# Bayesian multi-objective optimization with constraints

## Application to the design of a commercial aircraft environment control system

### Atelier du GdR MASCOT NUM, 13 mai 2016, IHP, Paris

Taking into acount stochastics in optimization problems

Paul FELIOT (IRT SystemX), Julien BECT (L2S), Emmanuel VAZQUEZ (L2S)

## 1 - The multi-objective optimization problem

Multi-objective optimization problem with constraints:

$$\begin{cases} \text{Minimize} & f(x) \\ \text{Subject to} & x \in \mathbb{X} \quad \text{and} \quad c(x) \leq 0 \end{cases}$$

with

- $\mathbb{X} \subset \mathbb{R}^d$ → search space
- $f = (f_j)_{1 \leq j \leq p}$ → vector of objective functions to be minimized
- $c = (c_i)_{1 \leq i \leq q}$ → vector of constraint functions

- Sequentially approximate the set of feasible and non-dominated solutions:

$$\Gamma = \{x \in \mathbb{X} : c(x) \leq 0 \text{ and } \nexists x' \in \mathbb{X} \text{ such that } c(x') \leq 0 \text{ and } f(x') \prec f(x)\}$$

Assumptions:

- Both $f$ and $c$ are smooth, nonlinear, expensive-to-evaluate functions (black-box)
- The available simulation budget is very limited ($\approx 10d$)

# 1 - Positioning and contributions

Positioning:

- We adopt a Bayesian approach to this optimization problem.
- We focus on highly constrained problems. Tipically it is assumed that no feasible observation is available at the start of the optimization procedure.

Main contributions:

- Proposal of an Expected Improvement (EI) criterion for the constrained multi-objective optimization problem.
- Study on the use of Sequential Monte Carlo techniques for the optimization of improvement based sampling criteria.
- Implementation of the associated algorithm, which we call BMOO.

# 1 - Publications

Associated publications:

- P. Feliot, J. Bect, and E. Vazquez. *A bayesian approach to constrained single- and multi-objective optimization*. Journal of Global Optimization, 2016.
- P. Feliot, Y. Le Guennec, J. Bect, and E. Vazquez. *Design of a commercial aircraft environment control system using Bayesian optimization techniques*. ENGOPT, 2016 (to appear).

01 Introduction

## 02 The BMOO algorithm

- Expected Hypervolume Improvement (EHVI) criterion
- Extending the EHVI to handle constraints
- Criterion calculation and optimization
- Structure of the algorithm
- Illustration

03 Environment control system

04 Optimization of the system

05 Conclusions

## 2 - Expected Hypervolume Improvement criterion

- We first consider the multi-objective optimization problem without constraints.

Expected Hypervolume Improvement (EHVI) sampling criterion:

- Given evaluation results $(f(X_1), \ldots, f(X_n))$, with $f(X_i) = (f_1(X_i), \ldots, f_p(X_i))$, $1 \leq i \leq n$, define:

$$\begin{cases} \mathbb{B} & = & \left\{y \in \mathbb{R}^p;\ y \leq y^{\text{upp}}\right\}, y^{\text{upp}} \in \mathbb{R}^p \\ H & = & \{y \in \mathbb{B}; \exists x \in \mathbb{X}, f(x) \prec y\} \\ H_n & = & \{y \in \mathbb{B}; \exists i \leq n,\ f(X_i) \prec y\} \end{cases}$$

- Loss function:

$$\varepsilon_n(\underline{X}, f) = |H \setminus H_n| \ ,$$

- Hypervolume improvement:

$$\begin{aligned} I_n\left(X_{n-1}\right) & = \varepsilon_n(\underline{X}, f) - \varepsilon_{n+1}(\underline{X}, f) \\ & = |H_{n+1}| - |H_n| \end{aligned}$$

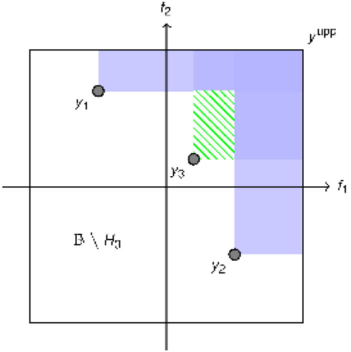



Figure: Hypervolume improvement yielded by the observation of $y_3$

## 2 - Expected Hypervolume Improvement criterion

Expected Hypervolume Improvement (EHVI) sampling criterion:

- Assume a vector-valued Gaussian random process model $\xi = (\xi_1, \ldots, \xi_p)$ of $f = (f_1, \ldots, f_p)$
- The EHVI[a] comes from taking the expectation of the improvement with respect to the posterior probability of $\xi$:

$$\begin{aligned} \mathrm{EI}_n(x) &= \mathbb{E}_n\left(|H_{n+1}| - |H_n|\right) \\ &= \mathbb{E}_n\left(\int_{\mathbb{B}\setminus H_n} \mathbb{1}_{\xi(x)\prec y}\,\mathrm{d}y\right) \\ &= \int_{\mathbb{B}\setminus H_n} \mathbb{P}_n\left(\xi(x) \prec y\right)\mathrm{d}y \end{aligned}$$

- The next sample is chosen as $X_{n+1} = \underset{x\in\mathbb{X}}{\operatorname{argmax}}\ \mathrm{EI}_n(x)$, which requires solving an auxiliary optimization problem at each iteration.

[a] References: Emmerich [2005], Emmerich et al. [2006]

## 2 - Constraint handling

- We now consider the multi-objective problem with constraints.

Constraint handling in the literature:[b]

- Emmerich [2005] and Couckuyt et al. [2014] propose to multiply the EHVI by $\mathbb{P}_n(\xi_c(x) \leq 0)$
  - This corresponds to $I_n = (|H_{n+1}| - |H_n|) \mathbb{1}_{c(x)<0}$ under the independence assumption.
  - The progress on constraint resolution is not measured.
- Gramacy et al. [2015] recently proposed to use the *Augmented Lagrangian* approach.
- Our proposal is to extend the Pareto domination rule to handle the constraints.

[b] References: See e.g. Conn et al. [1991], Sasena [2002], Parr et al. [2012], Gelbart [2015]

## 2 - Extended domination rule

Extended domination rule definition:[c]

- Denote $\mathbb{Y}_o = \mathbb{R}^p$, the objective space and $\mathbb{Y}_c = \mathbb{R}^q$ the constraint space, and let $\mathbb{Y} = \mathbb{Y}_o \times \mathbb{Y}_c$.
- We shall say that $y_1 \in \mathbb{Y}$ dominates $y_2 \in \mathbb{Y}$, which will be written as $y_1 \lhd y_2$, if $\psi(y_1) \prec \psi(y_2)$, where $\prec$ is the usual Pareto domination rule.

$$\begin{aligned} \psi : \quad \mathbb{Y}_o \times \mathbb{Y}_c \quad &\rightarrow \quad \overline{\mathbb{R}}^p \times \mathbb{R}^q \\ (y_o, y_c) \quad &\mapsto \quad \begin{cases} (y_o, 0) & \text{if } y_c \leq 0, \\ (+\infty, \max(y_c, 0)) & \text{otherwise.} \end{cases} \end{aligned}$$

- Properties of the extended domination rule:
  - Feasible solutions are compared wrt their objectives values.
  - Non-feasible solutions are compared wrt their component wise constraint violations.
  - Feasible solutions always dominate non-feasible ones.

[c] References: Ray et al. [2001], Fonseca and Fleming [1998], Oyama et al. [2007]

# 2 - Expected extended hypervolume improvement

- Similarly to the EHVI derivation, we can define the following sets.

$$\begin{cases} \mathbb{B} & = & \mathbb{B}_o \times \mathbb{B}_c, \text{ where } \mathbb{B}_o \subset \mathbb{Y}_o \text{ and } \mathbb{B}_c \subset \mathbb{Y}_c \text{ are bounded hyper-rectangles} \\ H & = & \{y \in \mathbb{B};\ \exists x \in \mathbb{X}\ (f(x), c(x)) \lhd y\} \\ H_n & = & \{y \in \mathbb{B};\ \exists i \leq n\ (f(X_i), c(X_i)) \lhd y\} \end{cases}$$

- Expected hypervolume improvement criterion with constraints:

$$\begin{aligned} \mathrm{EI}_n(x) &= \mathbb{E}_n\left(|H_{n+1}| - |H_n|\right) \\ &= \mathbb{E}_n\left(\int_{\mathbb{B}\setminus H_n} \mathbb{1}_{\xi(x) \lhd y}\, \mathrm{d}y\right) \\ &= \int_{\mathbb{B}\setminus H_n} \mathbb{P}_n\left(\xi(x) \lhd y\right) \mathrm{d}y, \end{aligned}$$

- where $\xi = (\xi_o, \xi_c)$, with $\xi_o = (\xi_{o,1}, \ldots, \xi_{o,p})$ and $\xi_c = (\xi_{c,1}, \ldots, \xi_{c,q})$.

# 2 - Criterion calculation and optimization

Criterion calculation:

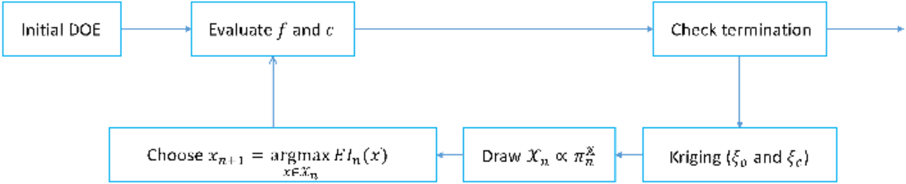
- The criterion takes the form of a multi-dimensional integral for which no closed form solution exists.
- The BMOO algorithm implements a Monte Carlo approximation of the integral using sequential Monte Carlo techniques.

Criterion optimization:

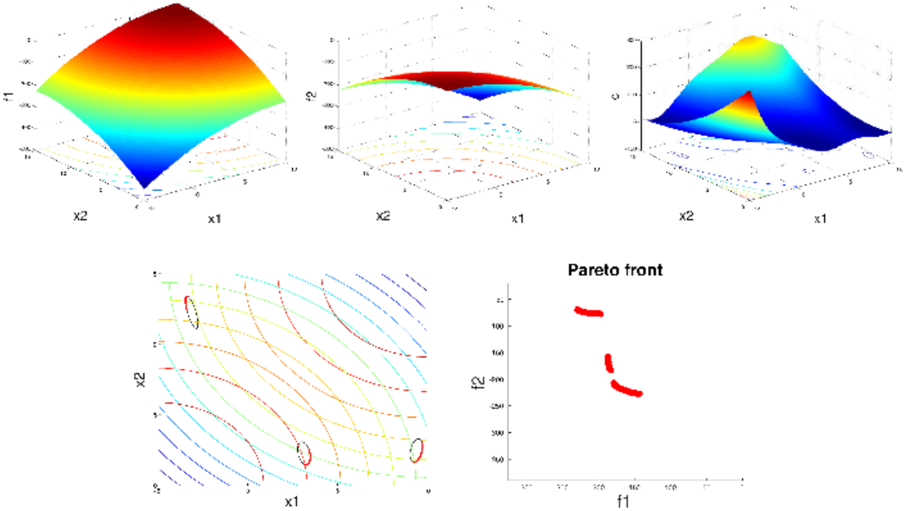
- The criterion optimization is performed using sequential Monte Carlo techniques as well.
- At time $n$, the next sample is chosen out of a population $\mathcal{X}_n$ of particles distributed according to a density of interest $\pi_n^{\mathbb{X}}$.

## 2 - Structure of the algorithm

Structure of the algorithm:

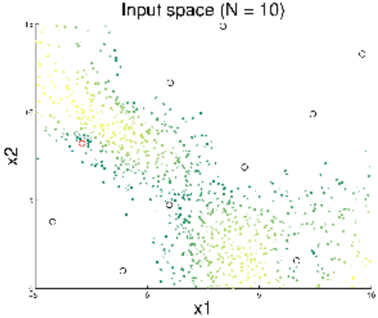


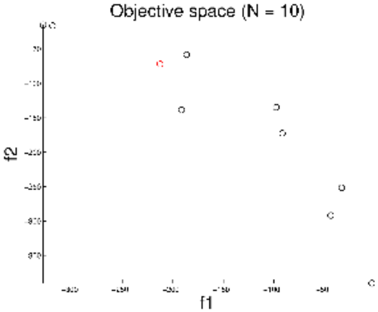
## 2 - Illustration on a toy example



# 2 - Illustration on a toy example

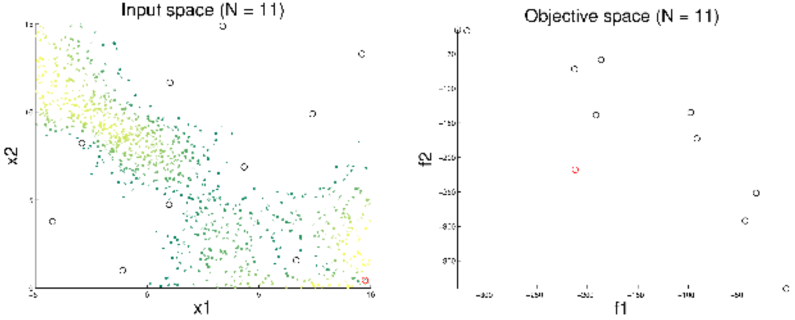
Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



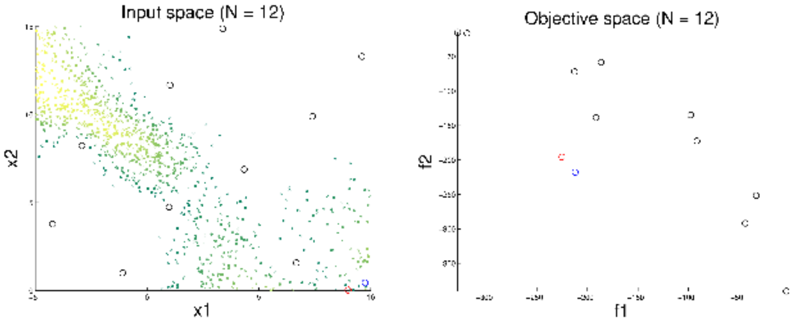
# 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



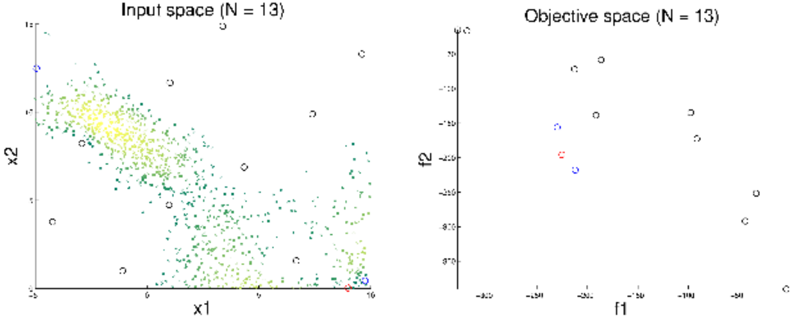
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



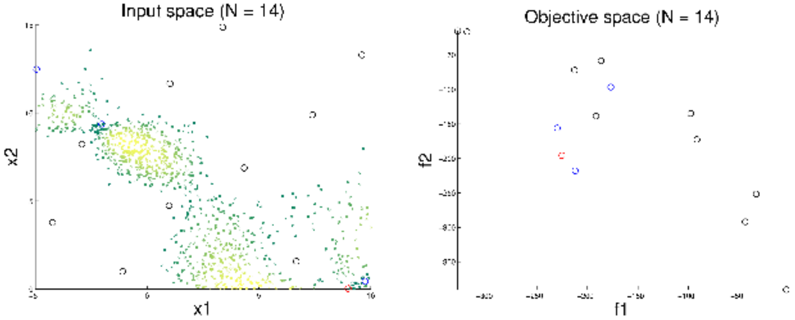
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



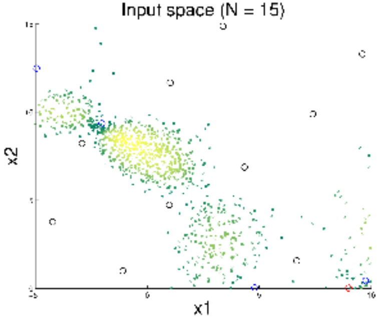
## 2 - Illustration on a toy example

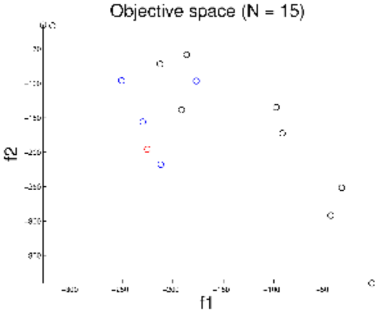
Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



## 2 - Illustration on a toy example

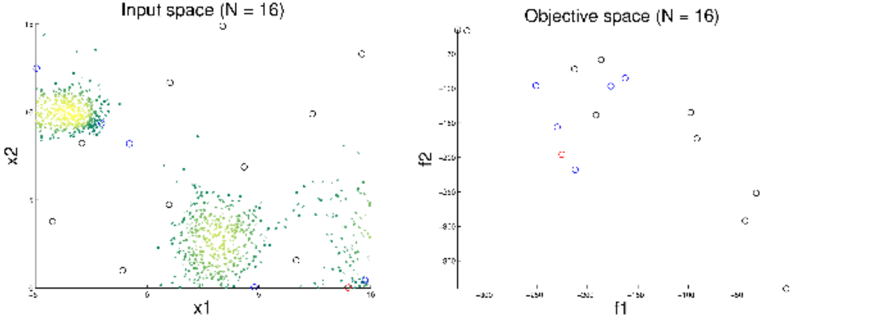
Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



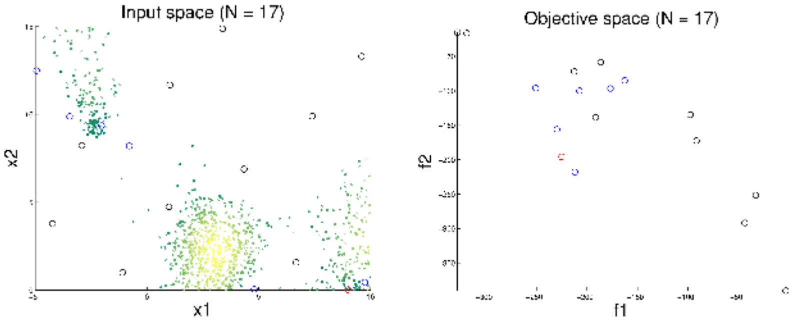
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



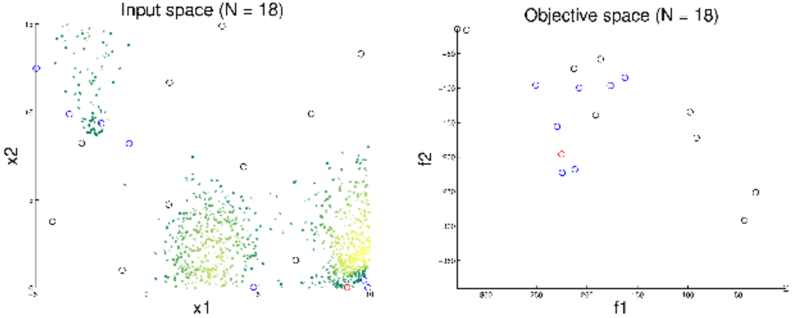
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



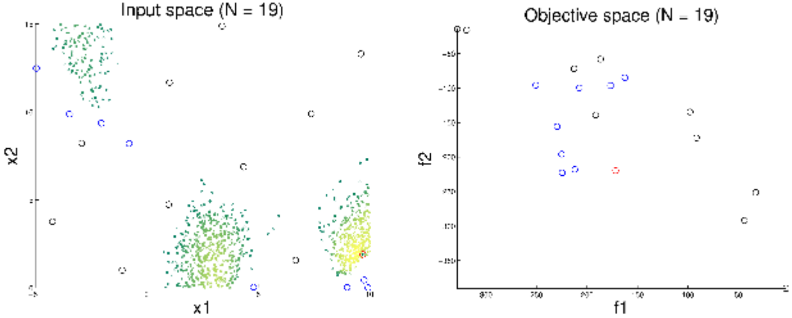
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



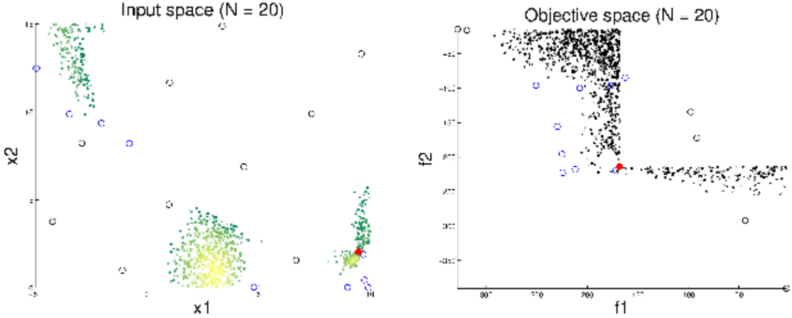
# 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



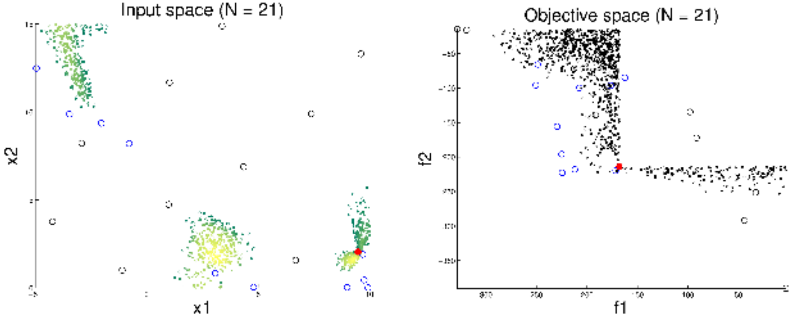
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



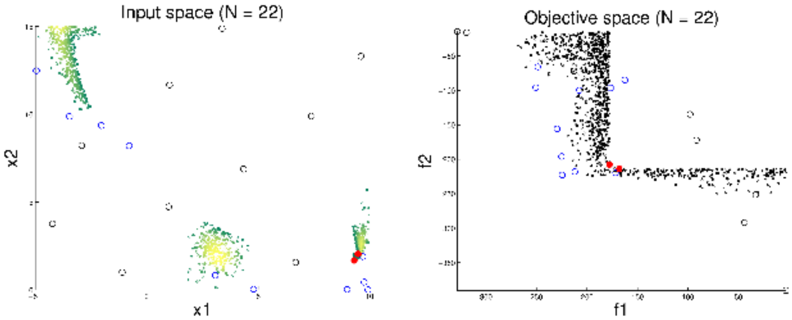
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



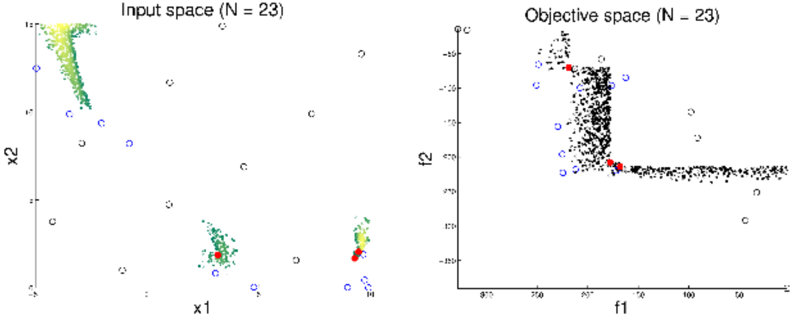
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



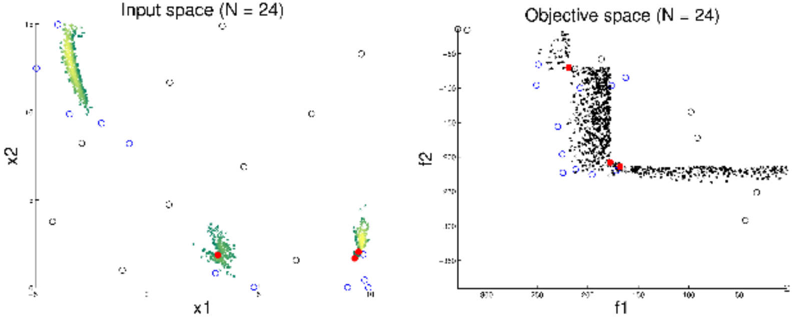
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



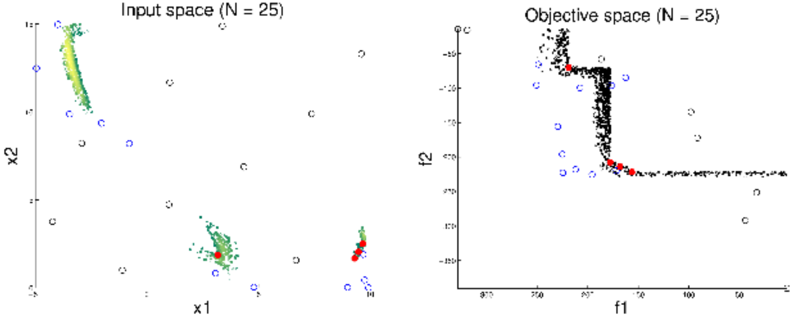
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



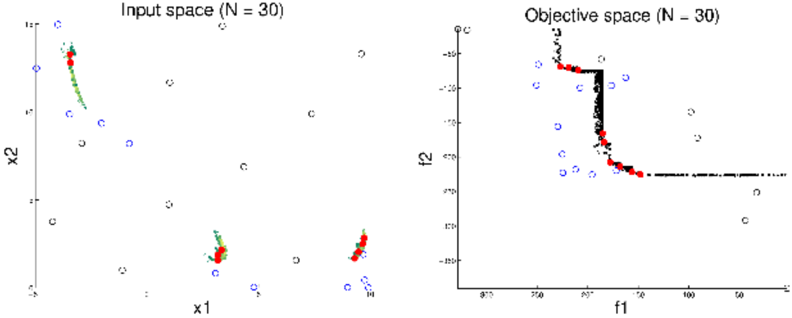
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



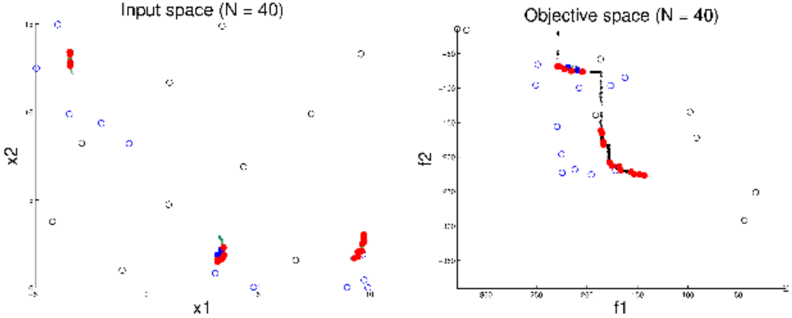
## 2 - Illustration on a toy example

Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



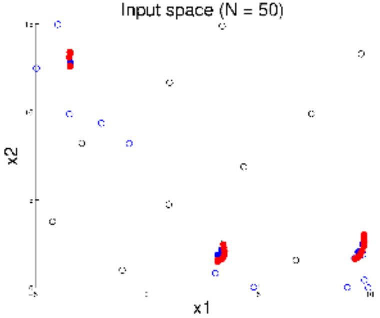
## 2 - Illustration on a toy example

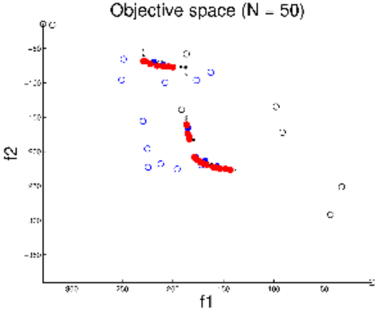
Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



## 2 - Illustration on a toy example

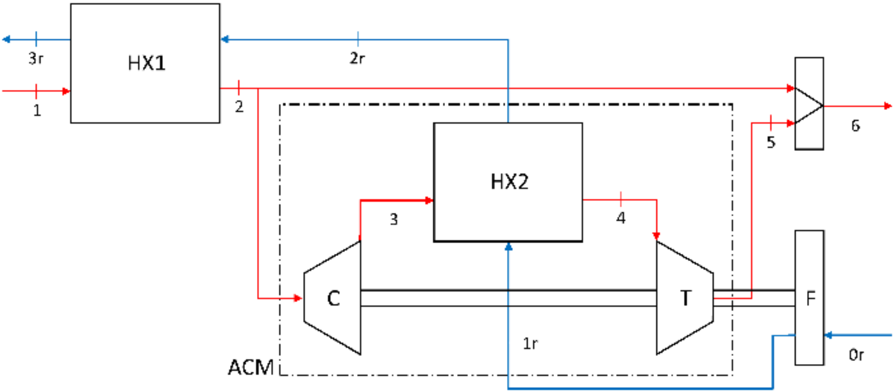
Illustration with $\pi_n^{\mathbb{X}}(x) \propto \mathbb{P}_n(\xi(x) \in (\mathbb{B} \setminus H_n))$ and $\pi_n^{\mathbb{Y}}(y) \propto L_2^{opt}(y)$



# 3 - Presentation of the problem

Architecture of the ECS:



## 3 - Presentation of the problem

Sizing scenario:

- Aircraft-on-ground, full of passengers, with 50°C outside.

$$\mathcal{P}_{HT} = \mathcal{P}_{out} + \mathcal{P}_{eq} + N_{pax}\mathcal{P}_{pax} + N_{crew}\mathcal{P}_{crew}$$

- Keep the cabin temperature at $T_c = 24^\circ$C.

$$\mathcal{P}_{HT} \leq \dot{m}c_p\left(T_c - T_5\right)$$

- Keep the cabin pressure close to the atmospheric pressure.

$$P_{min} \leq P_5 \leq P_{max}$$

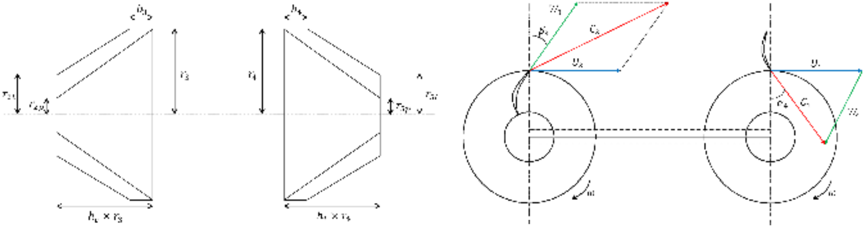
- The air injected into the cabin must lie between $T_{min} = 15^\circ$C and $T_{max} = 25^\circ$C.

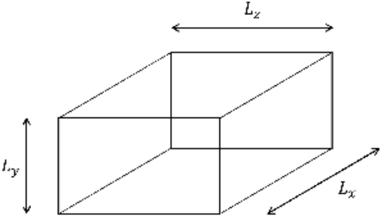
Objectives:

- Minimize the mass of the system.
- Minimize its entropy generation rate.

# 3 - One dimensional model of the system

Parametrization:





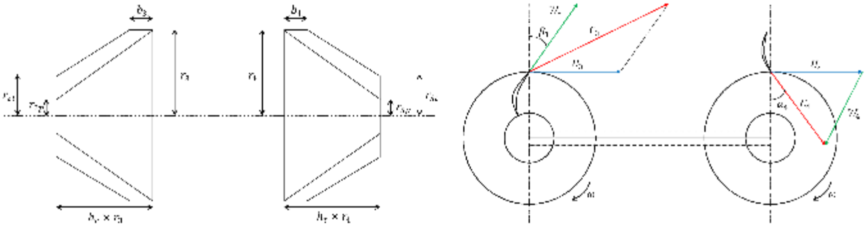

- 18 variables: $\dot{m}$, $\dot{m}_r$, $r_3$, $r_{2p}$, $r_{2t}$, $b_3$, $\beta_3$, $r_4$, $r_{5p}$, $r_{5t}$, $b_4$, $\alpha_4$, $L_{x1}$, $L_{y1}$, $L_{z1}$, $L_{x2}$, $L_{y2}$, $L_{z2}$.

# 3 - One dimensional model of the system

- The system is ruled by a non-linear system of 13 equations with 13 unknowns: $T_{t2}, T_{t3}, T_{t4}, T_{t5}, T_{t2r}, T_{t3r}, P_{t2}, P_{t3}, P_{t4}, P_{t5}, P_{t2r}, P_{t3r}$ and $\omega$.

$$\begin{cases} \dot{m}c_p(T_{t1} - T_{t2}) & = & \dot{m}_r c_p(T_{t3r} - T_{t2r}) \\ \dot{m}c_p(T_{t3} - T_{t4}) & = & \dot{m}_r c_p(T_{t2r} - T_{t1r}) \\ c_p(T_{t1} - T_{t2}) & = & \epsilon_1 c_p(T_{t1} - T_{t2r}) \\ c_p(T_{t3} - T_{t4}) & = & \epsilon_2 c_p(T_{t3} - T_{t1r}) \\ P_{t2} - P_{t1} & = & \Delta P_{HX} \\ P_{t4} - P_{t3} & = & \Delta P_{HX} \\ P_{t3} & = & P_{t2}\left(1 + \eta_C \frac{T_{t3} - T_{t2}}{T_{t2}}\right)^{\frac{\gamma}{\gamma-1}} \\ P_{t5} & = & P_{t4}\left(1 + \frac{1}{\eta_T}\frac{T_{t5} - T_{t4}}{T_{t4}}\right)^{\frac{\gamma}{\gamma-1}} \\ \dot{W}_C & = & \dot{m}\left(r_3^2\omega^2 - \frac{\dot{m}\tan(\beta_3)}{2\pi\rho b_3}\omega\right) \\ \dot{W}_T & = & -\frac{\dot{m}^2\tan(\alpha_4)}{2\pi\rho b_4}\omega \\ \dot{W}_C + \dot{W}_T + \frac{1}{\eta_F}\frac{\dot{m}_r^3}{2\rho^2 A_r^2} & = & 0 \\ \dot{W}_C & = & \eta_C \dot{m}c_p(T_{t3} - T_{t2}) \\ \dot{W}_T & = & \frac{1}{\eta_T}\dot{m}c_p(T_{t5} - T_{t4}) \end{cases}$$

# 3 - Formulation of an optimization problem



## Design constraints:

$$
\begin{array}{llllll}
d_1 & : & \dot{m} & \leq & \dot{m}_r\,, & \\
d_2 & : & b_3 & \leq & h_c r_3\,, & \\
d_3 & : & b_4 & \leq & h_t r_4\,, & \\
d_{4-5} & : & r_{2p} & \leq & r_{2t} & \leq \quad r_3\,, \\
d_{6-7} & : & r_{5p} & \leq & r_{5t} & \leq \quad r_4\,, \\
d_8 & : & \Delta & \geq & 0\,, & \\
d_9 & : & \dfrac{\tan(\beta_3)}{b_3} & \geq & -\dfrac{\tan(\alpha_4)}{b_4}\,. &
\end{array}
$$

# 3 - Formulation of an optimization problem

Simulation constraints:

$$
\begin{array}{lcccccc}
c_{1-2} & : & T_{min} & \leq & T_5 & \leq & T_{max}, \\
c_{3-4} & : & P_{min} & \leq & P_5 & \leq & P_{max}, \\
c_{5-6} & : & 0.5 & \leq & \epsilon_1 & \leq & 0.9, \\
c_{7-8} & : & 0.5 & \leq & \epsilon_2 & \leq & 0.9, \\
c_9 & : & C_2 & \leq & 0.95\sqrt{\gamma R T_2}, & & \\
c_{10} & : & C_3 & \leq & 0.95\sqrt{\gamma R T_3}, & & \\
c_{11} & : & C_4 & \leq & 0.95\sqrt{\gamma R T_4}, & & \\
c_{12} & : & C_5 & \leq & 0.95\sqrt{\gamma R T_5}, & & \\
c_{13} & : & r_3\omega & \leq & \sqrt{\gamma R T_3}, & & \\
c_{14} & : & r_4\omega & \leq & \sqrt{\gamma R T_4}, & & \\
c_{15} & : & \mathcal{P}_{\mathcal{H}\mathcal{T}} & \leq & \dot{m} c_p \left(T_c - T_5\right). & &
\end{array}
$$

# 3 - Formulation of an optimization problem

Summary:

- 18 variables: $\dot{m}$, $\dot{m}_r$, $r_3$, $r_{2p}$, $r_{2t}$, $b_3$, $\beta_3$, $r_4$, $r_{5p}$, $r_{5t}$, $b_4$, $\alpha_4$, $L_{x1}$, $L_{y1}$, $L_{z1}$, $L_{x2}$, $L_{y2}$, $L_{z2}$.
- Bound constraints on the variables.
- 9 design constraints: $d_{1-9}$.
- 15 simulation constraints: $c_{1-15}$.
- 2 objectives: $\mathcal{M}$ and $\dot{S}$.

Remarks:

- The design space is not hypercubic.
- Possibility of simulation failures (hidden constraints).

## 4 - Simulation failures management

Hidden constraints management:

- There are two cases where the model fails to produce a result:
  - The system can not be inverted.
  - The air flowrate becomes supersonic.
- This is taken into account in BMOO by multiplying the Expected improvement by a probability of observability[d].
  - A nearest-neighbour classifier is built on the observed/non-observed data to this end.
- The sampling density $\pi_n^{\mathbb{X}}$ is also multiplied by the probability of observability.

[c] Reference: Gramacy and Lee [2011]

# 4 - Handling non-hypercubic design spaces

Initial DoE:

- Denote $\mathbb{D}$ the domain delimited by the bound constraints and define:

$$\mathbb{X} = \mathbb{D} \setminus \{x \in \mathbb{D}; d(x) > 0\}$$

- If $|\mathbb{X}|/|\mathbb{D}|$ is not too small, a pseudo-maximin design on $\mathbb{X}$ can be achieved using the following process.
    - Sample a large population uniformly on $\mathbb{D}$.
    - Discard particles falling out of $\mathbb{X}$.
    - Prune the remaining particles to augment the maximin distance until the required number of particles remain.

During the optimization:

- The sampling density $\pi_n^{\mathbb{X}}$ is multiplied by $\mathbb{1}_{\mathbb{X}}$.

# 4 - Optimization results

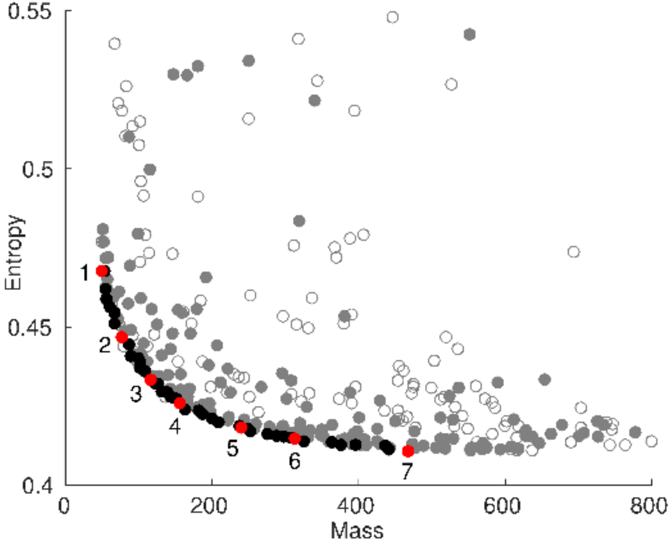
A few statistics:

- Initial DoE:
  - $N_{init} = 90$ samples.
  - 44 simulation failures.
  - No feasible observation

.

- Optimization process:
  - $N_{max} = 500$ samples.
  - 92 additional simulation failures.
  - First feasible point found after 25 iterations.

# 4 - Optimization results



# 4 - Optimization results

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $\dot{m}$ | 2.95 | 2.92 | 2.94 | 2.94 | 2.94 | 2.95 | 2.94 |
| $\dot{m}_r$ | **7.74** | **6.86** | **5.63** | **5.06** | **4.64** | **4.40** | **4.27** |
| $r_{2o}$ | 0.07 | 0.05 | 0.05 | 0.03 | 0.03 | 0.07 | 0.04 |
| $r_{2t}$ | 0.10 | 0.08 | 0.08 | 0.08 | 0.06 | 0.09 | 0.10 |
| $r_3$ | 0.10 | 0.11 | 0.10 | 0.10 | 0.12 | 0.12 | 0.13 |
| $b_3$ | 0.01 | 0.01 | 0.05 | 0.05 | 0.04 | 0.02 | 0.03 |
| $\beta_3$ | 0.36 | 0.74 | 0.97 | -0.16 | 0.61 | 0.94 | 0.48 |
| $r_{5o}$ | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| $r_{5t}$ | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| $r_4$ | 0.10 | 0.10 | 0.11 | 0.12 | 0.11 | 0.10 | 0.11 |
| $b_4$ | 0.02 | 0.02 | 0.04 | 0.02 | 0.04 | 0.03 | 0.03 |
| $\alpha_4$ | 1.04 | 0.50 | 0.89 | 1.01 | 0.44 | 0.79 | 0.30 |
| $L_{x1}$ | 0.67 | 0.65 | 0.68 | 0.68 | 0.63 | 0.69 | 0.70 |
| $L_{y1}$ | 0.65 | 0.68 | 0.61 | 0.67 | 0.67 | 0.66 | 0.65 |
| $L_{z1}$ | **0.03** | **0.04** | **0.07** | **0.12** | **0.17** | **0.20** | **0.32** |
| $L_{x2}$ | 0.66 | 0.69 | 0.66 | 0.66 | 0.70 | 0.68 | 0.69 |
| $L_{y2}$ | 0.69 | 0.53 | 0.68 | 0.65 | 0.65 | 0.68 | 0.65 |
| $L_{z2}$ | **0.03** | **0.06** | **0.09** | **0.10** | **0.17** | **0.25** | **0.36** |

# 5 - Conclusions

- BMOO: A Bayesian optimization algorithm for single- and multi-objective optimization.
  - Designed to address highly constrained problems.
  - Based on a generalized EHVI criterion defined using an extended domination rule.
  - SMC techniques for the criterion calculation and optimization.
  - Hidden constraints handling capability.
  - Non-hypercubic design spaces handling capability.
- The algorithm is applied to the design of a commercial aircraft environment control system with promising results.
- Directions for future work on this application case.
  - Taking into account uncertainties on key parameters of the simulation.
  - Sensitivity analysis on non-hypercubic design spaces.
  - Sensitivity analysis in constrained multi-objective optimization.

# 6 - References I

A. R. Conn, N. I. M. Gould, and P. Toint. A globally convergent augmented lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Numerical Analysis*, 28(2):545–572, 1991.

I. Couckuyt, D. Deschrijver, and T. Dhaene. Fast calculation of multiobjective probability of improvement and expected improvement criteria for pareto optimization. *Journal of Global Optimization*, 60(3):575–594, 2014.

M. Emmerich. *Single- and multiobjective evolutionary design optimization assisted by Gaussian random field metamodels*. PhD thesis, Technical University Dortmund, 2005.

M. Emmerich, K. C. Giannakoglou, and B. Naujoks. Single- and multi-objective evolutionary optimization assisted by Gaussian random field metamodels. *IEEE Transactions on Evolutionary Computation*, 10(4):421–439, 2006.

C. M. Fonseca and P. J. Fleming. Multiobjective optimization and multiple constraint handling with evolutionary algorithms. I. A unified formulation. *IEEE Transactions on Systems, Man and Cybernetics. Part A: Systems and Humans*, 28(1):26–37, 1998.

M. A. Gelbart. *Constrained Bayesian Optimization and Applications*. PhD thesis, Harvard University, Graduate School of Arts and Sciences, 2015.

# 6 - References II

R. B. Gramacy and H. Lee. Optimization under unknown constraints. In *Bayesian Statistics 9. Proceedings of the Ninth Valencia International Meeting*, pages 229–256. Oxford University Press, 2011.

R. B. Gramacy, G. A. Gray, S. Le Digabel, H. K. H. Lee, P. Ranjan, G. Wells, and S. M. Wild. Modeling an augmented lagrangian for blackbox constrained optimization. *Technometrics*, (just-accepted):00–00, 2015.

A. Oyama, K. Shimoyama, and K. Fujii. New constraint-handling method for multi-objective and multi-constraint evolutionary optimization. *Transactions of the Japan Society for Aeronautical and Space Sciences*, 50(167):56–62, 2007.

J. M. Parr, A. J. Keane, A. I. J. Forrester, and C. M. E. Holden. Infill sampling criteria for surrogate-based optimization with constraint handling. *Engineering Optimization*, 44(10): 1147–1166, 2012.

T. Ray, K. Tai, and K. C. Seow. Multiobjective design optimization by an evolutionary algorithm. *Engineering Optimization*, 33(4):399–424, 2001.

M. J. Sasena. *Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations*. PhD thesis, University of Michigan, 2002.

# 06 Additional Material

# 7 - Adaptive procedure to set $\mathbb{B}_c$ and $\mathbb{B}_o$

Adaptive procedure to set $\mathbb{B}_c$ and $\mathbb{B}_o$:

- Assume that $n$ evaluation results $\xi(X_i)$, $1 \leq i \leq n$, are available. Then, we define the corners of $\mathbb{B}_o$ by

$$\begin{cases} y_{o,i,n}^{\text{low}} &= \min\left(\min_{i\leq n} \xi_{o,i}(X_i),\ \min_{x\in\mathcal{X}_n} \hat{\xi}_{o,i,n}(x) - \lambda_o \sigma_{o,i,n}(x)\right), \\ y_{o,i,n}^{\text{upp}} &= \max\left(\max_{i\leq n} \xi_{o,i}(X_i),\ \max_{x\in\mathcal{X}_n} \hat{\xi}_{o,i,n}(x) + \lambda_o \sigma_{o,i,n}(x)\right), \end{cases}$$

for $1 \leq i \leq p$, and the corners of $\mathbb{B}_c$ by

$$\begin{cases} y_{c,j,n}^{\text{low}} &= \min\left(0,\ \min_{i\leq n} \xi_{c,j}(X_i),\ \min_{x\in\mathcal{X}_n} \hat{\xi}_{c,j,n}(x) - \lambda_c \sigma_{c,j,n}(x)\right), \\ y_{c,j,n}^{\text{upp}} &= \max\left(0,\ \max_{i\leq n} \xi_{c,j}(X_i),\ \max_{x\in\mathcal{X}_n} \hat{\xi}_{c,j,n}(x) + \lambda_c \sigma_{c,j,n}(x)\right), \end{cases}$$

for $1 \leq j \leq q$, where $\lambda_o$ and $\lambda_c$ are positive numbers.

# 7 - $L_2^{opt}$ density for EHVI calculation

$L_2^{opt}$ density for EHVI calculation:

- Let $(x_k)_{1 \leq k \leq m_{\mathbb{X}}}$ be a population of candidates on which we want to compute the EHVI value, using a population of particles $(y_i)_{1 \leq i \leq m_{\mathbb{Y}}}$ distributed according to some density $\pi$.

$$\hat{I}_k^{\pi} = \frac{1}{m_{\mathbb{Y}}} \sum_{i=1}^{m_{\mathbb{Y}}} \frac{\mathbb{P}_n(\xi(x_k) \prec y_i)}{\pi(y_i)}$$

- We want to achieve a good approximation for all particles $(x_k)_{1 \leq k \leq m_{\mathbb{X}}}$. Using the $L_2$ norm, we get the following:

$$\begin{aligned} \mathbb{E}\left(\left\|\hat{I}^{\pi} - I\right\|_2^2\right) &= \mathbb{E}\left(\sum_{k=1}^{m_{\mathbb{X}}} (\hat{I}_k^{\pi} - I_k)^2\right) \\ &= \frac{1}{m_{\mathbb{Y}}} \sum_{k=1}^{m_{\mathbb{X}}} \left(\int \frac{\mathbb{P}_n(\xi(x_k) \prec y)^2}{\pi(y)^2} \pi(y) \mathrm{d}y - I_k^2\right) \end{aligned}$$

$$\rightarrow L_2^{opt}(y) \propto \sqrt{\sum_{k=1}^{m_{\mathbb{X}}} \mathbb{P}_n(\xi(x_k) \prec y)^2}$$

# 7 - A density for heavily constrained problems

A density for problems with a lot of constraints:

- Suppose $q = d$, $\mathbb{X} = [-1/2, 1/2]^q$ and $c_j : x = (x_1 \dots, x_q) \mapsto |x_j| - \frac{\varepsilon}{2}$, $\varepsilon \in (0; 1]$.
- Thus, the feasible domain is $C = [-\varepsilon/2, \varepsilon/2]^q$ and the volume of the subset of $\mathbb{X}$ where exactly $k$ constraints are satisfied is

$$V_k \approx \binom{q}{k} \varepsilon^k (1 - \varepsilon)^{q-k}.$$

- Assume moreover that the Gaussian process models are almost perfect, i.e.,

$$\mathbb{P}_n\left(\xi_{c,j}(x) \le 0\right) \approx \begin{cases} 1, & \text{if } c_j(x) \le 0, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

- Further assume $n = 1$ with $X_1 = \left(\frac{1}{2}, \dots, \frac{1}{2}\right)$ so that the probability of improvement $\mathbb{P}_n\left(\xi(x) \in G_1\right)$ is close to one everywhere on $\mathbb{X}$.
- The expected number of particles satisfying exactly $k$ constraints is $m\,V_k$
- If $q$ is large, the particles thus tend to concentrate in regions where $k \approx q\varepsilon$.
- To compensate for the decrease of $V_k$, we suggest using the following modified sampling density:

$$\pi_n^{\mathbb{X}} \propto \mathbb{E}_n\left(K(x)!\, \mathbb{1}_{\xi(x) \in G_n}\right),$$

where $K(x)$ is the number of constraints satisfied by $\xi$ at $x$.

# 7 - Intermediate subsets

Parametric construction of intermediate subsets:

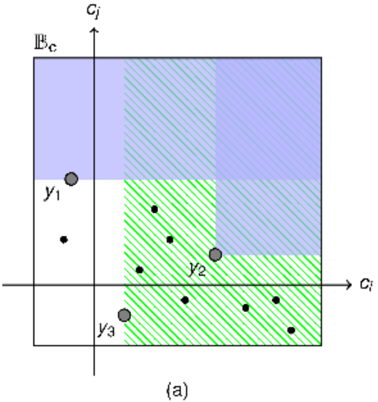



Figure: Procedure to construct intermediate subsets before the observation of a feasible solution.

# 7 - Intermediate subsets

Parametric construction of intermediate subsets:

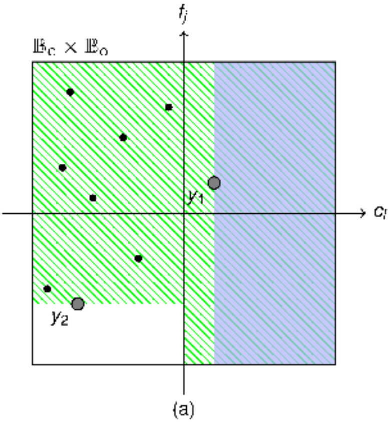



Figure: Procedure to construct intermediate subsets after the observation of a feasible solution.