Adaptive Sampling Strategies for Crashworthiness Applications

Pietro Lualdi
pietro.lualdi@dlr.de
Institute of Vehicle Concepts
German Aerospace Center (DLR)

Ralf Sturm
ralf.sturm@dlr.de
Institute of Vehicle Concepts
German Aerospace Center (DLR)

Abstract

In the context of surrogate metamodeling for crashworthiness applications, the implementation of adaptive sampling strategies holds great potential for overcoming the challenge of setting the optimal number of samples a priori. These adaptive strategies offer a significant advantage by avoiding the common pitfalls of under- and oversampling, making them attractive for expensive-to-evaluate functions such as those commonly encountered in crashworthiness applications. Despite their potential, most current research in this area relies predominantly on static sampling strategies. Recognizing this gap, our work explores the adaptation of innovative adaptive sampling methods specifically tailored to the needs of crashworthiness applications. In this context, we describe the Multi-Query Cross-Validation Voronoi (MQCVVor) method. This approach extends the traditional CVVor technique by integrating parallel processing, thus improving the efficiency and accuracy of surrogate models, especially for small scale multi-response systems. Our method demonstrates a significant improvement over conventional static Latin Hypercube Design (LHD) in terms of convergence speed and robustness. In addition to these results, we briefly discuss the potential limitations of adaptive sampling strategies and lay the groundwork for future research aimed at refining these techniques for more complex scenarios.

1 Introduction

The pursuit of efficiency and accuracy in crashworthiness optimization often encounters the prohibitive barrier of computational cost. Surrogate model (also known as metamodels or response surface model), have emerged as a promising solution to this challenge. These models aim to emulate the complex, non-linear functions characteristic of crashworthiness problems, where the analytical relationship between inputs and outputs remains unknown. The effectiveness of these surrogate models is highly dependent on the quality and quantity of the underlying data. A major concern, however, is the cost associated with collecting such simulation data. It is of utmost importance to minimize the time spent on response evaluation - a task that often takes several hours [5] - as this aspect significantly overshadows the time required for training and evaluation of surrogate models.

A critical approach to improving the efficiency of this process is to employ ideal sampling strategies, which aim to sample the domain according to specific optimality criteria. By carefully selecting samples, the information gained from each simulation run can be maximized, thereby increasing the overall efficiency. Traditionally, sampling strategies in crashworthiness applications are static, meaning that all required samples are selected in a one-shot, initial phase prior to fitting the surrogate models. Often referred to as space-filling or response-free strategies, these approaches aim to fully explore the variable domain without regard to the output values of the function.

In contrast, adaptive sampling strategies provide a dynamic approach that adapts to the specific characteristics of the function, such as gradients, peaks, and discontinuities. This adaptability allows for a more judicious allocation of resources, leading to efficient results [1].

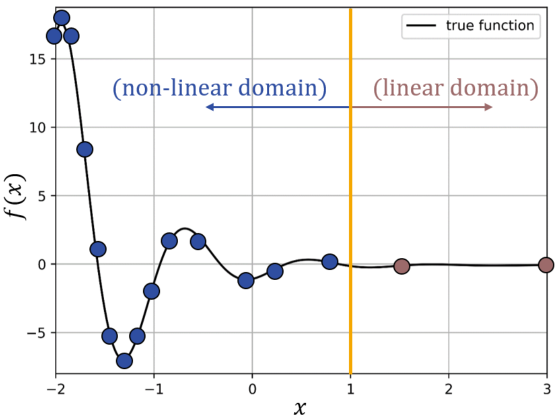


Figure 1: One-dimensional damped sinusoidal function (such as $f(x) = A^{-\lambda \cdot x} \cdot cos(\omega x))$ sampled using an adaptive sampling approach. The function is very simple and almost linear to approximate on the left side, but very non-linear on the right side. Intuitively, more samples should be allocated to the right side to compensate for this non-linearity.

The motivation for using adaptive strategies becomes clear when considering a function with varying characteristics, such as the damped sinusoidal wave function shown in Figure 1. This function exhibits nearly linear behavior on the left-hand side, while the right-hand side is characterized by strong non-linearity. In such scenarios, a fixed budget of observations would be better utilized by allocating more samples to capture the non-linear regions, since the linear portions require fewer samples for accurate representation. In the area of computational experimental design, adaptive sampling has been reported to be potentially more efficient than static sampling methods. This efficiency is evidenced by its ability to reduce the number of samples needed to achieve a given level of accuracy in surrogate models used to approximate black-box functions [4, 12].

Our study aims to identify an adaptive sampling strategy tailored for surrogate-based optimization in crashworthiness applications. We plan to explore its potential limitations, develop solutions to these problems, and test the approach with a crash simulation experiment.

2 Adaptive sampling for surrogate models

The main benefit of adaptive sampling lies in the sequential selection of samples, which effectively addresses the issues of under- and oversampling that are common in static sampling techniques. This sequential approach is balanced by a dual focus: exploration, which involves investigating unexplored areas of the variable domain, and exploitation, which focuses on detailed examination of areas with noteworthy features such as gradients, peaks, valleys, and discontinuities. To handle this trade-off, different data sources can be used, depending on the adaptive strategy being considered. Nevertheless, the availability of these sources depends on the type of surrogate models employed. In surrogate-based optimiaztion for crashworthiness applicarions, Gaussian process (GP), support vector regression (SVR), polynomial response function (PRF), and radial basis function (RBF) are among the most commonly used metamodels [2]. GPs are widely preferred for their accuracy and flexibility. However, their cubic training complexity ($\mathcal{O}(n^3)$) might become a limiting factor for high-dimensional problems. This drives us to search for an adaptive sampling method that not only minimizes the need for costly observations, but is also universally applicable across different surrogate models. This requirement, together with the specific needs of crashworthiness applications, which include prediction of nonlinear patterns, scalability, robust global metamodeling capability, and computational efficiency, led us to a thorough examination of the existing literature, including key findings from the reviews by Garud et al. [4] and Fugh et al. [3]. After careful cross-checking of this information, we identified the CVVor (Cross-Validation Voronoi) sampling method as a potential optimal choice. We will explain and discuss CVVor in more detail in the following section.

3 Cross-Validation Voronoi sampling

In this section, we review the CVVor method originally introduced by Xu et al. [13] and analyze its limitations. We then propose an extension of the CVVor method to adapt it to the requirements of crashworthiness problems.

3.1 Original method and limitations

The CVVor sampling method aims to handle the trade-off between exploration and exploitation by using the Voronoi algorithm and the Leave-One-Out Cross-Validation (LOOCV) method, respectively. Consider a black-box function f: $\mathbb{R}^d \rightarrow \mathbb{R}$ in a d-dimensional problem, where the initial exploration data set $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n\}$ and its corresponding output values $\mathbf{y} = \{f(\mathbf{p}_1), f(\mathbf{p}_2), ..., f(\mathbf{p}_n)\}$ is given and consists of $n = 5 \cdot d$ pairs of input/output observations. A surrogate model $\tilde{f} : \mathbb{R}^d \rightarrow \mathbb{R}$ approximating the real function f is trained on the initial samples.

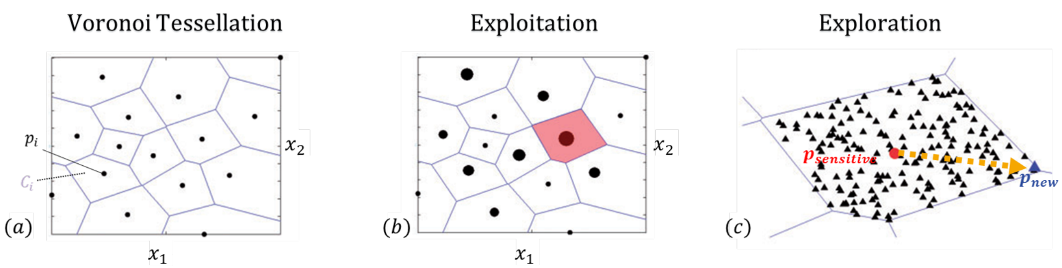


Figure 2: Main steps of the CVVor method: (a) The domain is partitioned into a set of Voronoi cells. (b) The sensitive Voronoi cell $C_{senstive}$ is identified by using Leave-One-Out cross-validation. (c) A new sample $\mathbf{p}_{new}$ is selected to be the farthest away from $\mathbf{p}_{sensitive}$ within the sensitive cell $c_{senstive}$. Image adapted from Xu et al. [13].

The first step of the method to partition the domain into n Voronoi cells - i.e. the region consisting of all points closer to that sample than to any other [12] – centering them on existing sample points, p_i. We denote these cells as $\mathbf{C} = \{c_1, c_2, ..., c_n\}$. Next, the LOOCV phase (exploitation step) identifies the most sensitive Voronoi cell, $c_{sensitive}$. This involves temporarily excluding each sample point, building a metamodel with the remaining points $\tilde{f}_{P \setminus p_i}$, and then evaluating the accuracy of the model against the excluded point. By repeating this for all points, the method identifies the cell where model improvement is most needed. The error of the point p_i is calculated according to Eq. (1):

$$e^i_{\text{LOO}} = |f(p_i) - \tilde{f}_{P \setminus p_i}(p_i)| \tag{1}$$

The final step (exploration step) is to strategically place a new sample point, $\mathbf{p}_{new}$, within the identified sensitive cell, maximizing the distance from the existing central point to explore less sampled regions. For a clear visual understanding of the method, these three main steps of the CVVor approach are illustrated in Figure 2.

Despite the attractive features of the CVVor method for handling the exploration-exploitation trade-off, two practical limitations must be addressed for its effective application in crashworthiness studies:

- **Single-response**: While the Voronoi method is response function independent, the LOOCV component is only applicable to single-objective problems without constraint functions. This scenario is rare in crashworthiness applications, which often require more complex multi-constrained analyses.

- **Single-query**: The fine granularity of the method could hinder efficiency, especially when parallel computing capabilities are available, such as on high-performance computing (HPC) clusters.

In the next section, we propose an extension of the CVVor method to handle multi-response problems and to implement a batch query point logic that allows for parallel computation of simulations.

3.2 Proposed extension

Let us now consider a multi-response system F consisting of t functions such as $F = \{f_1, f_2, ..., f_t\}$. Note that to keep the notation simple, f denotes a generic response function, not necessarily an objective function. Thus, constraint functions are also included in this notation. Inspired by the idea of the "most critical function" introduced by Liu et al. [6], we extend the CVVor approach to address the first limitation. This extension can be visualized in Figure 3 and summarized in five main steps:

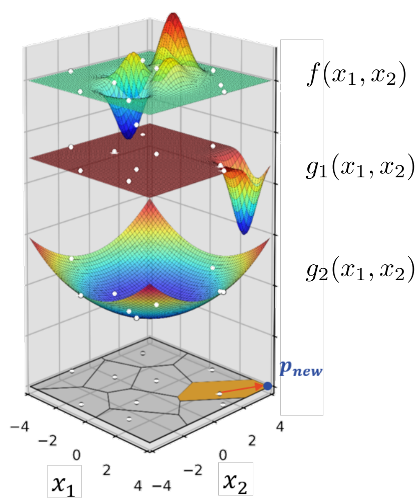


Figure 3: Detection of $c_{sensitive}$ and $\mathbf{p}_{new}$ in a multi-response system consisting of an objective function f_1 and two constraint functions, g_1 and g_2, respectively.

- Since response functions can produce responses with orders of magnitude of difference, we first normalize them based on the real responses collected. This can be done using the MinMaxScaler approach shown in Eq. (2):

$$\hat{f}_{j,norm}(x) = \frac{\hat{f}_j(x) - \min(\mathbf{y}_j)}{\max(\mathbf{y}_j) - \min(\mathbf{y}_j)}, \quad j \in \{1, 2, \ldots, t\} \tag{2}$$

- We generate the starting n exploration designs according to the Latin Hypercube Translational Propagation (TPLHD) algorithm developed by Diana et al. [10] and the associated Voronoi tessellation (i.e. the partition into Voronoi cells).

- According to the normalized generalized mean square cross-validation error (NGMSE), we calculate the weights w_j of the F functions based on their accuracy level (see Appendix A for further details about NGSME):

$$w_j = \frac{\text{NGSME}_j}{\sum_{i=1}^{t} \text{NGSME}_i}, \quad j \in \{1, 2, \ldots, t\}$$

The poorer the accuracy of a surrogate model $\hat{f}_j$, the greater its weight w_j.

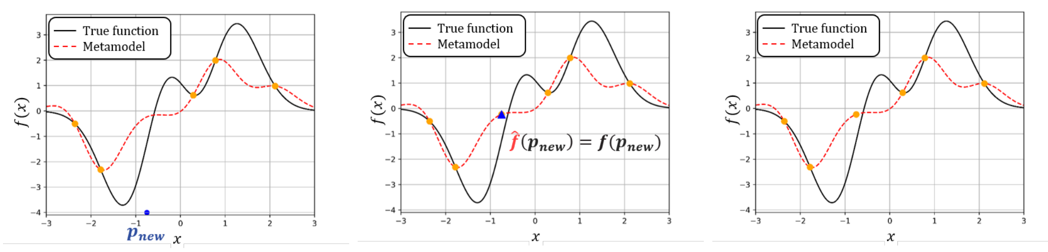
- Using these weights, we then rank the errors of the Voronoi cells according to the sum-weighted approach shown in Eq. (3):

$$e_i^{sys} = \sum_{j=1}^{F} w_j \cdot e_i^j, \quad i \in \{1, 2, \ldots, n\} \tag{3}$$

- Finally, analogous to the original method, we pick $\mathbf{p}_{new}$ within $c_{sensitive}$ (i.e. the cell with the highest error e_i^{sys}) that is furthest away from $\mathbf{p}_{sensitive}$.

With such a differential weighting method, errors from less accurate surrogate models have a more significant impact on the evaluation of cell errors, leading the algorithm to prioritize their improvement. As a result, this adaptation not only preserves the integrity of well-performing models, but also actively seeks to improve the performance of underperforming models.

To overcome the limitation of parallel processing, we now present the second part of the extension of the original method. Let the number of simultaneous function evaluations be denoted by n_{par}.This extension can be effectively summarized in three steps, which are clearly illustrated in Figure 4.



(a) Find $\mathbf{p}_{new}$ (b) $\hat{f}(\mathbf{p}_{new}) = f(\mathbf{p}_{new})$ (c) Retrain the metamodel $\hat{f}$

Figure 4: Extension of the CVVor method to a multi-query approach enabling parallel simulations.

- First, a new point, $\mathbf{p}_{new}$, is determined using the standard procedure described in the original method.
- Instead of directly evaluating the function at $\mathbf{p}_{new}$, we temporarily assume that the function at this point is known and equal to the true function value, so we set $\tilde{f}(\mathbf{p}_{new}) = f(\mathbf{p}_{new})$.
- This assumption allows us to include $\mathbf{p}_{new}$ and $f(\mathbf{p}_{new})$ provisionally in the dataset, allowing the metamodel to be updated without the need for immediate function evaluation.

This process is repeated n_{par} times, depending on the parallel execution capacity. While this technique does not guarantee that each new point will provide in-depth exploitation insights, the Voronoi tessellation ensures that the sampling remains well distributed over the domain. Since the computational cost of evaluating one function and n_{par} functions in parallel is comparable, this parallelized method provides a superior amount of information.

In the following, we will refer to our proposed extension as Multi-Query CVVor (MQCVVor) for ease of notation and explanation.

Example on analytic functions

In this section, our goal is to visually assess whether the performance of the proposed method is in line with our expectations. For this purpose, we consider the multi-system already observed in Figure 3, composed of the Peaks function, the Easom function and the Sphere function, representing

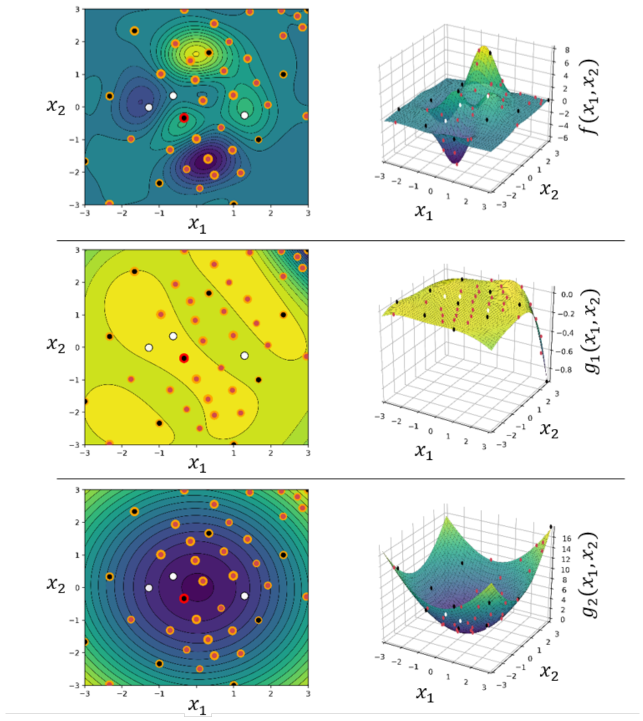


Figure 5: From top to bottom: the surrogate models for the Peaks function, the Easom function, and the Sphere function after ten iterations. The black dots represent the points from the initial design of the experiment. The red dots are the adaptive points added by the MQCVVor method, while the white dots represent the last three adaptive points added in the final iteration.

f, g_1 and g_2 respectively. For more details on these functions, we refer the reader to the work of Molga et al. [7].

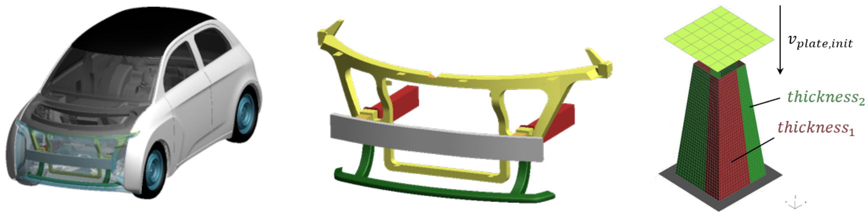
The results shown in Figure 5 are in line with our expectations. Besides the black points, which represent the purely exploratory points of the initial dataset, the adaptive points - marked in red - (with the last batch of three points marked in white) are mainly distributed to capture the gradients in the center and in the upper right corner of the domain. These gradients are due to the Peaks and Easom functions, respectively. The Sphere function, characterized by a uniform quadratic trend throughout the domain, appears to have a neutral effect on the sample distribution.

Further comparisons on the use of a multi-query strategy applied to the same multi-response problem show that, for example, using three parallel simulations instead of one can achieve the same accuracy target with about 43 % less computational time (see Appendix B for more details).

4 Experiment

Finally, we test the presented approach on a crash application. The crash-box of the front structure of the Urban Modular Vehicle (UMV) [8] - a battery-powered modular vehicle concept developed by our institute (see Figure 6) - is investigated.

This structure is analyzed within the NCAP Full Width Rigid Barrier load case scenario [9], which simulates a vehicle impacting a rigid barrier at 50 km/h. Our simplified finite element analysis (FEA) model of the crash-box is divided into two parts with different thicknesses. Since the crash-box, compressed by a rigid plate moving at the initial collision speed, is expected to absorb 30-40 % of the kinetic energy, we set the mass of the plate at 400 kg. In our multi-response system, we



(a) UMV Basic (b) Front structure (c) Crash-box

Figure 6: Extension of the CVVor method to a multi-query approach enabling parallel simulations.

consider three functions of interest: the intrusion (or crushing length), the internal energy (absorbed energy), and the total mass of the crash-box.

We compare our extension of the CVVor method to the maxmin Latin Hypercube Design (LHD), a static sampling strategy that maximizes the minimum distance between points in the design space, commonly used in crashworthiness optimization [11, 2]. Following the guidelines of Xu et al. [13] we start with a data set of size $n = 10$. The iteration process stops when the accuracy improves by 50 % compared to the accuracy measured on the initial dataset. Accuracy is measured by the normalized root mean square error (NRMSE) among the three response functions (i.e., the surrogate model with the worst accuracy). The NRMSE for the j-th response function is defined as shown in Eq. (4):

$$\mathrm{NRMSE}_j = \sqrt{\frac{1}{q}\sum_{i=1}^{q}\left(\frac{f_j(\mathbf{x}_i) - \hat{f}_j(\mathbf{x}_i)}{\max\{\mathbf{y}_{q,j}\} - \min\{\mathbf{y}_{q,j}\}}\right)^2} \tag{4}$$

where q is the number of test points, and $\mathbf{y}_{q,j}$ is the response vector calculate at these test points. It is noted that, typically, test points are not available in crash applications. For this reason, we use NGSME in the standard procedure based on cross-validation in our method. However, since the FEA model under consideration is a simplified and computationally inexpensive model, we take advantage of 100 test design points randomly distributed in the variable domain to more accurately measure the effectiveness of the method.

5 Results

A preliminary visual analysis of the surrogate models shown in Figure 7 looks very promising. The adaptive points (marked in white) seem to effectively identify the steep slopes of the internal energy and the intrusion of the crash box. These are located at diagonally opposite angles of the variable domain, leading to a splitting of the adaptive samples mainly in these two regions. The mass, as expected for a perfectly linear function, does not seem to negatively interfere with the adaptive sampling process.

As for the comparison between the convergence curves of the maxmin-LHD method and MQCVVor shown in Figure 8, it is important to note that our adaptive method iteratively adds a batch of 3 samples to an existing design, whereas the maxmin-LHD method creates a new one-shot design with three additional samples each time. Each test is repeated 25 times to account for the stochastic components of the two methods. The results show that the MQCVVor method outperforms the LHD method both in terms of convergence speed (reaching the target after 20 iterations) and robustness (narrower confidence intervals).

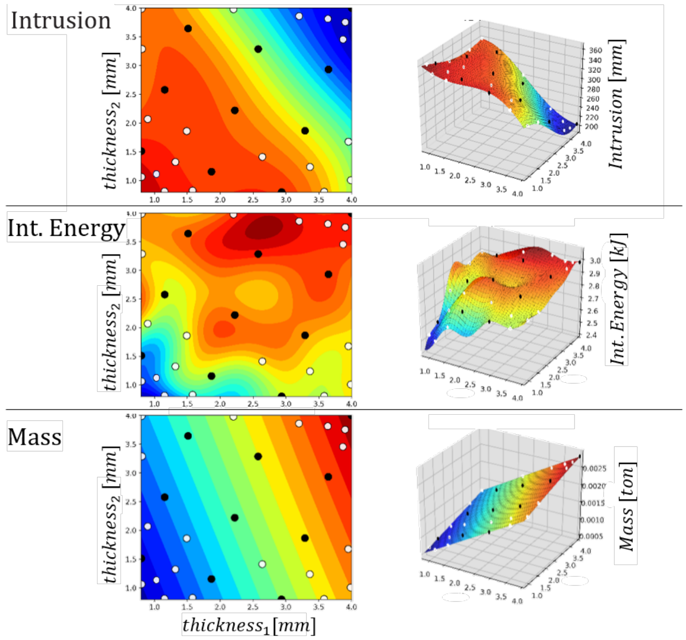


Figure 7: From top to bottom: the surrogate models for intrusion, internal energy, and mass of the crash-box after seven iterations. The black dots represent the points from the initial design of experiments, while the white dots are the points added adaptively by the MQCVVor method.

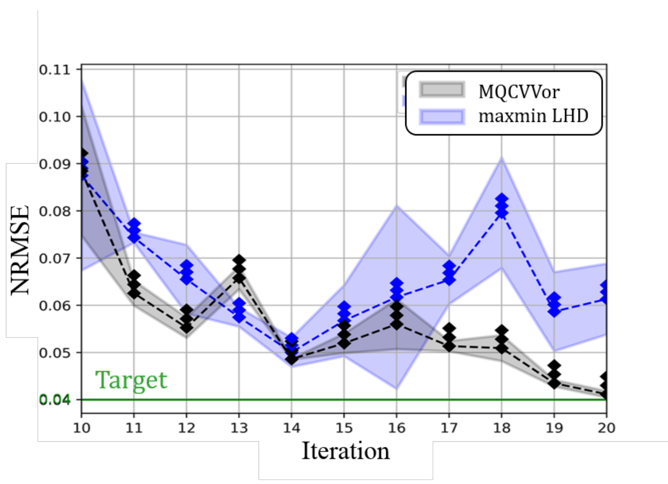


Figure 8: Comparison between the convergence curves of the MQCVVor method and maxmin-LHD. Three new samples are added at each iteration. The test repeated 25 times to generate 95 % confidence intervals

6 Conclusion and Outlook

In this study, we introduced the MQCVVor method as an extension of the existing CVVor approach, tailored to the specific needs of crashworthiness applications. By enabling parallel processing, MQCVVor efficiently enhances the adaptive sampling capacity, ensuring more robust and faster convergence compared to conventional methods. Our visual and quantitative analyses confirm that adaptive points are effectively placed within critical regions, improving model accuracy and robustness. Nevertheless, as the number of response functions increases, conflicting trade-offs have emerged, suggesting potential limitations in the effectiveness of the method for large multi-response systems. Future research will delve into these challenges and aim to identify the limits of the applicability of this approach. The potential need for a purely exploratory strategy will be explored in scenarios characterized by a larger number of response functions, where the complexity may diminish the benefits of the current adaptive strategy.

References

[1] Karel Crombecq, Dirk Gorissen, Dirk Deschrijver, and Tom Dhaene. A novel hybrid sequential design strategy for global surrogate modeling of computer experiments. *SIAM Journal on Scientific Computing*, 33(4):1948–1974, 2011. ISSN 1064-8275. doi: 10.1137/090761811.

[2] Jianguang Fang, Guangyong Sun, Na Qiu, Nam H. Kim, and Qing Li. On design optimization for structural crashworthiness and its state of the art. *Structural and Multidisciplinary Optimization*, 55(3):1091–1119, 2017. ISSN 1615-1488. doi: 10.1007/s00158-016-1579-y. URL https://link.springer.com/article/10.1007/s00158-016-1579-y.

[3] Jan N. Fuhg, Amélie Fau, and Udo Nackenhorst. State-of-the-art and comparative review of adaptive sampling methods for kriging. *Archives of Computational Methods in Engineering*, 28 (4):2689–2747, 2021. ISSN 1886-1784. doi: 10.1007/s11831-020-09474-6. URL https://link.springer.com/article/10.1007/s11831-020-09474-6#Fig5.

[4] Sushant S. Garud, Iftekhar A. Karimi, and Markus Kraft. Design of computer experiments: A review. *Computers & Chemical Engineering*, 106:71–95, 2017. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2017.05.010. URL https://www.sciencedirect.com/science/article/pii/S0098135417302090.

[5] Marcel Koch, Steffen Mattern, and Robert D. Bitsche. Facing future challenges in crash simulation engineering – model organization, quality and engineering. In *Proceedings of the 15th International LS-DYNA Users Conference*, Detroit, 2018.

[6] Haitao Liu, Shengli Xu, Xiaofang Wang, Shuhua Yang, and Jigang Meng. A multi-response adaptive sampling approach for global metamodeling. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 232(1):3–16, 2018. ISSN 0954-4062. doi: 10.1177/0954406216672250.

[7] Marcin Molga and Czesław Smutnicki. *Test functions for optimization needs*. 2005. URL https://marksmannet.com/robertmarks/classes/engr5358/papers/functions.pdf.

[8] Marco Münster, M. Schäffer, R. Sturm, and H. E. Friedrich. Methodological development from vehicle concept to modular body structure for the dlr ngc-urban modular vehicle. *16. Internationales Stuttgarter Symposium*, pp. 581–595, 2016. ISSN 2198-7440. doi: 10.1007/978-3-658-13255-2{\textunderscore}43. URL https://link.springer.com/chapter/10.1007/978-3-658-13255-2_43.

[9] Euro NCAP. Full width rigid barrier. https://www.euroncap.com/en/car-safety/the-ratings-explained/adult-occupant-protection/frontal-impact/full-width-rigid-barrier/, 2015. Accessed: December 2023.

[10] Felipe A. C. Viana, Gerhard Venter, and Vladimir Balabanov. An algorithm for fast optimal latin hypercube design of experiments. *International Journal for Numerical Methods in Engineering*, 82(2):135–156, 2010. ISSN 0029-5981. doi: 10.1002/nme.2750.

[11] Lin Wang, Qian Xiao, and Hongquan Xu. Optimal maximin l_1-distance latin hypercube designs based on good lattice point designs. *The Annals of Statistics*, 46(6B), 2018. ISSN 0090-5364. doi: 10.1214/17-AOS1674.

[12] Paul Willem Westermann and Ralph Evins. Adaptive sampling for building simulation surrogate model derivation using the lola-voronoi algorithm. In Vincenzo Corrado, Enrico Fabrizio, Andrea Gasparella, and Francesco Patuzzi (eds.), *Proceedings of Building Simulation 2019: 16th Conference of IBPSA*, Building Simulation Conference proceedings, pp. 1559–1563. IBPSA, 2020. doi: 10.26868/25222708.2019.211232.

[13] Shengli Xu, Haitao Liu, Xiaofang Wang, and Xiaomo Jiang. A robust error-pursuing sequential sampling approach for global metamodeling based on voronoi diagram and cross validation. *Journal of Mechanical Design*, 136(7), 2014. ISSN 1050-0472. doi: 10.1115/1.4027161.

Appendix

A Normalized generalized mean square cross-validation error

If no test points are available, the surrogate model accuracy can be assessed by the normalized generalized mean square cross validation error (NGMSE) as shown in Eq. (5) and Eq. (6):

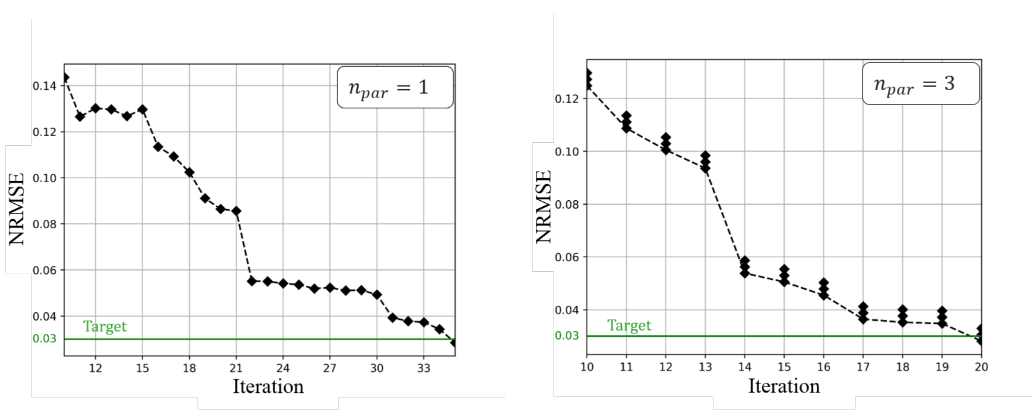
$$\mathrm{NGMSE}_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i^j)^2} \tag{5}$$

$$e_i^j = \frac{\left|f_j(\mathbf{x}_i) - \tilde{f}_{j,P\setminus p_i}(\mathbf{x}_i)\right|}{\max(\mathbf{y}_j) - \min(\mathbf{y}_j)} \tag{6}$$

where e_i^j is the normalized CV error of the i-th sampled point calculated by the j-th metamodel.

B Further results

In the multi-response system considered in Figure 5, with an NRSME set to 0.03, it is observed that by using n_{par} parallel simulations, the same objective can be achieved in 20 iterations (Figure 9a) instead of 35 (Figure 9b), resulting in a saving of about 43 % in computational time.



(a) One query point per iteration.

(b) Three query point per iteration

Figure 9: Convergence plots of the MQCVVor method for the multi-system shown in Figure 5 with varying query points per iteration.