010. 30, 32, 33
- Srinivas, N., A., K., Kakade, S., and Seeger, M. (2012). Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. IEEE Transactions on Information Theory, 58:3250–3265.
  provide rates of convergence for the sequential strategy GP-UCB (optimization)
- Bull, A. D. (2011). Convergence rates of efficient global optimization algorithms. Journal of Machine Learning Research, 12:2879–2904.
  provide rates of convergence for expected improvement. Here the function f to optimize is deterministic and belongs to the RKHS of k However in general $P(\xi \in RKHS(k)) = 0 \Longrightarrow$ problematic from a Bayesian point of view
- Vazquez, E. and Bect, J. (2010a). Convergence properties of the expected improvement algorithm with fixed mean and covariance functions. Journal of Statistical Planning and Inference, 140(11):3088–3095.
  prove the consistency of Expected Global Improvement. They work with covariance functions which are not too smooth and not degenerate (we will improve this point here)
- Julien Bect, François Bachoc, David Ginsbourger. "A supermartingale approach to Gaussian process based sequential design of experiments." Bernoulli, 25(4A) 2883-2919 November 2019.

Innover les énergies

Retrouvez-nous sur :

www.ifpenergiesnouvelles.fr

@IFPENinnovation

IFP Energies nouvelles