# ON THE USE OF GRADIENTS IN KRIGING SURROGATE MODELS

Selvakumar Ulaganathan
Ivo Couckuyt
Tom Dhaene
Eric Laermans

Joris Degroote

Ghent University - iMINDS
Department of Information Technology (INTEC)
Gaston Crommenlaan 8
Ghent, 9050, BELGIUM

Ghent University
Department of Flow, Heat and Combustion Mechanics
Sint-Pietersnieuwstraat 41
Ghent, 9000, BELGIUM

## ABSTRACT

The use of Kriging surrogate models has become popular in approximating computation-intensive deterministic computer models. In this work, the effect of enhancing Kriging surrogate models with a (partial) set of gradients is investigated. While, intuitively, gradient information is useful to enhance prediction accuracy, another motivation behind this work is to see whether it is worth including the gradients versus their computation time. Test results of two analytical functions and a fluid-structure interaction (FSI) problem from bio-mechanics show that this approach, known as Gradient Enhanced Kriging (GEK), can significantly enhance the accuracy of Kriging models even when the gradient data is only partially available.

## 1 INTRODUCTION

Despite a substantial advancement in computing power, the analysis and optimization of simulation codes for many problems in science and engineering require a significant investment of computational time and resources. This turned many researchers towards the alternative of using cheap approximations or surrogate models instead. In this respect, various attempts to enhance the accuracy of surrogate models with secondary information such as gradients, Hessian data etc. have been carried out in the recent years. For example, finite difference approximations of the gradients are used to improve the theoretical error estimates of the surrogate model in (Hinze and Volkwein 2005). (Ito and Ravindran 1998) incorporated the gradients directly in a finite-element-like framework, referred to as Hermite approach. These approaches share the idea of exploiting gradient data within the context of projection schemes and are statistically biased at the sample points. This problem is alleviated in (Zimmermann 2013) by directly introducing the gradients to enhance the proper orthogonal decomposition (POD) based interpolation, referred to as gradient-enhanced POD (GEPOD) approach.

Kriging surrogate models offer an efficient way to approximate deterministic and computation-intensive simulation codes (Sacks et al. 1989). Kriging surrogate models can take advantage of additional information, such as gradient data, Hessian data, multi-fidelity function data, etc. in order to enhance their accuracy. For example, (Yamazaki, Rumpfkeil, and Mavriplis 2010) incorporated first-order and second-order derivative data in Kriging surrogate models which resulted in models with improved accuracy over models based on function data only. The context of this work is providing a guideline to designers and engineers whether it is worthwhile to include gradient information based on several factors. While gradient data often improves the accuracy, this improvement may not be worth the additional computation or fitting cost.

Gradient enhancement in Kriging surrogate models is first introduced by (Morris, Mitchell, and Ylvisaker 1993). This approach is denoted as direct Gradient Enhanced Kriging (GEK) throughout this work. An alternative formulation of GEK, called indirect GEK, is further introduced by (Chung and Alonso 2002).

Indirect GEK is built from the same mathematical formulation of Kriging, but augments the training data with additional function data estimated from the gradient data. The authors argued that both direct GEK and indirect GEK are almost identical in performance. An integrated mean squared error-based GEK (IMSE-GEK) is proposed by (Liu 2003), but its performance is lower than indirect GEK.

The modeling efficiency of Kriging is largely determined by the ability of its correlation function to capture the actual local behavior of the function to be modeled. Moreover, GEK has an extra requirement that the correlation function must be twice differentiable in order to estimate the correlation between gradient data. An elaborate discussion on various differentiable correlation functions is given by (Rasmussen and Williams 2006).

In this paper, we are primarily interested in investigating the potential of including a (partial) set of gradients in Kriging for providing more accurate surrogate models with as few runs of the expensive simulation code as possible. To that end, analytical expressions for the derivatives of the Matérn $\frac{5}{2}$ correlation function with respect to design variables are derived. The GEK methodology is applied to two analytical benchmark functions and to one real-life problem from bio-mechanics. Based on the results of this investigation, an insight on choosing GEK over Ordinary Kriging (OK) with respect to surrogate model accuracy, surrogate model fitting cost and computational cost of estimating derivatives is given. In addition, a guideline is suggested to achieve a reasonably good trade-off between surrogate model fitting cost and surrogate model accuracy of GEK models.

The reminder of the paper is organized as follows: GEK is discussed in Section 2 after a brief overview of Kriging. The correlation function being essential to obtain accurate surrogate models, is also discussed in Section 2. The benchmark and the simulation-based examples are given in Section 3. Section 4 discusses the results of the benchmark and the simulation-based examples. Finally, Section 5 summarizes our conclusions.

## 2 GRADIENT ENHANCED KRIGING

The Kriging estimate $\hat{y}(\boldsymbol{x}^*)$ at a prediction point $\boldsymbol{x}^*$ is expressed as a summation of a constant trend function $\hat{\mu}$ and a realization of a stationary Gaussian random process:

$$\hat{y}(\boldsymbol{x}^*) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1}(\boldsymbol{y} - \boldsymbol{1}\hat{\mu}), \tag{1}$$

where $\boldsymbol{\psi}$ contains the correlation between the sample data and the prediction point $\boldsymbol{x}^*$, $\boldsymbol{y}$ contains the function values of the sample data and $\boldsymbol{\Psi}$ is the correlation matrix which contains the correlation between the sample data points. In the case of GEK, Equation (1) becomes,

$$\hat{y}(\boldsymbol{x}^*) = \hat{\mu} + \dot{\boldsymbol{\psi}}^T \dot{\boldsymbol{\Psi}}^{-1}(\dot{\boldsymbol{y}} - \boldsymbol{f}\hat{\mu}), \tag{2}$$

where

$$\dot{\boldsymbol{\Psi}} = \left\{ \begin{array}{ccccc} \boldsymbol{\Psi} & \frac{\partial \boldsymbol{\Psi}}{\partial x_1^{(i)}} & \cdots & \frac{\partial \boldsymbol{\Psi}}{\partial x_v^{(i)}} & \cdots & \frac{\partial \boldsymbol{\Psi}}{\partial x_k^{(i)}} \\ \frac{\partial \boldsymbol{\Psi}}{\partial x_1^{(j)}} & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(i)} \partial x_1^{(j)}} & \cdots & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(i)} \partial x_v^{(j)}} & \cdots & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(i)} \partial x_k^{(j)}} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \frac{\partial \boldsymbol{\Psi}}{\partial x_u^{(j)}} & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(j)} \partial x_u^{(i)}} & \cdots & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_u^{(i)} \partial x_v^{(j)}} & \cdots & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_u^{(i)} \partial x_k^{(j)}} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{\Psi}}{\partial x_k^{(j)}} & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(j)} \partial x_k^{(i)}} & \cdots & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_v^{(j)} \partial x_k^{(i)}} & \cdots & \frac{\partial^2 \boldsymbol{\Psi}}{\partial x_k^{(i)} \partial x_k^{(j)}} \end{array} \right\}, \tag{3}$$

$$\dot{\boldsymbol{\psi}} = \left( \boldsymbol{\psi}, \frac{\partial \boldsymbol{\psi}}{\partial x_1}, ..., \frac{\partial \boldsymbol{\psi}}{\partial x_k} \right)^T, \tag{4}$$

$$\dot{\boldsymbol{y}} = \left(\boldsymbol{y}, \frac{\partial \boldsymbol{y}}{\partial x_1}, ..., \frac{\partial \boldsymbol{y}}{\partial x_k}\right)^T, \tag{5}$$

$$\boldsymbol{f} = \left(1_1, ...1_{n_s}, 0_{n_s+1}, ..., 0_{(k+1)n_s}\right)^T, \tag{6}$$

where $k$ is the dimensionality, $n_s$ is the number of training samples, $\dot{\boldsymbol{\Psi}}$ is a $(k+1)n \times (k+1)n$ symmetric block matrix and contains the correlations of function and gradient data between the sample data points and the correlation vector $\dot{\boldsymbol{\psi}}$ contains the correlation of both function values and gradients between the sample data and $\boldsymbol{x}^*$. The vector $\dot{\boldsymbol{y}}$ contains both the function values and gradients of the sample data. The notations $\partial \boldsymbol{\Psi}/\partial x_u^{(j)}$ and $\partial^2 \boldsymbol{\Psi}/\partial x_u^{(i)}\partial x_v^{(j)}$ denote the correlation between function data and $u^{th}$ dimension gradients and correlation between $u^{th}$ dimension gradients and $v^{th}$ dimension gradients, respectively. The direction of differentiation is denoted by $i$ and $j$ with $x^{(i)}$ and $x^{(j)}$ denoting two different samples.

The Matérn $\frac{5}{2}$ correlation function, which is widely used in the machine learning context, is used in this work. The mathematical expressions for the correlation, gradient of the correlation and the Hessian of the correlation with respect to $\boldsymbol{x}$ for the Matérn $\frac{5}{2}$ correlation function are given in Appendix A. The constant trend function for GEK, $\hat{\mu}$, is calculated via the general least squares method as $\hat{\mu} = (\boldsymbol{f}^T\dot{\boldsymbol{\Psi}}^{-1}\boldsymbol{f})^{-1}\boldsymbol{f}^T\dot{\boldsymbol{\Psi}}^{-1}\dot{\boldsymbol{y}}$. The hyperparameters ($\theta_m, m = 1, ..., k$) of the GEK model are determined by maximizing the concentrated likelihood function.

## 3   TEST PROBLEMS

Two analytical examples and one simulation example are used as test problems. The surrogate model accuracy is estimated with two different error measures: A validation data set of $n_p = 500$ uniformly distributed pseudorandom points and $k$-fold Cross-Validation (CV) (Meckesheimer et al. 2002), both using the Normalized Root Mean Square Error (NRMSE) which can be expressed as,

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n_p}\left(y_t^i - \hat{y}^i\right)^2}{n_p}}}{max(\boldsymbol{y}_t) - min(\boldsymbol{y}_t)}, \tag{7}$$

where $\boldsymbol{y}_t$ is the vector of true function values and $\hat{\boldsymbol{y}}$ is the vector of predicted function values.

### 3.1 Analytical Examples

Two widely used benchmark functions, summarized in Table 1, are employed as test functions.

Table 1: Analytical examples

| Benchmark Functions | Number of design variables ($k$) | Properties |
|---|---|---|
| Peaks | 2 | Multi modal |
| Sphere | 5 | Unimodal and convex |

### 3.2 Simulation Example

A numerical simulator (Degroote et al. 2013) that determines the difference between a given wall displacement and a calculated wall displacement for a given stiffness distribution along the length of an artery is used as the simulation example. The numerical simulator uses a simplified fluid-structure interaction model to identify the stiffness distribution along the length of an artery. The fluid-structure interaction model is one-dimensional in an axisymmetric $(r, \phi, z)$ coordinate system, as depicted in Figure 1. It consists of

$k-1$ elastic segments ($k$ represents the number of design variables), each with its own stiffness. Inside the artery, there is an incompressible blood flow. Furthermore, the interaction between this blood flow and the elastic wall is taken into account. For more details the reader is referred to (Degroote et al. 2013).
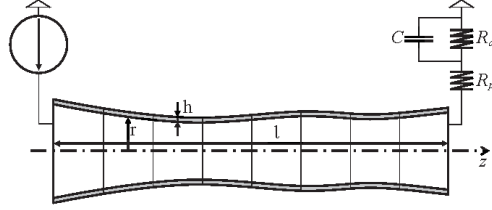


Figure 1: The one-dimensional and axisymmetric model for blood flow in an artery with the prescribed velocity at the inlet (left) and the Windkessel model at the outlet (right). The segments, radius $r$, wall thickness $h$ and length $\ell$ are indicated

## 4    RESULTS AND DISCUSSION

### 4.1 Analytical Examples

Figure 2 shows the evolution of NRMSE as a function of $n_s$ for the benchmark test functions utilizing the complete gradient data. As expected, the gradient enhancement significantly reduces the number of training samples required for GEK models to achieve the similar accuracy level of OK models. GEK models achieve a tentative 40% and 80% reduction in the number of training samples ($n_s$) to reach the equivalent accuracy level of OK models for the Peaks and Sphere functions, respectively. The performance of GEK models over OK models is more pronounced as the number of training sample points and the dimensionality ($k$) of the problem increases. The cross-validation error measure also compares in favor of GEK models, see Table 2.



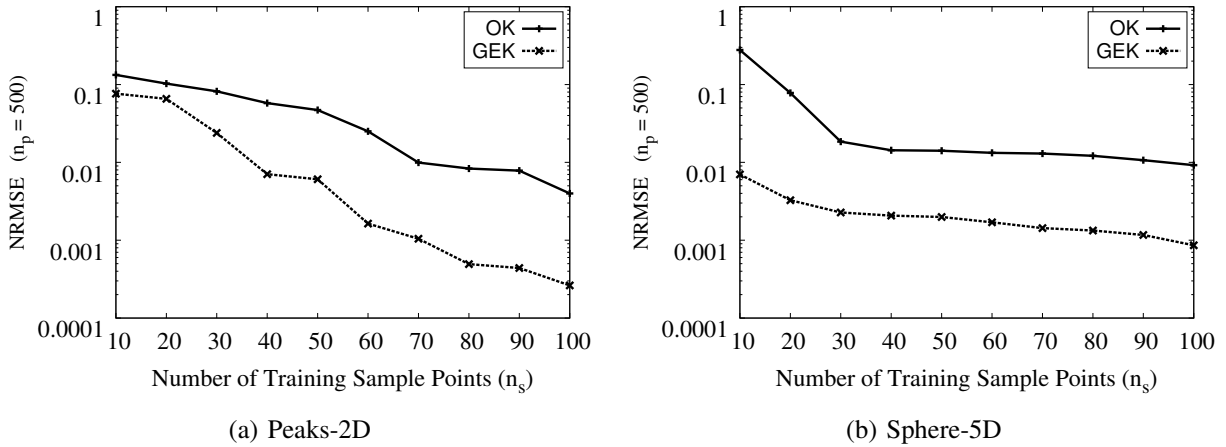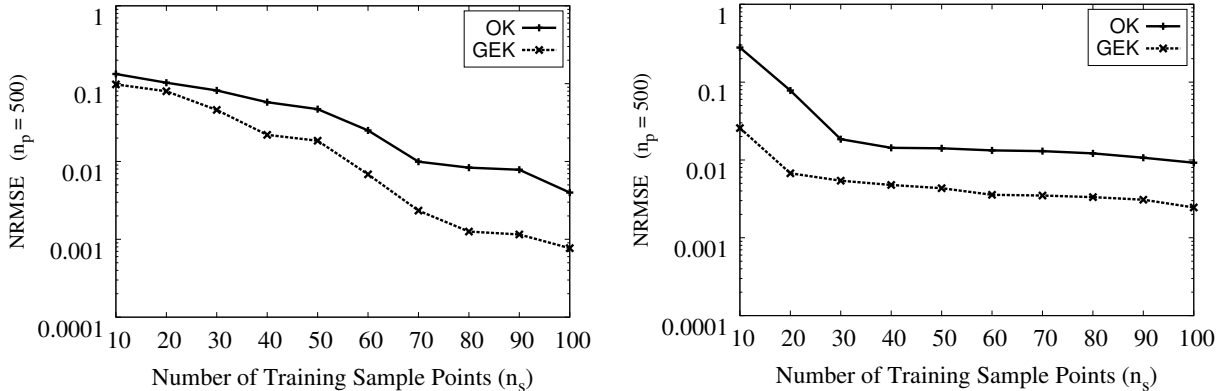(a) Peaks-2D                    (b) Sphere-5D

Figure 2: Evolution of NRMSE as a function of the number of training samples. GEK incorporates $n_s$ gradients in all the dimensions

The accuracy of GEK models is also assessed by using only a partial set of gradients in the construction of the GEK models. GEK results in more accurate surrogate models than OK even when the gradient data in some of the dimensions is completely left out, see Figure 3. Moreover, incorporating only a partial set of gradients while constructing the GEK models reduces the size of $\dot{\Psi}$ to $(n_s + (n_s \times k')) \times (n_s + (n_s \times k'))$ with $k'$ being the number of dimensions in which the partial set of gradients is incorporated. Although the

Table 2: Efficiency of the GEK models based on the cross-validation error (CVE) measure. The cross-validation is carried out with 10 data folds. Each data fold contains 11 samples

| Benchmark Functions | OK CVE | GEK CVE | % of Improvement |
|---|---|---|---|
| Peaks-2D | 0.0166 | 0.0036 | 78% |
| Sphere-5D | 0.0111 | 0.0020 | 82% |

dynamics of the left out gradients influence the surrogate model accuracy, this feature offers the possibility of reducing the computational cost associated with the inversion of $\boldsymbol{\Psi}$ by leaving out the least influencing gradients. This, in turn, can significantly reduce the surrogate model fitting cost. The least influencing gradients, although, in general, subject to the problem at hand, are usually the gradients which correspond to the lower hyperparameter ($\theta$) valued dimensions, see Table 3. If the value of a hyperparameter is low as compared to the values of all the other hyperparameters, then the function will be smoother in the associated dimension as compared to all the other dimensions (i.e., higher correlation among the sample points). Tables 3 and 4 give the individual and the cumulative accuracy improvement achieved by the GEK models when gradients in more than one direction are successively incorporated during surrogate model construction. From the benchmark results, only incorporating a partial set of gradients in $k/2$ dimensions is observed to be a good trade-off between the GEK model fitting cost and model accuracy as the accuracy of GEK models with $k/2$ partial set of gradients is relatively close to that of GEK models with a complete $k$ set of gradients, see Figures 2 and 3 and Tables 3 and 4.



(a) Gradients in the second dimension only. (Peaks-2D)

(b) Gradients in the first and second dimensions only. (Sphere-5D)

Figure 3: Effect of a partial set of gradients. Evolution of NRMSE as a function of the number of training samples

## 4.2 Simulation Examples

Similar to the analytical benchmark examples, using GEK leads to a significant 60% to 70% reduction in number of training sample points to reach the equivalent accuracy level of OK models, see Figure 4. Moreover, for the same number of training samples, adding gradient information leads to a more than 50% accuracy improvement in GEK models over OK models, see Table 5. However, the surrogate model fitting cost grows significantly in GEK as the incorporation of additional gradient information leads to a substantial growth in the size of $\dot{\boldsymbol{\Psi}}$, see Figure 5. The size of $\dot{\boldsymbol{\Psi}}$ is directly related to the cost of Cholesky decomposition of $\dot{\boldsymbol{\Psi}}$ which is carried out multiple times during the hyperparameters estimation. Hence for

Table 3: Individual and cumulative accuracy improvement of the GEK models over the OK models when gradients in more than one direction are successively incorporated. The cross-validation is carried out with 10 data folds. Each data fold contains 11 samples. The hyperparameter values are obtained from an OK model with $n_s = 100$ (Peaks-2D)

| Dimension | value of $\theta$ | % of individual Improvement (over OK) | % of cumulative Improvement (over OK) |
|---|---|---|---|
| 2 | 0.56 | 58% | 58% |
| 1 | 0.50 | 45% | 78% |

Table 4: Individual and cumulative accuracy improvement of the GEK models over the OK models when gradients in more than one direction are successively incorporated. The cross-validation is carried out with 10 data folds. Each data fold contains 11 samples. The hyperparameter values are obtained from an OK model with $n_s = 100$ (Sphere-5D)

| Dimension | value of $\theta$ | % of individual Improvement (over OK) | % of cumulative Improvement (over OK) |
|---|---|---|---|
| 1 | 0.01 | 42% | 42% |
| 2 | 0.01 | 34% | 55% |
| 3 | 0.01 | 41% | 67% |
| 4 | 0.01 | 38% | 75% |
| 5 | 0.01 | 38% | 82% |

a fair comparison regarding the fitting cost, the correlation matrix of OK models is augmented with more function values so that its size becomes equal to that of GEK models. This way, OK and GEK models can be compared subject to equal surrogate model fitting cost. Results from the simulation examples show that the OK models with augmented correlation matrix outperforms GEK models, see Figure 6. Unsurprisingly, this confirms the intuitive fact that a function value is more informative than a gradient value. However, the computational cost spent on acquiring the additional function values required to augment the correlation matrix of OK models is significantly higher than that of estimating function and gradient values for GEK models of equivalent correlation matrix size, see Figure 7.

Table 5: Efficiency of the GEK models based on the cross-validation error (CVE) measure. The cross-validation is carried out with 10 data folds. Each data fold contains 11 samples

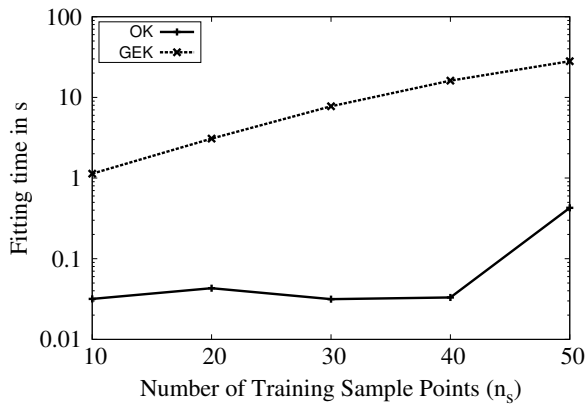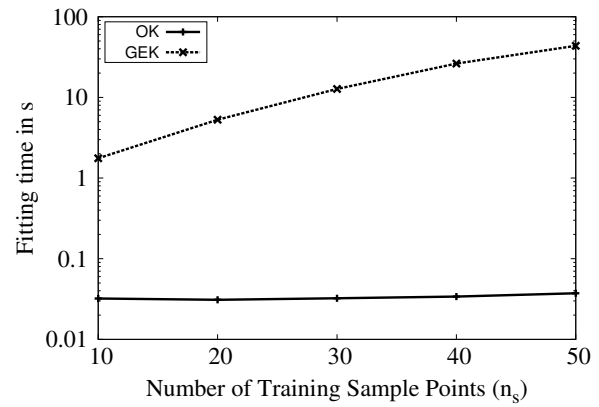| FSI Functions | OK CVE | GEK CVE | % of Improvement |
|---|---|---|---|
| 7D | 0.1849 | 0.0684 | 63% |
| 8D | 0.1591 | 0.0775 | 51% |
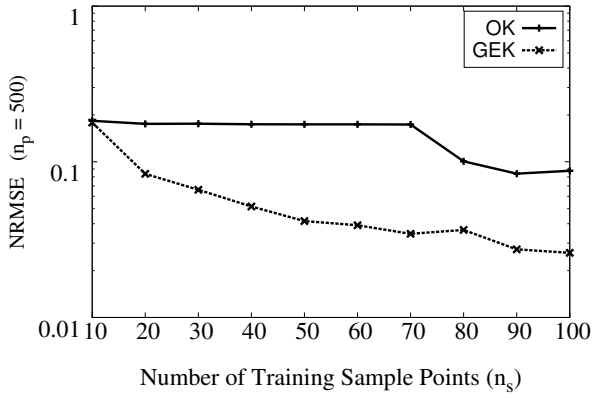
(a) FSI function-7D
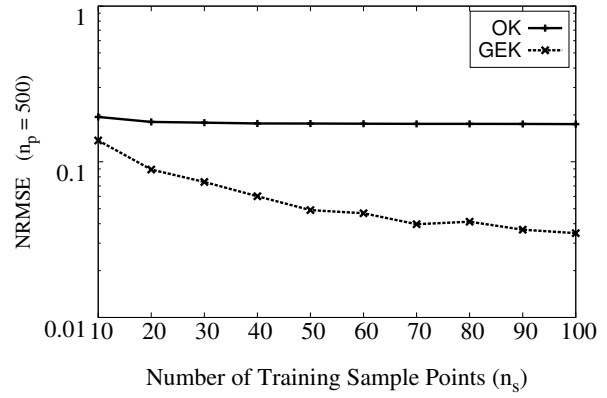
(b) FSI function-8D

Figure 4: Evolution of NRMSE as a function of number of samples



(a) FSI function-7D

(b) FSI function-8D

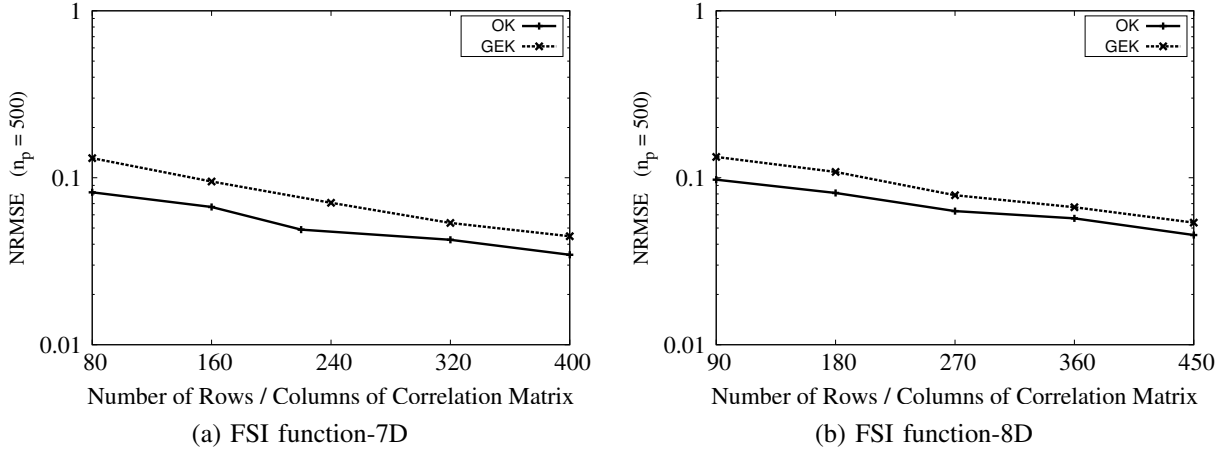Figure 5: Surrogate model fitting cost of OK and GEK models

(a) FSI function-7D

(b) FSI function-8D

Figure 6: Efficiency of OK models with $\mathbf{\Psi}$ being equal in size with $\dot{\mathbf{\Psi}}$ of GEK models (i.e., equal surrogate modeling cost). OK models with $n_s + k \times n_s$ function values. GEK model with $n_s$ function values and $k \times n_s$ gradient values
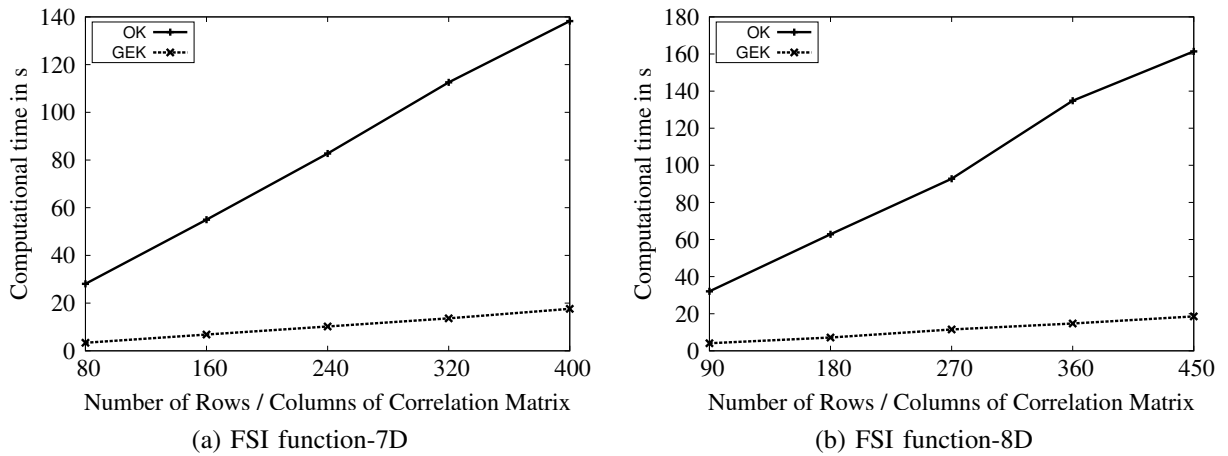


(a) FSI function-7D

(b) FSI function-8D

Figure 7: Computational time spent of acquiring $n_s + (k \times n_s)$ function values for OK and $n_s$ function and $k \times n_s$ gradient values for GEK

## 5 CONCLUSIONS

We have validated that gradient enhancement in Kriging can significantly reduce the computational cost of acquiring training data with substantially improved surrogate model accuracy over the Ordinary Kriging (OK) models. We demonstrate that using only a partial set of gradients already results in more efficient surrogate models in terms of prediction capability than the OK models, while potentially reducing the computational complexity of constructing the GEK models significantly. As a result, a guideline is proposed to achieve a reasonably good trade-off between surrogate model fitting cost and surrogate model accuracy when the surrogate model fitting cost is comparable to the computational cost of acquiring training data. Interestingly, OK outperforms GEK when the training data of OK models are augmented with additional function data in order to equal the surrogate model fitting cost of GEK models; but, at the computational cost of acquiring additional function data which is significantly higher than that of acquiring gradient data for GEK. Moreover, the gradient enhancement can also be successfully carried out during the modeling of time-varying data if the gradient data is available at various time steps (e.g., dynamic Kriging).

## ACKNOWLEDGMENTS

## A CORRELATION FUNCTION DERIVATIONS

The Matérn $\frac{5}{2}$ correlation function is expressed as,

$$\psi_{v=\frac{5}{2}}(d) = (1 + \sqrt{5}a + \frac{5a^2}{3})exp\left(-\sqrt{5}a\right), \tag{8}$$

where $a = \sqrt{\sum_{m=1}^{k} \theta_m d_m^2}$ and $d = |x_m^i - x_m^j|$.

The gradient of the Matérn $\frac{5}{2}$ correlation function with respect to $\boldsymbol{x}$ (i.e., cross-correlation) is derived as,

$$\frac{\partial \boldsymbol{\Psi}^{(i,j)}}{\partial x^{(j)}} = \frac{5\theta d(\sqrt{5}a+1)exp\left(-\sqrt{5}a\right)}{3}. \tag{9}$$

Finally, the Hessian of the Matérn $\frac{5}{2}$ correlation function with respect to $\boldsymbol{x}$ (i.e., cross-correlation) is derived as,

$$\frac{\partial^2 \boldsymbol{\Psi}^{(i,j)}}{\partial x_u^{(i)} \partial x_v^{(j)}} = \begin{cases} \frac{-25\theta_u\theta_v d_u d_v exp\left(-\sqrt{5}a\right)}{3} & \text{if } u \neq v \\ \left[\frac{-25\theta^2 d^2 + 5\theta(\sqrt{5}a+1)}{3}\right] exp\left(-\sqrt{5}a\right) & \text{if } u = v. \end{cases} \tag{10}$$

## References

Chung, H.-S., and J. J. Alonso. 2002. "Using Gradients to Construct Cokriging Approximation Models for High-Dimensional Design Optimization Problems". In *Problems, 40th AIAA Aerospace Sciences Meeting and Exhibit, AIAA*, 2002–0317. Reno, NV.

Degroote, J., M. Hojjat, E. Stavropoulou, R. Wüchner, and K.-U. Bletzinger. 2013. "Partitioned solution of an unsteady adjoint for strongly coupled fluid-structure interactions and application to parameter identification of a one-dimensional problem". *Structural and Multidisciplinary Optimization* 47 (1): 77–94.

Hinze, M., and S. Volkwein. 2005. "Proper Orthogonal Decomposition Surrogate Models for Nonlinear Dynamical Systems: Error Estimates and Suboptimal Control". In *Dimension Reduction of Large-Scale Systems*, Volume 45 of *Lecture Notes in Computational Science and Engineering*, 261–306. Springer Berlin Heidelberg.

Ito, K., and S. Ravindran. 1998. "A reduced-order method for simulation and control of fluid flows". *Journal of Computational Physics* 143 (2): 403–425.

Liu, W. 2003. "Development of Gradient-Enhanced Kriging approximations for multidisciplinary design optimisation". Ph. D. thesis, Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, Indiana. http://etd.nd.edu/ETD-db/theses/available/etd-07012003-215221/[Accessed July 15,2014].

Meckesheimer, M., A. J. Booker, R. R. Barton, and T. W. Simpson. 2002. "Computationally Inexpensive Metamodel Assessment Strategies". *AIAA* 40 (10): 2053–2060.

Morris, M. D., T. J. Mitchell, and D. Ylvisaker. 1993. "Bayesian Design and Analysis of Computer Experiments: Use of Gradients in Surface Prediction". *Technometrics* 35 (3): 243–255.

Rasmussen, C. E., and C. K. I. Williams. 2006. *Gaussian processes for machine learning*. Cambridge, MA, USA: The MIT Press.

Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn. 1989. "Design and Analysis of Computer Experiments". *Statistical Science* 4 (4): 409–423.

Yamazaki, W., M. P. Rumpfkeil, and D. J. Mavriplis. 2010. "Design optimization utilizing Gradient/Hessian enhanced surrogate model". In *28th AIAA Applied Aerodynamics Conference*. Chicago, Illinois, USA.

Zimmermann, R. 2013. "Gradient-enhanced surrogate modeling based on proper orthogonal decomposition". *Journal of Computational and Applied Mathematics* 237 (1): 403–418.

## AUTHOR BIOGRAPHIES

**SELVAKUMAR ULAGANATHAN** is a PhD student in the Department of Information Technology (INTEC), Ghent University, Belgium. He holds a Bachelor of Engineering (BE) degree in Aeronautical Engineering from Anna University Chennai, India and a M.Sc. degree in Computational Fluid Dynamics from Cranfield University, UK. His research deals with surrogate modeling, machine learning and optimization problems. His email address is selvakumar.ulaganathan@ugent.be.

**IVO COUCKUYT** is a postdoctoral fellow in the INTEC Broadband Communication Networks (IBCN) research group of the Department of Information Technology (INTEC), Ghent University, Belgium. His research activities include machine learning, bioinformatics, optimal design, numerical analysis techniques, surrogate modeling, surrogate-based optimization and inverse modeling of time-consuming problems. He received his M.Sc. degree in Computer Science from University of Antwerp (UA), Belgium in 2007. He received his Ph.D. degree in Engineering from Ghent University, Belgium, in 2013. His email address is ivo.couckuyt@ugent.be.

**TOM DHAENE** is a professor in Distributed Scientific Computing within the department of information technology (INTEC-IBCN) of the faculty of Engineering at Ghent University, Belgium. In 1993, after his Ph.D. on electromagnetic modeling at INTEC, Ghent University, he joined the EDA start-up company Aphabit, later acquired by Hewlett-Packard, and now part of Agilent Technologies. In 2000, he went back to academia. His current research interests include machine learning, surrogate modeling, bioinformatics and distributed computing. He has published over 200 papers in international journals and conference proceedings, and he is the holder of 5 U.S. patents. His email address is tom.dhaene@ugent.be.

**JORIS DEGROOTE** is a professor in the Department of Flow, Heat and Combustion Mechanics at Ghent University, Belgium. He received a Masters degree and a Ph.D. degree both in electromechanical engineering from Ghent university, Belgium in 2006 and 2010, respectively. His research mainly deals with fluid-structure interaction problems. His email address is Joris.Degroote@ugent.be.

**ERIC LAERMANS** is an assistant professor in the IBCN research group at Ghent University, Belgium since 2003. He received a Masters degree in Engineering Physics and a Ph.D. degree in Electrical Engineering from Ghent University, Belgium in 1994 and 1999, respectively. His research domain has evolved from the electromagnetic modeling of high-speed interconnection structures (with special attention to via holes) and reverberation chambers to data analysis and machine learning, more specifically, surrogate modeling and experimental design. His email address is eric.laermans@ugent.be.