Targeting Well-Balanced Solutions in Multi-Objective Bayesian Optimization under a Restricted Budget

David Gaudrie, Rodolphe Le Riche, Victor Picheny, Benoît Enaux, Vincent Herbert

Targeting Well-Balanced Solutions in Multi-Objective Bayesian Optimization under a Restricted Budget

David Gaudrie[1,2,3], Rodolphe le Riche[2], Victor Picheny[3], Benoît Enaux[1], Vincent Herbert[1]

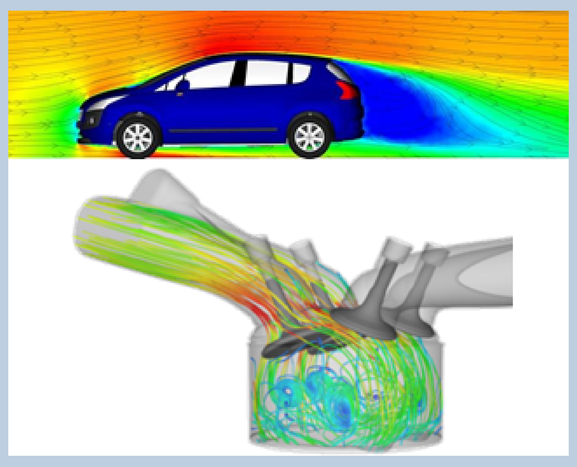
[1]Groupe PSA, [2]CNRS, Mines Saint-Étienne, [3]MIAT, Université de Toulouse, INRA

Industrial context

Multi-objective optimization of high dimensional systems (d up to 40)

$$\min_{\mathbf{x}\in X\subset\mathbb{R}^d} (f_1(\mathbf{x}),..,f_m(\mathbf{x}))$$

Computationally expensive CFD codes (24 hours per simulation) ⇒ optimization under restricted budget ⇒ metamodel-based optimization: Multi-Objective EGO [1]



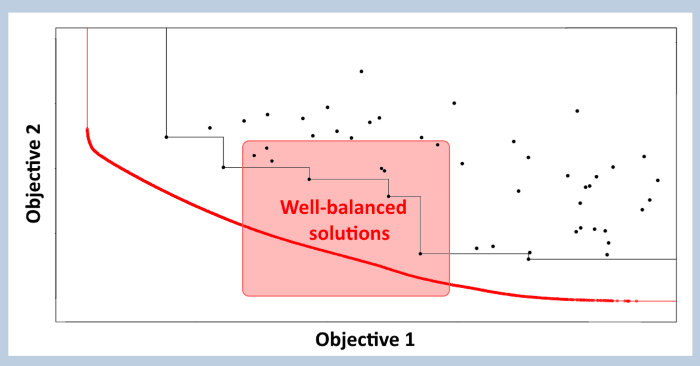
Multi-criteria decision-aid: choice among the optimal solutions made by a Decision Maker

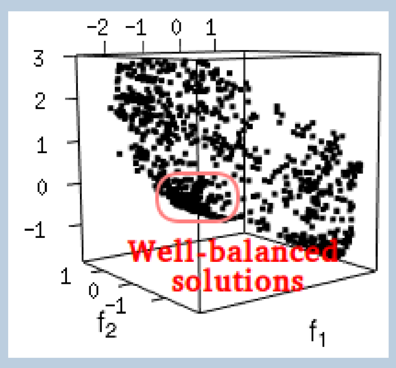
Very tiny budget (≈ 100 evaluations), many objectives ($m \approx$ 6-8) ⇒ impossible for classical MO-EGO approaches to uncover the Pareto Front (growing size of $\mathcal{P_Y}$ with m)

How to obtain several optimal trade-off solutions in spite of the extremely parsimonious use of the computer code, and the multiple conflicting objectives?

Targeting: motivations

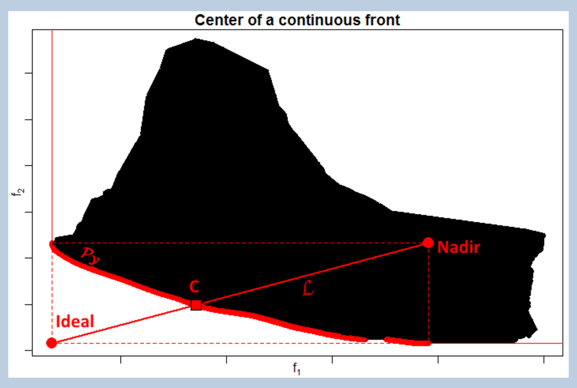
- Restricted budget and large number of objectives ⇒ Uncovering the ~~whole~~ Pareto Front $\mathcal{P_Y}$ in a "region of interest"
- Shrink search to a smaller subset ⇒ faster convergence
- Emphasize solutions that *equilibrate* the objectives: (unknown) central part of the Pareto Front (PF) ⇒ interesting solutions for Decision Makers



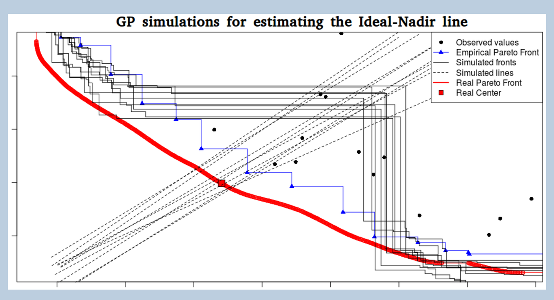


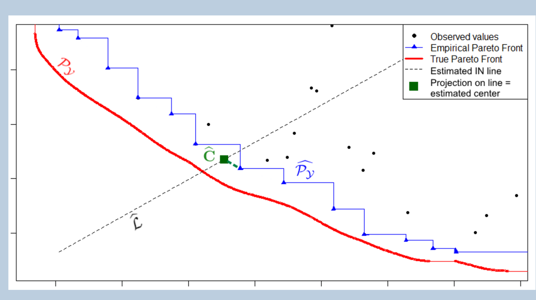
Center of the Pareto Front

Center $\mathbf{C}$ = Projection of closest non-dominated point on Ideal-Nadir line $\mathcal{L}$



- Low dependence to variations of $\mathbf{I}$ and $\mathbf{N}$: $|\frac{\partial C_i}{\partial I_i}|$ and $|\frac{\partial C_i}{\partial N_i}| < 1$ for a continuous front
- Insensitive to a linear scaling of the objectives in a bi-objective case, and when $\mathcal{L}$ intersects $\mathcal{P_Y}$
- Computationally cheap, even for large m
- Estimation: GP simulations emphasizing the edges of the PF ⇒ estimated Ideal, Nadir and $\widehat{\mathcal{L}}$ ⇒ estimated center $\widehat{\mathbf{C}}$



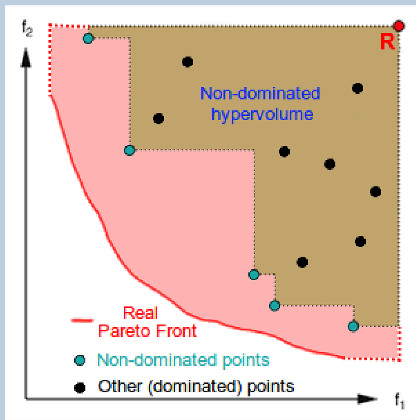


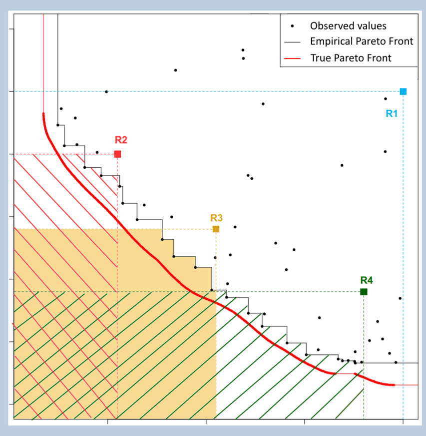
Infill Criteria for targeting the center

$IC = f(Y_1(\cdot),..,Y_m(\cdot);\mathbf{x};\mathbf{\Theta})$: directs the search towards attractive new designs $\mathbf{x}^*$

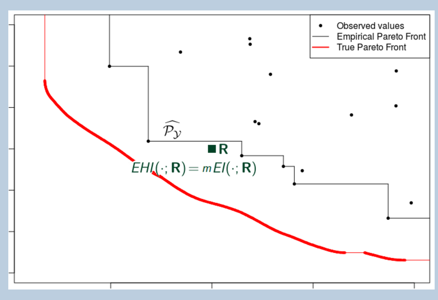
Modify existing IC through $\mathbf{\Theta}$ to direct the search towards the estimated central area

- Hypervolume Indicator [2]: $H(\widehat{\mathcal{P_Y}};\mathbf{R}) = \Lambda\left(\bigcup_{\mathbf{y}\in\widehat{\mathcal{P_Y}}}\{\mathbf{z}:\mathbf{y}\preceq\mathbf{z}\preceq\mathbf{R}\}\right)$
- EHI: Expected Improvement of the Hypervolume Indicator [3] (relatively to $\mathbf{R}$), if adding design $\mathbf{x}$



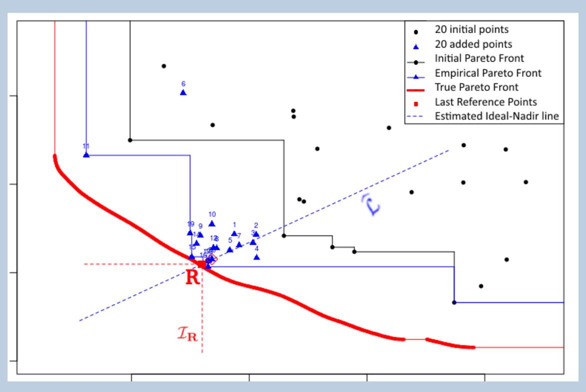


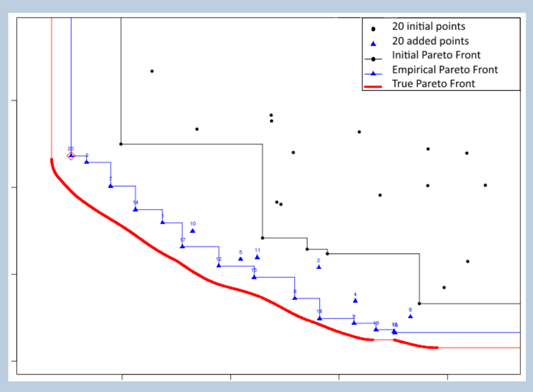
- Subspace targeted by $\mathbf{R}$: $\mathcal{I}_\mathbf{R} = \{\mathbf{y}\in\mathbb{R}^m : \mathbf{y}\preceq\mathbf{R}\}$
- Product of Expected Improvement w.r.t. $\widehat{\mathbf{C}}$: $\text{mEI}(\mathbf{x};\widehat{\mathbf{C}}) = \prod_{i=1}^m \text{EI}_i(\mathbf{x};\widehat{C}_i)$, $\text{EI}_i(\cdot,\widehat{C}_i)$: EI in objective i considering $\widehat{C}_i$ as the current minimum
- If $\mathbf{R} \npreceq \widehat{\mathcal{P_Y}}$, EHI$(\cdot;\mathbf{R})$ =mEI$(\cdot;\mathbf{R})$



- Still cheap for large m
- Analytical expression
- Parallelizable

Use the estimated center $\widehat{\mathbf{C}}$ as reference point $\mathbf{R}$ ⇒ Optimization directed towards the center

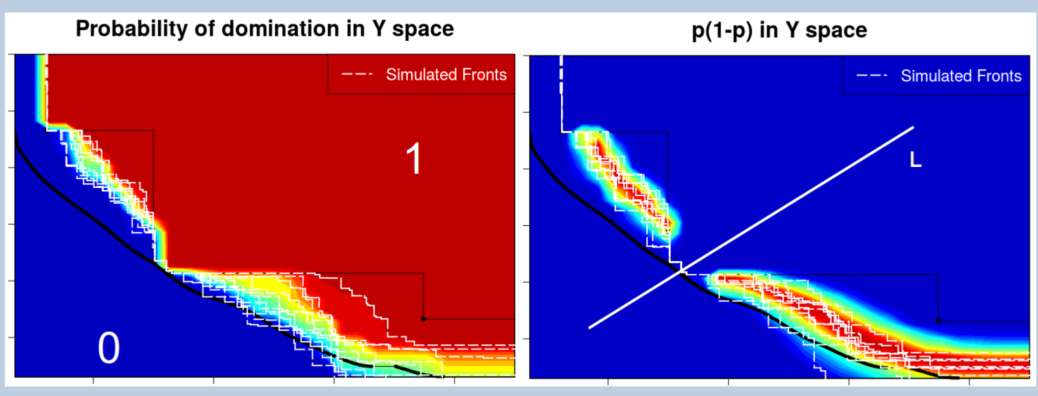




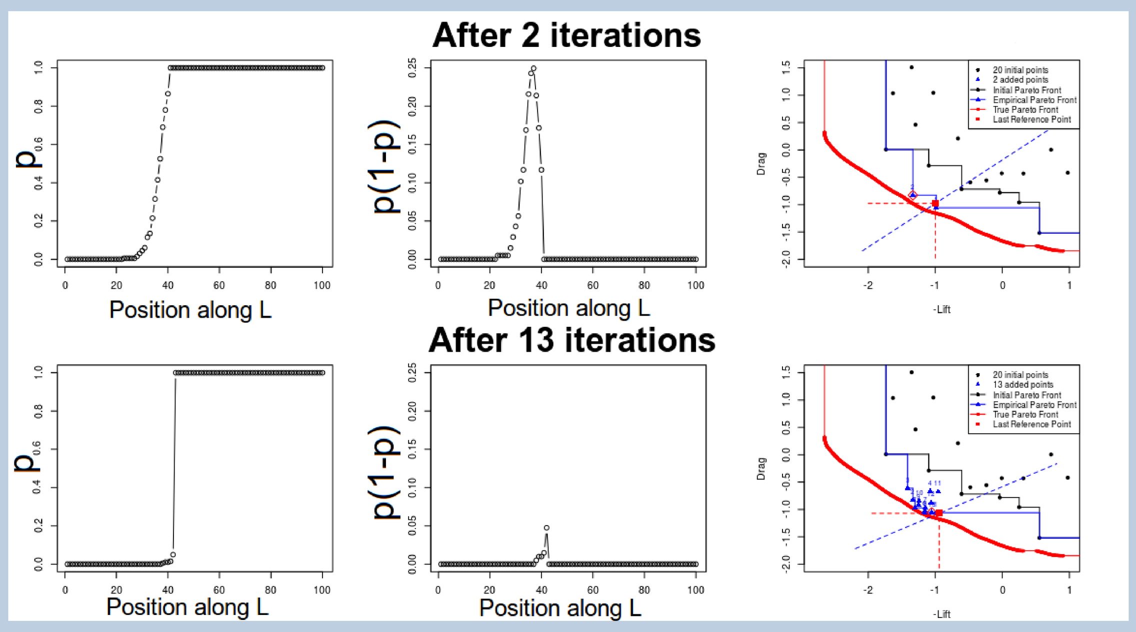
Convergence towards the center of the PF

When to stop the targeting of the center?

- Probability of domination $p(\mathbf{y})$: probability that objective vector $\mathbf{y}$ can be dominated by any $(f_1(\mathbf{x}),..,f_m(\mathbf{x}))$, $\mathbf{x}\in X$
- Estimated using Pareto Fronts from GP draws: $p(\mathbf{y}) = \frac{1}{n_{sim}}\sum_{i=1}^{n_{sim}} \mathbb{1}_{\widetilde{\mathcal{P_Y}}^{(i)}\preceq\mathbf{y}}$
- Information about (local) uncertainty and convergence towards the PF



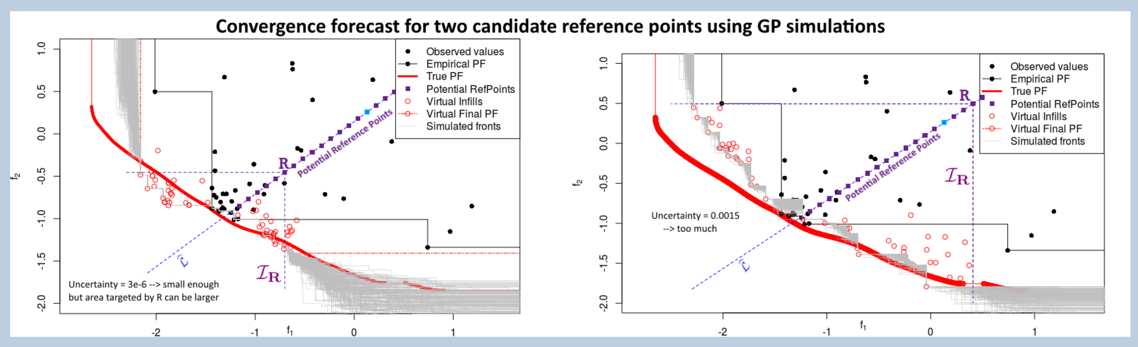
Assume local convergence towards the central part of the PF when $\int_{\widehat{\mathcal{L}}} p(\mathbf{y})(1-p(\mathbf{y}))d\mathbf{y} \le \varepsilon$

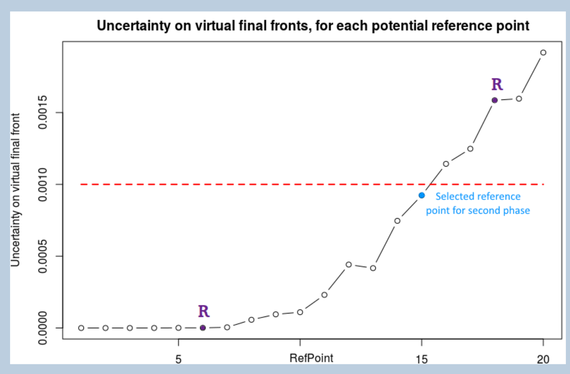


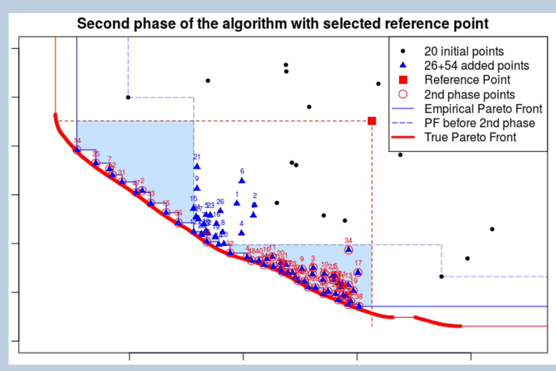
After local convergence: expansion of the PF

b iterations remaining ⇒ what to do next?

- Local convergence detected
- Use the remaining budget b to converge towards $\mathcal{P_Y}$ in a wider but attainable central part
- EHI$(\cdot,\mathbf{R})$, with $\mathbf{R}$ on $\widehat{\mathcal{L}}$ ⇒ focus on central part of $\mathcal{P_Y}$, size of targeted subspace depending on distance between $\mathcal{P_Y}$ and $\mathbf{R}$
- Anticipate the algorithm's behavior in the next iterations and the final PF via *virtual infills* that depend on $\mathbf{R}$: forecast the width of the PF that can be accurately discovered in the last b steps
- *Virtual infills*: either through a Kriging Believer strategy (anticipated $\mathbf{y}$'s are the kriging mean) or through GPs realizations (anticipated $\mathbf{y}$'s are samples)
- Small uncertainty on virtual final Pareto Front = convergence in $\mathcal{I}_\mathbf{R}$ ⇒ choose farthest $\mathbf{R}$ such that $\frac{1}{Vol(\mathbf{I},\mathbf{N})}\int_{\mathbf{y}\preceq\mathbf{R}} p(\mathbf{y})(1-p(\mathbf{y}))d\mathbf{y} \le \varepsilon$



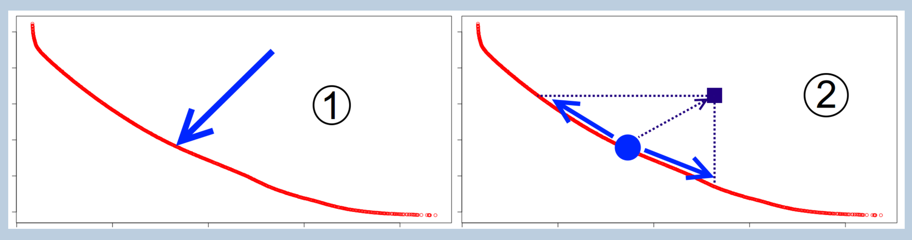




Summary

A two-step algorithm for targeting well-balanced solutions within a few iterations:

- Estimate the Ideal-Nadir line $\widehat{\mathcal{L}}$, on which the expected center of the PF is located
- Define a reference point $\mathbf{R}$ combining $\widehat{\mathcal{L}}$ and the current approximation front $\widehat{\mathcal{P_Y}}$
- Target the estimated central part of the PF using $\mathbf{R}$ and a targeting infill criterion
- When convergence is detected, widen the region of interest $\mathcal{I}_\mathbf{R}$ by a backward step of $\mathbf{R}$ calculated through virtual infills.



References

[1] D. R. Jones, M. Schonlau, and W. J. Welch, Efficient global optimization of expensive black–box functions, Journal of Global Optimization, 13, 455–492 (1998)

[2] E. Zitzler and L. Thiele, Multiobjective Optimization Using Evolutionary Algorithms - A Comparative Case Study, Conference on Parallel Problem Solving from Nature (PPSN V), pages 292–301 (1998)

[3] M. Emmerich, K. Giannakoglou, B. Naujoks, Single and Multiobjective Evolutionary Optimization Assisted by Gaussian Random Field Metamodels, IEEE Transactions on Evolutionary Computation, 10 (2006)