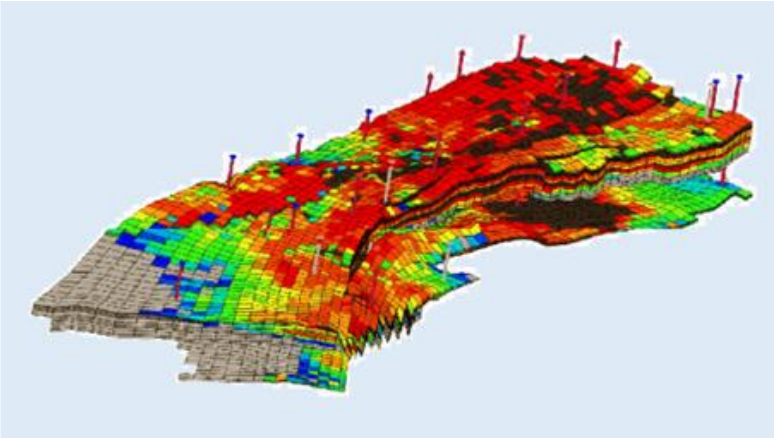
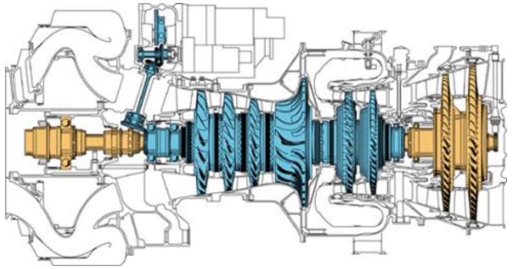
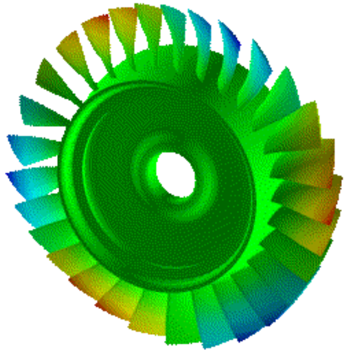
DERIVATIVE FREE OPTIMIZATION AND APPLICATIONS

Delphine Sinoquet (IFPEN)

COURSE 3: SOME CHALLENGES IN DFO







https://www.ifpenergiesnouvelles.fr/page/delphine-sinoquet

DERIVATIVE FREE OPTIMIZATION AND APPLICATIONS

- Course 1: main DFO methods
- Course 2: various applications of DFO
- **Course 3: some challenges in DFO**

CHALLENGES IN DFO

- **Dealing with mixed continuous and discrete variables**
  - Extension of trust region derivative free method
  - Extension of surrogate optimization based on kriging (EGO or Bayesian Optimization)
- **Dealing with uncertainties**
  - 1 practical example
  - Didier Lucor’s course: introduction to optimization under uncertainty

DFO WITH MIXED CONTINUOUS AND DISCRETE VARIABLES

- **Motivations:** some optimal design problems lead to optimization problem with discrete variables as, for instance,
  - the number of componants – **integer** variables,
  - the type of materials – **categorical** variables, often non ordered variables,
  - the presence or not of some componants – **binary** variables.

2 main classes of applications of optimization with mixed continuous and discrete variables

- **with engineering models**
  - depend typically on a limited number of variables of interest (~several dozens),
  - most of the variables are continuous and a few are discrete,
  - require to solve complex systems of equations (e.g. PDE).
- **with heterogeneous models with numerous sub-systems**
  - with a very large number of discrete variables,
  - for instance, number of units of a chemical process or of a network, or binray choices on/off …

DFO WITH MIXED CONTINUOUS AND DISCRETE VARIABLES

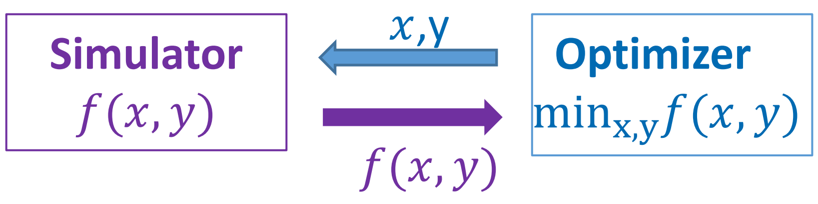
- **Motivations:** some optimal design problems lead to optimization problem with discrete variables as, for instance,
  - the number of componants – **integer** variables,
  - the type of materials – **categorical** variables, often non ordered variables,
  - the presence or not of some componants – **binary** variables.

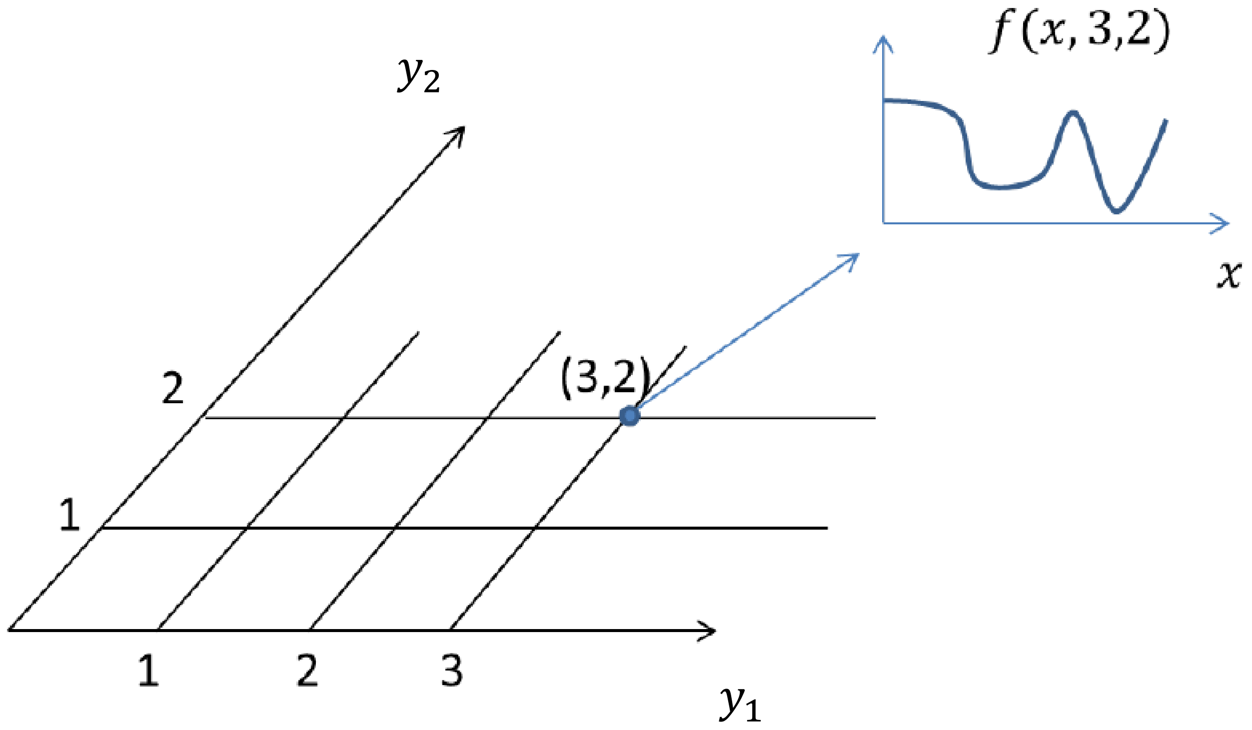
2 main classes of applications of optimization with mixed continuous and discrete variables

- **with engineering models**
  - depend typically on a limited number of variables of interest (~several dozens),
  - most of the variables are continuous and a few are discrete,
  - require to solve complex systems of equations (e.g. PDE).
- **with heterogeneous models with numerous sub-systems**
  - with a very large number of discrete variables,
  - for instance, number of units of a chemical process or of a network, or binray choices on/off …

DFO WITH MIXED CONTINUOUS AND DISCRETE VARIABLES

$$\min_{x \in \mathbb{R}^n, y \in I} f(x, y)$$





- $I_i = \{1, \dots, m_i\}$ numerical representation of the discrete variable y_i
- $I = \prod_{i=1}^{q} I_i$ the set of discrete variables

SOME APPLICATIONS OF MIXED DISCRETE DFO

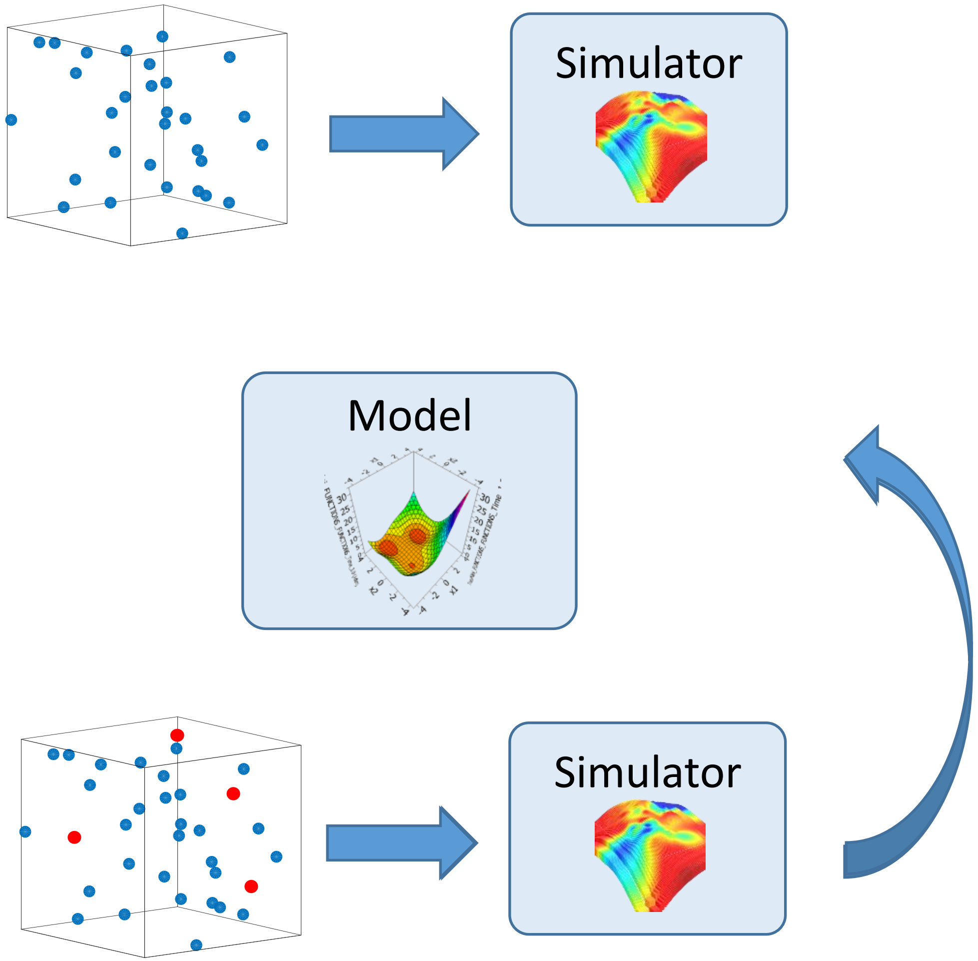
- Optimize the well placement and opening schedule for oil/gaz production fields
  Lizon et al. (2014)
- Maximize a turbo-machine efficiency and minimize the vibrations by modifying the blades shapes for a helicopter engine
  Tran et al. (2021)
- Design offshore wind turbine platform, optimize wind turbine layout
  on-going work (ANR project Samourai)

➢ All these industrial problems involve costly to evaluate simulators solving hydro-dynamic, aero-dynamic, fluid flow, solid mechanics equations

MODEL-BASED DFO METHODS

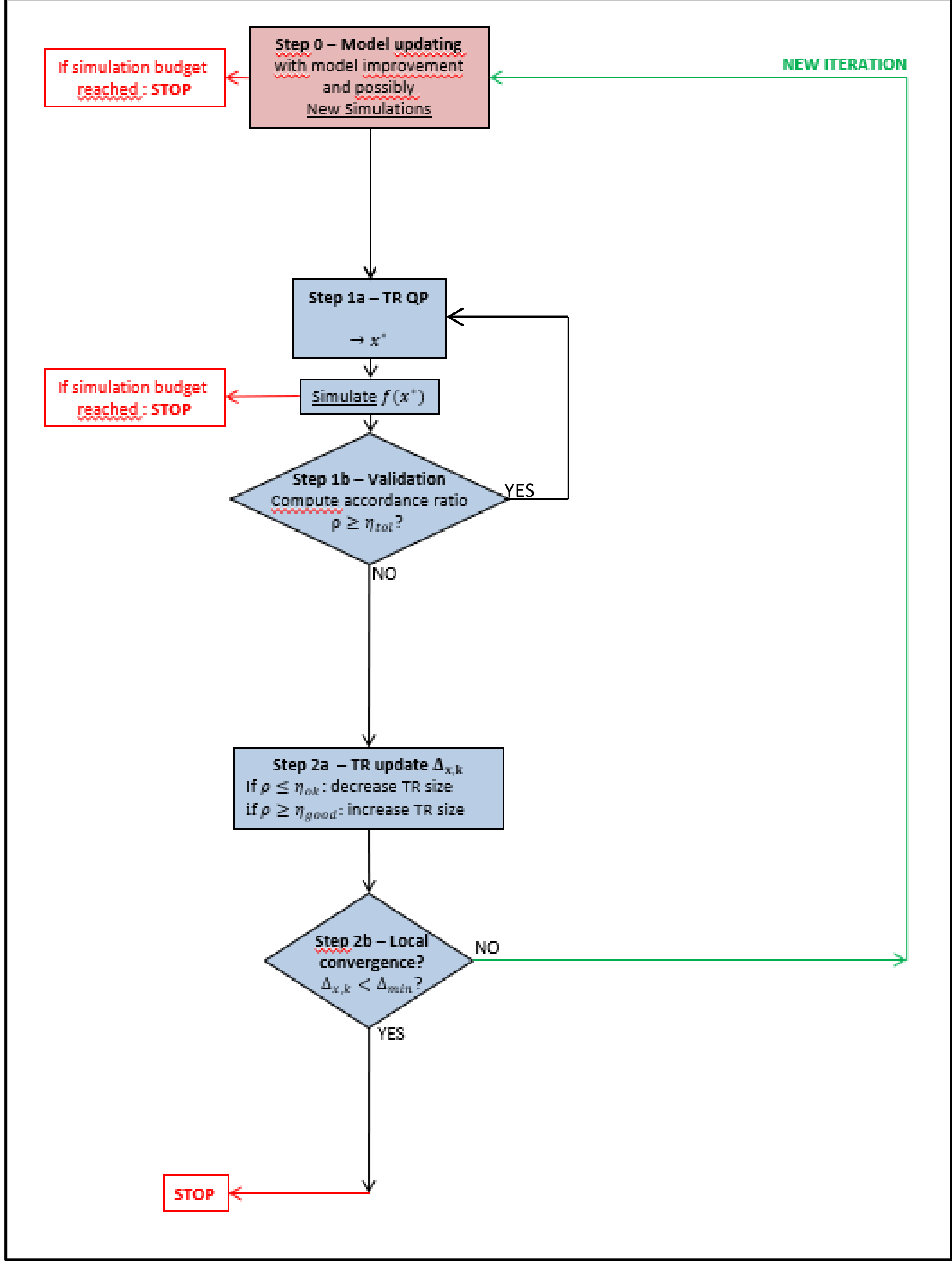
Remember 1st course on DFO

- Generate an initial set of points (DoE)
- Build a local or global model
- Use an improvement criteria based on the model to propose a new point to evaluate
- Improve the model (exploration)
- Go toward optimum (exploitation)
- Add the point to the DoE, update the model and iterate until simulation budget or other stopping criteria are reached



TRUST REGION DFO (LOCAL MODEL)

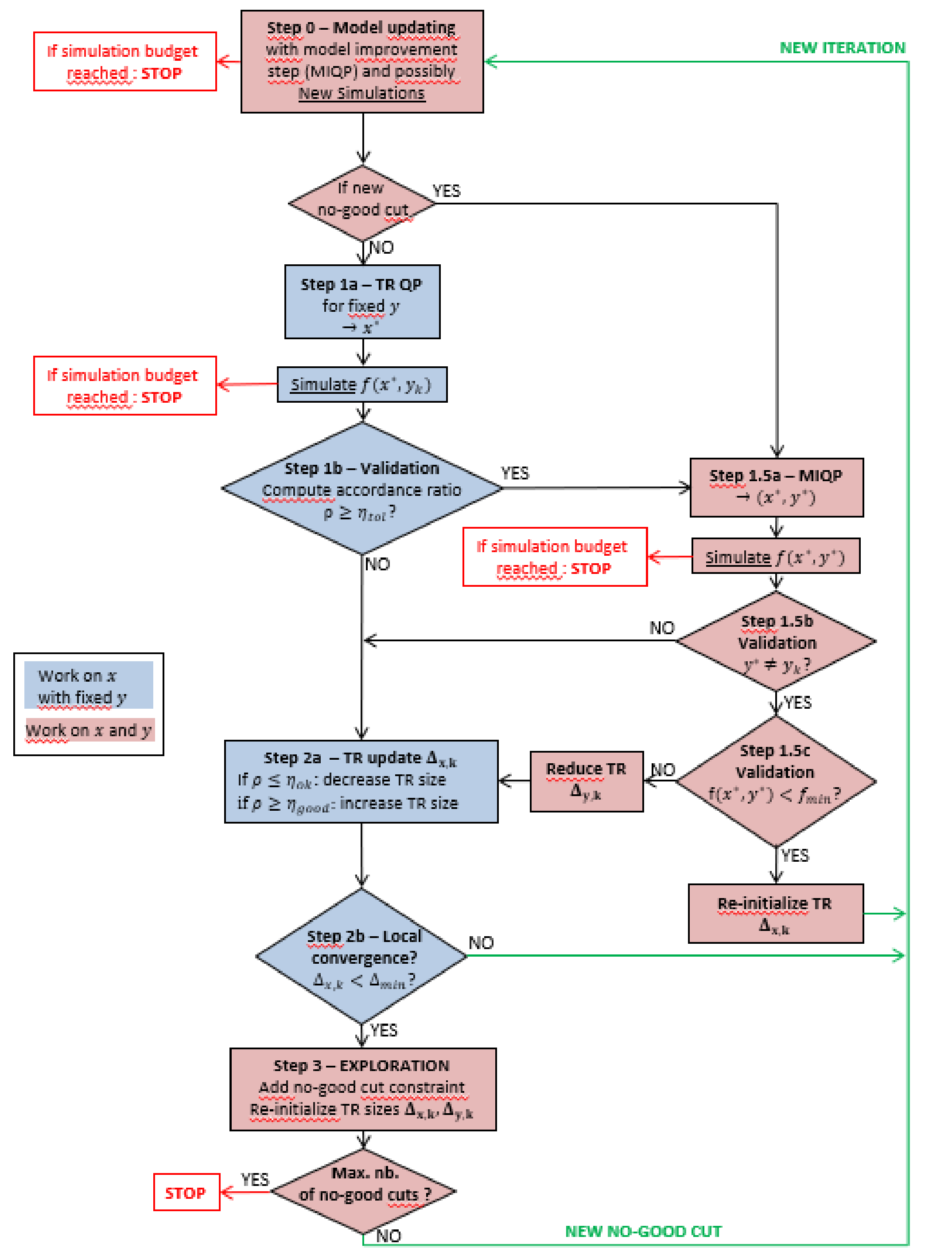
- Initial interpolation step
- Build/update a quadratic interpolation model + Model improvement step
- Solve QP problem in the current trust region
- Add new simulations
- Validation according to $\rho_k = \dfrac{f(x^k) - f(x^k + s^k)}{f(x^k) - \hat{F}_k(s^k)}$
- Update the trust region, the current TR center
- Stop when the TR size is too small



TRUST REGION DFO

Extended to mixed continuous and binary variables (Conn et al, 2016)

- Initial interpolation set in mixed space
- Build a quadratic interpolation model w.r.t both mixed variables
- Model improvement step (MIQP)
- **First minimization problem** w.r.t. continuous variables (binary variables are temporary fixed)
  - QP problem within the TR $\mathcal{B}(x^k; \Delta_\mathrm{x}^k)$
- If previous step is successful
  second minimization problem w.r.t. both variables
  - MIQP problem within the TR $\mathcal{B}(x^k; \Delta_\mathrm{x}^k) \times \mathcal{B}(y^k; \Delta_\mathrm{y}^k)$



TRUST REGION DFO

Extended to mixed continuous and binary variables (Conn et al, 2016)

- A trust region $\mathcal{B}(y^k; \Delta_y^k)$ is introduced for binary variables – **the local branching constraint**

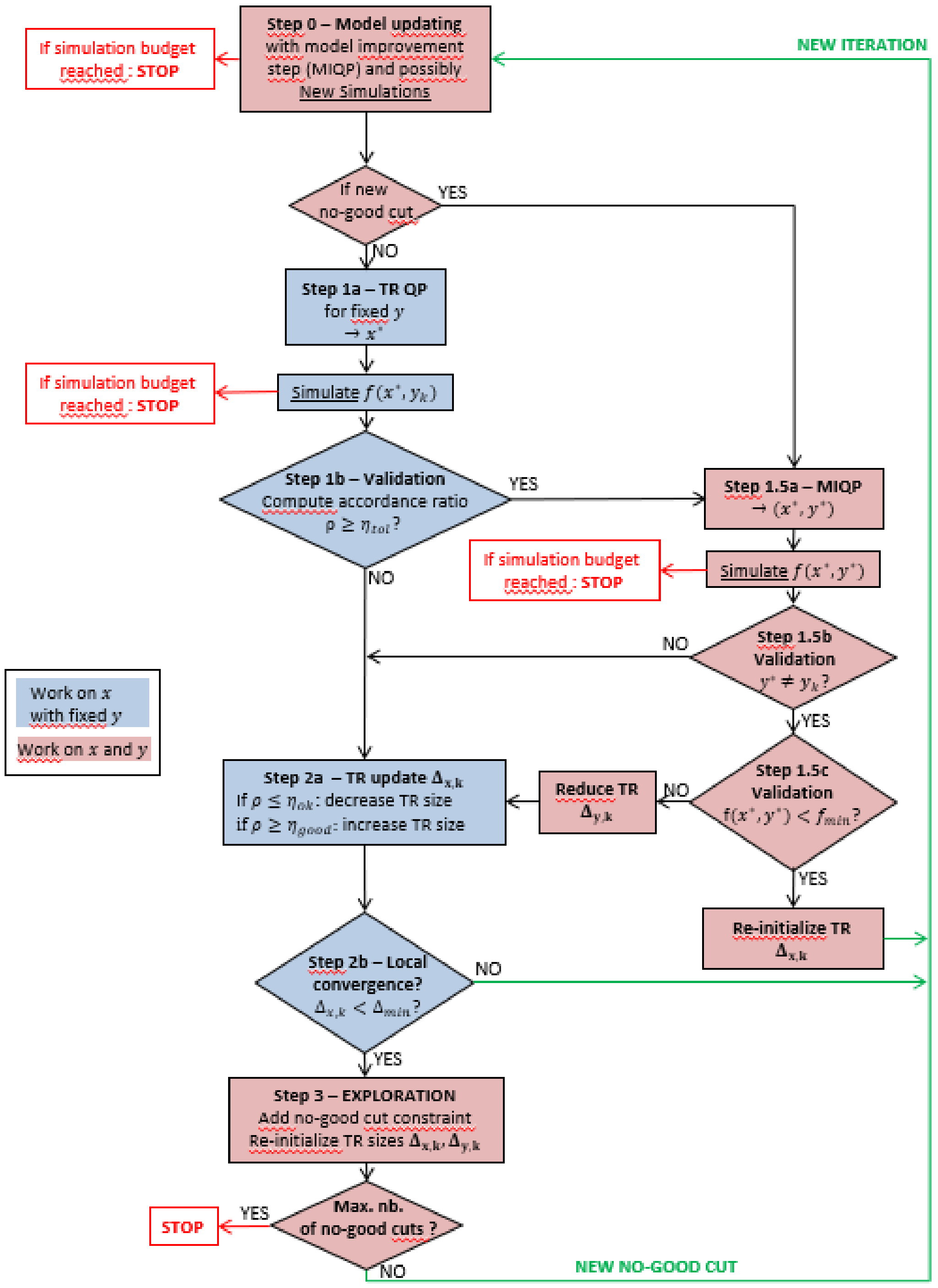
$$\left\| y - y^k \right\|_{Hamming} \le \Delta_y^k$$

which defines a neighbourhood of size Δ_y^k around the current TR center y^k (limits number of flips)

- **Exclusion constraints** allow to mimic the pruning process in branch & bound method to exclude explored regions w.r.t. binary variables and thus force exploration

$$\left\| y - y^k \right\|_{Hamming} \ge K'$$

(y^k, K') represent the center and the radius of the region we consider as sufficiently explored



GLOBAL SURROGATE OPTIMIZATION

Extended to mixed continuous and categorical variables

- Surrogate model based on Gaussian process (kriging) with **adapted kernel for mixed variables**

GLOBAL SURROGATE OPTIMIZATION

Remember course 1 for continuous variables :

SURROGATE OPTIMIZATION METHODS

Global models: Gaussian process (kriging)

- **Assumption:** the objective function is assumed to be a realization of a Gaussian random process (GP) with parametric mean function and stationary covariance function

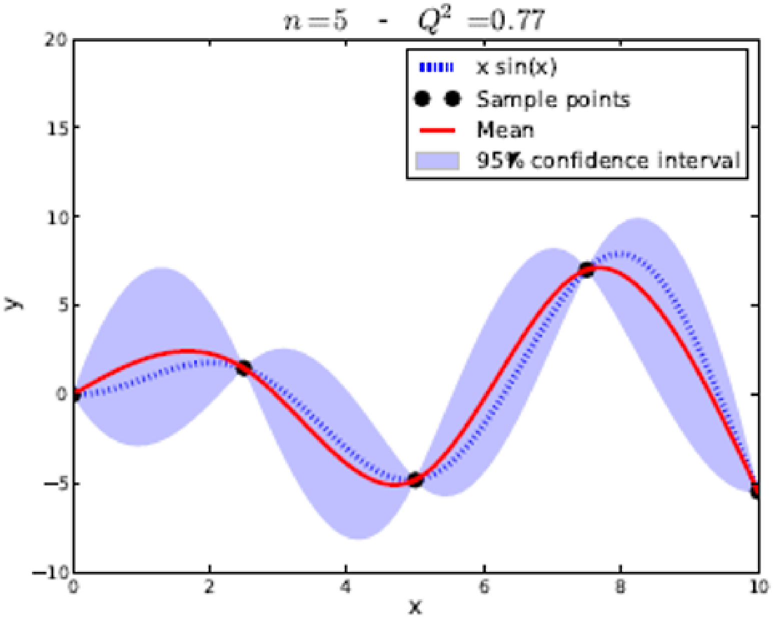
$$F(x) = \beta^T r(x) + Z(x)$$

- **The surrogate model** is the conditional expectation of the GP

$$\hat{F}(x) = E\left(F(x) | \left(x_i, f(x_i)\right)_{i=1,\dots,p}\right)$$
$$= \beta^T\ r(x) + k^T(x) K^{-1}\left(Y_p - R\beta\right)$$

- The variance of GP are used as **error indicators**

$$\sigma^2(x) = \sigma^2 - k^T(x) K^{-1} k^T(x)$$



avec $R = \left(r_j(x_i)\right)_{i,j}$, $K = \left(\rho(x_i, x_j)\right)_{i,j}$, $k(x) = \left(\rho(x, x_1), \dots, \rho(x, x_p)\right)$

© 2020 IFPEN

GLOBAL SURROGATE OPTIMIZATION

Extended to mixed continuous and categorical variables
(Munoz Zuniga and Sinoquet, 2020)

- Surrogate model based on Gaussian process (kriging) with **adapted kernel for mixed variables**

$$k((x,y),(x',y')) = k_{cont}(x,x') \times k_{catego}(y,y')$$

with

$$k_{catego}(y,y') = \prod_{l=1}^{m} T_{y_l,y'_l}$$

with T the correlation matrix between two binary vectors, T_{y_l,y'_l} is the correlation between the two levels y_l, y'_l of the l^{th} binary variables.

GLOBAL SURROGATE OPTIMIZATION

Extended to mixed continuous and categorical variables
(Munoz Zuniga and Sinoquet, 2020)

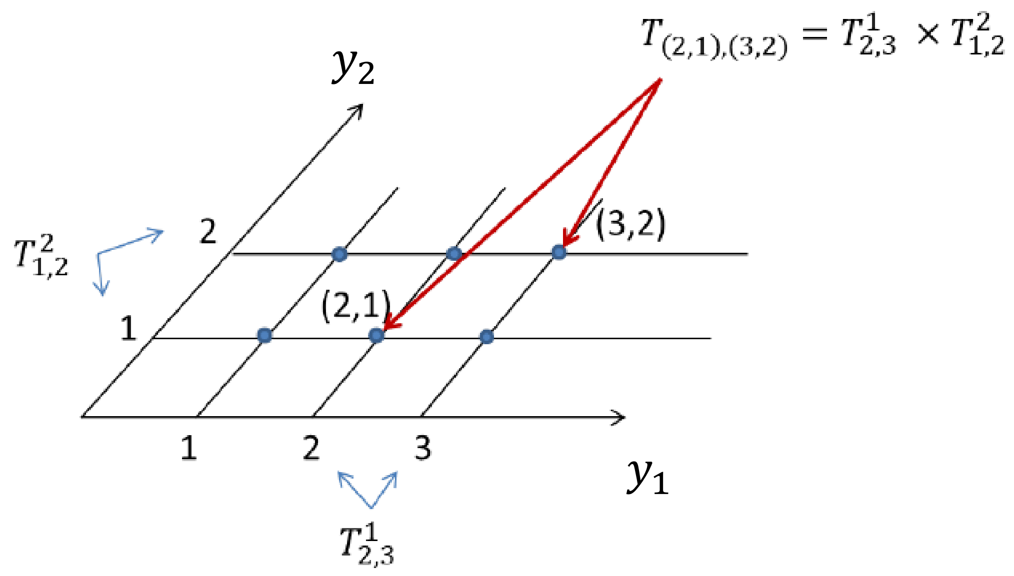
- Surrogate model based on Gaussian process (kriging) with **adapted kernel for mixed variables**

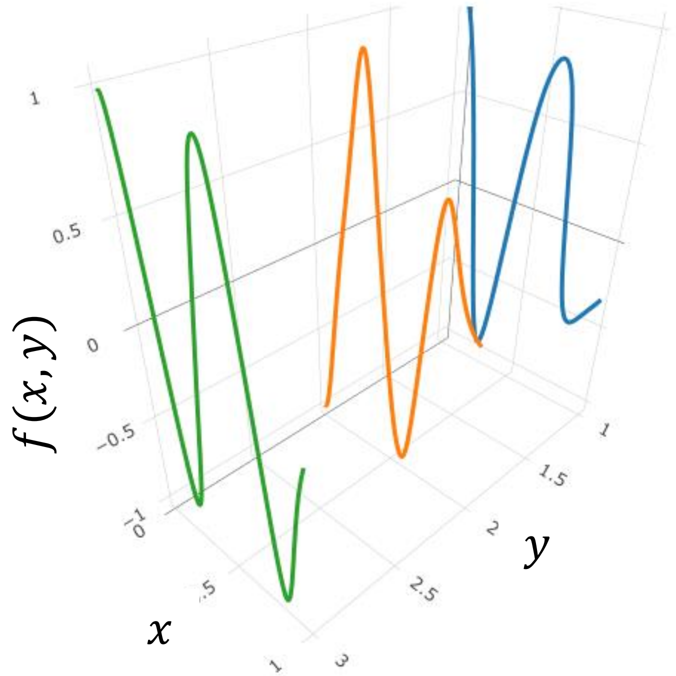
$$k((x,y),(x',y')) = k_{cont}(x,x') \times k_{catego}(y,y')$$

with

$$k_{catego}(y,y') = \prod_{l=1}^{m} T_{y_l,y_l'}$$

with T the correlation matrix between two binary vectors, $T_{y_l,y_l'}$ is the correlation between the two levels y_l, y_l' of the l^{th} binary variables.





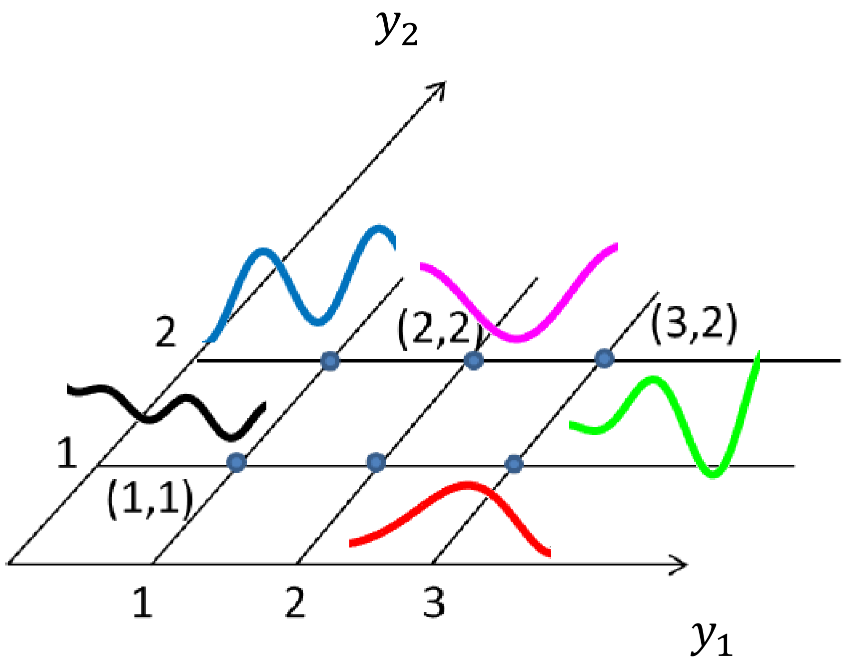
GLOBAL SURROGATE OPTIMIZATION

Extended to mixed continuous and categorical variables
(Munoz Zuniga and Sinoquet, 2020)

- The maximization problem for the sampling criterion Expected Improvement (EI) is now a **mixed continuous and categorical problem**

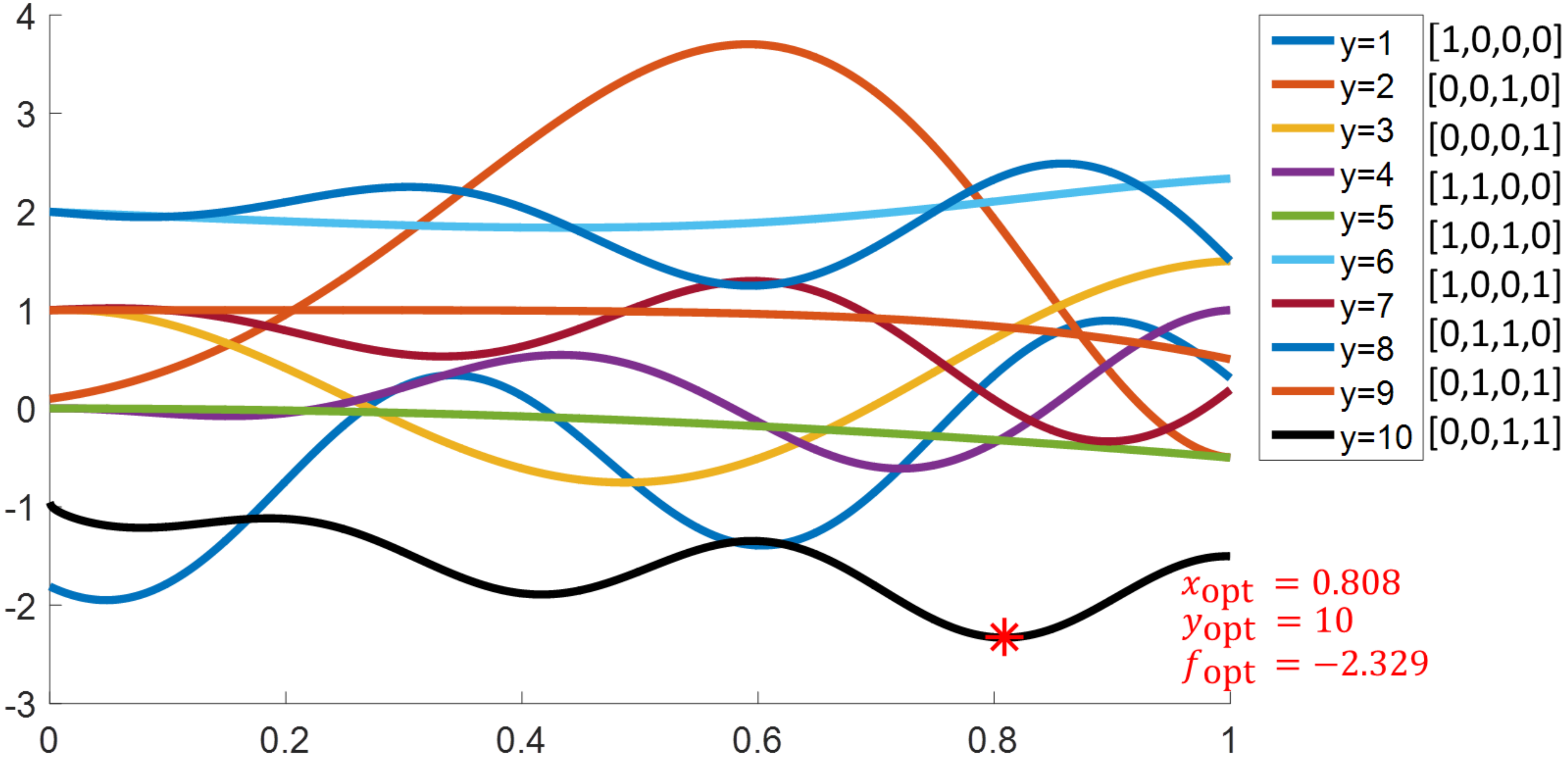
$$\operatorname{argmax}\big(EI(x,y)\big) = E\big(I(x,y)\big) = E\big(\max(0, f_{min} - \hat{F}(x,y))\big)$$

- EI is not expensive to compute (closed-form formula)
- Use of an appropriate optimizer NOMAD (Mesh Adaptive Direct Search) with an adapted exploration of the categorical space (poll step):
  - considering the correlations learnt by the GP
  - the potentially interesting levels for minimization (levels with small values of f)



APPLICATION TO A TOY PROBLEM

- A function of 2 variables : $x \in [0;1]$ and $y \in \{1,2,\ldots 10\}$, integer converted into 4 binaries



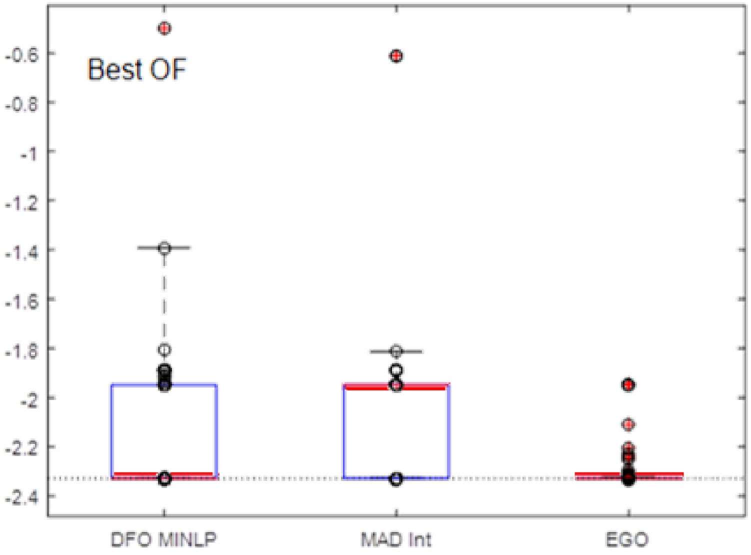
APPLICATION TO A TOY PROBLEM

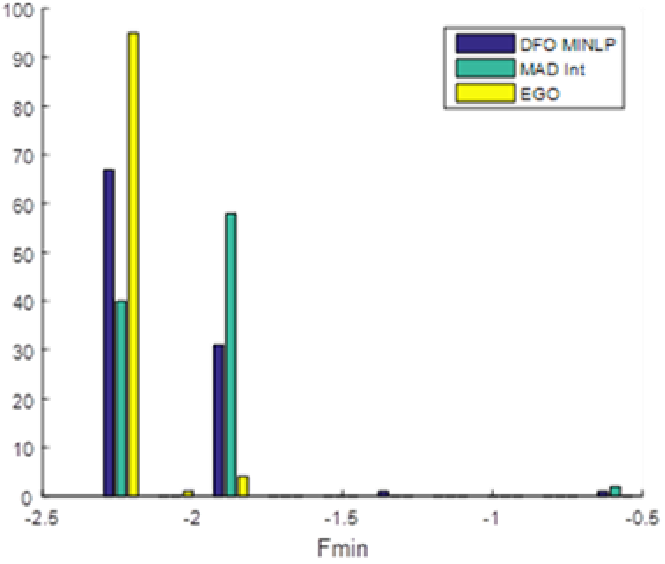
- A function of 2 variables : $x \in [0;1]$ and $y \in \{1,2,\ldots 10\}$,
  integer converted into 4 binaries for DFO Trust Region method for mixed binary variables

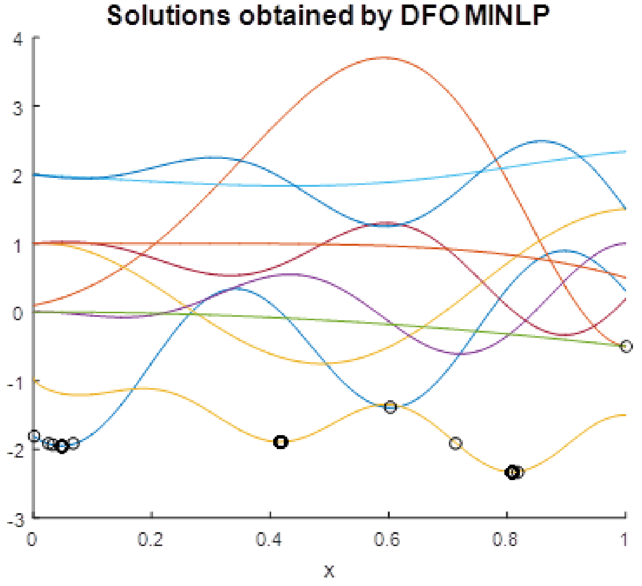
- Comparison of TR DFO method with
  - NOMAD : mesh adaptive direct search method for mixed integer variables
  - EGO : Efficient Global Optimization based on adaptive Gaussian Process surrogate models (Kriging) adapted to mixed categorical variables

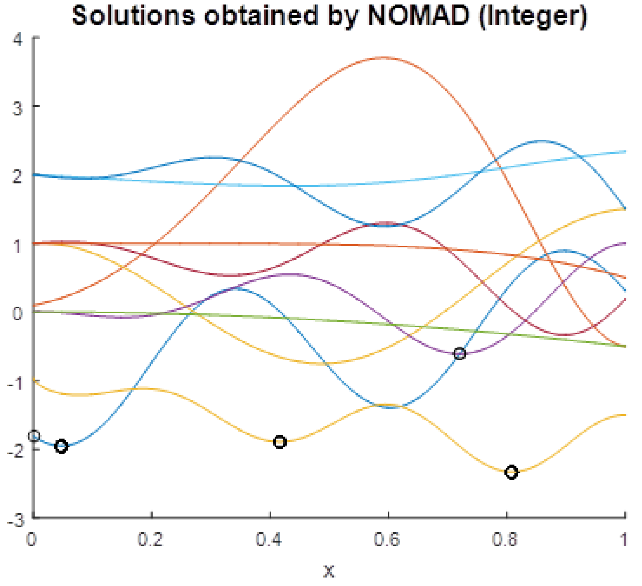
APPLICATION TO A TOY PROBLEM

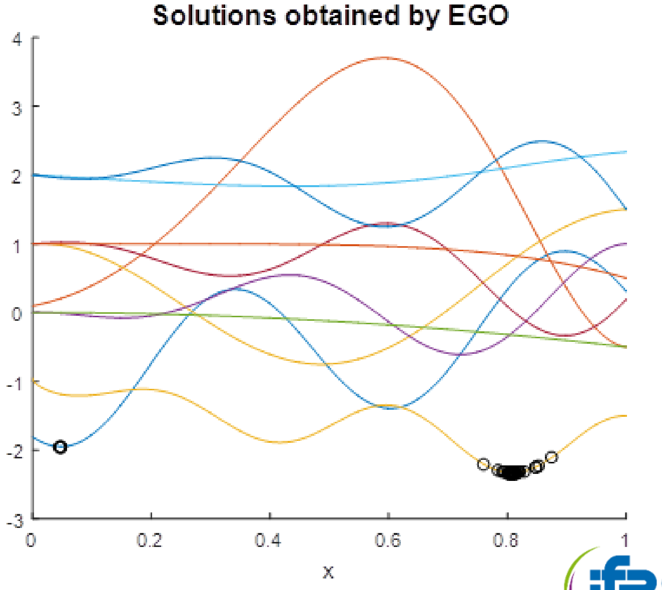
Initial Points:
5 randomly chosen points
100 repetitions











APPLICATION TO A TOY PROBLEM : SUMMARY

TR DFO method is able to

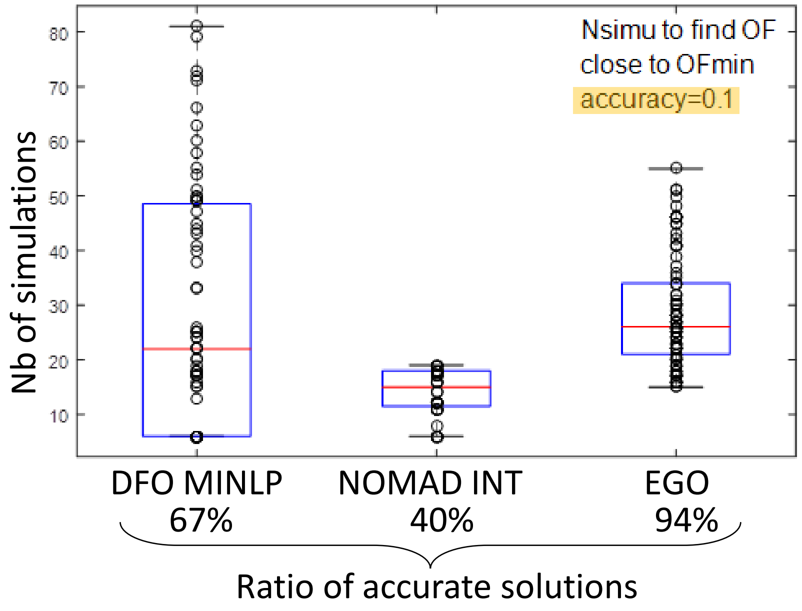
- Ensure a convergence to local minima with a controlled accuracy (thanks to the TR management for continuous variables)
- Explore the binary variable domain thanks to exclusion constraints
- Reach a good compromise between simulation cost (thanks to surrogate models) and accuracy (thanks to <u>local</u> models)

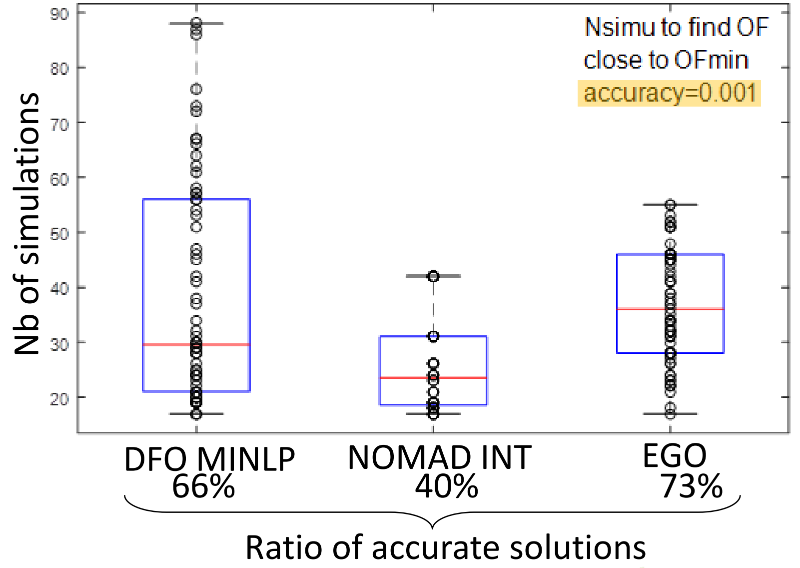
- Obtain more often the global optimum
- But with more simulations

Compared to NOMAD

- Less global because of our local models compare to global models of EGO
- But more accurate
- Require less simulations

Compared to EGO





DFO WITH MIXED CONTINUOUS AND DISCRETE VARIABLES

- An active research subject: phD subjects and publications

- A very studied application: hyperparameter optimization for machine learning (type of activation functions, number of layers, …)

- Challenges:
  - higher dimension (# of variables and # of levels for discrete variables)
  - Complex structures: be able to manage efficiently graph structures
    e.g. the number of continuous variables depends on the value taken by a discrete variable

CHALLENGES IN DFO

- **Dealing with mixed continuous and discrete variables**
  - Extension of trust region derivative free method
  - Extension of surrogate optimization based on kriging (EGO or Bayesian Optimization)
- **Dealing with uncertainties**
  - 1 practical example
  - See Didier Lucor’s course: introduction to optimization under uncertainty

ROBUST/RELIABLE DESIGN

Uncertain variables
u

Controllable variables
= design variables

x

Simulateur

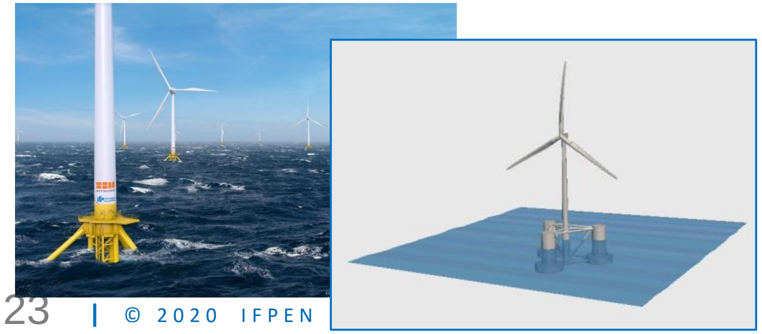
$f(x,u)$
$g(x,u)$

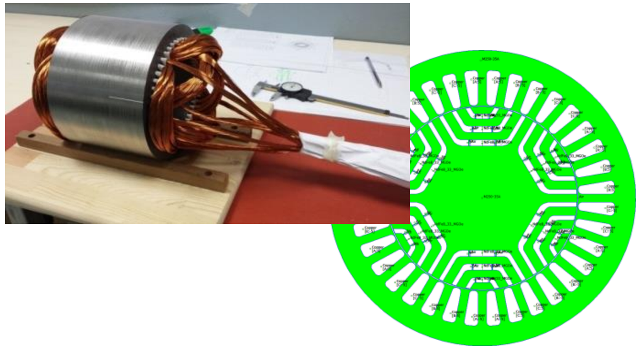
Quantities of Interest

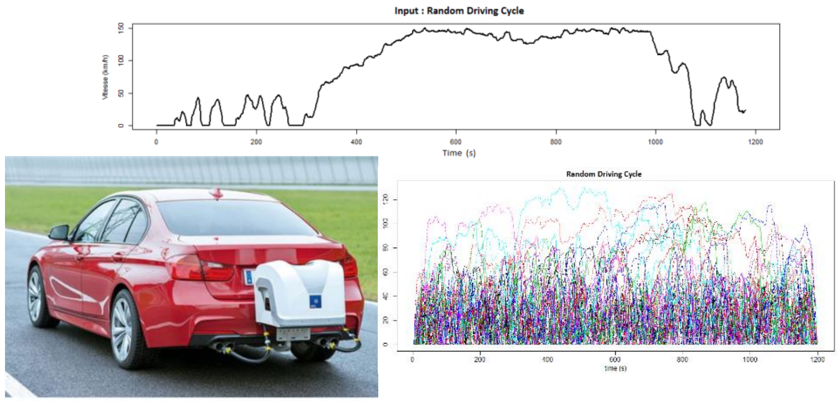
minimize/maximize
constrain

Applications in optimal design

- Reliability w.r.t. environnemental conditions (*e.g.* wind, wave)
- Robustness to the dispersions of the design parameters (manufacturing), to the componant characteristics (*e.g.* electomagnetic characteristics of magnets), …







ROBUST/RELIABLE DESIGN

x: design variables, u: uncertain variables
f, g : performances, costs, damages

$\min_x f(x)$ $\Rightarrow$	$\min_x \max_u f(x,u)$	Worst case
	ou $\min_x \begin{pmatrix} \mathbb{E}_u f(x,u) \\ \mathbb{V}_u f(x,u) \end{pmatrix}$	Gain/risk compromise
$\min_x f(x)$ s.t. $g(x) \leq 0$ $\Rightarrow$	$\min_x \mathbb{E}_u f(x,u)$ s.t. $\mathbb{P}_u(g(x,u) > 0) \leq \epsilon$	Probability of violating the constraint
$\Gamma = \{x \mid f(x) \leq S\}$ $\Rightarrow$	$\Gamma = \{x \mid \mathbb{E}_u f(x,u) \leq S\}$	Inversion: all the designs matching the constraint on average

ROBUST/RELIABLE DESIGN

- **Difficulties**
  - Simulation cost !!!
  - Approximation of statistical moments $\mathbb{E}_u\, f(x,u), \mathbb{V}_u\, f(x,u), \mathbb{P}_u(f(x,u)$
  - Complex inputs/outputs : high dimension, functionnal, mixed (discrete and continuous)

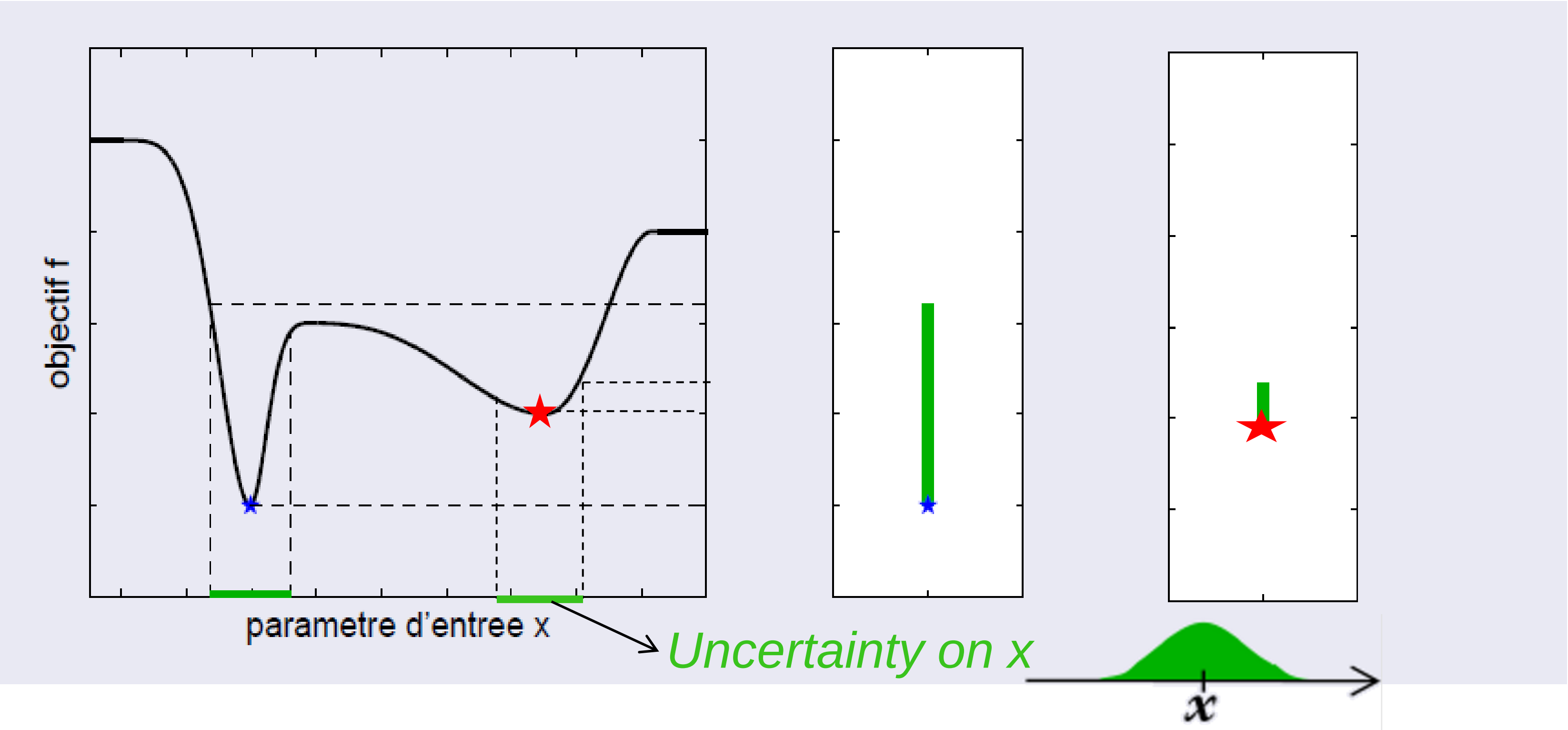
- **Solutions**
  - Meta-models and adaptive sampling of the simulation points dedicated to the formulations
  - Dimension reduction
  - « Goal-oriented » sensitivity analysis

➢ One practical example of robust optimization

➢ Course of Didier Lucor for more details

ROBUST OPTIMIZATION

Problem with one controllable variable



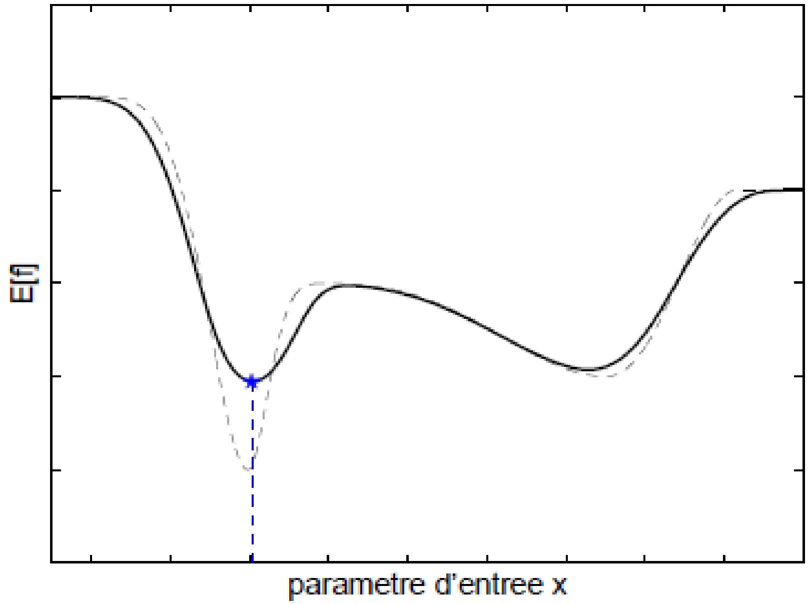
ROBUST OPTIMIZATION CRITERIA

Determine the values of controllable variables that minimize

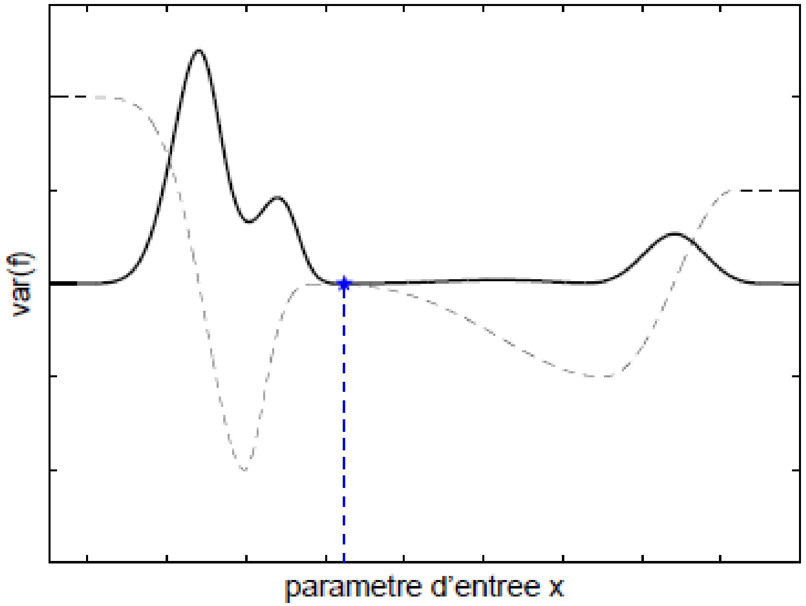
- **Mean of the objective w.r.t. the uncertain variables**
  (≠ minimize the objective with uncertain variables fixed to their mean value)
- **Variance of the objective w.r.t. the uncertain variables**
  (minimal risk approach)
- **Worst case of the objective w.r.t. the uncertain variables**
- **A quantile of the objective**
  (*e.g.* 75% of the smallest values of the objective)

ROBUST OPTIMIZATION CRITERIA

Mean



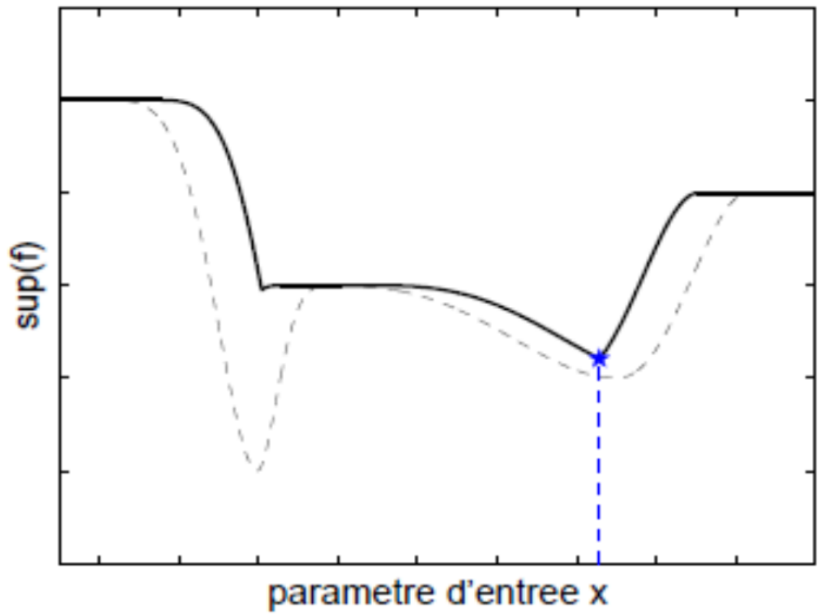
Variance



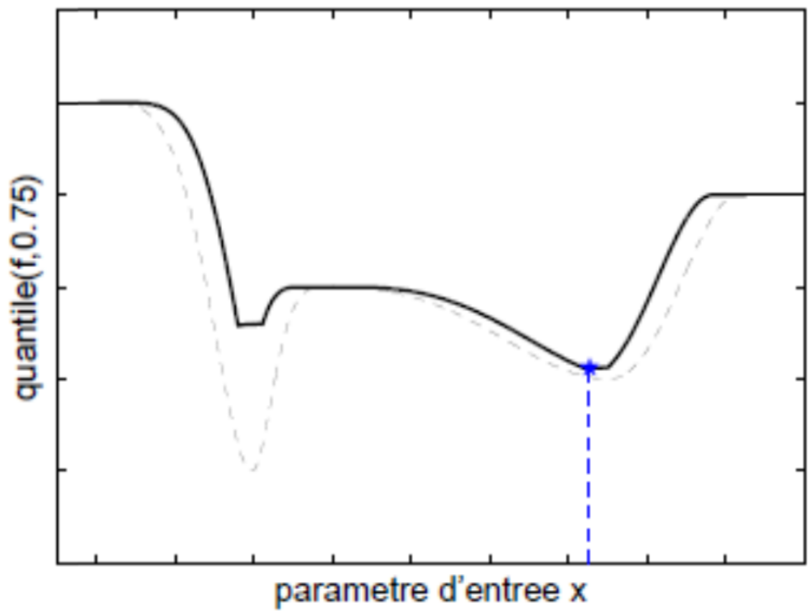
- Aggregated objective: mean + variance: compromise **gain/risk**

ROBUST OPTIMIZATION CRITERIA

Worst case



Quantile (75%)



- ➢Guaranteed value: no risk !
- ➢Often too restricted in practice

- ➢Quantile: a controlled compromise
- ➢But more complex too compute

ROBUST OPTIMIZATION

- General formulation

Robust criterion

Quantity of interest (simulator output)

$$\min_x G(f(x,u))$$

Controllable variables

Uncertain parameters

- Robust criterion computation requires often a sampling of uncertain variables (no closed-form expression in general)

  → Monte-Carlo sampling = expensive in simulations

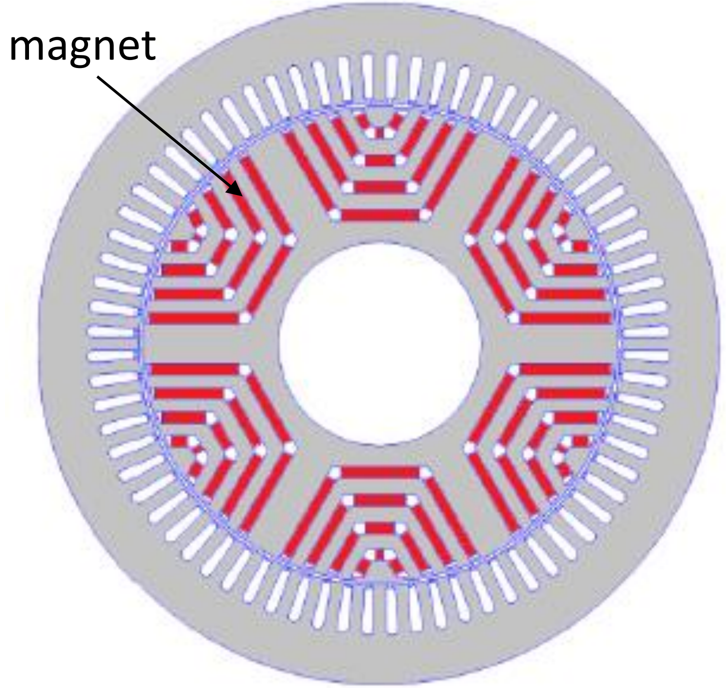
$$G(f(x,.)) = E_u[f(x,u)] \approx \frac{1}{N}\sum_{i=1}^{N} f(x,u_i)$$

- Define approximations
  → meta-models / response surfaces to reduce the simulation cost and possibly even obtain a closed-form expression of G.

CHALLENGES IN DFO

- **Dealing with mixed continuous and discrete variables**
  - Extension of trust region derivative free method
  - Extension of surrogate optimization based on kriging (EGO or Bayesian Optimization)
- **Dealing with uncertainties**
  - 1 practical example
  - See Didier Lucor's course: introduction to optimization under uncertainty

DFO WITH UNCERTAINTIES



Rotor of an electric engine

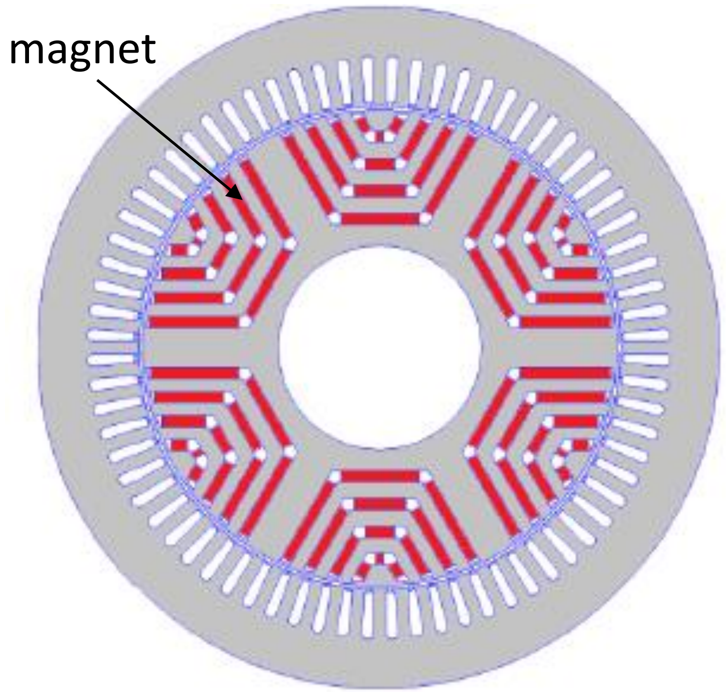
Study by A. Reyes Reyes et al.

- Electric machines for automotive application:
  study of the impact of material deterioration for optimal design

- Magnet characteristics may vary and these variations
  should be taken into account during the design of the machine

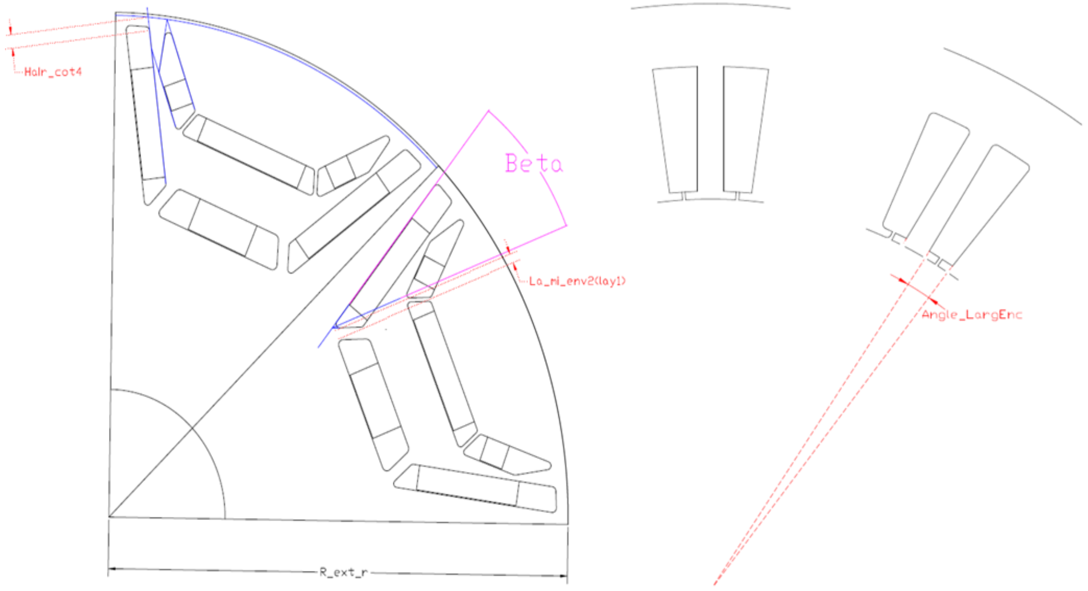
- Optimize the geometry parameters of the machine to maximize power and torque and limit the torque ripple
  - Multi-objective constrained optimization

DFO WITH UNCERTAINTIES: A CASE STUDY



Rotor of an electric engine

- ~10 design parameters: angles, lengths, …
- Objectives to be optimized
  - Maximal torque
  - Ripple of maximal torque
  - Maximal power for all rotation speeds
  - Power at a given rotation speed (14000 rpm)
- Simulation via FEMM software
  (Finite Element Method Magnetics)
  ~ 5-20mn per simulation points



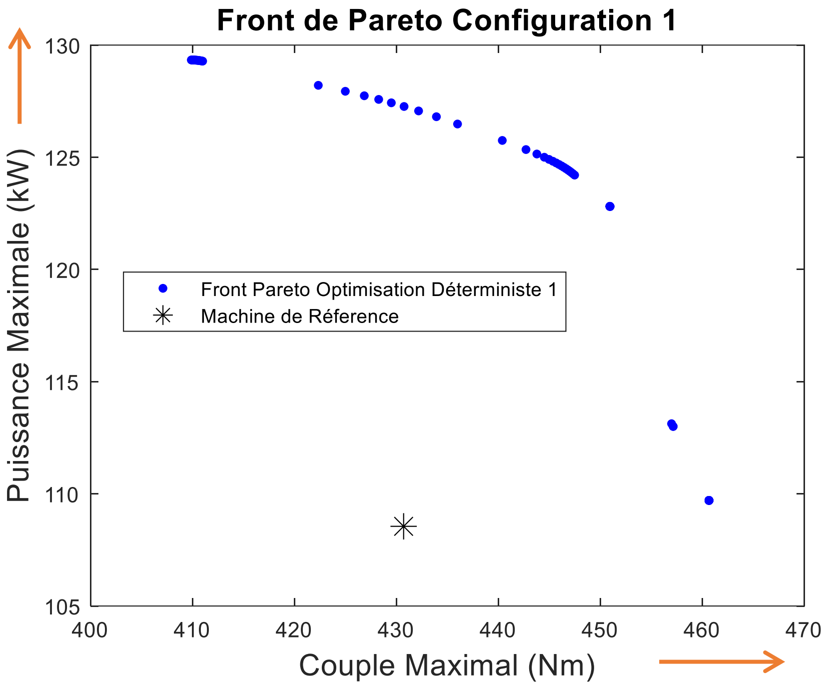
APPLIED METHODOLOGY

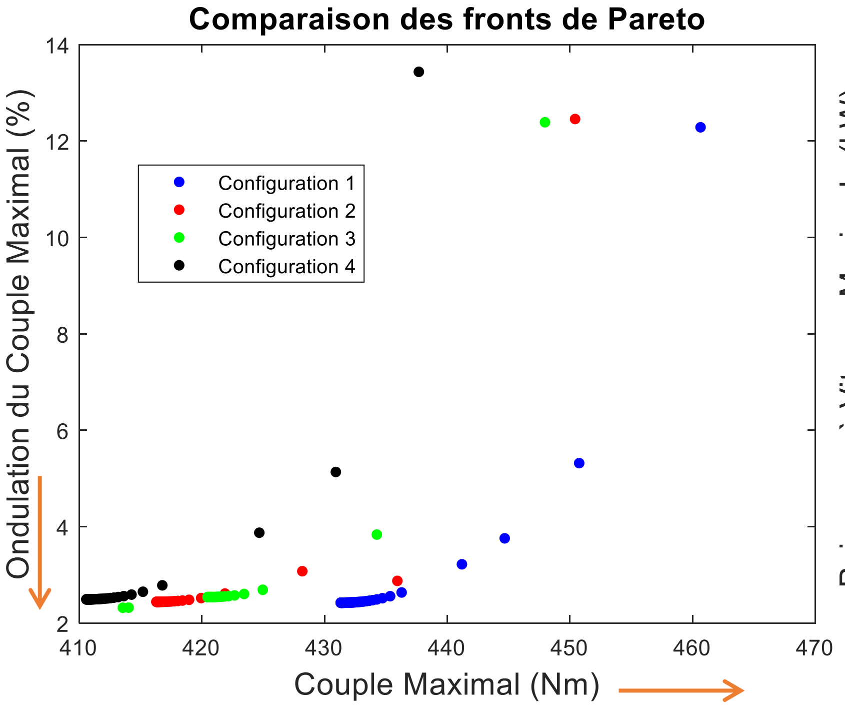
- First study: deterministic optimizations with 4 configurations associated with 4 different magnet characteristics

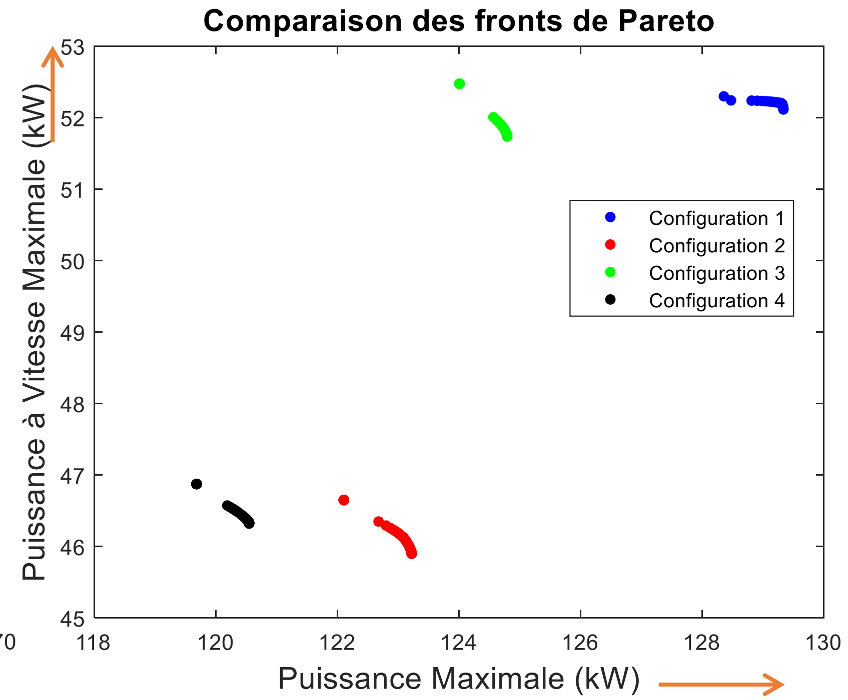
Methodology applied by the engineers

- “Space filling” design of experiments (Latin Hypercube Sampling):
  ~150 simulated points (training points)
  ~80 (validation)
- Build response surfaces of the simulator outputs with Gaussian process models
  Check the prediction errors with the validation points
- Multi-objective optimization based on these response surface
- Validation of the optima with final simulations

4 DETERMINISTIC OPTIMIZATIONS







- Compromise between torque and power
- Test 4 configurations with different deteriorated materials
  - The maximal power is very sensitive to the material degradation (- 10 %)

ROBUST OPTIMIZATION

- Introduce additional random variables in the optimization
  - **Robust configuration 1 :** $\alpha \sim \mathcal{U}(0,1)$ et $\beta \sim \mathcal{U}(0,6.5)$.
  - **Robust configuration 2 :** PontRad1$\sim \mathcal{U}(2,6, \ \ 2,64)$, PontRad2 $\sim \mathcal{U}(0,9 , \ \ 0,94)$, PontRad 3 $\sim \mathcal{U}(0,5, \ \ 0,54)$ et PontTang $\sim \mathcal{U}(0,5, \ \ 0,54)$.
  - **Robust configuration 3 :** $\alpha \sim \mathcal{U}(0,1)$, $\beta \sim \mathcal{U}(0,6.5)$, PontRad1$\sim \mathcal{U}(2,6, \ \ 2,64)$, PontRad2 $\sim \mathcal{U}(0,9 , \ \ 0,94)$, PontRad 3 $\sim \mathcal{U}(0,5, \ \ 0,54)$ et PontTang $\sim \mathcal{U}(0,5, \ \ 0,54)$

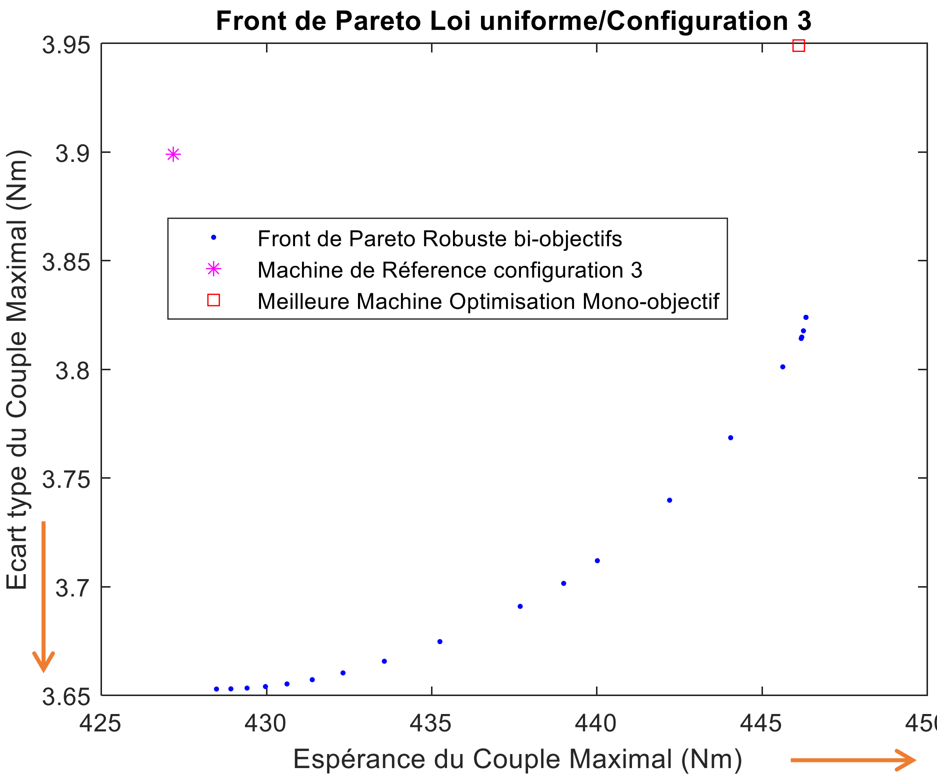
- Robust optimization

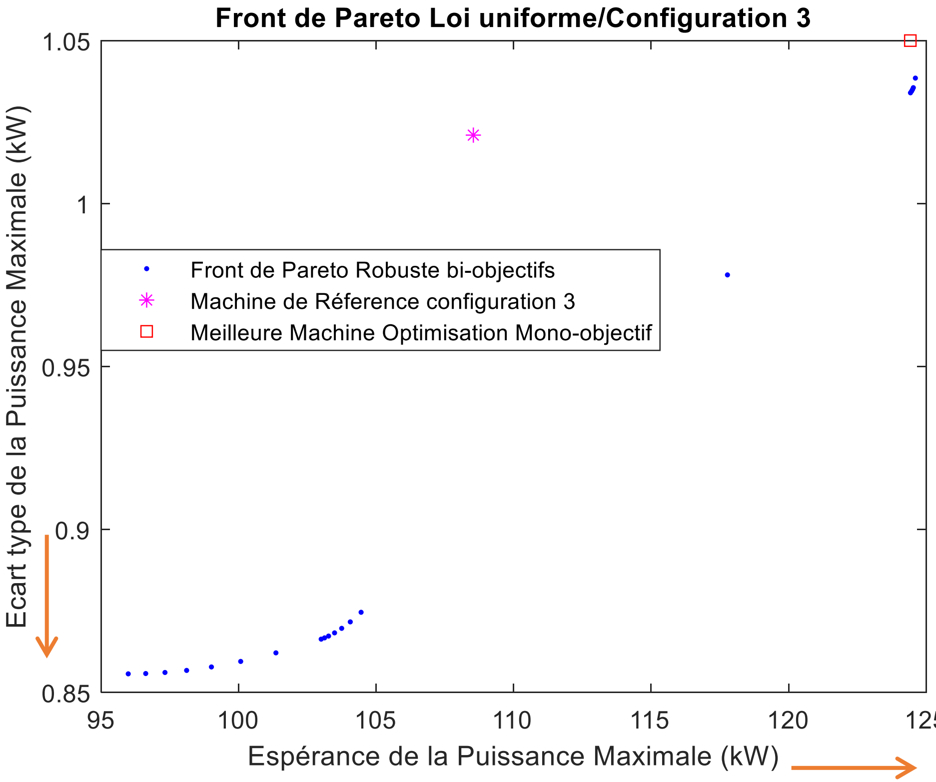
$$\min_{x \in X} \left[\mathbb{E}_U[f_i(x,u)], \sqrt{\mathbb{V}_U[f_i(x,u)]} \right]$$

with u the vector of uncertain variables

- Minimize the expectation of the objective with respect to the uncertain variables together with minimizing the variance of the objective
- Same model-based methodology as deterministic optimizations

ROBUST OPTIMIZATION





- Compromise between optimizing the expectation and minimizing the variance of the objective
- The robust optimization provides solutions that have better performances and that are more robust to material degradations compared to reference solution
- Next steps on this application: EGO-type optimization (adaptive design) with multi-objectives and constraints

DFO WITH UNCERTAINTIES

Other formulations of optimization with uncertainties

- RBDO (reliability based design optimization) or chance constrained optimization

$$\min_{x \in \mathrm{X}} \mathbb{E}_U[f(x,u)]$$
$$\text{s.t. } P_U[g(x,u) \leq S] \geq 1 - \varepsilon$$

- Robust inversion: find feasible solutions

$$\Gamma = \{x \in \mathrm{X} \mid \mathbb{E}_U[g(x,u)] \leq S\}$$
$$\Gamma = \{x \in \mathrm{X} \mid P_U[g(x,u) \leq S] \geq 1 - \varepsilon\}$$

- Computation of $\mathbb{E}_U$ and P_U is expensive in terms of simulations (Monte-Carlo simulations) (in general, no closed-form formula)
  - ➢ Use surrogate models of the simulator with respect to both variables (x, u)

Thank you for your attention !

Thanks to the organizers !

Баяртай

phD positions at IFPEN
https://www.ifp-school.com/en/theses

delphine.sinoquet@ifpen.fr
https://www.ifpenergiesnouvelles.fr/page/delphine-sinoquet