On some new [French] recipes for surrogate modeling, Bayesian and topology

Optimization

If we have time ...

Prof. Joseph Morlier









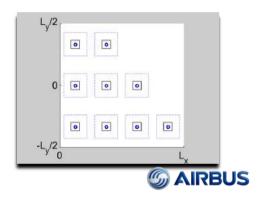


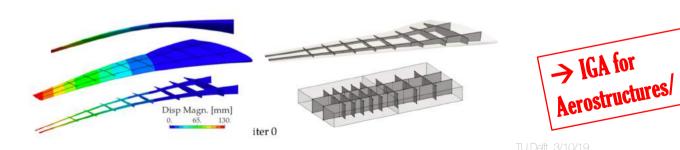
## My Research Group

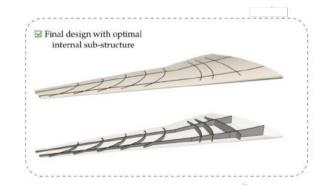
http://www.institut-clement-ader.org/pageperso.php?id=jmorlier

• 4 PhDs, 1 postdoc, 1 research assistant, 4 MsCs



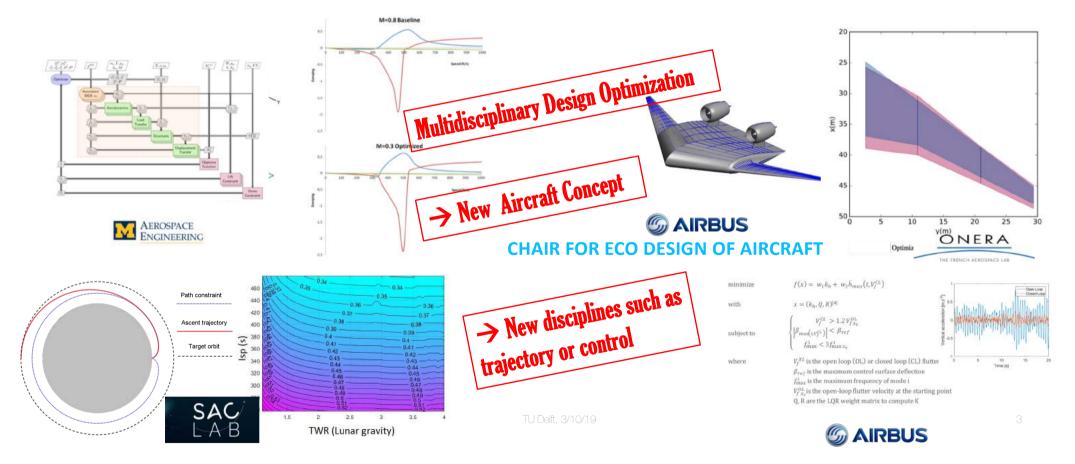




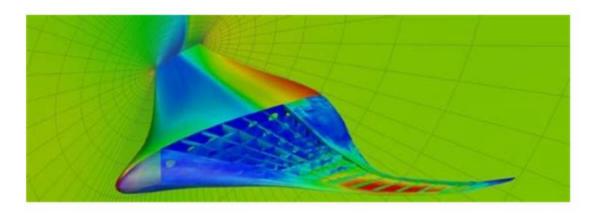


## My Research Group (Joint research with ONERA on MDO)

• 4 PhDs, 1 postdoc, 1 research assistant, 4 MsCs



## Popularization



http://mdolab.engin.umich.edu

### **Optimization [MDO] for connecting** people?

Publié le 14 février 2019



Modifier l'article



∀ Voir les stats



joseph morlier Professor in Structural and Multidisciplinary Design Optimization, ... any idea? 2 articles







https://www.linkedin.com/pulse/optimization-mdo-connecting-people-joseph-

# Our Goal: Embed efficient optimization algorithms in the design process of Aerostructures

- Reduce in a « smart way » the computational time of optimization for coupled simulations
- ullet Global Optimization using surrogate modeling ullet fixed budget (enriching process) to deal with INDUSTRIAL problems
- Specialized surrogate models for HD (engineering) problems and UQ
- Taking into account different levels of fidelity
- → Methods applied to AD Aircraft Design: Put the aircraft structure / aeroelasticity in the loop at the early stage of MDO process
- → Methods applied to SLD Space Launcher Design: Put the control/trajectory/propulsion in the preliminary design loop
- → compatible with



## Outlines for today

multidisciplinary Design optimization

multidisciplinary optimization

- 1. Surrogate modeling
- 2. Bayesian Optimization
- 3. Topology Optimization

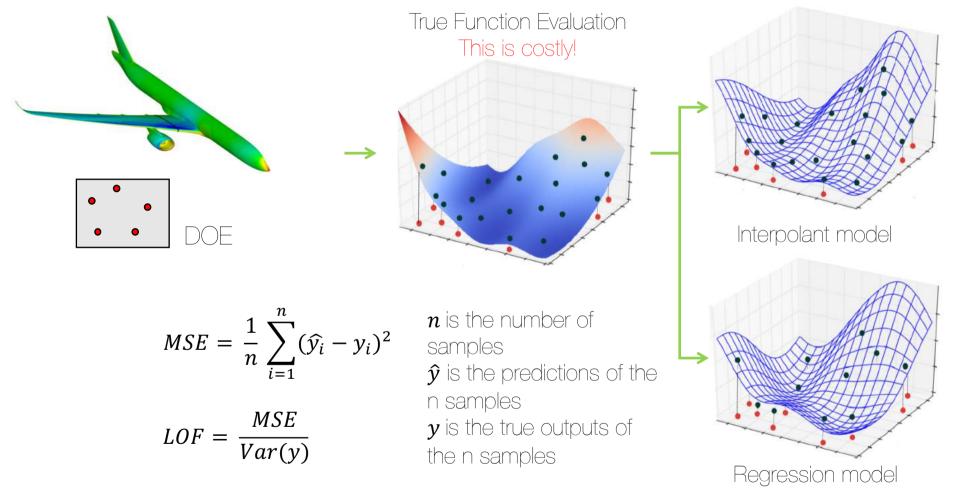
I,2: Common researches with N. Bartoli, T. Lefebvre (ONERA/DTIS)

## Outlines for today

# 1. Surrogate modeling

- 2. Bayesian Optimization
- 3. Topology Optimization

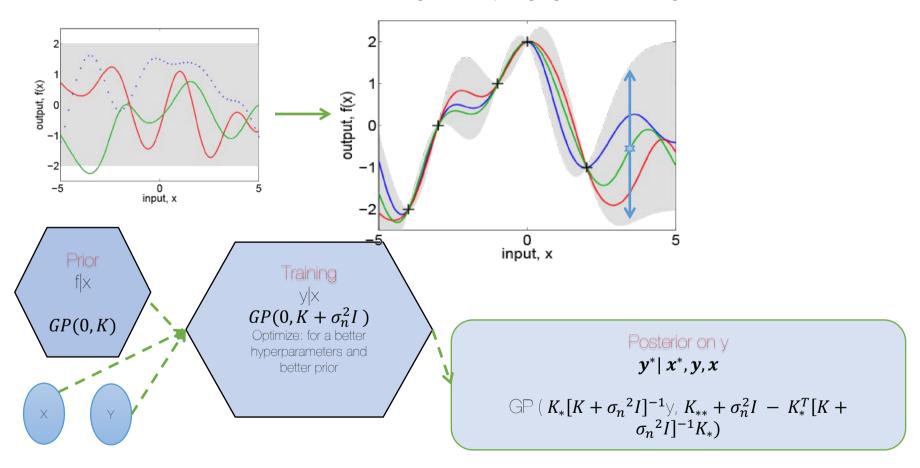
## Surrogate modeling Recipes



TUDelft 3/10/19

## Gaussian Process (aka Kriging)

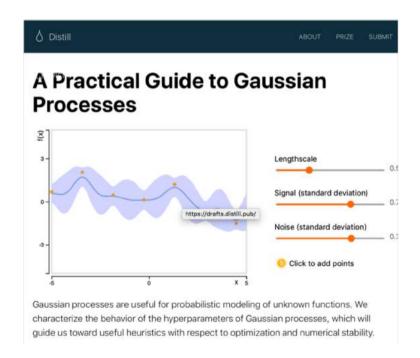
Image Source: http://mlg.eng.cam.ac.uk/teaching/4f13/1314/



TUDelft 3/10/19

## A good starting point x<sub>0</sub>=Rasmussen's book

https://drafts.distill.pub/gp/



C. E. Rasmussen & C. K. I. Williams, Gaussian Processes for Machine Learning, the MIT Press, 2006, ISBN 026218253X. © 2006 Massachusetts Institute of Technology. www.GaussianProcess.org/gpml

Gaussian Processes for Machine Learning

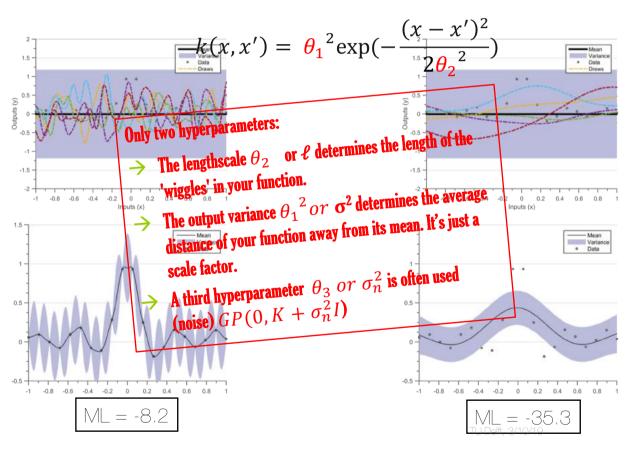
TUDelft 3/10/19

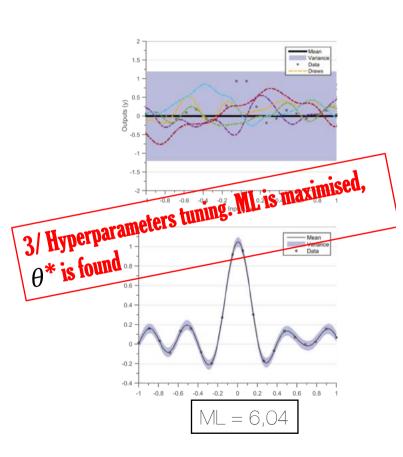
## 3/ Choose a Kernel/Construct Kxx Matrix view of Gaussian Process and Hyperparameters tuning 1/ Get your inputs/outputs data $\sqrt{2/Y_{ou}}$ wan to predict at $x^*$ $k(x,x') = \frac{\theta_1^2 \exp(-\frac{(x-x')^2}{2\theta_2^2})}{\left[\text{Kex}\right]}$ = $[K \times X]$ [KXX]-1 [KXX]-1 and variance of estimate $m(x_*) = K_*[Kxx]^{-1}$ A compute mean $var(x_*, x_*') = K_{**} - K_*^T [Kxx]^{-1} K_*$

## Optimizing Marginal Likelihood (ML)

$$\mathsf{ML} = log(p(y|X,\theta)) = -\frac{1}{2}y^{\mathsf{T}}K^{-1}y - \frac{1}{2}log|K| - \frac{n}{2}log(2\pi)$$

• It is a combination of data-fit term, a complexity penalty term and a normalization term





## Surrogate Model Toolbox: SMT



#### **SMT: Surrogate Modeling Toolbox**

The surrogate model toolbox (SMT) is an open-source Python package consisting of libraries of surrogate modeling methods (e.g., radial basis functions, kriging), sampling methods, and benchmarking problems. SMT is designed to make it easy for developers to implement new surrogate models in a well-tested and well-document platform, and for users to have a library of surrogate modeling methods with which to use and compare methods.

The code is available open-source on GitHub.

#### Focus on derivatives

SMT is meant to be a general library for surrogate modeling (also known as metamodeling, interpolation, and regression), but its distinguishing characteristic is its focus on derivatives, e.g., to be used for gradient-based optimization. A surrogate model can be represented mathematically as

$$y = f(\mathbf{x}, \mathbf{xt}, \mathbf{yt}),$$

where  $\mathbf{xt} \in \mathbb{R}^{nDAIX}$  contains the training inputs,  $\mathbf{yt} \in \mathbb{R}^{nt}$  contains the training outputs,  $\mathbf{x} \in \mathbb{R}^{nx}$  contains the prediction inputs, and  $y \in \mathbb{R}$  contains the prediction outputs. There are three types of derivatives of interest in SMT:

- 1. Derivatives (dy/dx): derivatives of predicted outputs with respect to the inputs at which the model is evaluated.
- Training derivatives (dyt/dxt): derivatives of training outputs, given as part of the training data set, e.g., for gradient-enhanced kriging.
- Output derivatives (dy/dyt): derivatives of predicted outputs with respect to training outputs, representing how the prediction changes if the training outputs change and the surrogate model is re-trained.

Not all surrogate modeling methods support or are required to support all three types of derivatives; all are optional.



https://github.com/SMTorg/SMT

M.-A. Bouhlel, J. T. Hwang, N. Bartoli, R. Lafage, J. Morlier, J. R.R.A Martins (2019), A Python surrogate modeling framework with derivatives, Advances in Engineering Software

Prediction derivatives (dy/dx) are derivatives of predicted outputs with respect to the inputs at which the model is evaluated. These are computed together with the prediction outputs when the surrogate model is evaluated. These are required for gradient-based optimization when the surrogate models.

Training derivatives  $(dy_t/dx_t)$  are derivatives of the training outputs with respect to the corresponding inputs. These are provided by the user and are used to improve the model corresponding inputs. These are provided by the user and are used to improve the model corresponding inputs. When the adjoint method is used to compute training derivatives, a accuracy in GE-KPLS. When the adjoint method is used to compute training derivatives, a high-quality surrogate model can be constructed with a low relative cost, because the adjoint high-quality surrogate model can be constructed with a low relative cost, because the adjoint method computes these derivatives at a cost independent of the number of inputs.

Output derivatives  $(dy/dy_t)$  are derivatives of predicted outputs with respect to training outputs, which is a measure of how the prediction changes with a change in training outputs, accounting for the re-training of the surrogate model. These post-training outputs, accounting for the re-training of the surrogate within an optimization iteration.

Given its focus on derivatives, SMT is synergistic with the OpenMDAO framework, which is a software framework for gradient-based multidisciplinary analysis and optimization. SMT can provide the derivatives that OpenMDAO requires from its components to compute the coupled derivatives of the multidisciplinary model.

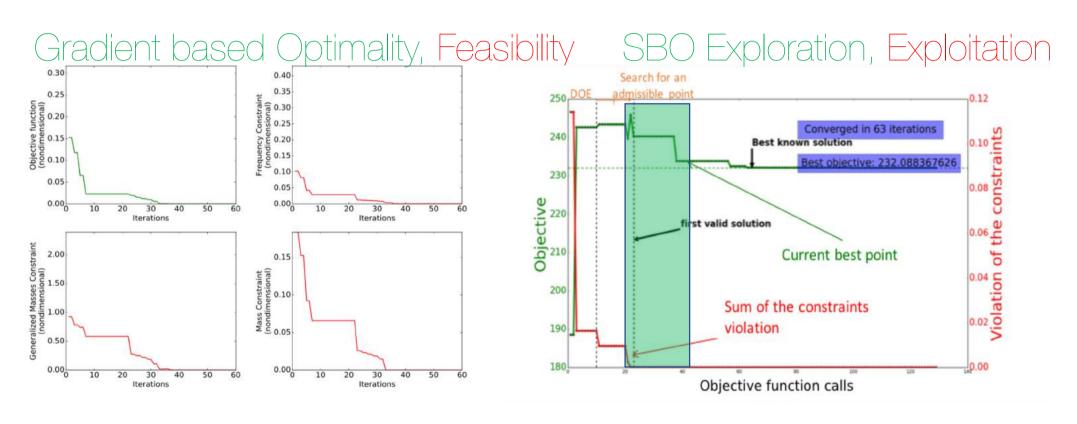
## Outlines for today

1. Surrogate modeling

# 2. Bayesian Optimization

3. Topology Optimization

## New paradigm for Surrogate Based Optimization (SBO)



Stopping criteria: tolfun, tolx, maxiter

Stopping criteria: Max Budget (Function calls)

## A good starting point X<sub>0</sub>=Forrester's book

## **Engineering Design via Surrogate Modelling**

**A Practical Guide** 

Alexander I. J. Forrester, András Sóbester and Andy J. Keane

University of Southampton, UK

## The goal is: find min of f(x) by sampling + and Kriging

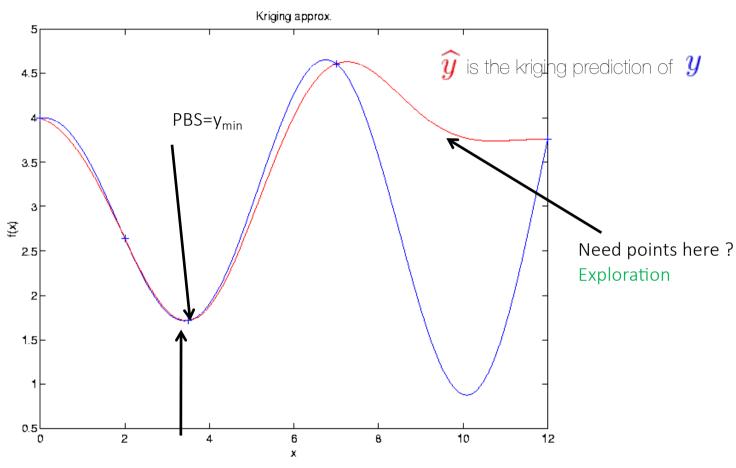
updating



We note the present best solution  $(PBS=y_{min})$ 

At every x there is some chance of improving on the PBS.

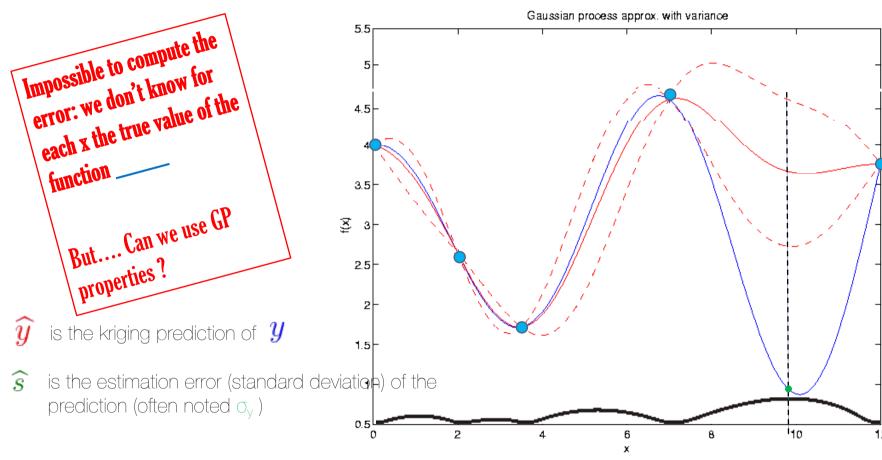
Then we ask: Assuming an improvement over the PBS, where is it likely be largest?



Exploitation may drive the optimization to a local optimum

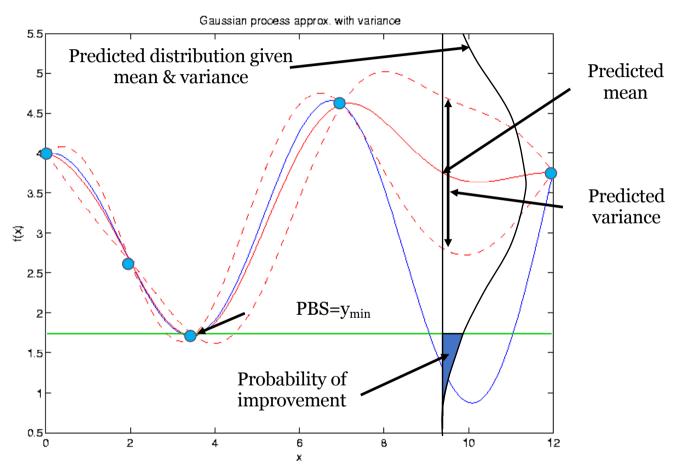
TU Delft. 3/10/19

## In supervised mode ... have a look to max(RMSE)



PBS=y<sub>min</sub>

## Probability of improvement



## Improvement ... explicitely

• Improvement : 
$$I(\mathbf{x}) = \max \left( y_{min} - \hat{Y}(\mathbf{x}), 0 \right)$$

• Expected Improvement :

$$EI(x) = E[\max(0, y_{\min} - \hat{y}(x))]$$

$$E[I(\mathbf{x})] = \int_{-\infty}^{y_{min}} (y_{min} - \hat{y}) \varphi \left( \frac{y_{min} - \mu_{\hat{Y}}(\mathbf{x})}{\sigma_{\hat{Y}}(\mathbf{x})} \right) d\hat{y}$$

$$E[I(\mathbf{x})] = (y_{min} - \mu_{\hat{Y}}(\mathbf{x}))\Phi\left(\frac{y_{min} - \mu_{\hat{Y}}(\mathbf{x})}{\sigma_{\hat{Y}}(\mathbf{x})}\right) + \sigma_{\hat{Y}}(\mathbf{x})\varphi\left(\frac{y_{min} - \mu_{\hat{Y}}(\mathbf{x})}{\sigma_{\hat{Y}}(\mathbf{x})}\right)$$

**Exploitation** 

**Exploration** 

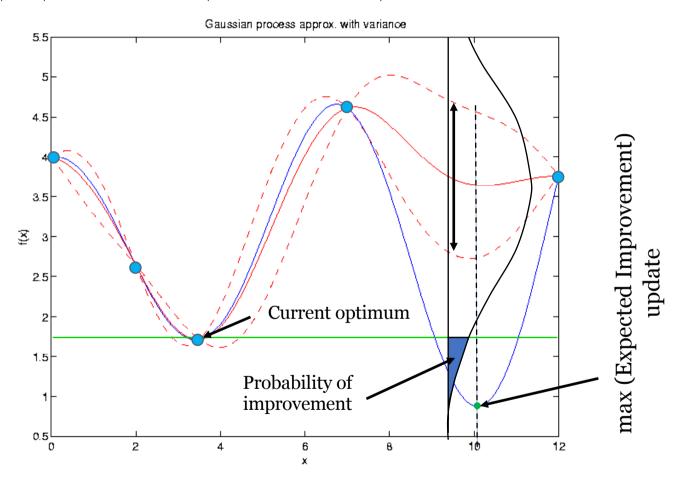
 $\mathcal{N}(0,1)$ 

global optimum can be found because P[I(x)] = 0 when s = 0 so that there is no probability of improvement at a point which has already been sampled  $\rightarrow$  guarantees global convergence

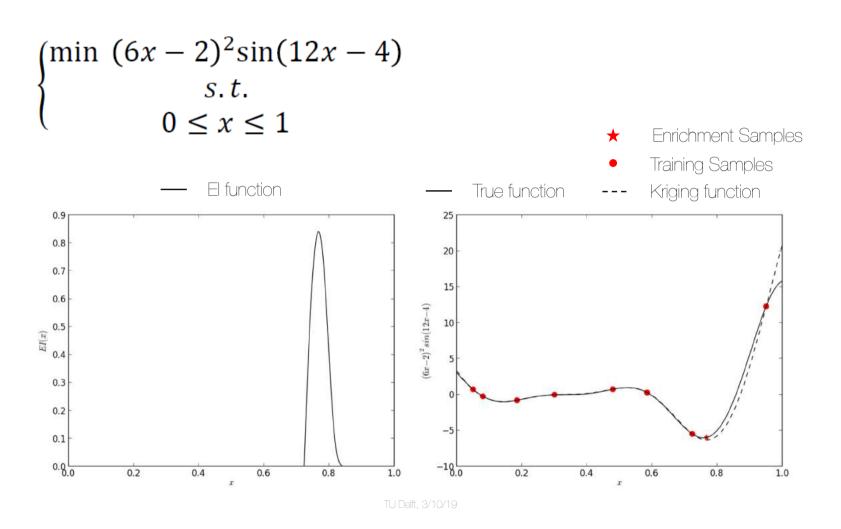
 $\Phi$ : cumulative distribution function  $\mathcal{N}(0,1)$   $\phi$ : probability density function

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

## Infill Criteria: max(Expected improvement)



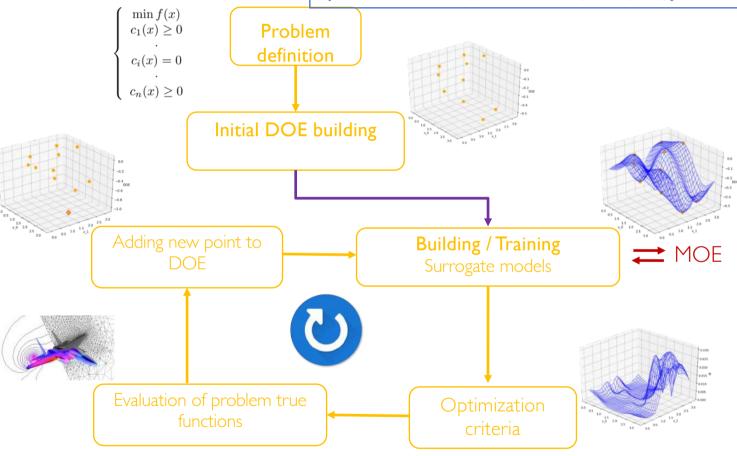
## Illustration on 1D example



## SEGOMOE algorithm

N. Bartoli, T. Lefebvre, S. Dubreuil, R. Olivanti, N. Bons, J.R.R.A. Martins, M.-A. Bouhlel, J. Morlier, "An adaptive optimization strategy based on mixture of experts for wing aerodynamic design optimization", 18th AIAA/ISSMO Multidisciplinary Analysis

Optimization Conference, AIAA-2017-4433, Denver, USA, June 2017



### New formulation

N. Bartoli, T. Lefebvre, S. Dubreuil, R. Olivanti, N. Bons, J.R.R.A. Martins, M.-A. Bouhlel, J. Morlier, "An adaptive optimization strategy based on mixture of experts for wing aerodynamic design optimization", 18th AIAA/ISSMO Multidisciplinary Analysis

Adapted from Super EGO

### Costly initial problem

$$\begin{cases} \min f(x) \\ c_1(x) \ge 0 \\ \vdots \\ c_i(x) = 0 \\ \vdots \\ c_n(x) \ge 0 \end{cases}$$

### Cheap enrichment problem

$$\Rightarrow \begin{cases} \max_{x \in \mathbb{R}^d} & \operatorname{EI}(x)/\operatorname{WB2}(x)/\operatorname{WB2s}(x) \\ \text{s.t.} \\ \hat{c}_1(x) \ge 0 \\ \vdots \\ \hat{c}_i(x) = 0 \end{cases} \qquad \text{n+1} \\ \hat{c}_i(x) = 0 \\ \vdots \\ \hat{c}_n(x) \ge 0$$

Optimization Conference, AIAA-2017-4433, Denver, USA, June 2017

Possibly Multimodal

Global optimization method

Multimodal

## ADODG6 \* testcase (R. Olivanti, R. Priem MsC)

CFD guys know very well the multimodality of this problem...

Wing drag minimization problem (subsonic, Euler equations)

	Function/variable	Description	Quantity	Range
minimize	$C_D$	Drag coefficient	1	
with respect to	α	Angle of attack	1	[-3.0, 6.0] (°)
	$\theta$	Twist	8	[-3.12, 3.12] (°)
	δ	Dihedral	8	[-0.25, 0.25] (unit of chord)
		Total variables	17	F (2)
subject to	$C_L = 0.2625$	Lift coefficient	1	
	50065	Total constraints	1	

Can SEGOMOE help us to reach the global optimum? Is it less dependant on  $X_0$  compared to SNOPT £?

\*AIAA, Aerodynamic Design Optimization Discussion Group <a href="http://mdolab.engin.umich.edu/content/aerodynamic-design-optimization-workshop">http://mdolab.engin.umich.edu/content/aerodynamic-design-optimization-workshop</a>

£ https://web.stanford.edu/group/SOL/snopt.htm

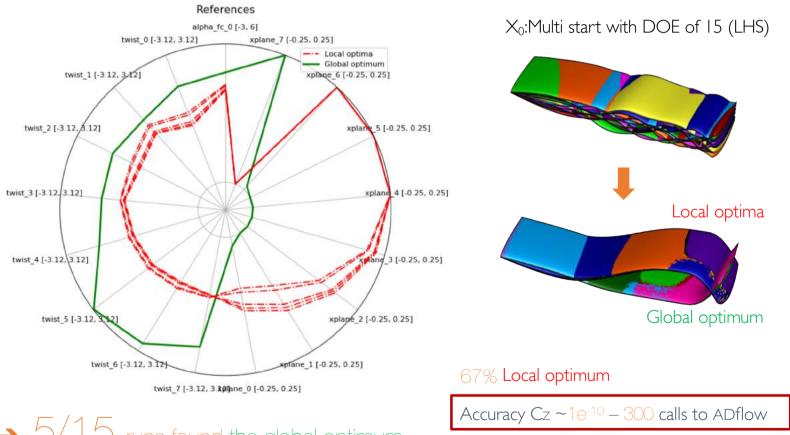
## ADODG\* 6 TOOLS





## Multimodal optimization problem (SNOPT Benchmark)

Wing drag minimization problem (subsonic, Euler equations with ADFlow solver) (Mesh 180K cells)



→ 5/15 runs found the global optimum

### Initial DOE= 68 points (4xd)

..... DOE=17 n\_runs=18

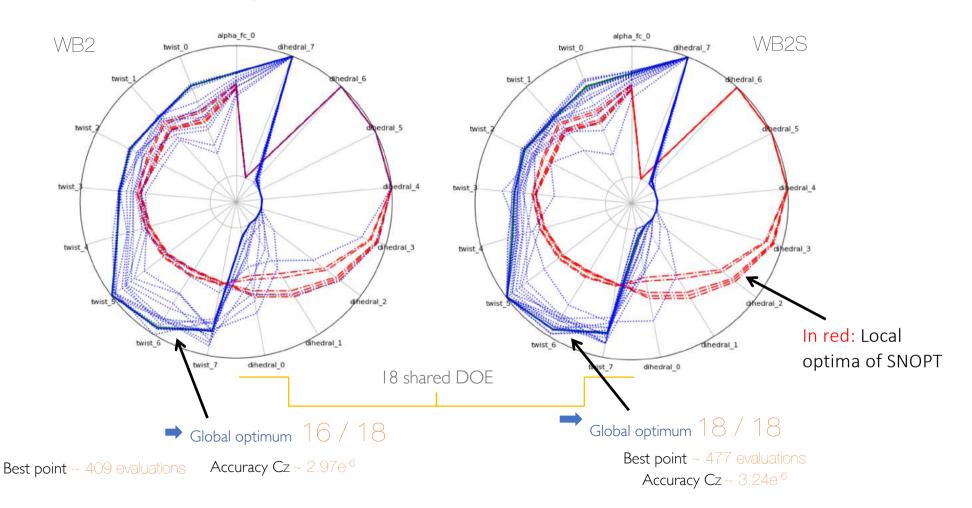
--- Local optima

Global optimum

## Multimodal optimization problem (SEGOMOE 2)

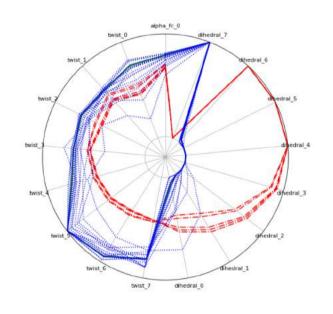
Frozen budget: 500 evaluations

Surrogate models: KPLSK

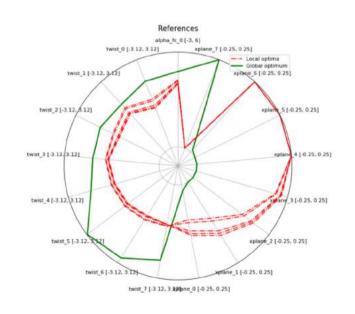


## Idea: 2 steps method

- 1. Start with SEGOMOE, stop with maxiter
- x° with high confidence near x\* (so we avoid to get stuck in local minima!)

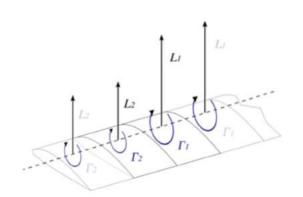


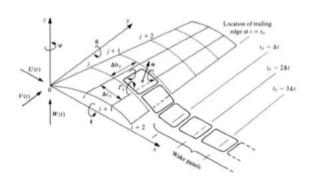
## 2. Use $\mathbf{x}^{\circ}$ as $\mathbf{x}_0$ in SNOPT to reach rapidly the Global Optimum $\mathbf{x}^{*}$

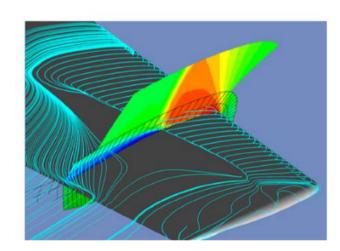


## What if?

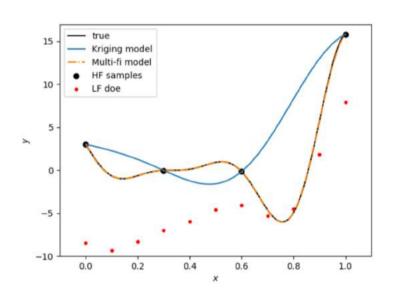
- Several levels of fidelity of the same simulation are available
- → For example, in aerodynamics: Liflting line theory, Vortex lattice method, and RANS CFD code

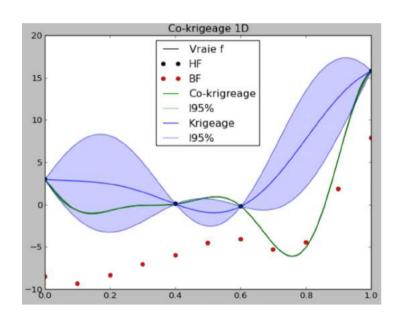






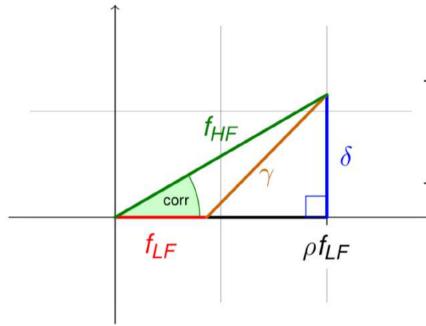
# How to use low-fidelity information to speed up the optimization?





→ It is also away to learn the difference between HF & LF ...

## Co Kriging



Additive formulation [Lewis 2000]

$$f_{HF}(x) = f_{LF}(x) + \gamma(x)$$

- Kennedy-O'Hagan [Kennedy 2001]

$$f_{HF}(x) = \rho f_{LF}(x) + \delta(x)$$
  
 $f_{LF}(\cdot) \perp \delta(\cdot)$ 

The addition of the term p makes the multi-fidelity learning more robust to poor correlation as well as differences in modelization.

\$Alexandrov, N., Lewis, R., Gumbert, C., Green, L., & Newman, P. (2000, January). Optimization with variable-fidelity models applied to wing design. In 38th Aerospace Sciences Meeting and Exhibit (p. 841).

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3), 425-464.

### MFEGC

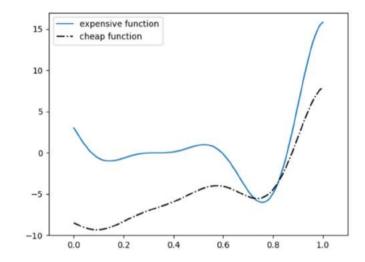
#### 2 step approach

• Most promising point: El-based criterion

$$\mathbf{x}^* = \arg\max_{\mathbf{x}} \left( \mathrm{EI}(\mathbf{x}) \right)$$

• Choice of levels of enrichment: trade off information gain/cost

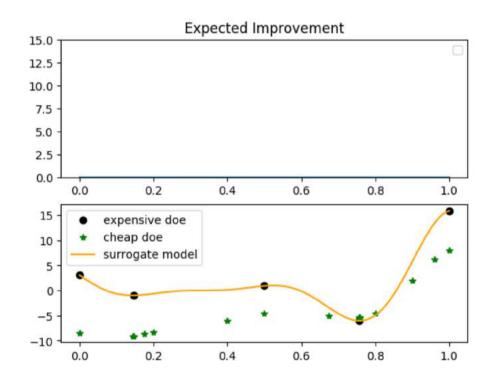
$$k^* = \arg\max_{k \in (0,\dots,\ell)} \quad \frac{\sigma_{\mathrm{red}}^2(k, \mathbf{x}^*)}{\mathrm{cost}_{\mathrm{total}}(k)^2}$$



- ⇒ By using low-fidelity to reduce the uncertainty we reduce the Exploration contribution to the El criterion
- ⇒ High-fidelity is used for Exploitation and model enhancement

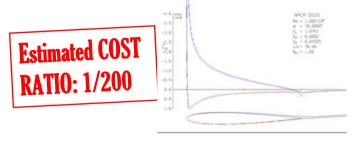
$$f_{HF}(x) = (6x - 2)^2 \times \sin(2(6x - 2))$$
  
 $f_{LF}(x) = 0.5f_{HF} + 10(x - 0.5) - 5$ 

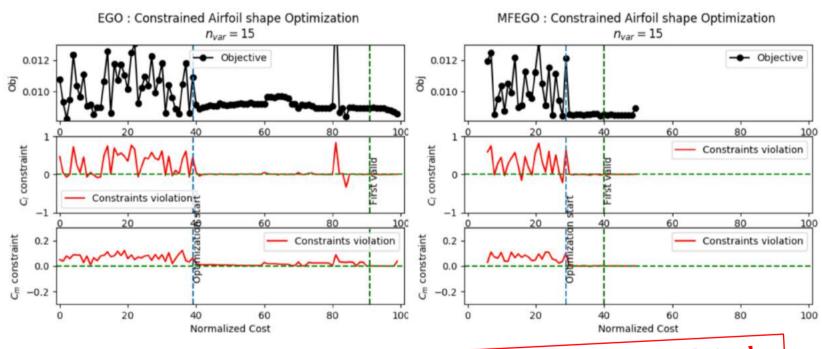
## Results (Toy problem)



Cost ratio: 1/1000						
	HF	LF	Cost			
MFEGO	3+2	6+9	5.015			
EGO	4+11		15			

## Second application: Constrained Optimization





\*https://web.mit.edu/drela/Public/web/xfoil/ \$ http://mdolab.engin.umich.edu MFEGO can speed up the Optimization process by reducing the calls to HF expensive code!

TU Delft 3/10/19

## Conclusions

- Our new Bayesian Optimization offers about the same efficiency than standard Gradient Based Method <u>even</u> for HD problems <u>wlithout</u> the influence of starting point (Exploration-Exploitation trade off) on a pure aerodynamic shape optimization problem

  EGO on SMT
- The multifidelity / Mixture of experts (MOE) options help us to speed the process (ongoing work)

  MFK, MOE on SMT
- Still have some Matlab codes to translate in python such as Sparse Physics-Based GP (AIRBUS Flight Physics)
- SBO widely used in collaborative projects (A CFD group can work with a CSM group independently) ... the Global optimizer is asking the good point to be computed for each group
- A good example:



### Recent Papers on this topic



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





Bouhlel, M., Bartoli, N., Regis, R. G., Otsmane, A., & Morlier, J. (2018). Efficient global optimization for high-dimensional constrained problems by using the Kriging models combined with the partial least squares method. Engineering Optimization, 1-16.

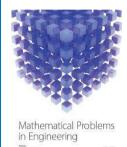


Bouhlel, M. A., Hwang, J. T., Bartoli, N., Lafage, R., Morlier, J., & Martins, J. R. (2019). A Python surrogate modeling framework with derivatives. Advances in Engineering Software, 102662.













### MDO courses & seminars



- NB: Since 2013 new course at SUPAERO: MDO [Structural&Multidisciplinary Design Optimization, 2\*30H] (MsC level) with ONERA/AIRBUS. Since 2016 one MDO seminar per year (open to PhDs and researchers)
- Since 2017 we offer some fund to SUPAERO students to do research with us in order to be « PhD ready ». Part of this presentation has been made by SUPAERO MsC Students

# Thanks

### to My co-workers:

Joaquim Martins, Nathalie Bartoli, Thierry Lefevbre, Mohamed Bouhlel, Emmanuel Benard, Claudia Bruni, John Hwang, Joan Mas Colomer, Peter Schmolgruber, Youssef Diouane, Sylvain Dubreuil, Christian Gogu, Stephanie Lisy-Destrez, Miguel Charlotte and

PhDs: Pierre-Jean Barjhoux, Simone Coniglio, Alessandro Sgueglia, Laurent Beauregard, Emmeline Faisse, Edouard Dui

## Outlines for today

- 1. Surrogate modeling
- 2. Bayesian Optimization

## 3. Topology Optimization

## A good starting point $x_0$ =MsC thesis of J. Overvelde, TU Delft, 2012

### Optimization variables:

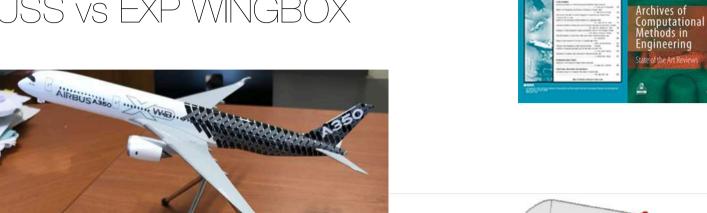
- Positions (x,y)
- Orientation (θ)
- Dimensions (Lx,Ly)
- Mass (M)



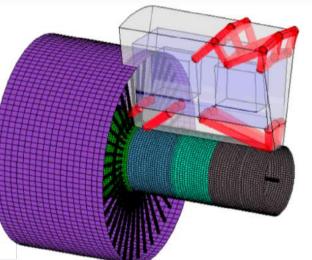
Structural Members engineering bricks like: beam, plate, geometric primitives

Moving Node Approach (MNA) 6 variables per node

## Bionic SIMP vs EXP TRUSS vs EXP WINGBOX







AIAA JOURNAL



#### https://github.com/mid2SUPAERO MID2

Multidisciplinary optimization for Innovation: Design and Data

SUPAERO

Repositories 38

Packages

People 37

Teams





ONERA

THE FRENCH AEROSPACE LAB









Prof. J. Morlier

joseph.morlier@isae-supaero.fr

40+ Students In MDO Courses At Master Level

In 2019 We Developed New Methodologies (or Applications) In:

#### Computational Structural Mechanics

#Gaussian Processes For Linear Elasticity

#Geometric Projection In Topology Optimization (AIRBUS and ICA)

#Topology Optimization For 3Dprinting (AIRBUS and ICA)

#High Resolution Topology Optimization (AIRBUS and ICA)

#Levet Set For Automatic Fiber Placement

#Eco Material Selection

#1D Refined FE Model In Dynamics

#IsoGeometric Analysis (LAMCOS and IMT)

#### Multidisciplinary Design Optimization

#HALE Ecodesign (CEDAR Chair)

#Reusable Launchers (SACLAB Chair)

#Multifidelity Method with Gaussian Processes (ONERA and MDOlab)

#MDA Acceleration

#Codesign For Robust Flutter (AIRBUS)

**#Trajectory Control** 

#Gaussian Processes and POD for coupled problem (ONERA)

#Aeroelastic for Scaled Aircraft (CEDAR Chair, ONERA and MDOlab)

#Hybrid Optimization (AIRBUS And IRT)

#BWB (CEDAR Chair, ONERA)

Thanks To All Supaero's Students (MAE, PIR, PhDs And Postdoc)

https://smt.readthedocs.io/en/latest/