Science Arts & Métiers (SAM)

is an open access repository that collects the work of Arts et Métiers Institute of Technology researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: https://sam.ensam.eu
Handle ID: .http://hdl.handle.net/10985/13417

To cite this version :

Reda EL BECHARI, Stéphane BRISSET, Jean-Claude MIPO, Stephane CLENET - Enhanced Meta-model Based Optimization under Constraints using Parallel Computations - IEEE Transactions on Magnetics - Vol. 54, n°3, p.1-4 - 2017

Any correspondence concerning this service should be sent to the repository Administrator : scienceouverte@ensam.eu

Arts et Métiers Institute of Technology

Enhanced Meta-model Based Optimization under Constraints using Parallel Computations

Reda El Bechari[1,2], Stéphane Brisset[1], Stéphane Clénet[1], and Jean-Claude Mipo[2]
[1]Université Lille, Centrale Lille, Arts et Métiers ParisTech, HEI, EA 2697 - L2EP - Laboratoire d'Electrotechnique et d'Electronique de Puissance, F-59000 Lille, France, reda.el-bechari@centralelille.fr
[2]Valeo Equipements Electriques Moteur, Créteil, France

Meta-models proved to be a very efficient strategy for optimization of expensive black-box models, e.g. Finite Element simulation for electromagnetic devices. It enables to reduce the computational burden for optimization purposes. Kriging is a popular method to build meta-model. Its statistical properties were firstly used in efficient global optimization for unconstrained problems. Afterwards many extensions were introduced in the literature to deal with constrained optimization. This paper presents a comparative study of some infill criteria for constraints handling and a new strategy for parallelization of the expensive computations of models. TEAM workshop problem 22 is taken as an electromagnetic test problem.

***Index Terms*—Constrained optimization, Expensive simulation, Kriging, Infill criteria, Parallelization strategy.**

I. Introduction

META-MODELS [1] are used in many fields, mainly to replace expensive black-box models [2] [3]. In an optimization problem the objective function and/or constraints are not always cheaply available data, thus these surrogate models aim to give a model able to approximate the expensive black-box models from a limited number of solutions. Optimization using kriging meta-models were first introduced in [4] to tackle unconstrained optimization. Its main advantage is the reduction of the number of calls to the expensive model. However, for problem with high number of parameters the number of evaluations arises exponentially (curse of dimensionality). Thus, the purpose of this paper is to compare methods to handle constraints and propose a new strategy for parallelization, which enables to run several evaluations at each iteration. A brief review of meta-model based optimization and infill criteria for constrained optimization is presented. Then, the parallelization strategy is presented and tested on an analytical model and the TEAM workshop benchmark problem 22.

II. Meta-model Based Optimization

An optimization problem can be formulated as follows

$$\begin{aligned} \min_{x} \quad & y(x) \\ s.t. \quad & g_i(x) \leq 0, \quad i = 1 \dots n_c \end{aligned} \tag{1}$$

where x are the design variables, $y(x)$ is the objective function and $g_i(x)$ are n_c constraints. In the case of electromagnetic modeling, finite element models are widely used, but the simulations can be expensive in terms of computational time. An alternative consists in replacing them by cheaper models. These meta-models are constructed with a relatively small size sample obtained by the Finite Element model at first, but refined with other samples found using an infill criterion. Furthermore, the numerical noise due to discretization present a problem for gradient based algorithm. The optimization in this case may become difficult or even impossible.

Meta-model based optimization flowchart is presented in Fig. 1. The first step aims to determine initial set of parameter values (initial design) using a design of experiments, e.g. Latin Hypercube Sampling (LHS). The full (expensive) model is solved for each set of parameters. Afterwards a meta-model is built based on the initial design and the output data. Kriging is well suited for building the meta-model due to its statistical properties. The most important part in the process, on which the paper will focus, is to find the infill point which improves the actual best solution and increases the meta-model precision. This point will be evaluated using the Finite Element model in the next iteration. In Fig. 1, the arrow that goes out and comes back into step "Find Infill point" means that a sub-optimization problem is solved to find infill points. Finally some stopping criteria are evaluated to terminate the optimization.

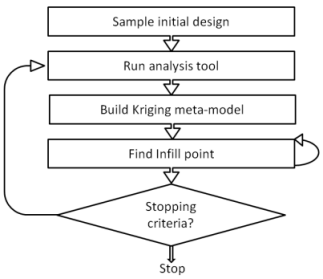

Fig. 1. Flowchart of Kriging meta-model based optimization.

A. Kriging

Kriging [5] is an interpolation method based on a regression term and a stochastic term. The stochastic term aims to eliminate the error due to regression and is constructed based on the location of the sampled points. Furthermore Kriging characterizes the variance, or the precision, of the prediction. Thus, Kriging considers the response a normally

distributed random variable with given expected value and standard deviation.

The Kriging method, also called Gaussian process regression, was originally developed by D. Krige a mining engineer in South Africa. Afterwards the method was introduced into the field of numerical design.

In [6] (Dace a Matlab Kriging Toolbox) an exhaustive presentation and the implementation of kriging predictor are detailed. The developed toolbox is used in the numerical evaluation to construct a response surface of the objective function and of the constraints from a set of sampled points.

III. Infill Criteria

A. Expected Improvement

The most used infill criterion to deal with unconstrained optimization is Expected Improvement (EI) criterion, presented in [4]. This criterion enables a trade-off between exploitation and exploration of the design space.

$$EI(x) = (y_{min} - \hat{y}(x))\Phi(u(x)) + \hat{s}_y(x)\phi(u(x)) \quad (2)$$

where $u(x) = \frac{y_{min}-\hat{y}(x)}{\hat{s}_y(x)}$, $\hat{y}(x)$ and $\hat{s}_y(x)$ are the expected value and the standard deviation of the kriging predictor for the objective function y and y_{min} is the smallest sampled value of y.

Maximizing $EI(x)$ leads to the point x^* with the highest probability of improvement, either by sampling toward the optimum or improving the approximation of the meta-model. These characteristics can be justified by the fact that the derivative of $EI(x)$ with respect to $\hat{y}(x)$ is negative, meaning that the smaller $\hat{y}(x)$ the higher $EI(x)$ (exploitation) and the derivative $EI(x)$ with respect to $\hat{s}_y(x)$ is positive, meaning that the bigger $\hat{s}_y(x)$ the higher $EI(x)$ (exploration).

There exist other infill criteria, that often reveal striking similarities. An exhaustive set of infill criteria was presented in [7] for unconstrained and constrained optimization.

B. Constraint handling

The probability of feasibility (PF) criterion is widely used for constrained optimization. It quantifies the probability that a constraint is satisfied

$$PF(x) = \Phi\Big(\frac{-\hat{g}(x)}{\hat{s}_g(x)}\Big) \quad (3)$$

where $\hat{g}(x)$ and $\hat{s}_g(x)$ are the expected value and the standard deviation of the kriging predictor for a constraint.

To sample point that improves the actual solution and respects constraints both EI and PF should be considered. Table I summarizes the main formulations that aim to satisfy these requirements.

The formulation (A) has a statistical derivation (assuming that the objective function and the constraint are statistically independent). It aims to sample points that maximize the expectation of improvement and fulfill constraints. However the product of EI and PF reveals high modality and algorithms maximizing this criterion often fail to find the global optimum.

Thus, the formulation (B) was proposed in [8] to consider the infill criterion as bi-objective to reduce the modality. From Pareto front, the point chosen is the one that maximizes the product of EI and PF.

TABLE I
Infill criteria for constrained optimization

Formulation	infill point determination
A	$\max_x EI(x).PF(x)$
B	$\max_x(EI(x), PF(x))$
C	$\max_x EI(x)$ $s.t. PF(x) \geq P_{tol}$

In the case of many constraints, both formulations consider the global PF as the product of the probability of feasibility of each constraint. However, this impacts the search close to the constraints boundary, so if the optimum lies on the constraint boundaries these infill criteria may fail to find it. This is due to the maximization of PF that leads to point inside the feasible region and less likely on the constraint boundaries.

The formulation (C) was proposed in [7] and considers the problem as a constrained one to reduce the modality of the infill criterion and to gain in precision of the solution in case of the optimum activates some constraints. $P_{tol} = 0.95$ was recommended but it has an effect on the precision, it fails to locate the points on the constraint boundaries. The authors opinion is that $P_{tol} = 0.5$ seems more reasonable because $PF = 0.5$ when $\hat{g}(x) = 0$. This formulation considers each constraint independently and calculates their respective probabilities of feasibility, ending up with the same number of constraints as the original problem.

IV. Parallelization strategy

In the sequential process, only one point is found by maximizing EI. The aim of the parallelization strategy is to find multiple promising points at each iteration to take advantage of clusters by running distributed expensive evaluations.

In [9] the straightforward extension of EI from sequential to parallel was presented. It computes an analytic expression for two points. For more than two points, the authors propose its estimation through Monte Carlo simulation.

In [10] a hybrid method was proposed. It adds artificially the point found at each iteration to a subset, evaluates it with the kriging predictor, and reconstruct the meta-model. Afterwards, the next point is found until reaching the number of required points. When the number of evaluated points increases, the construction of the meta-model become a time consuming step which penalize the whole process.

The proposed method is based mainly on the multimodal behavior of EI and searches the points that maximize EI and exclude its vicinity to find another point. The process is repeated until the number of excluded points equals the number of required parallel evaluations. The exclusion area is defined by the distance from that point.

In the infill criterion, the exclusion area is defined by the constraint $\big(\frac{x-x_i}{range(x)}\big)^2 \geq \frac{1}{4n^2}$, where $range(x)$ is the range of variable x and n is the number of parallel evaluations wanted, this constraint defines an area around the points already found in the current iteration to look for other promising regions in the design space.

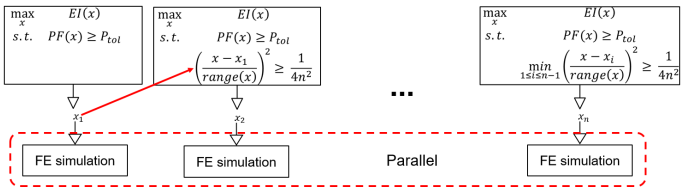

Fig. 2. Infill point determination

The advantages of this strategy over the two above-mentioned is that we have an analytical form for the infill criterion and the reconstruction of the meta-model is not needed multiple times at the same iteration. However, it presents a small drawback, at the last iterations of the optimization EI becomes 0 in a large part of the design space. The points found, at an iteration, are close to each other.

This drawback will be further investigated, an initial idea is to vary the size of the exclusion area, or using another infill criterion (e.g. $\hat{y}(x) - k\hat{s}_y(x)$ [7] [11] $k = 2.5, 3$) which keeps the multimodal behavior.

V. Enhancements

A. Pre-Optimization

Frequently, in the case of constrained optimization the initial sample (generated by LHS) has no feasible point. Thus, before starting the optimization. An initial infill criterion is used for this purpose, it aims to find feasible points.

$$\max_x \quad PF(x) \tag{4}$$

Maximizing PF means that we look for the point that is most likely feasible.

In the case were multiple feasible points are wanted, another infill criterion is proposed

$$\max_x \quad PF(x) \times D(x) \tag{5}$$

where the distance $D(x) = \min_{d \in X_{feas}} \|x - d\|$ and X_{feas} is the set of feasible sampled points found until the current iteration.

This infill criterion enables to sample points that have the highest probability of feasibility while being distant from previously sampled points.

B. Cheap Constraints handling

Cheap constraints are the constraints that can be evaluated rapidly without calling the FE model. Generally these of constraints are explicit functions of design variables. Sometimes these constraints express the feasibility of design, it means that if this constraint is not fulfilled the modeled phenomena have no physical meaning (e.g. overlapping constraint of SMES Device). These constraints are handled differently, they are embedded in the infill criterion directly.

Given $g(x) \leq 0$ a cheap constraint, the infill criterion problem is written

$$\begin{aligned} \max_x \quad & EI(x) \\ s.t. \quad & PF(x) \geq P_{tol} \\ & g(x) \leq 0 \end{aligned} \tag{6}$$

Adopting this procedure ensures that the point found by the infill criterion satisfy the constraint.

VI. Analytic test

The analytic problem from [12] is taken for comparison purpose.

$$\begin{aligned} \min_x \quad & -\frac{(x_1+x_2-10)^2}{30} - \frac{(x_1-x_2+10)^2}{120} \\ s.t. \quad & 1 - \frac{x_1^2 x_2}{5} \leq 0 \\ & 1 - \frac{(x_1+x_2-5)^2}{30} - \frac{(x_1-x_2-12)^2}{120} \leq 0 \\ & 1 - \frac{80}{x_1^2+8x_2+5} \leq 0 \\ & 0 \leq x_1, x_2 \leq 10 \end{aligned} \tag{7}$$

To compare formulations, 10 different initial designs of experiments are generated by the LHS are used. Furthermore, SQP algorithm with 10 start points uniformly distributed in the design space, is presented.

The averages of the results are shown in Table II. The metrics used for comparison are the convergence rate ($C.R.$) that is the percentage of the 10 tests that converge to the known solution for less than 1% of the range of variable. $dist$ is the normalized Euclidean distance to the known solution for the tests that converged, $evals$ is the average number of evaluations of the exact (supposed expensive) model, and $iters$ is average number of iterations, this metric is only considered for parallel.

TABLE II
Comparison of optimization results

Formulation	$C.R.$	$dist$	$evals$	$iters$
SQP	0.3	1.00e-6	31.5	31.5
A	0.8	3.26e-2	47.7	47.7
B	1.0	4.09e-2	55.3	55.3
C ($P_{tol} = 0.95$)	0.4	3.47e-2	32.5	32.5
C ($P_{tol} = 0.5$)	0.9	1.42e-4	37.4	37.4
Parallel	1.0	2.11e-5	52.5	17.5

The table shows that SQP and the formulation (C) with $P_{tol} = 0.95$ has the lowest convergence rate. The formulations (A) and (B) have good convergence rates but the number of evaluations is higher. Due to the multimodality of the problem, the algorithm often fails to find the global optimum. The formulation (C) with $P_{tol} = 0.5$ shows the best results among the sequential formulations.

For the parallelization strategy, three evaluations of the exact model at each iteration were done. The results show that the number of exact model evaluations has increased by 40%, however the number of iterations has decreased by 52%.

VII. Team Workshop problem 22

The SMES device [13] consists of two concentric superconducting coils fed with currents that flow in opposite directions. The inner coil is used for storing magnetic energy E, while the outer one has the role of diminishing the magnetic stray field B_{stray}.

The goal of the optimization problem is to find the design configurations (8 parameters) that give a specified value of

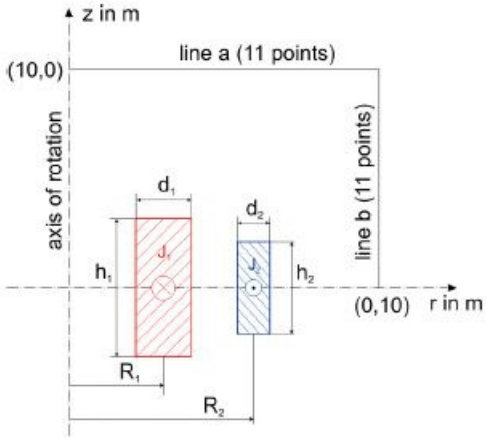

Fig. 3. SMES Device [13]

TABLE III
OPTIMIZATION RESULTS

	Reference	C ($P_{tol} = 0.5$)	Parallel
$Enegy(MJ)$	178.73	180.14	180.0
$B_{stray}(\mu T)$	14.56	8.81	8.93
OF	0.0123	0.0029	0.0020
$dist$	0	2.85e-3	3.41e-3
$evals$	-	3403	4128
$iters$	-	3403	516
$time(s)$	-	10254	6541

stored magnetic energy and a minimal magnetic stray field. Mathematically, this is formulated as

$$\begin{aligned} \min_{x} \quad & OF(x) = \frac{B^2_{stray}(x)}{B^2_{norm}} + \frac{|E(x)-E_{ref}|}{E_{ref}} \\ s.t. \quad & |\mathbf{J}| + 6.4|\mathbf{B}| - 54 \leq 0 \\ & R_1 - R_2 + \tfrac{1}{2}(d_1 + d_2) < 0 \end{aligned} \tag{8}$$

where $E_{ref} = 180MJ$, $B_{norm} = 200\mu T$ and x are the design variables $x = (R_1, R_2, h_1/2, h_2/2, d_1, d_2, J_1, J_2)$.

The FE simulation is considered as a Black-Box model with inputs and outputs and response surfaces using Kriging for the objective function and the constraints are constructed.

Thus, three response surfaces are constructed, one for the objective function, and two for the quench constraint. These last two response surfaces represent each one the quench condition in a coil (coil 1 and coil 2). The third constraint is handled as a cheap constraint because it depends directly on the design variables.

The results of optimization are summarized in Table III. The comparison is done between the reference [13] and two formulation of infill criterion, sequential formulation C with $P_{tol} = 0.5$ and the parallel formulation (also with $P_{tol} = 0.5$) to find $n = 8$ points at each iteration.

The number of evaluations ($evals$) has increased by 21% on the other hand the overall time for the optimization has decreased 36%. These performances are due to two reasons. Firstly, the parallelization of evaluations. Secondly, the number of reconstructions of the kriging meta-model (3403 times for sequential, and 516 for parallel).

VIII. CONCLUSION

In this communication, we have developed a strategy of optimization based on kriging meta-model and the parallelization of the computations of the full model. The results obtained on an analytical example were promising.

This strategy was assessed on the TEAM Workshop problem 22 with 8 variables, the results show that the proposed parallelization strategy enables to reduce significantly the computational time (a speedup of 1.56). However, the expected speedup was to be 8 (8 points in parallel), this will be further investigated, to enhance the performances.

REFERENCES

[1] R. Jin, W. Chen, and T. W. Simpson, "Comparative Studies of Metamodeling Techniques Under Multiple Modeling Criteria," *Structural and Multidisciplinary Optimization*, vol. 23, no. 1, pp. 1–13, 2001.

[2] M. Li, F. Gabriel, M. Alkadri, and D. A. Lowther, "Kriging-assisted multi-objective design of permanent magnet motor for position sensorless control," *IEEE Transactions on Magnetics*, vol. 52, pp. 1–4, March 2016.

[3] J. P. C. Kleijnen, W. Van Beers, and I. Van Nieuwenhuyse, "Expected improvement in efficient global optimization through bootstrapped kriging," *Journal of Global Optimization*, vol. 54, no. 1, pp. 59–73, 2012.

[4] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, vol. 13, pp. 455–492, 1998.

[5] J. D. Martin and T. W. Simpson, "Use of Kriging Models to Approximate Deterministic Computer Models," *AIAA Journal*, vol. 43, no. 4, pp. 853–863, 2005.

[6] S. N. Lophaven, J. Søndergaard, and H. B. Nielsen, "DACE A Matlab Kriging Toolbox," *IMM Informatiocs and Mathematical Modelling*, pp. 1–28, 2002.

[7] M. J. Sasena, P. Papalambros, and P. Goovaerts, "Exploration of Metamodeling Sampling Criteria for Constrained Global Optimization," *Engineering Optimization*, vol. 34, no. 3, pp. 263–278, 2002.

[8] C. Durantin, J. Marzat, and M. Balesdent, "Analysis of multi-objective kriging-based methods for constrained global optimization," *Computational Optimization and Applications*, vol. 63, no. 3, pp. 903–926, 2016.

[9] D. Ginsbourger, R. L. Riche, and L. Carraro, "Kriging Is Well-Suited to Parallelize Optimization," *Computational Intelligence in Expensive Optimization Problems*, no. 2, pp. 131–162, 2010.

[10] M. Schonlau, "Computer experiments and global optimization," *Computer*, p. 130, 1997.

[11] D. D. Cox and S. John, "A statistical method for global optimization," in *Systems, Man and Cybernetics, 1992., IEEE International Conference on*, vol. 2, pp. 1241–1246, 1992.

[12] D. W. Kim, N. S. Choi, K. K. Choi, and D. H. Kim, "A single-loop strategy for efficient reliability-based electromagnetic design optimization," *IEEE Transactions on Magnetics*, vol. 51, p. 3, March 2015.

[13] P. Alotto, U. Baumgartner, F. Freschi, M. Jaindl, A. Kostinger, C. Magele, and W. Repetto, "Smes optimization benchmark: Team workshop problem 22," *Graz, Austria*, 2008.