Quantifying the uncertainties of a multidimensional and multifidelity nonlinear simulation, application to the calculation of the deformation of fuel assemblies in a pressurized water reactor

Ali Abboud[†,1,2], Josselin Garnier[§,1], Bertrand Leturcq[§,2], Olivier Fandeur[§,2]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2022 – Sep. 2025

[1] CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France
{ali.abboud, josselin.garnier}@polytechnique.edu

[2] Université Paris-Saclay, CEA, SEMT, 91191 Gif-sur-Yvette, France
{ali.abboud, olivier.fandeur, bertrand.leturcq}@cea.fr

Abstract

In the core of nuclear reactors, fluid-structure interaction [1] and intense irradiation lead to progressive deformation of fuel assemblies [2]. When this deformation is significant, it can lead to additional costs and longer fuel unloading and reloading operations. Therefore, it is preferable to adopt a fuel management that avoids excessive deformation and interactions between fuel assemblies. However, the prediction of deformation and interactions between fuel assemblies is uncertain [2]. Uncertainties affect neutronics, thermohydraulics and thermomechanics parameters. Indeed, the initial uncertainties are propagated over several successive power cycles of 12 months each through the coupling of non-linear, nested and multidimensional thermal-hydraulic and thermomechanical simulations.

The objective of the thesis is to propose a methodology to manage uncertainties in the calculation of deformations of assemblies in a nuclear reactor core. The methodology implemented should deal with the multidimensional and nonlinear character of the physical phenomena to be modeled/simulated, and take advantage of the different levels of simulation available to improve the computational cost/uncertainty trade-off.

Short biography (PhD student)

During my studies in master of applied mathematics at the University of Reims Champagne Ardenne, I decided to specialize in the field of partial differential equations and the finite element method. For this reason, I chose to do an internship at the CEA in the reactor study and applied mathematics department, for a subject entitled "development and comparison of finite element bases for a transport equation on a hexagonal mesh". I had the opportunity to apply and develop my theoretical as well as my computer skills and to gain a culture in neutronics. This internship led me to pursue a thesis in applied mathematics in the field of statistics and probability theory applied to the coupled simulation of thermohydraulics and thermomecanics entitled "Quantifying the uncertainties of a multidimensional and multifidelity nonlinear simulation, application to the calculation of the deformation of fuel assemblies in a pressurized water reactor". This Phd takes place at the CEA in collaboration with the CMAP of the Ecole Polytechnique and Framatom.

References

[1] Stanislas de Lambert, Jérome Cardolaccia, Vincent Faucher, Olivier Thomine, Bertrand Leturcq, and Guillaume Campioni. Semi-analytical modeling of the flow redistribution upstream from the mixing grids in a context of nuclear fuel assembly bow. *Nuclear Engineering and Design*, 371:110940, January 2021.

[2] Andreas Wanninger, Marcus Seidl, and Rafael Macián-Juan. Mechanical analysis of the bow deformation of a row of fuel assemblies in a pwr core. *Nuclear Engineering and Technology*, 50(2):297–305, 2018. Special issue on the water reactor fuel performance meeting 2017 (WRFPM 2017).

Estimation of (small) reliable sets using a sequential Bayesian strategy

Romain Ait Abdelmalek-Lomenech[†,1], Julien Bect[‡,1], Vincent Chabridon[2], Emmanuel Vazquez[‡,1]

[†] PhD student (presenting author). [‡] PhD supervisor

PhD expected duration: Oct. 2021 – Sep. 2024

[1] Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 91190 Gif-sur-Yvette, France
{romain.ait, julien.bect, emmanuel.vazquez}@centralesupelec.fr

[2] EDF R&D, 6 Quai Watier, 78401 Chatou, France
vincent.chabridon@edf.fr

Abstract

When dealing with a numerical model of a physical phenomenon or a system, one is often interested in estimating the set of input parameters leading to a given output range. In the case of an *expensive-to-evaluate* function, the objective is to reach a good approximation of the set of interest using a small number of evaluations. The choice of the evaluation points must then be conducted carefully.

In this communication, we focus on a particular robust formulation of the set inversion problem that we call "Reliable Set Inversion" (RSI) [1], in which the function of interest has both deterministic and uncertain inputs. In this formulation, the objective is to estimate the set of deterministic inputs such that the probability (w.r.t. the distribution of the uncertain inputs) that the output variables belong to a critical region is below a threshold.

More precisely, given a vector-valued function $f : \mathbb{X} \times \mathbb{S} \mapsto \mathbb{R}^q$, a critical region $C \subset \mathbb{R}^q$, and a threshold $\alpha \in (0, 1)$, the object of interest in the RSI problem is the set

$$\Gamma(f) = \{x \in \mathbb{X} \ : \ \mathbb{P}(f(x, S) \in C) \leq \alpha\} \,,$$

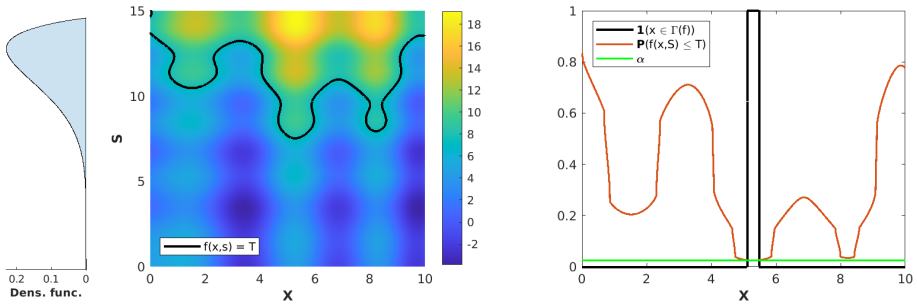

Figure 1: Scalar-valued example function (middle), probability density function associated to the distribution $\mathbb{P}_S$ of the uncertain input variable (left), and set of interest obtained with critical region $C = (-\infty, T]$ and threshold $\alpha = 0.025$ (right).

where $\mathbb{X}$ and $\mathbb{S}$ are respectively the spaces of design and uncertain input variables, and S is a random vector with known probability distribution $\mathbb{P}_S$. A simple scalar-valued example of this framework is shown in Figure 1.

To address this problem, while taking into consideration the expensive-to-evaluate nature of the given function, we propose a sequential Bayesian strategy based on the *Stepwise Uncertainty Reduction* (SUR) principle; see, e.g., [4] and references therein. SUR-based strategies are Bayesian methods for selecting the evaluation points by minimizing the expected future uncertainty about the quantity or object of interest, measured by a particular metric. Such strategies have been used with success in a diverse range of problems, such as function fitting, optimization [5], and set inversion problems [2, 3].

In this work, the object of interest is the set $\Gamma(f)$. We improve the method proposed in [1] to tackle more difficult problems where the set is "small" relatively to the full input domain $\mathbb{X}$. We illustrate the performance of our method by applying it to several test functions.

Short biography (PhD student)

With a backgroung in probability theory, statistics, machine learning and quantitative finance, Romain Ait Abdelmalek-Lomenech began his PhD in October 2021 with CentraleSupélec, under the supervision of Emmanuel Vazquez and Julien Bect. His work aims at developing new Bayesian methods for the optimization/inversion of expensive-to-evaluate functions in presence of uncertain input variables.

The present work is part of a PhD thesis funded by the French National Research Agency (ANR) in the context of the SAMOURAI project (ANR-20-CE46-0013).

References

[1] R. Ait Abdelmalek-Lomenech, J. Bect, V. Chabridon, and E. Vazquez. Bayesian sequential design of computer experiments to estimate reliable sets. *arXiv preprint arXiv:2211.01008*, 2022.

[2] D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, and Y. Richet. Adaptive design of experiments for conservative estimation of excursion sets. *Technometrics*, 63(1):13–26, 2019.

[3] J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3): 773–793, 2012.

[4] J. Bect, F. Bachoc, and D. Ginsbourger. A supermartingale approach to Gaussian process based sequential design of experiments. *Bernoulli*, 25(4A):2883–2919, 2019.

[5] J. Villemonteix, E. Vazquez, and E. Walter. An informational approach to the global optimization of expensive-to-evaluate functions. *J. of Global Optimization*, 44(4):509–534, 2009.

Bayesian Optimization in High-dimension via a Combination of Kriging sub-models

Tanguy Appriou[1, 2, †], Didier Rullière[1, §], David Gaudrie[2, §]

[†] PhD student (presenting author) § PhD supervisor

PhD expected duration: Nov. 2021 – Nov. 2024

[1] Mines Saint-Etienne, Univ Clermont Auvergne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol, F - 42023 Saint-Etienne France.
tanguy.appriou@emse.fr
drulliere@emse.fr

[2] Stellantis, Centre Technique Velizy, F-78943 Vélizy-Villacoublay France.
david.gaudire@stellantis.com

Abstract

Numerical optimization has been widely used to solve design engineering problems in recent years. However, due to the large number of numerical simulations required, the optimization process involved is often computationally expensive. In order to accelerate the optimization process, strategies based on surrogate models are often used. One such approach, efficient global optimization (EGO) [1], based on Kriging surrogate models [2], has been successfully applied to a number of real-worlds applications in low-dimensions (less than 20).

However, in engineering design optimization, the designs are often parametrized by more than 50 shape parameters. The ordinary Kriging method scales poorly for high-dimensional problems, and building a good Kriging model is met with various setbacks. One of the main challenges is related to length-scale hyperparameter optimization. Most Kriging models consider one length-scale hyperparameter per dimension which all need to be optimized simultaneously. This is typically done by maximizing the log-likelihood of the model. For high-dimensional problems, this optimization is problematic due to the exponential growth of the search space with the dimension, to the shape of the log-likelihood function and to its computational cost, and to over-fitting issues when there are too few observations. Several papers address these difficulties by reducing the dimension of the problem, for instance by embedding the design space into a lower dimensional space [3,4], or by considering simplifying hypothesis such as additive models [5].

In this paper, we propose a new method for high-dimensional Kriging models which bypasses the length-scales optimization by combining Kriging sub-models with fixed length-scales. Contrarily to other approaches, this does not rely on dimension reduction and preserves the correlation between all input variables. Furthermore, the combination has a closed-form expression and does not require any inner optimization. We also describe a novel approach to obtain suitable bounds for the length-scales of the sub-models, and we compare different weighting schemes for the approximation of high-dimensional test functions. Finally, we present a way to obtain the variance of the combination for any weighting methods and we apply our combined model to high-dimensional EGO. We show that the classical Kriging approach using maximum likelihood estimation fails to properly optimize the length-scale hyperparameters and that our combination of sub-models with fixed length-scales successfully build more accurate surrogate models for EGO.

Short biography (PhD student)

Tanguy Appriou received an engineering degree for Ecole Centrale de Lyon and a master in aerospace engineering from Tohoku University in 2021. He is currently pursuing a Ph.D. degree in Stellantis and Mines Saint-Etienne. This thesis was partly funded by a CIFRE grant (convention #2021/1284) established between the ANRT and Stellantis.

This thesis was conducted with the support of the consortium in Applied Mathematics CIROQUO, gathering partners in technological and academia in the development of advanced methods for Computer Experiments.

References

[1] Jones, D.R., Schonlau, M. and Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, *13*(4), pp.455-492.

[2] Rasmussen, C.E. and Williams, C.K., 2006. *Gaussian processes for machine learning*. Cambridge, MA: MIT press.

[3] Constantine, P.G., 2015. *Active subspaces: Emerging ideas for dimension reduction in parameter studies*. Society for Industrial and Applied Mathematics.

[4] Bouhlel, M.A., Bartoli, N., Otsmane, A. and Morlier, J., 2016. Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction. *Structural and Multidisciplinary Optimization*, *53*(5), pp.935-952.

[5] Durrande, N., Ginsbourger, D. and Roustant, O., 2012. Additive covariance kernels for high-dimensional Gaussian process modeling. In *Annales de la Faculté des sciences de Toulouse: Mathématiques* (Vol. 21, No. 3, pp. 481-499).

Multi-scale composite reliability-based design optimization for aeroelastic applications

R. Ballester[†,1], N. Fabbiane[§,1], C. Fagiano[1], C. Julien[1], L. Coelho[1], D. Lucor[2]

[†] PhD student (presenting author). [§] PhD supervisor.

PhD expected duration: Oct. 2022 – Sep. 2025

[1] ONERA, The French Aerospace Lab, Chatillon, France
roger.ballester_claret@onera.fr

[2] Université Paris-Saclay, CNRS, LISN, Orsay, France
didier.lucor@limsi.fr

Abstract

The inherent incertitudes linked to the composite manufacturing process (e.g., material properties, thickness, ply orientations) call for a reliability-based approach to composite structural optimization [3]. Those composite optimisation problems require a dual design space to efficiently compute the optimum stacking sequence, being one the lower scale space where the heterogeneous material (ply scale) is defined and the second design space being a homogenised material where the design variables are noticeably reduced. Uncertainty is often introduced at lower scales of the composite material, while the macroscopic scale is the preferred design space for optimization.

This work proposes an iterative methodology based on [1], that combines both the low-dimensional macroscopic design space with the high-dimensional lower-scale space. Within this algorithm gradient information is used to perform quick and accurate optimization within the macroscopic space while within the lower scale space the uncertain design variables are modeled and upscaled for each iteration. An inverse problem, governed by genetic algorithms, is solved at each iteration of the optimization process to identify the low-scale material configuration that meets the homogenized parameters in terms of statistical description. The study of incertitude propagation throughout the design scales becomes viable thanks to efficient metamodel upscaling: a particular orthonormal basis is constructed with Fourier chaos expansion, which provides a very efficient closed-form expression of the macroscopic design variable statistics.

This approach is applied to a composite plate optimization with uncertain ply angles such as the ones presented in [4]. Furthermore, due to the multimodal nature of the aeroelastic response (i.e., the flutter velocity), a surrogate model strategy based on Gaussian process classification for the reliability analysis is proposed to further analyse the mechanical aeroelastic behaviour of the problems at study. The design variables within the study have been fixed with the introduction of manufacturing constraints, leaving finally a three-dimensional space where not only the in plane bending properties of the composite are optimised, but also the number of plies in the composite laminate. All of this has been done analyzing all the respective constraints and the effect of those constraints within the reliability of the optimisation algorithms.

Within this study, the main aim is to reduce the wing's weight while remaining reliable with respect to the flutter phenomenon. Moreover, the reduction of the thickness on the plate may lead to configurations where the design speed of the plate is over the limit speed the plate can withstand. Therefore, the introduction of the thickness as a design variable has created the need

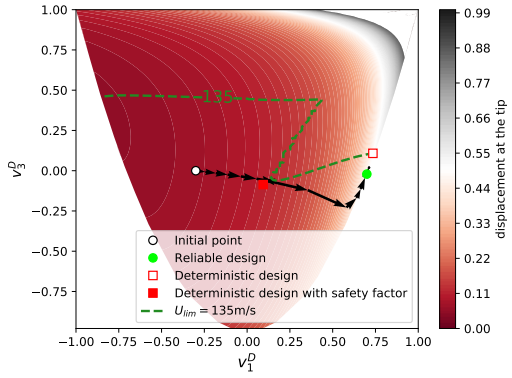

Figure 1: Optimization path of the Reliability-Based Design Optimization at constant thickness. Final macroscopic design variables are compared with deterministic optimization.

for mechanical constraints, such as the Tsai-Wu or Tsai-Hill failure criteria for composites, which have been implemented into the optimisation process following [2], and all the corresponding limit state equations are taken into consideration within the optimisation process. The results show a good convergence of this optimization approach with a significant improvement in the reliability compared to the deterministic optimized design (cf. Fig. 1) and a significant computational gain compared to the approach of directly optimizing ply angles via an evolutionary algorithm as in [4].

Short biography (PhD student)

Engineer with a mechanics MSc from UPC and a aeronatics MSc by ECN, with industry experience having worked in *HP* and *Safran Aircraft Engines*. Since October 2022 working at *ONERA* as a PhD student, carrying out a study in *Robust Optimisation of Aeronautical Structures under Aeroelastic Constraints*, founded entirely by *ONERA*. The aim of the thesis being the definition of a methodology for aeroelastic tailoring to be used in the next generations of civil aircrafts.

References

[1] Ludovic Coelho, Nicolo Fabbiane, Christian Fagiano, Cedric Julien, and Didier Lucor. Reliability-based design optimization of a composite plate through a dual design space. In *VIII Conference on Mechanical Response of Composites*, CT13 - Optimisation, 2021.

[2] François Xavier Irisarri, Cédric Julien, Dimitri Bettebghor, Florian Lavelle, Yannick Guerin, and Kevin Mathis. A general optimization strategy for composite sandwich structures. *Structural and Multidisciplinary Optimization volume*, 63:3027–3044, 2021.

[3] C.L. Pettit. Uncertainty quantification in aeroelasticity: Recent results and research challenges. *Journal of Aircraft*, 41(5):1217–1229, 2004.

[4] C. Scarth and J.E. Cooper. Reliability-based aeroelastic design of composite plate wings using a stability margin. *Structural and Multidisciplinary Optimization*, 57(4):1695–1709, 2018.

Adaptive Importance Sampling in high dimension

J. Beh[†,1,2], F. Simatos[§,1], J. Morio[§,2]

[†] PhD student (presenting author). [§] PhD supervisor.

PhD expected duration: Jan. 2023 – Dec. 2025

[1] ISAE-SUPAERO and Université de Toulouse, Toulouse, France
kaixun5183@gmail.com
florian.simatos@isae-supaero.fr

[2] ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France
jerome.morio@onera.fr

Abstract

Important sampling (IS) is a sampling-based method which allows the study of a probability density function f on $\mathbb{R}^d$ using random independent samples $(Y_1, \dots, Y_n)$ generated from an auxiliary density g, by associating each sample with an importance weight $w_j = f(Y_j)/g(Y_j)$. One of the applications is the estimation of the rare event (failure) probability P of a given system of dimension d, represented by the rare event $\mathcal{F} = \{x \in \mathbb{R}^d; \varphi_d(x) \geq 0\}$, in which the function φ_d characterizes the system. The IS estimator $\widehat{P}$ of the probability P is then given by

$$\widehat{P} = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\mathcal{F}}(Y_j) w_j,$$

where $\mathbb{1}_{\mathcal{F}}$ represents the indicator function of the event $\mathcal{F}$.

The zero-variance density of the IS estimator is given by the density $g_* = \mathbb{1}_{\mathcal{F}} f/P$ but it is unusable in practice since P is unknown. The adaptive importance sampling (AIS) approach instead determines an auxiliary density in a parametric family by estimating the parameter satisfying a theoretical optimality criterion, for instance that which minimizes its Kullback-Leibler (KL) divergence D with the zero-variance density. A representation of such family is the family of Gaussian densities g_θ parameterized by $\theta = (\mu, \Sigma)$, where $\mu \in \mathbb{R}^d$ is the mean and $\Sigma \in \mathbb{R}^{d \times d}$ is the covariance matrix. The Gaussian density which minimizes the KL divergence with g_* has been theoretically proven to be the Gaussian density g_{θ_*} with the same mean and covariance matrix as g_*. In this case, AIS translates into the empirical estimation $\widehat{\theta_*}$ of these optimal mean and covariance matrix, by either using n_{p} number of samples drawn directly from g_*, or using an iterative algorithm such as the classical *Cross-Entropy*.

The precision of the AIS approach deteriorates in high dimension due to two reasons. Firstly, a weight degeneracy problem occurs, which is the concentration of all the mass into one weight, with the other importance weights almost being zero. Secondly, the estimation of a covariance matrix requires the estimation of roughly $d^2/2$ number of terms. The number of samples n_{p} for the covariance matrix estimation should be at least $d^2/2$.

The weight degeneracy phenomenon has been mainly studied theoretically in the context of particle filter [1]. Bengtsson and al. [1] proved that the number of samples n has to be at least e^d to avoid weight degeneracy. In the context of IS, Chatterjee and Diaconis [2] proved that the L_1 error of the IS estimator is bounded if the number of samples n is greater than e^D. Therefore,

one of the aims of this thesis will be to prove similar results for AIS, by potentially establishing the link between weight degeneracy and Kullback-Leibler divergence. Several numerical results seemed to show that the L_1 error can be small even when weight degeneracy occurs, as long as the number of samples n is greater than e^D. Furthermore, e^D decreases with n_{p}, therefore it is crucial to find a balanced budget between n and n_{p} to ensure $n \geq e^D$.

On the other hand, several recent numerical strategies [5] [3] try to reduce the dimension of the problem by projecting the parameters into a subspace of lower dimension. These strategies have yielded significant numerical improvements to the precision of AIS. The aim of the thesis is also to prove theoretical results quantifying these improvements and understand the gains we can expect.

In the case of Gaussian parametric family, one of the projection choices on the covariance matrix is to only estimate its diagonal terms. Preliminary result showed that the number of samples n_{p} required, if generated directly from g_*, is no longer $d^2/2$ but only d. Ongoing work seeks to extend this result on the *Cross-Entropy* algorithm, potentially corroborating the numerical improvements that have been observed in [4].

Short biography (PhD student)

I hold an M. Sc. in engineering from ISAE-Supaero in Toulouse. Then, I completed my final-year internship at ISAE-Supaero which led to my current PhD thesis co-funded by EUR MINT and ONERA under the supervision of F. Simatos (ISAE-Supaero) and J. Morio (ONERA). The main goal of the PhD thesis is to improve the performances of adaptive importance sampling in high dimension.

References

[1] T. Bengtsson, P. Bickel, and B. Li. "Curse-of-Dimensionality Revisited: Collapse of the Particle Filter in Very Large Scale Systems". In: *Institute of Mathematical Statistics Collections*. 2008, pp. 316–334.

[2] S. Chatterjee and P. Diaconis. "The Sample Size Required in Importance Sampling". In: *The Annals of Applied Probability* 28.2 (Apr. 1, 2018). ISSN: 1050-5164.

[3] M. El Masri, J. Morio, and F. Simatos. "Improvement of the Cross-Entropy Method in High Dimension for Failure Probability Estimation through a One-Dimensional Projection without Gradient Estimation". In: *Reliability Engineering & System Safety* 216 (Dec. 2021), p. 107991. ISSN: 09518320.

[4] M. El Masri, J. Morio, and F. Simatos. "Optimal Projection to Improve Parametric Importance Sampling in High Dimension". In: preprint arXiv:2107.06091 (2022).

[5] F. Uribe et al. "Cross-Entropy-Based Importance Sampling with Failure-Informed Dimension Reduction for Rare Event Simulation". In: arXiv:2006.05496 (June 9, 2020).

Deterministic optimisation in Deep Learning from a Lyapunov point of view

B. Bensaid[†,1], G.Poette (CEA-CESTA)[†,2], R.Turpault (Bordeaux-INP)[§,3],

1[†] PhD student (presenting author). 2, 3[§] PhD supervisors

PhD expected duration: Oct. 2021 – Oct. 2024

[1] CEA/CESTA and Université de Bordeaux
bilel.bensaid@u-bordeaux.fr

[2] CEA/CESTA
gael.poette@cea.fr

[3] Bordeaux-INP
rodolphe.turpault@u-bordeaux.fr

Abstract

Deep learning models have recently been used to substitute parts of simulation codes with neural networks [4]. In this context, huge networks do not seem appropriate from a computational point of view. With this in mind, this work aims at building efficient shallow networks. The guiding principle of this research starts from this assessment: some recent papers [5] point the optimizer out as a lever to improve performances. There are stochastic and deterministic optimizers. Before analysing stochastic algorithms we suggest studying first the deterministic ones (as stochastic optimizers are often based on deterministic ones).

In a previous work (accepted after revisions [1]), we have analysed some optimizers of the literature such as Gradient Descent (GD), Momentum [6] and Adam [3]. This study pointed out the lack of convergence properties and Lyapunov stability of the classical optimizers as illustrated by figure 1. On these maps for each minimum we assign a colour to all the initial points that converge to this minimum and the brown part refers to divergent trajectories. Starting very close to a given minimum, GD can diverge whereas a variant of Adam called AWB converges to a point far away from the initial guess. This is a very concerning issue because the optimizer can deteriorate the initialization and the a priori knowledge integrated in the neural networks. To understand how to fix this problem it has been useful to introduce some ODE approximations in the sense that the iterates converge to the solution of the ODE when the learning rate (time step) tends to 0. In this framework, it is possible to build some functions that decrease along the trajectories of the ODE. By choosing an adaptive time step in order to decrease the discretisation of these Lyapunov functions, it comes up a new way to design algorithms that have better stability properties and performances without tuning hyperparameters.

In this work, we go a step further by preserving the continuous dissipation rate (given by the ODEs) of Lyapunov functionals in a weak form. Imposing this supplementary inequality in addition to a convergent discretisation of the ODE gives **much faster algorithms with better predictive performances** compared to the state of the art, on many classical machine learning tasks such as MNIST, FASHION-MNIST,... Combining ideas from Lyapunov control theory and dissipation ODEs [2], this new scheme enables to **prove stability and convergent properties and to compute convergence rates in the non-convex** setting.

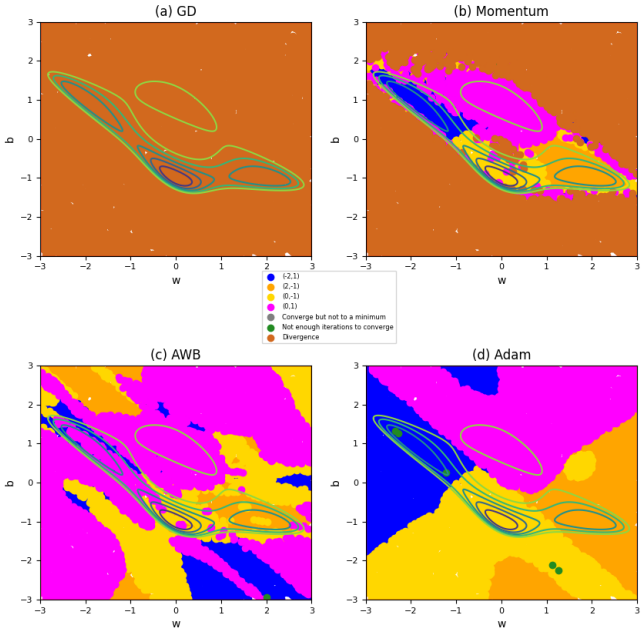

Figure 1: Sensitivity analysis of the classical optimizers to the initial point

Short biography (PhD student)

Engineer student from Mines Saint-Etienne specialized in Data Science I am also graduated of a master degree in Applied Mathematics (MAEA from Université Lyon Bernard). My thesis at Institut de Mathématiques de Bordeaux is financed by CEA-CESTA within the scope of LRC-Anabase.

References

[1] R.Turpault B.Bensaid, G.Poette. Deterministic optimisation in deep learning from a continuous and energy point of view. *Journal of Scientific Computing*, 2023.

[2] Alain Haraux and Mohamed Jendoubi. *The Convergence Problem for Dissipative Autonomous Systems: Classical Methods and Recent Advances*. 02 2015.

[3] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations*, 12 2014.

[4] G. Kluth, K. D. Humbird, B. K. Spears, J. L. Peterson, H. A. Scott, M. V. Patel, J. Koning, M. Marinak, L. Divol, and C. V. Young. Deep learning for nlte spectral opacities. *Physics of Plasmas*, 27(5):052707, 2020.

[5] Paul Novello, Gaël Poëtte, David Lugato, and Pietro M Congedo. Explainable Hyperparameters Optimization using Hilbert-Schmidt Independence Criterion. working paper or preprint, June 2021.

[6] B.T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964.

A moment approach to parameter-dependent hyperbolic conservation laws

C. Cardoen[†,1], A. Nouy[§,1], S. Marx[§,2], N. Seguin[§,3]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Sep. 2022 – Aug. 2025

[1] Nantes Université, École Centrale de Nantes, Laboratoire de Mathématiques Jean Leray, CNRS UMR 6629, France
{clement.cardoen,anthony.nouy}@ec-nantes.fr

[2] Nantes Université, École Centrale de Nantes, LS2N, CNRS UMR 6004, France
swann.marx@ls2n.fr

[3] Inria, Centre de l'Université Côte d'Azur, antenne de Montpellier, Imag, UMR CNRS 5149, Université de Montpellier, France
nicolas.seguin@inria.fr

Abstract

Non-linear hyperbolic conservation laws model numerous physical phenomena in fluid mechanics, traffic flow or non-linear acoustics [1, 9]. The numerical computation of such equations often proves to be a challenge - in particular when the solution has discontinuities. It is even more challenging to solve parameter-dependent hyperbolic conservation laws, for which classical dimension reduction or surrogate modelling approaches fail.

Here, we propose a new surrogate modelling method that is an extension of [6] to parameter-dependent or random conservation laws. Whereas it is classical to seek weak solutions to hyperbolic conservation laws, we are interested in so-called measure-valued (mv) solutions introduced by DiPerna in [2], that are Young measures $\nu_{(t,x,\xi)}$ indexed by time, space and parameters respectively. In practice, a mv solution is a Dirac measure $\delta_{u(t,x,\xi)}$ supported on the graph of the weak solution $u(t,x,\xi)$.

The framework we consider is similar to the one presented in [8, 7], but our starting point is a weak-parametric formulation of the problem. Under the assumption that data are polynomial, this formulation provides constraints on the moments of the measure. This allows us to consider the problem as a generalized moment problem (GMP), an infinite-dimensional optimization problem over sequences of moments of measures, where both the cost and the constraints are linear in the moments of the measures. Powerful results from real algebraic geometry allow to reformulate the constraint that a sequence is a moment sequence into tractable semi-definite constraints. This problem is then solved using Lasserre's (moment sum-of-squares) hierarchy [4], which consists in solving a sequence of convex semi-definite programs of increasing size to approximate the moments of mv solutions.

Once an approximation of moments is obtained, several post-treatments are possible. First, we can obtain statistics of variables of interest that are functions of the moments of the solution. Also, the graph of the solution $u(t,x,\xi)$ can be recovered using a localization property of the Christoffel-Darboux kernel of the approximate measure $\nu_{(t,x,\xi)}$, following the methodology proposed in [5]. Following [3], one can also have access to more general quantities of interest.

Short biography (PhD student)

I have studied applied mathematics in Centrale Nantes. I have been an intern with Swann Marx, Anthony Nouy and Nicolas Seguin for 6 months, working on these problems. My PhD thesis is about model reduction of conservation laws, using moment approaches, high-dimensional approximation and optimal transport methods. This thesis is funded by NExT (Nantes Excellence Trajectory), and is a part of the ANR project AiBy4.

References

[1] C. M. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics; 3rd ed.* Grundlehren der mathematischen Wissenschaften: a series of comprehensive studies in mathematics. Springer, Dordrecht, 2010.

[2] R. J. Diperna. Measure-valued solutions to conservation laws. *Archive for Rational Mechanics and Analysis*, 88:223–270, 1985.

[3] D. Henrion and J.B. Lasserre. Graph Recovery from Incomplete Moment Information. *Constructive Approximation*, 56:165–187, 2022.

[4] Jean Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1. World Scientific, 2009.

[5] S. Marx, E. Pauwels, T. Weisser, D. Henrion, and J.B. Lasserre. Semi-algebraic approximation using Christoffel-Darboux kernel. *Constructive Approximation*, April 2021.

[6] S. Marx, T. Weisser, D. Henrion, and J.B. Lasserre. A moment approach for entropy solutions to nonlinear hyperbolic PDEs. *Mathematical Control and Related Fields*, 10(1):113–140, January 2020.

[7] S. Mishra, N.H. Risebro, C. Schwab, and S. Tokareva. Numerical solution of scalar conservation laws with random flux functions. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):552–591, 2016.

[8] S. Mishra and C. Schwab. Sparse tensor multi-level monte carlo finite volume methods for hyperbolic conservation laws with random initial data. *Mathematics of Computation*, 81(280):1979–2018, 2012.

[9] G. B. Whitham. *Linear and nonlinear waves*, volume 42. John Wiley & Sons, 2011.

Gaussian process regression for high dimensional graph inputs

Raphaël Carpintero Perez[†,1,2], Sébastien Da Veiga[§,3], Josselin Garnier[§,2], Brian Staber[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Nov. 2022 – Nov. 2025

[1] Safran Tech, Digital Sciences & Technologies Department
{raphael.carpintero-perez, brian.staber}@safrangroup.com

[2] CMAP, Ecole polytechnique, Institut Polytechnique de Paris
{raphael.carpintero-perez, josselin.garnier}@polytechnique.edu

[3] ENSAI, CREST
sebastien.da-veiga@ensai.fr

Abstract

In the design phases of new product parts, numerical simulation is now omnipresent to predict the behavior and performance of such parts, thus allowing to find the right characteristics (geometry, material types, etc) to achieve the desired specifications.
Identifying these characteristics involves solving an optimization problem, which implies to call upon the numerical simulation code many times.

To reduce the computational cost of the optimization in the design phase, the most used approach is based on the construction of statistical regression models commonly called meta-models or response surfaces. Once trained, the meta-model is used to predict quantities of interest for new combinations of the unsimulated inputs (for the purpose of optimization, uncertainty propagation, etc.). In solid and fluid mechanics, the dimension of the input variables is very high because they are often derived from the parameterization of a 3D shape. One approach to manage such variables is to consider directly the 3D shape in the form of a graph .

The goal of this thesis is to improve the predictivity of Gaussian process meta-models on high dimensional graph spaces (in the sense of the number of vertices). Despite all their advantages, Gaussian processes have practical limitations related to the curse of dimensionality.
The definition of kernels between graphs is not new, and there are a number of comprehensive reviews on the subject [1, 3, 2]. Unfortunately, state of the art kernels have an extremely penalizing complexity with respect to the number of nodes and are mainly applied to graphs with discretely labeled vertices/edges. There are also many papers on kernels working on a space of signals on graphs sharing their nodes and edge sets (sometimes called the domain of the graph) but here the problem is wider as several inputs can have different edges and even a different number of nodes.

The first objective of the thesis will therefore be to propose new approaches to build kernels of low complexity with respect to the number of vertices and able to consider vertices with continuous attributes (such as 3D coordinates).

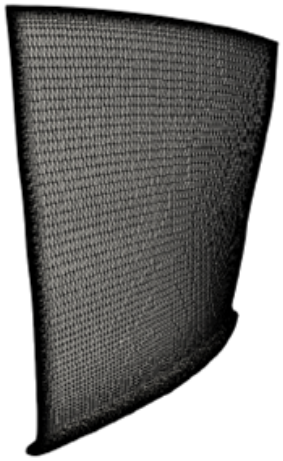

Figure 1: Example of a turbine blade with 30 000 vertices

Short biography (PhD student)

I have a double Master's degree in Mathematics and Computer Science for Data Science (MIDS) from the University Paris Cité (formerly Diderot).
This thesis is a joint work between the Ecole polytechnique and Safran Tech.
It is part of the ANR SAMOURAI (ANR-20-CE46-0013).

References

[1] Karsten Borgwardt, Elisabetta Ghisu, Felipe Llinares-López, Leslie O'Bray, Bastian Rieck, et al. Graph kernels: State-of-the-art and future challenges. *Foundations and Trends® in Machine Learning*, 13(5-6):531–712, 2020.

[2] Nils M Kriege, Fredrik D Johansson, and Christopher Morris. A survey on graph kernels. *Applied Network Science*, 5(1):1–42, 2020.

[3] Giannis Nikolentzos, Giannis Siglidis, and Michalis Vazirgiannis. Graph kernels: A survey. *Journal of Artificial Intelligence Research*, 72:943–1027, 2021.

Gradient-based Coupled Parameter-Data Dimension Reduction for Bayesian Inverse Problems

Q. Chen[†,1], E. Arnaud[§,1], R. Baptista[2], O. Zahm[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2021 – Sep. 2024

[1] Laboratoire Jean-Kuntzmann, Université Grenoble Alpes, Inria Grenoble
{qiao.chen,elise.arnaude,olivier.zahm}@inria.fr

[2] California Institute of Technology
rsb@caltech.edu

Abstract

Dimension reduction is essential for applications such as Bayesian inference that require numerous simulations with different parameters and large potentially correlated data sets. In the inverse problems community, much research has been dedicated to dimension reduction for the typically high-dimensional parameter spaces [2, 3]. Recently, there has been increasing interest in reducing the data dimension to derive optimal experimental designs for Bayesian inverse problems [4, 5].

We propose a novel method to reduce both the dimension of the parameter and data space of high-dimensional Bayesian inverse problems in a coupled manner. The work is inspired by [6], where a gradient-based method for dimension reduction in Bayesian inference is introduced that can control the posterior approximation error. We also build on the work in [1], where the authors develop a gradient-based parameter-data dimension reduction method albeit treating both spaces separately. The new method couples the two dimension reduction problems and potentially exhibits better approximations at lower dimensions than [1] since choosing a low-dimensional parameter subspace influences which data subspace is informative. Vice versa, a fixed low-dimensional data subspace determines which parameter subspace is informed.

Our method computes two projection matrices for the parameter and output space of the nonlinear forward operator G. Assuming a Gaussian likelihood, the Kullback-Leibler divergence of the joint parameter-data distribution equals the L^2 error of the approximate forward operator. We show that the optimal projection matrices minimising this L^2 error are given by an optimisation problem involving a conditional expectation of G, which is hard to compute in practice. Nevertheless, assuming the prior satisfies a subspace Poincaré inequality, we derive an upper bound of the L^2 error, which replaces the conditional expectations with gradients of G. We assume access to gradient evaluations in our applications e.g. through the Adjoint method. It is regularly implemented in the models of our target fields including oceanography, atmospheric and climate sciences. To optimise the deduced gradient-based error bound, we propose an efficient and straightforward two-block Gauss-Seidel algorithm that alternates between the computation of two singular value decompositions.

We demonstrate the proposed method on a conditioned diffusion model, where we use the reduced input-output spaces to compute a surrogate model for G. Furthermore, we apply it to compute optimal sensor placements for a large-scale Bayesian inverse problem in ocean modelling based on the shallow-water equation.

Short biography (PhD student)

I graduated with a MSc in mathematics from Heidelberg University and am currently a second-year PhD student in the inria team AIRSEA at the Université Grenoble Alpes. My thesis aims to develop dimension reduction methods for Bayesian inverse problems that are applicable for ocean and atmosphere science. It is funded by the Université Grenoble Alpes.

References

[1] Ricardo Baptista, Youssef Marzouk, and Olivier Zahm. Gradient-based data and parameter dimension reduction for bayesian models: an information theoretic perspective. *arXiv preprint arXiv:2207.08670*, 2022.

[2] Paul G Constantine, Eric Dow, and Qiqi Wang. Active subspace methods in theory and practice: applications to kriging surfaces. *SIAM Journal on Scientific Computing*, 36(4):A1500–A1524, 2014.

[3] Tiangang Cui, James Martin, Youssef M Marzouk, Antti Solonen, and Alessio Spantini. Likelihood-informed dimension reduction for nonlinear inverse problems. *Inverse Problems*, 30(11):114015, 2014.

[4] Loïc Giraldi, Olivier P Le Maître, Ibrahim Hoteit, and Omar M Knio. Optimal projection of observations in a bayesian setting. *Computational Statistics & Data Analysis*, 124:252–276, 2018.

[5] Keyi Wu, Peng Chen, and Omar Ghattas. A fast and scalable computational framework for goal-oriented linear bayesian optimal experimental design: Application to optimal sensor placement. *arXiv preprint arXiv:2102.06627*, 2021.

[6] Olivier Zahm, Tiangang Cui, Kody Law, Alessio Spantini, and Youssef Marzouk. Certified dimension reduction in nonlinear bayesian inverse problems. *Mathematics of Computation*, 91(336):1789–1835, 2022.

Adaptive importance sampling for stochastic processes on graphs

G. Chennetier[†,1,2], J. Garnier[§,2], H. Chraibi[1], A. Dutfoy[1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Apr. 2021 – Mar. 2024

[1] EDF Lab Paris-Saclay, Boulevard Gaspard Monge, 91120 Palaiseau, France
{guillaume.chennetier,hassane.chraibi,anne.dutfoy}@edf.fr

[2] CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau Cedex, France
{guillaume.chennetier,josselin.garnier}@polytechnique.edu

Abstract

Piecewise deterministic Markov processes (PDMPs) are a class of stochastic processes that can be used to model complex dynamical industrial systems [3]. These processes are composed of a continuous part called position and a discrete part called mode. The position follows a deterministic trajectory given by a differential equation that depends on the current mode of the process. At deterministic and/or random times, the process jumps to a new random mode. The different modes form a graph in which two modes are connected if the process can jump from one to the other one.

The counterpart of the modeling capability of PDMPs is their simulation cost, which makes reliability assessment untractable with standard Monte Carlo methods. Indeed, the failure of a complex system is a rare event and estimating the probability of its occurrence using a Monte-Carlo method requires the simulation of a very large number of trajectories of the underlying process. A significant variance reduction can be obtained with an adaptive importance sampling (AIS) method based on a cross-entropy (CE) procedure [1]. It is known that the optimal distribution for importance sampling depends explicitly on the committor function of the PDMP [2]. At any time during the process trajectory, the committor function returns the probability that this trajectory realizes the rare event of interest (in our case, the failure of the system) knowing its state at this time.

In the case of coherent industrial systems, the graph formed by all the modes of the PDMP modeling the system has a tree structure (called fault tree) [4]. We have proposed in [1] a methodology to approximate the committor function of the PDMP based on the concepts of minimal paths and minimal cuts from fault tree analysis.

In this work, we extend the spectrum of applications to PDMPs whose graphs formed by the modes are large and do not have any particular exploitable structure except sparsity (i.e., far from fully connected). This generalization is based on the definition of a relevant family of "distances" between a vertex and a subpart of the graph. We thus propose a new framework for rare event simulation for stochastic processes with values in large graphs.

Short biography (PhD student)

After a master's degree in statistics at Sorbonne University, Guillaume Chennetier did his internship at EDF R&D Saclay in rare event simulation and sensitivity analysis for piecewise deterministic Markov processes in order to access the reliability of hybrid multi-component systems involved in the operation of nuclear and hydraulic power plants. He is now pursuing this work in a PhD thesis with Ecole Polytechnique and still with EDF R&D Saclay.

References

[1] Guillaume Chennetier, Hassane Chraibi, Anne Dutfoy, and Josselin Garnier. Adaptive importance sampling based on fault tree analysis for piecewise deterministic markov process. *arXiv preprint arXiv:2210.16185*, 2022.

[2] Hassane Chraibi, Anne Dutfoy, Thomas Galtier, and Josselin Garnier. On the optimal importance process for piecewise deterministic markov process. *ESAIM: Probability and Statistics*, 23:893–921, 2019.

[3] François Dufour, Benoîte de Saporta, and Huilong Zhang. *Numerical methods for simulation and optimization of piecewise deterministic Markov processes: application to reliability*. John Wiley & Sons, 2015.

[4] Enno Ruijters and Mariëlle Stoelinga. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. *Computer science review*, 15:29–62, 2015.

Identification of the excitation of trains from on-board mesures

M. Chihaoui[†,1], D. Duhamel[§,1], G. Perrin[§,2], C. Funfschilling[§,3]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Feb. 2023 – Feb. 2026
malek.chihaoui@ensta-paris.fr

[1] Laboratoire Navier, Ecole Nationale des Ponts et Chaussées
denis.duhamel@enpc.fr

[2] COSYS, Université Gustave Eiffel
guillaume.perrin@univ-eiffel.fr

[3] SNCF, Direction Technologies, Innovation et Projets Groupe
christine.funfschilling@sncf.fr

Abstract

Nowadays, the challenges of the railway sector are diverse and the railway system as a whole is subject to more and more constraints. With the new arriving competing european companies and the stricter economical and environmental standards, *SNCF* is pushed to find more innovative and competitive ways to operate.

In this context, the understanding and the mastery of the mechanical behavior of the trains on the railway tracks is key [1]. In particular, the knowledge of the axle-rail contact forces and the tracks irregularities is essential to study both the *dynamic response* and the mechanical damage of the train and the tracks. Such information would be an input to many computational models (Rail and Wheel damage, Track settlement, Wheel wear...) and would considerably enhance the optimization of future conception and maintenance. Nevertheless, these are relatively poorly known due to the difficulty in measuring them on commercial trains. The wheel/rail loads are thus measured during on-track test campaign with measuring wheelsets and the track irregularities with special trains which are costly and thus their usage is not as frequent as we want them to be.

The present work focuses on the reconstruction of the rails geometry and the contact forces between the train axles and the tracks with the help of on-board sensors which can be used on commercial trains. For now, we intend to use accelerometers but later, strain gauges on the rails may also be considered. For now, we have numerical simulation softwares which allow us to estimate the dynamic response and the kinematics of the elements of the train from a measure of the tracks irregularities. The inverse problem is however complex to solve [2, 3]. Our task is to find, using the same software, the excitation imposed on the trains that would produce a similar response thus making our problem an optimization problem. For this, we intend to use the Covariance Matrix Adaptation - Evolution Strategy (CMA-ES) algorithm.

The first difficulty we encounter lies in the proper mathematical definition of the identification of this excitation which must be robust to the *uncertainties* to which the train is subject (Wind, Humidity, Load...) but also to the measure, model & numerical errors [4, 5]. Secondly, it must be

said that the track-train system is highly non-linear which constitutes a non-negligible difficulties. Thirdly, for the inverse problem, a functional excitation must be found for the optimization problem. *Dimensionnality reduction* and use of *meta-models* must thus be considered to achieve the optimisation..

References

[1] K. Matsuoka and H. Tanaka, "Drive-by detection estimation method for simple support bridges based on track irregularities measured on a traveling train," *Mechanical Systems and Signal Processing*, p. 182, 2023.

[2] D. Lebel, C. Soize, C. Funfschilling, and G. Perrin, "Statistical inverse identification for nonlinear train dynamics using a surrogate model in a bayesian framework," *Journal of Sound and Vibration*, pp. 158–176, 2019.

[3] D. Lebel, C. Soize, C. Funfschilling, and G. Perrin, "high-speed train suspension health monitoring using computational dynamics and acceleration measurements," *Vehicle Systems Dynamics*, pp. 911–932, 2020.

[4] G. Schueller and H. J. Pradlwarteren, "Uncertain linear systems in dynamics : Retrospective and recent developments by stochastic approaches," *Engineering Structures*, pp. 2507–2517, 2009.

[5] A. Nouy, "Proper generalized decomposition and separated representations for the numerical solution of high dimensional stochastic problems," *Archives of Computational Methods in Engineering*, pp. 4030–4034, 2010.

Short biography (PhD student)

Malek Chihaoui is currently an PhD student at l'École nationale des ponts et chaussées. He is supervised by Prof Denis Duhamel, Prof Guillaume Perrin and Dr Christine Funfschilling. His thesis focuses on the reconstuction of the irregularities of the tracks and the axle-rails contact forces with the help of on-board sensors on the train. The work is supported by the French railway company SNCF.

Efficient estimation of multiple expectations with the same sample by adaptive importance sampling and control variates

Julien Demange-Chryst[†,1,2], François Bachoc[§,2], Jérôme Morio[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2021 – Sep. 2024

[1] ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France
{julien.demange-chryst,jerome.morio}@onera.fr

[2] Institut de Mathématiques de Toulouse, UMR5219 CNRS, 31062 Toulouse, France
francois.bachoc@math.univ-toulouse.fr

Abstract

Many physical systems are schematically described by a relation of the form $Y = \phi(\mathbf{X})$, where the d-dimensional input vector $\mathbf{X}$ is random and where the output Y is determined through the deterministic function ϕ, whose complexity makes it most of the time impossible to study analytically. Moreover, calls to the code are supposed to be expensive and can therefore only be made in limited number. Some classical problems, such as the method of moments [2] or the estimation of some sensitivity indices like the Sobol' indices [8] for example, require the estimation of multiple expectations related to ϕ. Having accurate estimations of each of them with the classical Monte Carlo method can be costly.

More precisely, for a family of $J \geq 2$ functions $(\phi_j)_{j\in[\![1,J]\!]}$ and its corresponding family of input distributions $(f_j)_{j\in[\![1,J]\!]}$, the main goal of the proposed algorithm is to efficiently estimate the family of expectations $\left(I_j = \mathbb{E}_{f_j}\left(\phi_j(\mathbf{X})\right)\right)_{j\in[\![1,J]\!]}$ with the same N-sample. The quality of the estimation of one expectation can be evaluated with the variance for unbiased estimators. When estimating J expectations, a natural criterion to evaluate the quality of the common estimation is the weighted sum of the individual variance of each estimator [6]. To define this criterion, let us consider a family of positive weights $(w_j)_{j\in[\![1,J]\!]} \in \mathbb{R}_+^J$. Then, for any $j \in [\![1, J]\!]$, let us denote $\widehat{I}_j$ an estimator of the expectation I_j such that all the estimators $\widehat{I}_1, \ldots, \widehat{I}_J$ are based on the same N-sample. The criterion to minimize is then:

$$\sum_{j=1}^{J} w_j \mathbb{V}\left(\widehat{I}_j\right). \tag{1}$$

The weights $(w_j)_{j\in[\![1,J]\!]}$ are used to adjust the importance given to each expectation to estimate.

Importance sampling (IS) [1] and control variates (CV) [7] are two well-known and deeply investigated variance-reduction techniques for estimating a single expectation. Both methods can also be paired and satisfy as a result some very interesting properties [6]. Indeed, when the IS sampling distribution is a mixture of some proposal distributions, using them to create CV allows to bound the variance of the resulting estimator by a quantity proportional to the variance of the best proposal distribution to estimate the expectation of interest. However, to

the best of our knowledge, these methods have not been adapted for jointly estimating multiple expectations with the same sample.

In this communication [4], we first show that there exists a family of estimators combining both IS and CV that make the criterion to minimize in Equation (1) equal to 0. However, these optimal estimators cannot be used in practice because they require the knowledge of the values of $(I_j)_{j\in[\![1,J]\!]}$. Motivated by the form of these optimal IS-CV estimators and the previous properties, we therefore propose an adaptive algorithm called ME-aISCV for jointly estimating J expectations with the same N-sample. In the same way as other adaptive IS algorithms [3], the general idea is to adaptively update the parameters of the estimators for approaching the optimal ones until a stopping criterion is reached. The main goal of this adaptive IS-CV procedure is to minimize as much as possible the criterion in Equation (1). We suggest a quantitative stopping criterion that exploits the trade-off between approaching these optimal parameters and having a sufficient budget left. At last, the left budget is used to draw a new independent sample according to the final sampling distribution which allows to get unbiased estimators by IS and CV of the J expectations to estimate. Finally, we illustrate and discuss the practical interest of the proposed algorithm. We first address the estimation of the even moments of the standard Gaussian distribution. Then, we show that the suggested ME-aISCV algorithm is generally applicable to sensitivity analysis, both on the input parameters and also on their uncertainty distribution, and this is applied to the physical cantilever beam problem. The code to reproduce the numerical experiments is publicly available at: `https://github.com/Julien6431/Multiple_expectation_estimation.git`.

Short biography (PhD student)

I graduated in 2021 from the graduate engineering school ISAE-SUPAERO in Toulouse and I also obtained a MSc degree in applied mathematics from Toulouse III - Paul Sabatier University. Then, I performed my final-year internship at *ONERA* in Toulouse which led to my current PhD thesis co-funded by *ONERA* and *Toulouse III - Paul Sabatier University*. The first research topic of my PhD thesis was the estimation of the target Shapley effects for variance-based reliability-oriented sensitivity analysis with dependent inputs, and it leads to a journal publication [5].

References

[1] James Bucklew. *Introduction to rare event simulation*. Springer Science & Business Media, 2004.

[2] George Casella and Roger L Berger. *Statistical inference*. Cengage Learning, 2021.

[3] Jean-Marie Cornuet, Jean-Michel Marin, Antonietta Mira, and Christian P Robert. Adaptive multiple importance sampling. *Scandinavian Journal of Statistics*, 39(4):798–812, 2012.

[4] Julien Demange-Chryst, François Bachoc, and Jérôme Morio. Efficient estimation of multiple expectations with the same sample by adaptive importance sampling and control variates. *arXiv preprint arXiv:2212.00568*, 2022.

[5] Julien Demange-Chryst, François Bachoc, and Jérôme Morio. Shapley effect estimation in reliability-oriented sensitivity analysis with correlated inputs by importance sampling. *Accepted in International Journal for Uncertainty Quantification*, 2022.

[6] Hera Y He and Art B Owen. Optimal mixture weights in multiple importance sampling. *arXiv preprint arXiv:1411.3954*, 2014.

[7] Barry L Nelson. Control variate remedies. *Operations Research*, 38(6):974–992, 1990.

[8] Ilya M Sobol. Sensitivity analysis for non-linear mathematical models. *Mathematical modelling and computational experiment*, 1:407–414, 1993.

Interpretable multi-step ahead time series forecasting

A. Dhaou[†,1,2], J. Garnier[§,1], E. Le-Pennec[§,1], A. Bertoncello[§,2]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Apr. 2021 – Apr. 2024

[1] CMAP, Ecole Polytechnique, Institut Polytechnique de Paris Palaiseau, France
{amin.dhaou,josselin.garnier,erwan.le-pennec}@polytechnique.edu

[2] TotalEnergies, Palaiseau
antoine.bertoncello@totalenergies.com

Abstract

Many real-life phenomena that one seeks to predict are governed by equations involving several variables and depend on time trends. From multivariate time series, one of the current challenges is to find the underlying dynamics, which can, for example, be a physical equation governing the evolutions. Although these dynamics generally have nonlinear behaviors, it is often possible to approximate them locally by linear models. Thus, one of the objectives is to learn the invariant properties and characteristics from the data.

The goal is therefore to learn a simple, interpretable model that is able to learn the underlying dynamics from temporal data. To do so, we evaluate the ability of our model to learn this dynamics by making predictions at several time steps.

Our algorithm builds on DAgger [1] (and Data as Demonstrator [3]),an imitation learning algorithm, by proposing a meta-model to improve multi-step performance using a one-step predictor, the algorithm iteratively increases the data set with the help of the prediction errors during the training phase. Our work is, therefore, an extension and a consolidation of this approach. It allows not only to improve the performances of multi-step predictions but also to learn the dynamics of multivariate time series by adding an interpretability constraint. Moreover, we develop the theoretical guarantees in a more thorough way.

In our work, we propose a new method, a multi-step prediction model being interpretable, coherent and mitigating the error propagation during recursion. In this work, we propose and study several versions of our prediction method in order to make the best use of the available data set and to learn the underlying dynamics. Furthermore, we establish the theoretical guarantees of the method, seen as a Follow-The Leader (FTL) method [2]. Finally, we show on a variety of synthetic datasets generated from the equation of motions of real dynamic systems that this approach can be easily implemented and gives good results compared to the state of-the-art.

Short biography (PhD student)

Amin Dhaou started his CIFRE PhD thesis, funded by TotalEnergies R&D and Ecole Polytechnique in 2021. It focuses on causality for time series. The main objective is to develop a data-driven approach that extracts causal links between time series variables and finds counterfactual results.

References

[1] Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 627–635. JMLR Workshop and Conference Proceedings, 2011.

[2] Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

[3] Arun Venkatraman, Martial Hebert, and J Andrew Bagnell. Improving multi-step prediction of learned time series models. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

Multi-output Gaussian Process Regression for inversion

C. Duhamel [†,1], C. Prieur [§,1], C. Helbert [§,2], M. Munoz Zuniga [§,3], D. Sinoquet [§,3]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2020 – Jun. 2024*

[1] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK
clement.duhamel@inria.fr
clementine.prieur@univ-grenoble-alpes.fr

[2] Univ. Lyon, Ecole Centrale de Lyon, CNRS UMR 5208, Institut Camille Jordan,
celine.helbert@ec-lyon.fr

[3] IFP Energies Nouvelles, Rueil-Malmaison
delphine.sinoquet@ifpen.fr
miguel.munoz-zuniga@ifpen.fr

Abstract

Nowadays, many industrial issues are related to excursion set estimation, for instance to find feasible complex optimal designs. Given an input/output model, we define an excursion set as the set of input values such that a quantity of interest defined from the model output remains, e.g., below a threshold. In general, the output is computed from a black box numerical model, and each model evaluation is computationally expensive. Excursion set estimation is also known as inversion problem [3]. For example, the application to a vehicle pollution control system, allowing compliance with pollutant emission norms can be mentioned [6].

An effective way to solve an inversion problem is to model the costly black-box function of interest as a realization of a Gaussian process. This surrogate model is built thanks to a sequential Design of Experiments (DoE), whose points are chosen accordingly to the optimization of an acquisition criterion (see for example [1] and [8]). Acquisition criteria suitable for inversion include: deviation number [5], Bichon criterion also known as Expected Feasibility Function [2], and Ranjan criterion [9]. In addition, there exists a more elaborate and in general more efficient class of criteria that anticipates the impact of adding new points to the DoE: the so-called SUR strategies, for Stepwise Uncertainty Reduction (see [1]). For example, SUR strategy coupled to the Vorob'ev's theory introduced in [3], is particularly suitable for inversion framework.

The main objective of our latest work was to propose a SUR version of Bichon criterion [4], which is easier to implement and more robust than SUR Vorob'ev, and we highlighted the performances of this new criterion on analytical examples.

In the present work, we tackle the estimation of excursion sets for several outputs. We propose an extension of Bichon criterion to multi-output Gaussian process regression for inversion. The problem is much more complicated with a multi-dimensional output because of the correlation between outputs (see for example [7]). In this kind of problem, the multi-dimensional output $\boldsymbol{G} : \mathbb{X} \subset \mathbb{R}^d \to \mathbb{R}^p$ is represented by a multi-output Gaussian process $\boldsymbol{\xi}(\mathbf{x}) \sim \mathcal{N}(M_n(\mathbf{x}), \Sigma_n(\mathbf{x}))$ with $M_n(\mathbf{x}) \in \mathbb{R}^p$, $\Sigma_n(\mathbf{x}) \in \mathbb{R}^{p\times p}$ positive-definite symmetric, p dimension of $\boldsymbol{\xi}(\mathbf{x})$. By noting

*I benefit from an adaptation of the thesis (part-time 80% + extension of the contract) for health reasons.

$\boldsymbol{T} \in \mathbb{R}^p$ fixed thresholds of on the multi-dimensional output $\boldsymbol{G}$ and $\mathcal{E}_n$ the event $\boldsymbol{\xi}(\chi_n) = \boldsymbol{G}(\chi_n)$ with χ_n the n first points of the DoE, we propose the extension of Bichon criterion

$$\mathbf{x}^{(n+1)} := \underset{\mathbf{x}\in\mathbb{X}}{\operatorname{argmax}} \left\{ \det\big(\Sigma_n(\mathbf{x})\big)^{\frac{1}{p}} \, \mathbb{E}\left[\Big(\kappa^2 - \|\boldsymbol{T} - \boldsymbol{\xi}(\mathbf{x})\|^2_{\Sigma_n(\mathbf{x})} \Big)^+ \,\middle|\, \mathcal{E}_n \right] \right\}. \tag{1}$$

with $\mathbf{x}^{(n+1)}$ the new point to add to the DoE, $(.)^+$ the positive value function, κ^2 fixed and $\|.\|_{\Sigma_n(\mathbf{x})} := (.)^{\mathrm{T}}\Sigma_n^{-1}(.)$ the Mahalanobis distance associated to $\Sigma_n(\mathbf{x})$. The first term with the determinant allows to force the exploration ability. The second term permits the selection of a $\mathbf{x}$ such that $\boldsymbol{\xi}(\mathbf{x})$ is near the threshold $\boldsymbol{T}$ with high probability $1-\alpha$, κ being the quantile of order $1-\alpha$ of a chi-squared distribution.

A SUR version of the criterion (1) is also considered in a similar way to the one-dimensional formulation proposed in [4]. To do this, we generalize the fast kriging update formulas [3], to multi-output quantities of interest. Efficient implementations of the various integrations and optimizations are proposed. Numerical testing will also be provided to illustrate the performances of the proposed criterion.

Short biography (PhD student)

I'm a third year PhD student in AIRSEA team of Laboratoire Jean Kuntzmann at Univ. Grenoble Alpes. My thesis project initiated in October 2020 is about robust inversion for vectorial data and an application to the design of wind turbines. This project is funded by INRIA, and is part of a collaboration with IFPEN. Before this project, I did a master's degree on Applied Mathematics and Statistics at Univ. Claude Bernard Lyon I.

References

[1] 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, 2012.

[2] Barron J Bichon, Michael S Eldred, Laura Painton Swiler, Sandaran Mahadevan, and John M McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal*, 46(10):2459–2468, 2008.

[3] Clément Chevalier. *Fast uncertainty reduction strategies relying on Gaussian process models.* PhD thesis, Universität Bern, 2013.

[4] Clément Duhamel, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, and Delphine Sinoquet. A SUR version of the bichon criterion for excursion set estimation. 2022.

[5] Benjamin Echard, Nicolas Gayton, and Maurice Lemaire. Ak-mcs: an active learning reliability method combining kriging and monte carlo simulation. *Structural Safety*, 33(2):145–154, 2011.

[6] Mohamed Reda El Amri. *Analyse d'incertitudes et de robustesse pour les modèles à entrées et sorties fonctionnelles.* Theses, Université Grenoble Alpes, April 2019.

[7] Haitao Liu, Jianfei Cai, and Yew-Soon Ong. Remarks on multi-output gaussian process regression. *Knowledge-Based Systems*, 144:102–121, 2018.

[8] Maliki Moustapha, Stefano Marelli, and Bruno Sudret. A generalized framework for active learning reliability: survey and benchmark. *arXiv preprint arXiv:2106.01713*, 2021.

[9] Pritam Ranjan, Derek Bingham, and George Michailidis. Sequential experiment design for contour estimation from complex computer codes. *Technometrics*, 50(4):527–541, 2008.

Sensitivity analysis and causal inference: Application to the calibration of innovative sensors for monitoring pollutants in uncontrolled environments

Marine Dumon[†,1], Bérengère Lebental[§,1], Guillaume Perrin[§,1],

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2022 – Oct. 2025

[1] COSYS/IMSE, Université Gustave Eiffel
marine.dumon@univ-eiffel.fr

Abstract

Environmental pollution causes over 6.5 million deaths per year worldwide. To limit exposure to toxic gases and to limit the risk of exposure to water contamination, it is necessary to deploy detection solutions. This thesis focuses on the calibration of new innovative and low-cost nanomaterial-based sensors for air and water pollution monitoring.

Nanomaterial-based sensors [3] have been proposed for their high sensitivity to different chemical species in air and water and their ability to detect them even in very low concentrations. However, if in laboratory these sensors are able to show encouraging results, the passage in real conditions can pose difficulties. Indeed, the results can be degraded because of the relative humidity and temperature that we cannot control in an outdoor environment. Moreover, the high sensitivity of these sensors can also make them sensitive to other chemical species than those expected.

In order to model the calibration law of the sensors, the resolution of an inverse problem is conducted. The key point of this formalism is the choice of inputs and outputs. The difficulty is to choose the really influential parameters. A sensitivity analysis must be conducted to find the parameters, among all the possible chemical species measured, having a causal impact on the sensor outputs. Moreover, these pollutants are highly correlated and interact with each other through a relationship that is not trivially known. Causal inference [1] allows us to mathematically understand the relationships between the pollutants and the dependencies or independencies between the pollutants and the sensor outputs. We can then improve the calibration model by limiting the number of input variables.

A model error is introduced in the calibration law for the prediction of the pollutant concentrations, this error is modeled from the pollutant measurements used for the tests. We also add in the model an uncertainty term on the input and output data measurements. This leads to formulate the calibration and prediction problem under a Bayesian formalism [2] from a priori knowledge of the pollutants. This formalism allows us to estimate the a posteriori density for the concentration of chemical species by taking into account all the uncertainties.

Short biography (PhD student)

With a mathematical background, Marine Dumon is currently a PhD student at Gustave Eiffel University under the direction of Bérengère Lebental and Guillaume Perrin. Her thesis focuses

on innovative calibration methods for air and water quality sensors and is funded by innovative sensor design and deployment projects.

References

[1] Naoufal Acharki. Statistical learning and causal inference for energy production. *PhD thesis, Institut Polytechnique de Paris*, 2022.

[2] Guillaume Perin Bérengère Lebental. Uncertainty-based calibration method for environmental sensors - application to chlorine and ph monitoring with carbon nanotube sensor array. *submitted to IEEE sensors journal.*

[3] Weichun Huang Lanping Hu Yanfeng Tang Zhinan Guo Zhengbiao Ouyang Deepika Tyagi, Huide Wang and Han Zhang. Recent advances in two-dimensional-material-based sensing technology toward health and environmental monitoring applications. *Nanoscale*, 12:3535–3559, 2005.

Ensemble-based data assimilation for meshless simulations

M. Duvillard[†,1,2], L. Giraldi[§,1], O. Le Maître[§,2]

[†] PhD student (presenting author). [§] PhD supervisors

PhD expected duration: Oct. 2021 – Sep. 2024

[1] CEA, DES, IRESNE, DEC, SESC, LMCP, F-13108 Saint-Paul-lez-Durance, France
{marius.duvillard,loic.giraldi}@cea.fr

[2] CNRS, Inria, Centre de Mathématiques Appliquées, Ecole Polytechnique, IPP, Route de Saclay, 91128, Palaiseau Cedex, France
olivier.le-maitre@polytechnique.edu

Abstract

Meshless methods are simulation approaches relying on Lagrangian representations that can accommodate complex geometries with large deformations and changes in the shape of a continuum (fragmentation, free-surface flow,...). These methods discretize the continuous fields and operator using an ensemble of particles that move according to a velocity field. The Smooth Particles Hydrodynamics method (SPH) and the Material Point Method (MPM) are two popular simulation techniques of this kind [3, 7].

This work aims to propose new data assimilation methods adapted for meshless simulations. Data assimilation concerns the update of the model state using sequential observations [2, 5]. Generally, the assimilation problem is formulated either with a variational approach (minimization of a cost function), a bayesian approach (estimation of the model state's posterior distribution), or a hybridization of the two previous approaches. Variational methods define the best possible estimate of the state by a weighted least squares problem (3DVar, 4DVar methods). Bayesian approaches approximate distributions usually through a sequential scheme based on the assumption of linear equations and Gaussian noises (Kalman Filter) or a Monte-Carlo approximation (Particle filter). Hybrid approaches combine the benefit of variational analyses with the flexibility of ensemble methods for posterior estimation.

Ensemble approximations of the Kalman filter are called Ensemble Kalman Filters (EnKF) [4]. EnKF methods propagate an ensemble of states to estimate the covariance of the Gaussian forecast distribution (before assimilation) and compute the associate Kalman gain. EnKF methods are applied to high-dimensional non-linear systems without suffering from the curse of dimensionality (thanks to low-rank approximations of the covariance matrices), in contrast to Particle Filters.

Standard EnKF methods use an identical discretization (computational grid) for all members, which enables simple linear combinations of members to define the corrections. This restriction calls for an adaptation of the EnKF schemes for their application to generic meshless representations, particularly when using particle representations with member-dependent numbers of particles and positions. Besides the difficulty in estimating the forecast mean and covariance from the ensemble, the assimilation scheme must determine each member's updated particle discretization. Further, one would like to incorporate physics constraints during the assimilation.

We propose a hybrid approach to perform state assimilation in meshless simulations, combining variational and ensemble-based techniques. The state update relies on a Maximum A Poste-

riori estimation problem, assuming a gaussian process prior for each member. The mean and covariance functions of the *prior* distribution are approximated via the ensemble of forecasts. The posterior distribution is computed thanks to the Randomized Maximum Likelihood (RML) method [5]. This procedure yields, for each member j of the ensemble, a minimization problem of the form

$$\min_{\hat{f}\in\mathcal{V}} \quad \|y_j - H(\hat{f})\|^2_{R^{-1}} + \frac{1}{N}\sum_{i=1}^{N} \|\hat{f}(X_i^\star) - f_j^p(X_i^\star)\|^2_{k(X_i^\star,X_i^\star)^{-1}}$$

where y_j denotes perturbed observations, R^{-1} is the observation precision matrix, H the observation operator, the $X_i^\star$ are the error control vector points, f_j^p is the j^{th} forecast member, k the covariance function of the prior gaussian process and finally $\mathcal{V}$ the particle discretization space.

In this contribution, we will discuss the selection of the set of error control points $X_i^\star$ and investigate different strategies to reduce the cost of the optimization problem by adapting the size and location of the set of control points. The impact of the selected particle discretization space $\mathcal{V}$ and its adaptation for each member will be analyzed. Finally, we will introduce constraints on the particle discretization to enforce some physical constraints in the variational approximation (e.g., positivity, conservation [6, 1]). The proposed methods will be illustrated on one or two-dimensional transport problems.

Short biography (PhD student)

I'm a PhD student in the CEA center of Cadarache and the Platon team at the Inria Saclay Center (CMAP). I'm currently working on the development of assimilation methods that would be adapted to a grinding mill facilities involved in the fuel manufacturing process.

References

[1] David J Albers, Paul-Adrien Blancquart, Matthew E Levine, Elnaz Esmaeilzadeh Seylabi, and Andrew Stuart. Ensemble Kalman methods with constraints. *Inverse Problems*, 35(9):095007, September 2019.

[2] Mark Asch, Marc Bocquet, and Maëlle Nodet. *Data assimilation: methods, algorithms, and applications*. Number 11 in Fundamentals of algorithms. Society for Industrial and Applied Mathematics, Philadelphia, 2016.

[3] Alban de Vaucorbeil, Vinh Phu Nguyen, Sina Sinaie, and Jian Ying Wu. Material point method after 25 years: Theory, implementation, and applications. In *Advances in Applied Mechanics*, volume 53, pages 185–398. Elsevier, 2020.

[4] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.

[5] Geir Evensen, Femke C. Vossepoel, and Peter Jan van Leeuwen. *Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem*. Springer Textbooks in Earth Sciences, Geography and Environment. Springer International Publishing, 2022.

[6] Tijana Janjić, Dennis McLaughlin, Stephen E. Cohn, and Martin Verlaan. Conservation of Mass and Preservation of Positivity with Ensemble-Type Kalman Filter Algorithms. *Monthly Weather Review*, 142(2):755–773, February 2014.

[7] Chi Zhang, Yu-jie Zhu, Dong Wu, Nikolaus A. Adams, and Xiangyu Hu. Smoothed particle hydrodynamics: Methodology development and recent achievement. *Journal of Hydrodynamics*, 34(5):767–805, October 2022.

Sensitivity Analysis on (excursion) sets

N. Fellmann[†,1,2], C. Blanchet[§,1], C. Helbert[§,1], A. Spagnol[§,2], D. Sinoquet[§,2]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Dec. 2021 – Nov. 2024

[1] MI/ICJ, Ecole Centrale de Lyon
[2] IFP Energies Nouvelles

Abstract

In the black-box framework, Sensitivity Analysis (SA) consists in studying and quantifying the influence of the inputs on the variability of the output. It can be used in applications as robust optimization under constraints defined by:

$$x^* = \arg\min_{x \in \mathcal{X}} \mathbb{E}[f(x,U)] \quad \text{s.t. } \mathbb{P}[g(x,U) \leq 0] \geq P_{target}.$$

In this context, Sensitivity Analysis is Goal Oriented (GOSA) in order to find the inputs that have the most impact on the constraints compliance. However GOSA methods are mostly applied on the deterministic inputs x for screening purpose in order to reduce the input dimension [5]. But in presence of uncertain variables U, it might also be interesting to study their impact on the optimization. For instance, such information could be used to reduce the dimension of the uncertain space which could be useful to reduce the cost of meta-modeling before the optimization. That's why our aim is to conduct GOSA on the uncertain inputs.

To do so we propose to carry out GOSA with set-valued outputs $\Gamma = \phi(U)$ where Γ is a random excursion set i.e. the set where the constraints are respected: $\Gamma = \{x \in \mathcal{X}, g(x,U) \leq 0\}$.

However, most SA methods deal with scalar or vectorial outputs. A set-valued output space makes the adaptation of classical sensitivity analysis methods nontrivial. We propose three main ideas to carry out this adaptation. The first relies on random sets theory [4]. The second adapts universal indices from [3] defined for general metric output space. The last one uses kernel-based sensitivity indices applied on sets.

The latter is the most suited to set-valued output as it only requires a kernel on sets. That's why our first contribution is to propose a kernel k_{set} on sets: $k_{set}(\gamma_1, \gamma_2) := e^{-\frac{\mu(\gamma_1 \Delta \gamma_2)}{2\sigma^2}}$ for any closed sets γ_1, γ_2 where Δ stands for the symmetric difference and μ for the volume.

Having this kernel, we can define kernel-based sensitivity indices with set-valued outputs. The main sensitivity index derived from kernels is based on the Hilbert Schmidt Independence Criterion (HSIC). It measures the dependence between an input and the output by measuring the distance between the distribution of the couple (input,output) and the product of the marginals. With the kernel k_{set}, $HSIC_{k_{set}}(U_i, \Gamma)$ defines a sensitivity index which quantifies the influence of the input U_i on the set-valued output Γ. This index is adapted to screening through independence testing, and it has an ANOVA-like decomposition under some hypotheses which also makes it suited for ranking [2].

This provides a method to conduct sensitivity analysis on any set-valued output, and in particular on excursion sets in the context of optimization.

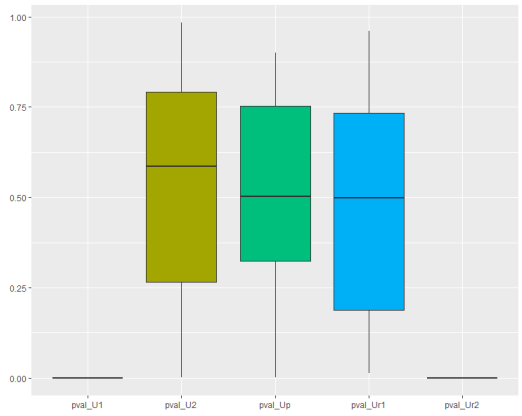

Figure 1: p-value of the HSIC-based index for the constraint $g \leq 0$

We give a toy example from [1] of excursion sets on which we compute kernel-based indices. The constraint $g \leq 0$ is given by:

$$g(\boldsymbol{x}, \boldsymbol{U}) = U_5 - \max_{t \in [0,T]} \mathcal{Y}''(x_1 + U_1, x_2 + U_2, U_3; t),$$

where $\mathcal{Y}(a, b, c)$ is the solution of an harmonic oscillator defined by:

$$a\mathcal{Y}''(t) + c\mathcal{Y}'(t) + b\mathcal{Y}(t) = \eta(t),$$

where η is a given strength. We compute the p-value of the independence tests $HSIC_{k_{set}}(U_i, \Gamma) = 0$, for $i = 1, ..., 5$, see figure 1. We obtain that three of them are higher than 0.05, which mean that U_2, U_3 and U_4 have no influence on the excursion set $\Gamma = \{\boldsymbol{x} \in \mathcal{X}, g(\boldsymbol{x}, \boldsymbol{U}) \leq 0\}$, i.e. on the constraint compliance.

Short biography (PhD student)

After 3 years at the engineering school Ecole Centrale de Lyon, and a Master in Applied Mathematics (MeA Lyon 1), I started a PhD thesis with Ecole Centrale de Lyon and IFP Energies nouvelles. My work is to use sensitivity analysis methods to simplify robust optimization problems. I'm supervised by Céline Helbert, Christophette Blanchet from ECL and Adrien Spagnol, Delphine Sinoquet from IFPEN. My thesis is part of the consortium CIROQUO.

References

[1] Alexis Cousin et al. "A Two-Step Procedure for Time-Dependent Reliability-Based Design Optimization Involving Piece-Wise Stationary Gaussian Processes". In: *Structural and Multidisciplinary Optimization* (Apr. 2022). Springer.

[2] Sébastien Da Veiga. "Kernel-based ANOVA decomposition and Shapley effects - Application to global sensitivity analysis". Jan. 2021.

[3] Jean-Claude Fort, Thierry Klein, and Agnès Lagnoux. "Global Sensitivity Analysis and Wasserstein Spaces". In: *SIAM/ASA Journal on Uncertainty Quantification* (2021).

[4] Ilya Molchanov. *Theory of random sets*. Vol. 19. 2. Springer, 2005.

[5] Adrien Spagnol. "Kernel-based sensitivity indices for high-dimensional optimization problems". Theses. Université de Lyon, July 2020.

Central limit theorem for the stratified selection mechanism

R. Flenghi[†,1], B. Jourdain[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Jan. 2021 – Jan. 2024

[1] Cermics, École des Ponts, INRIA, Marne-la-Vallée, France
{roberta.flenghi,benjamin.jourdain}@enpc.fr

Abstract

Given a sequence of random variables $(X_m)_{1\leq m\leq M}$ with associated positive random weights $(W_m)_{1\leq m\leq M}$ such that $\sum_{m=1}^{M} W_m = 1$, a resampling scheme defines the resampled sequence $(X_{A_m})_{1\leq m\leq M}$ where A_m is a random index in $\{1,\cdots,M\}$ such that, not formally,

$$\frac{1}{M}\sum_{m=1}^{M}\delta_{X_{A_m}} = \sum_{m=1}^{M} W_m \delta_{X_m}. \tag{1}$$

Depending on the definition of the random index A_m, several resampling schemes can be considered. The most widely used resampling techniques are: multinomial resampling, residual resampling, stratified resampling (see [5]) and systematic resampling (see [3]).
Douc et al. [9] provide a general description of the different methods and their properties. They show that residual and stratified resampling improve over multinomial resampling in the sense of having lower conditional variance (with respect to the σ-algebra generated by $(X_m)_{1\leq m\leq M}$) and they show, by providing a counter-example, that the same property does not hold for systematic resampling.
The resampling schemes have been largely studied in the literature and it is generally not easy to understand their asymptotic behaviour because of the complicated structure in the definition of the possible A_m.
One of the last contributions is given by Gerber et al. [6]. Using the notion of negative association [4], they provide a general consistency result and in particular they provide the proof of almost sure weak convergence of $\frac{1}{M}\sum_{m=1}^{M}\delta_{X_{A_m}}$ in the stratified resampling method. Moreover they provide a counter-example to almost sure weak convergence for the systematic resampling method.
One of the main application of the resampling methods is in particle filtering also known as Sequential Monte Carlo methods (see for instance [8], [1]). Indeed particle filtering can be seen as a combination of sequential importance sampling [2] and resampling, we refer to the introduction of [9] for a brief and more precise description of the basic method.
In 2022 Chopin et al. [7] study the resampling schemes applied to particle filters in the weakly informative regime that is when the weights tend to be close to uniform. In this case the performance of the method can be highly affected by the choice of the resampling method.

In this paper we focus our attention on the stratified resampling scheme and our purpose is to study the asymptotic behaviour of the method as the number of particles $M \to \infty$. We first

compute the asympotic variance of $\frac{1}{\sqrt{M}} \sum_{m=1}^{M} \delta_{X_{A_m}}$ showing a convergence to a finite limit and then, knowing the asymptotic variance, we show that a Central Limit Theorem holds.
As said before, from a computational point of view it is not straightforward to deal with the resampling schemes in particular with the stratified one and that's why part of the demonstrations are quite technical. As far as we know, such results appear for the first time in the literature.

Short biography (PhD student)

I obtained my bachelor's and master's degrees in pure and applied mathematics at University of Rome Tor Vergata. I then started working in an energy company in Italy that I left to start a M2 in mathematical finance (EX. DEA-Lamberton) at University Gustave Eiffel.
Now I am a PhD student under the supervision of Benjamin Jourdain at Ecole des Ponts. I work in the field of probability and numerical probability.

References

[1] N. de Freitas A. Doucet and N.J. Gordon. Sequential monte carlo methods in practice. *Springer-Verlag*, 2001.

[2] J. Handschin and D. Mayne. Monte carlo techniques to estimate the conditionnal expectation in multi-stage non-linear filtering. *International Journal of Control*, 9:547–559, 1969.

[3] P. Clifford J. Carpenter and P. Fearnhead. Improved particle filter for nonlinear problems. *IEE Proceedings Radar, Sonar Navigation*, 146(1):2–7, 1999.

[4] K. Joag-Dev and F. Proschan. Negative association of random variables with applications. *The Annals of Statistics*, 11(1):286–295, 1983.

[5] G. Kitagawa. Monte carlo filter and smoother for non-gaussian state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996.

[6] Nick Whiteley Mathieu Gerber, Nicolas Chopin. Negative association, ordering and convergence of resampling methods.

[7] T. Soto N. Chopin, S.S. Singh and M. Vihola. On resampling schemes for particle filters with weakly informative observations. *ArXiv*, 2022.

[8] D. Salmond N. Gordon and A. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F (Radar and Signal Processing)*, 140(2):107–113, 1993.

[9] O. Cappé R. Douc and E. Moulines. Comparison of resampling schemes for particle filtering. *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, pages 64–69, 2005.

Gaussian Processes for Molecules: Performance Assessments, Comparisons and Sequential Design.

Athénaïs Gautier[†,1], Anna Broccard[2] David Ginsbourger[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Nov. 2018 - June 2023

[1] IMSV, university of Bern
{athenais.gautier,david.ginsbourger}@unibe.ch

[2] SBB (was a Master student at IMSV, University of Bern at the time of the contribution)

Abstract

Recently, statistical machine learning methods have shown great potential to assist with scientific discovery in natural sciences [1]. These methods can be used to analyze and interpret large amounts of data, identify patterns and relationships, and make predictions about phenomena that are difficult or costly to study experimentally. In the field of chemistry and molecular science, Gaussian Processes (GPs) have been applied to a wide range of problems, including predicting chemical reactivity, analyzing molecular properties, and designing new molecules. In this work, we will focus on molecular property prediction and GP's potential in accelerating exploration of chemical space.

Several recent developments in the domain [3] leverage the so-called Tanimoto kernel. The molecule are first encoded as binary-valued vectors based on their structural properties, resulting in their Morgan fingerprints.

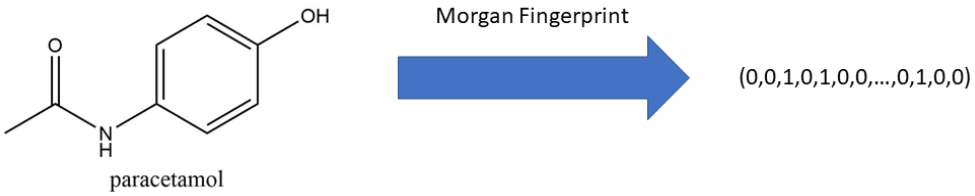

Figure 1: Basic principle of Morgan fingerprints

Then, similarity between fingerprints is then assessed through:

$$k_{\text{Tanimoto}}(x, x') := \begin{cases} \sigma^2 & \text{if } \langle x, x\rangle = \langle x', x'\rangle = 0 \\ \sigma^2 \frac{\langle x, x'\rangle}{||x||^2 + ||x'||^2 - \langle x, x'\rangle} & \text{else} \end{cases} \tag{1}$$

Despite its apparent simplicity, this kernel yields great predictive performances, hence justifying its popularity.

However, while it is well established that one of the key advantages of GPs is their ability to handle uncertainty both when integrating data or making predictions, our literature review

highlighted that most contributions assess their models mostly through the scope of predictive performance without accounting for the probabilistic nature of the predictions. It is natural to question whether neglecting to fully assess the model's calibration can hinder performances in active learning. Our work aims at addressing this question.

Our contribution is twofold: First, we review existing literature on principled diagnostics and state-of-the-art level-set estimation methods to make it accessible and ready to use for chemists and applied scientists. Second, we conduct a benchmark on several molecule datasets: Photoswitch, [4], ESOL, FreeSolv and Lipophilicity [5], HCEP [2] to assess whether well-calibrated models systematically lead to better exploration of molecular spaces.

Short biography (PhD student)

Athénaïs Gautier graduated with an engineering degree from Mines de Saint Etienne (2015-2018) as well as a MSc in applied mathematics from University Paris Dauphine (2018). Her PhD takes place within the framework of the Swiss National Science Foundation project number 178858 on "Uncertainty quantification and efficient design of experiments for data and simulation-driven inverse problem solving".

References

[1] Rafael Gómez-Bombarelli, Jennifer N. Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, Ryan P. Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, 2018. PMID: 29532027.

[2] Johannes Hachmann, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S Sánchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M Brockway, and Alán Aspuru-Guzik. The harvard clean energy project: large-scale computational screening and design of organic photovoltaics on the world community grid. *The Journal of Physical Chemistry Letters*, 2(17):2241–2251, 2011.

[3] Henry B. Moss and Ryan-Rhys Griffiths. Gaussian process molecule property prediction with flowmo. *CoRR*, abs/2010.01118, 2020.

[4] Aditya R. Thawani, Ryan-Rhys Griffiths, Arian Jamasb, Anthony Bourached, Penelope Jones, William McCorkindale, Alexander A. Aldrick, and Alpha A. Lee. The photoswitch dataset: A molecular machine learning benchmark for the advancement of synthetic chemistry. 2020.

[5] Zhenqin Wu, Bharath Ramsundar, Evan N Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learning. *Chemical science*, 9(2):513–530, 2018.

Grey-box modelling using multifidelity surrogate models to combine noisy measurements and computer simulations

A. Giannoukou[†,1], S. Marelli[§,2], B. Sudret[§,3]

[†] PhD student [§] PhD supervisor

PhD expected duration: Sep. 2021 – Sep. 2025

[1] Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich
katerina.giannoukou@ibk.baug.ethz.ch

[2] Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich
marelli@ibk.baug.ethz.ch

[3] Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich
sudret@ibk.baug.ethz.ch

Abstract

Predicting the output of complex systems can be very important when designing engineering systems that are safe and reliable. Nowadays, engineers and scientists have more tools than ever to deal with this challenge, one of these tools being different classes of predictive models. More precisely, computer simulations, or white-box models, are widely used to provide accurate predictions for systems where an underlying physical model exists, by modelling the system properties based on existing knowledge. However, these models are often hard to obtain, and their modelling accuracy may be limited due to the assumptions imposed. A second class of predictive models consists in data-driven approaches, or black-box models. These can provide great flexibility, but are not necessarily always in line with the underlying physical reality [3].

Engineering systems are often described by one or more white-box computational models, while additional data about the system output can be obtained from experiments. In this case, approaches that combine the white- and the black-box modelling frameworks are desirable, as they can overcome the limitations that each of them poses separately. Grey-box modelling aims at achieving this combined approach, and is explored and developed in depth within the context of the Horizon 2020 GREYDIENT project [1]

In this work, we propose the class of multifidelity surrogate models (MFSMs) as a way to perform grey-box modelling. MFSMs fuse information from multiple models of varying fidelity inside a single surrogate model. As model fidelity, we refer to the accuracy of the model, or the degree to which this model represents reality. Generally, as the fidelity of a model increases, the cost of obtaining data from it increases too. In multifidelity (MF) surrogate modelling, typically, a very small high-fidelity (HF) data set is supplemented with one or more larger lower-fidelity ones [1, 2]. In our grey-box setting, the available experimental data are considered as the high fidelity, whereas the low fidelity is provided by the white-box computational models.

Most of the available literature on MF surrogate modelling focus on noise-free models. However, in the real world, experimental data are contaminated by measurement noise, due to the generally limited precision of measurement devices. We treat this noise as aleatory uncertainty.

[1] https://www.greydient.eu/.

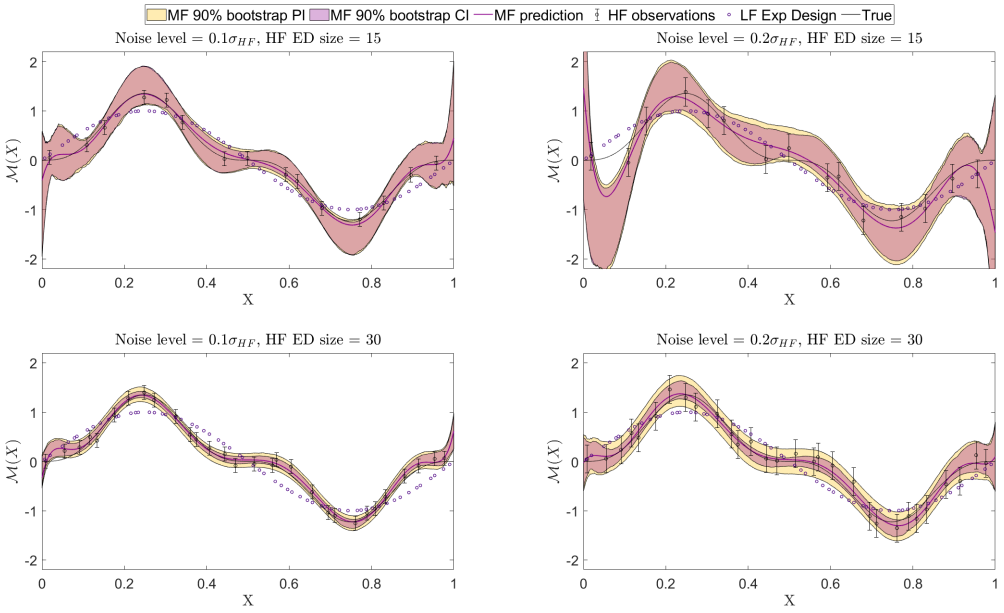

Figure 1: Multifidelity prediction with 90% confidence and prediction intervals for an analytical 1-D toy example.

In addition, the low-fidelity models we consider in our setting are in principle expensive-to-evaluate, and therefore they are represented by surrogate models trained on a limited set of model evaluations. This step introduces epistemic uncertainty in the MF model predictions.

This work proposes a novel framework for grey-box modelling with noisy data expressed as a multifidelity surrogate modelling problem. The goal of our MFSMs is to accurately model the response of the underlying deterministic HF model, or simply, to denoise the noisy HF data. Polynomial chaos expansions are employed as a surrogate for the MFSM construction. Moreover, we aim at quantifying the different types of uncertainty in our MFSM prediction. For this purpose, we use bootstrap to compute the confidence intervals for the mean MF model predictions, and the prediction intervals for future HF observations. Our methodology is validated on different use-cases, one of them being an analytical 1-D toy example, the results for which are shown in Figure 1 for different HF experimental design sizes and different noise levels.

Short biography

Aikaterini (Katerina) Giannoukou is a second-year PhD student at the Chair of Risk, Safety and Uncertainty Quantification, at ETH Zurich. Her project is part of the Marie Skłodowska-Curie Innovative Training Network GREYDIENT (No 955393), and her work focuses on multifidelity modelling.

References

[1] M. C. Kennedy and A. O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 2000.

[2] L. Le Gratiet and J. Garnier. Recursive co-Kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.

[3] H. J. A. F. Tulleken. Grey-box modelling and identification using physical knowledge and Bayesian techniques. *Automatica*, 29(2):285–308, 1993.

Parametric Model Order Reduction for the Statistical Finite Element Method

L. Hermann[†,1], U. Römer[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Dec. 2021 – Dec. 2024

[1] Institut für Dynamik und Schwingungen, Technische Universität Braunschweig,
Langer Kamp 19, 38106 Braunschweig, Germany
{l.hermann,u.roemer}@tu-braunschweig.de

Abstract

Finite Element (FE) simulations of physical models are widely used in the design phase and, more recently, for the monitoring of technical systems. In a monitoring context, data have to be incorporated into the FE approach, where an additional challenge stems from the fact that the FE predictions most often do not match exactly the given measurement data of a real system. Girolami et al. [2] have recently proposed their statistical FEM (statFEM) approach for incorporating measured data into the FE simulation, while accounting for model error. Unlike traditional Bayesian inversion techniques, not the model parameters but the model state is subjected to Bayesian inference and the model error is explicitly quantified as part of the inference in form of a Gaussian Process (GP).

In previous studies, we have adopted the method to assimilate sparse but accurate sensor data with a FE model based on the Helmholtz equation [3]. However, finding the model response for a fine discretization in frequency domain involves solving the model at many frequency steps, which is computationally expensive. Here, we replace the FEM model with a reduced order model (ROM) and therefore propose a new generating model for data assimilation in the frequency domain. The equations for the ROM-based data approximation, which we call statROM, are based on the assumption that the model state and the model misspecification are described as GPs. Starting with a ROM prior, induced by uncertainty in the state equations, Bayesian inference yields a posterior model output by conditioning on data. Carrying out ROM-based data assimilation enables an accurate approximation of the inference problem at a much lower computational cost.

For reducing the order of the model, we employ projection based approaches and compare both a ROM based on proper orthogonal decomposition and a second order Arnoldi moment matching method. Thereby, the ROM approximation error is explicitly quantified and accounted for in the data generating model, resulting in an improved accuracy of the inferred state. We discuss different means of retaining the parametric dependency of the model in the ROM approximation.

We start by presenting the theoretical basis for statROM as well as a simple example to illustrate the used methods and to show their convergence behavior. The model under consideration is a 1D discretization of the Helmholtz equation with Neumann and impedance boundary conditions. Unlike the original statFEM paper [2], we make use of an unscented transform to forward-propagate uncertainties in e.g. material parameters to the statROM prior density. Here, statROM is able to reproduce the statFEM results to a high accuracy while solving much smaller systems. In Figure 1, we compare a full order statFEM with 100 degrees of freedom

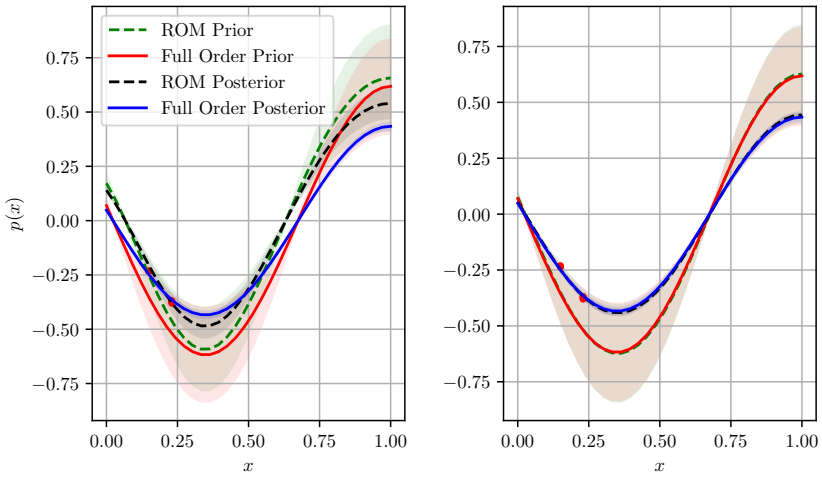

Figure 1: statROM solution for the 1D Helmholtz equation at 212Hz. Small ROM basis ($L = 3$) on the left vs. a larger one ($L = 6$) on the right. In both cases the reduced system is much smaller than the original FEM system (100 elements).

to statROM with a basis of size 3 on the left and of size 6 on the right. A certain size of the basis is needed for accurate results, but above that the accuracy is very high and the difference is barely visible. We derive an upper bound and conduct convergence studies to quantify the ROM approximation error, which depends on both the size of the ROM basis and the chosen wave number. The error is set in relation to the other sources of errors as discussed in [4] for the standard statFEM case. As a second example we consider a more complex Helmholtz problem in 3D. We approach emerging scalability issues by adapting a low-rank Bayesian update as proposed by Duffin et al. [1]. Several numerical studies will be given to illustrate the performance and accuracy of the method.

Short biography (PhD student)

Lucas Hermann studied Mechanical Engineering and received a B.Sc. from HAW Hamburg and a M.Sc. from TU Braunschweig. Since 2021, he is part of the Research Training Group RTG2075 at TU Braunschweig as a PhD student. His research topic is the statistical Finite Element Method, primarily for solving the Helmholtz equation. This research receives funding by the DFG (German Research Foundation) with the project number 255042459.

References

[1] C. Duffin, E. Cripps, T. Stemler, and M. Girolami. Low-rank statistical finite elements for scalable model-data synthesis. *Journal of Computational Physics*, 463, 2022.

[2] M. Girolami, E. Febrianto, G. Yin, and F. Cirak. The statistical finite element method (statfem) for coherent synthesis of observation data and model predictions. *Computer Methods in Applied Mechanics and Engineering*, 375, 2021.

[3] L. Hermann, H.K. Sreekumar, S.C. Langer, and U. Römer. A statistical approach for the fusion of data and finite element analysis in vibroacoustics. In *Fortschritte der Akustik - DAGA 2022, 48. Jahrestagung für Akustik, Stuttgart*, 2022.

[4] Toni Karvonen, Fehmi Cirak, and Mark Girolami. Error analysis for a statistical finite element method. arXiv:2201.07543v1, 2022.

Uncertainty quantification and global sensitivity analysis for a steam generator clogging simulation code

E. Jaber[†,1,3], É. Remy[‡,1], V. Chabridon[‡,1], M. Baudin[‡,1], Q. Feng[2]
M. Mougeot[§,3], D. Lucor[§,4]

[†] PhD student (presenting author) [‡] PhD industrial advisors [§] PhD supervisors

PhD expected duration: Feb. 2023 – Feb. 2026

[1] EDF R&D, PRISME (Performance, Industrial Risk, Monitoring for Maintenance and Operation) department, 6 quai Watier, 78401 Chatou, France
[2] EDF R&D, MFEE (Fluid Mechanics, Energy and Environment) department, 6 quai Watier, 78401 Chatou, France
[3] Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, 91190 Gif-sur-Yvette, France
[4] Laboratoire Interdisciplinaire Des Sciences Du Numérique, LISN-CNRS, F-91403 Orsay, France

Abstract

This work deals with the study of the clogging phenomenon observed in some steam generators of the French nuclear power plants. Clogging is due to corrosion in the secondary circuit of the power plants out of which iron oxide impurities are produced. These particles are then transported to the steam generator and, due to the vaporisation and the modification of the flow geometry, clogging deposits start to appear. Over time, these deposits may affect the steam generator in several ways: for instance, through a modification of the heat transfer between both primary and secondary circuits, or through vibrations induced by the flow redistribution. Therefore, the clogging rate is periodically measured with cameras during outages and, if necessary, chemical cleaning can be performed. A good maintenance planning is thus crucial for EDF, which explains the R&D focus to better understand this complex phenomenon.

A fine modeling of the most elementary mechanisms at stake in the clogging phenomenon has been proposed in the work of [2], as well as the related simulation tools. The predominant thermohydraulic and chemical phenomena for the modelling of clogging are the *vena contracta* and *flashing* mechanisms and the equations obtained is a combined system of both partial and ordinary differential equations, more precisely with two transport equations and one nonlinear ordinary differential equation. Based on this modelling, two simulation codes have been developed at EDF R&D: the DEPOTHYC code, used for the clogging predictions on short time scales (1 call $\approx$ 15 minutes) and the computational chain called THYC-Puffer-DEPOTHYC, used to predict clogging evolution on long time scales (1 call $\approx$ 4 hours).

The proposed work aims at applying the uncertainty quantification methodology to these costly codes, motivated by the fact that a number of physical input parameters are random and the model also contains a calibration parameter. The problem is rather low-dimensional (less than ten random input variables).

To this end, a coupling of these codes to the OpenTURNS library [1] has been built in order to perform a first set of Monte Carlo simulations. Then, a polynomial chaos expansions metamodel [3] is built to carry out global sensitivity analysis by computing Sobol' indices. Another complementary analysis is performed using various formulations (i.e., global, target, conditional) of

the HSIC indices [4]. Finally, a calibration study is performed with respect to the previously mentioned calibration parameter. As a result, this work enables to identify the most influential input variables on the clogging phenomenon, to reduce some input uncertainties and to better explore the inner behavior of the computational tool, as well as getting a predictive metamodel whose cost is negligible compared to the true simulator.

This study is a preliminary step in a longer study whose global aim is to couple this complex simulation tool with other strategies based on machine learning used to predict clogging according to a large panel of data.

Short biography (PhD student)

Recently graduated engineer from CentraleSupélec with a major in Applied Mathematics and an M2 in analysis of PDE's from Université Paris-Saclay. After an end-of-studies internship in UQ at EDF R&D, PRISME department, I am currently pursuing a PhD at EDF R&D jointly with the Centre Borelli of ENS Paris-Saclay and the LISN on hybrid approaches between statistical and deterministic models applied to the industrial issue of clogging in steam generators.

References

[1] M. Baudin, A. Dutfoy, B. Iooss, and A.-L. Popelin. OpenTURNS: An Industrial Software for Uncertainty Quantification in Simulation. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Handbook of Uncertainty Quantification*, pages 2001–2038. Springer International Publishing, Cham, Switzerland, 2017.

[2] T. Prusek. *Modélisation et simulation numérique du colmatage à l'échelle du sous-canal dans les générateurs de vapeur*. PhD thesis, Université Aix-Marseille, 2012.

[3] T. J. Sullivan. *Introduction to Uncertainty Quantification*. Texts in Applied Mathematics 63. Springer-Verlag, Berlin, Heidelberg, 2015.

[4] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis: Theory and Practice in R*. SIAM, Computational Science and Engineering, 2021.

Robust Design Optimization of electrical machines for electric and hybrid vehicles

A. Reyes-Reyes[1, 2†], S. Hlioui[1, §], D. Sinoquet[2, §], A. Nasr[2, §]

† PhD student (presenting author) § PhD supervisor

PhD expected duration: Dec. 2021 – Nov. 2024

[1] Université Paris-Saclay
adan.reyes-reyes@universite-paris-saclay.fr
sami.hlioui@ens-paris-saclay.fr

[2] IFP Energies nouvelles
delphine.sinoquet@ifpen.fr
andre.nasr@ifpen.fr

Abstract

Today, in order to reduce the environmental impact of our daily activities, electric and hybrid vehicles (EVs, EHVs) are presented as an alternative to combustion vehicles to reduce greenhouse gas emissions into the atmosphere. One of the key components in these vehicles is the electrical machine. To design an electrical machine, we often use Finite Element Method (FEM) codes, which are expensive in computational time. Dimensioning an electric machine for a vehicle presents several challenges , such as increasing torque and power densities, decreasing raw material prices, and increasing overall efficiency, hence the need to propose optimal design methodologies.

Today, most approaches do not consider any uncertainty in optimizing electrical machines despite the existence of several uncertainties related to manufacturing and assembly tolerances as well as material properties. These uncertainties may have a considerable impact on the quantities of interest in an electrical machine's optimization, so they must be considered in the optimization procedure in terms of a robustness metric, such as expectations, variances, or probabilities.

However, as mentioned before, the optimization of an electrical machine requires high computational time to compute all the samples needed to run an optimization algorithm based on FEM simulations. This problem is even more important if we consider the number of simulations required to obtain a good estimate of a robustness metric of a quantity of interest. Therefore, in a previous study we optimized an electrical machine using fixed surrogate models built with a sample of FEM simulations, to replace the objective functions [1]. Although the results obtained in this study are satisfactory, there are differences between the results predicted by the surrogate models and the FEM simulator. This is why we are currently working on adaptive optimization based on substitution models, in particular Bayesian Optimization (BO).

In this work, we propose to optimize an electrical machine via 3 different BO procedures:
The first of them solves a multi-objective optimization problem without uncertainty with two different infill criteria : q-step Expected Hypervolume Improvement (qEHVI) [2] and q-step ParEGO [3] which is a multi-objective extension of the well-known Efficient Global Optimization (EGO) algorithm [4]. These optimizations were performed to compare non-robust solutions to robust solutions. The second procedure takes uncertainty into account by adding a robustness metric known as robustness hypervolume [5] as an objective function. This metric calculates the joint degradation caused by uncertainties around the values of the non-perturbed outputs (when no

uncertainty is considered). At each iteration of the algorithm, we estimated the robustness hypervolume at each iteration at the observed non-perturbed data with an internal optimization algorithm. The multi-objective optimization algorithms used in this procedure are the same as for the first one.

The final proposed method aims at solving a multi-objective optimization problem where the objectives are the expectations of the quantities of interest with respect to uncertainties. This procedure is an extension of the Expected Feasible Improvement Stepwise Uncertainty Reduction algorithm (EFISUR) [6].

Short biography (PhD student)

Adan Reyes Reyes completed his *Diplôme d'Ingénieur* with a specialization in applied mathematics in Nantes, France. He joined IFPEN in the electric motor development team in collaboration with the Applied Mathematics Department. He continues this research work at IFPEN in the framework of a PhD thesis at Paris-Saclay University and is currently in his second year.

References

[1] Reyes, Adán; Sinoquet, Delphine; Nasr, André; Hlioui, Sami (2022) Robust design optimization taking into account manufacturing uncertainties of a permanent magnet assisted synchronous reluctance motor. 4th IEEE International Conference on Electrical Sciences and Technologies in Maghreb, Tunisia 2022 Online : https://arxiv.org/pdf/2211.16136.

[2] Daulton, Samuel; Balandat, Maximilian; Bakshy, Eytan (2020) Differentiable Expected Hypervolume Improvement for Parallel Multi-Objective Bayesian Optimization. In : Advances in Neural Information Processing Systems 33. Online : http://arxiv.org/pdf/2006.05078v3.

[3] J. Knowles (ParEGO) Parego a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. IEEE Transactions on Evolutionary Computation, 10(1):50–66, 2006.

[4] Jones, Donald; Schonlau, Matthias (1998) Efficient global optimization of expensive black-box functions. Journal of Global Optimization, 13, 1998.

[5] Krasopoulos, Christos; Beniakar, Minos; Kladas, Antonios (2017) Robust optimization of high-speed pm motor design. IEEE Transactions on Magnetics, 53(6):1–4, 2017.

[6] Amri, Reda El; Le Riche, Rodolphe; Helbert, Céline; Blanchet-Scalliet, Christophette; Da Veiga, Sébastien (2021) A sampling criterion for constrained Bayesian optimization with uncertainties. Online : http://arxiv.org/pdf/2103.05706v3.