Budgeted Bayesian Multiobjective Optimization

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

[1]Groupe PSA
[2]CNRS LIMOS, École Nationale Supérieure des Mines de Saint-Étienne
[3] Prowler.io

Journées de la Chaire Oquaido, Saint-Étienne, November 28th 2019

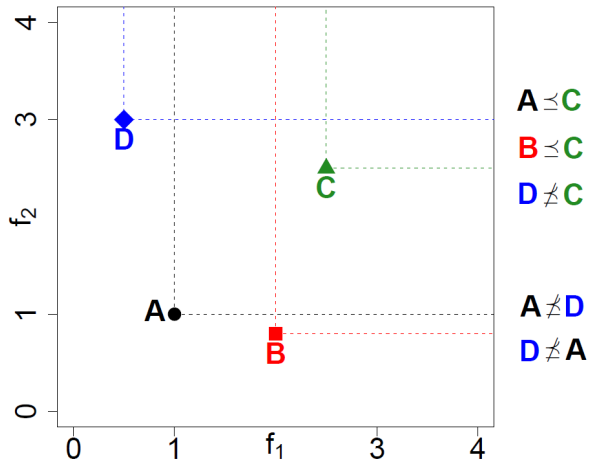
Context

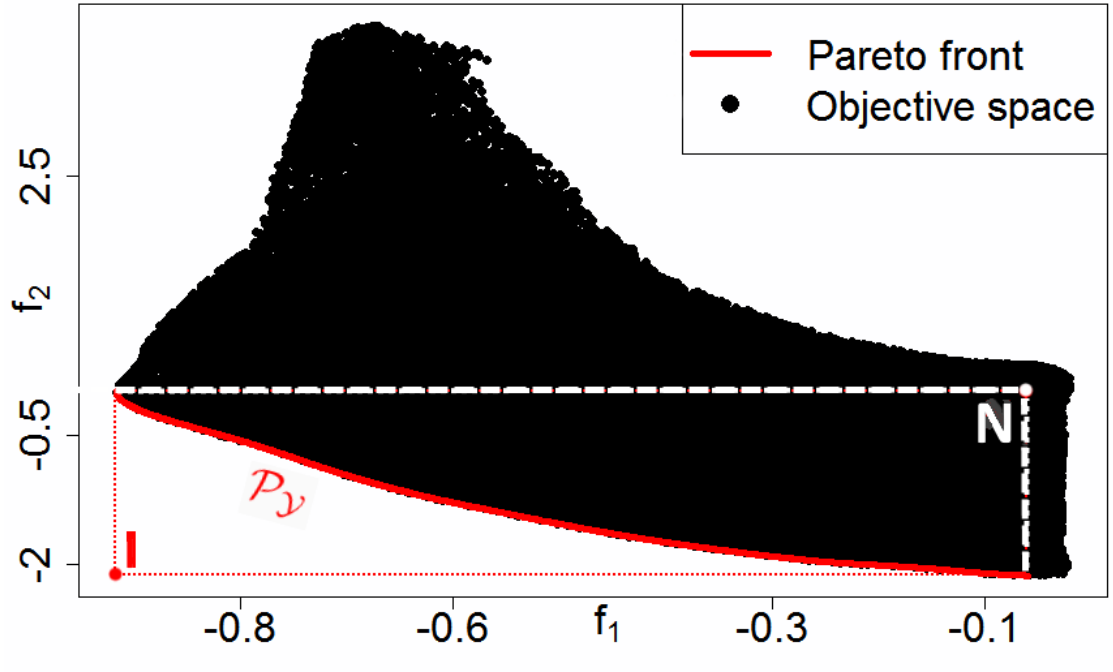
- Optimization of parametric systems.

Context

- Optimization of parametric systems.
- Multiobjective optimization problems:

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

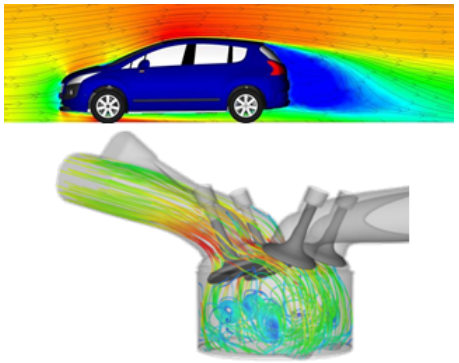




Context

- Optimization of parametric systems.
- Multiobjective optimization problems:

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

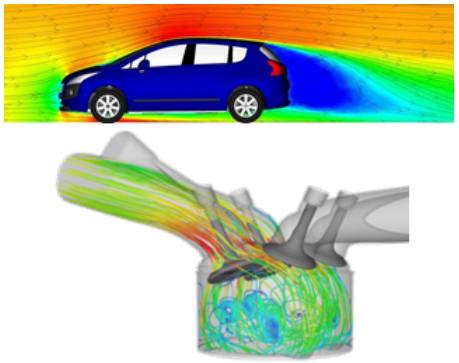


- Complex systems and physics $\Rightarrow$ expensive simulations $\Rightarrow$ restricted budget (≈ 100 evaluations).
- How to obtain *optimal* and *relevant* solutions in spite of the extremely parsimonious use of the computer code?

Context

- Optimization of parametric systems.
- Multiobjective optimization problems:

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



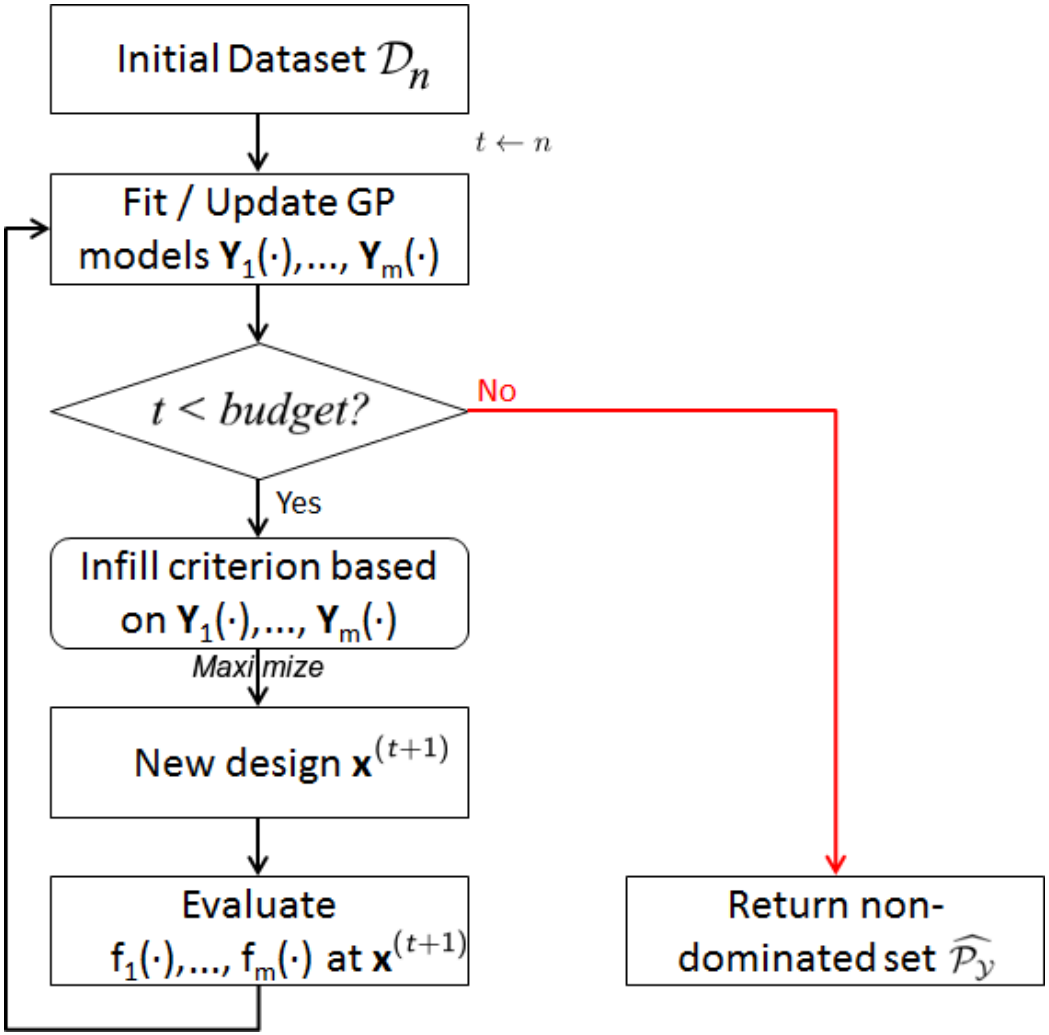
- Complex systems and physics $\Rightarrow$ expensive simulations $\Rightarrow$ restricted budget (≈ 100 evaluations).
- How to obtain *optimal* and *relevant* solutions in spite of the extremely parsimonious use of the computer code?

Bayesian Multiobjective Optimization

- Extension of EGO [Jones et al., 1998] to multiple objectives.

Bayesian Multiobjective Optimization

- Extension of EGO [Jones et al., 1998] to multiple objectives.

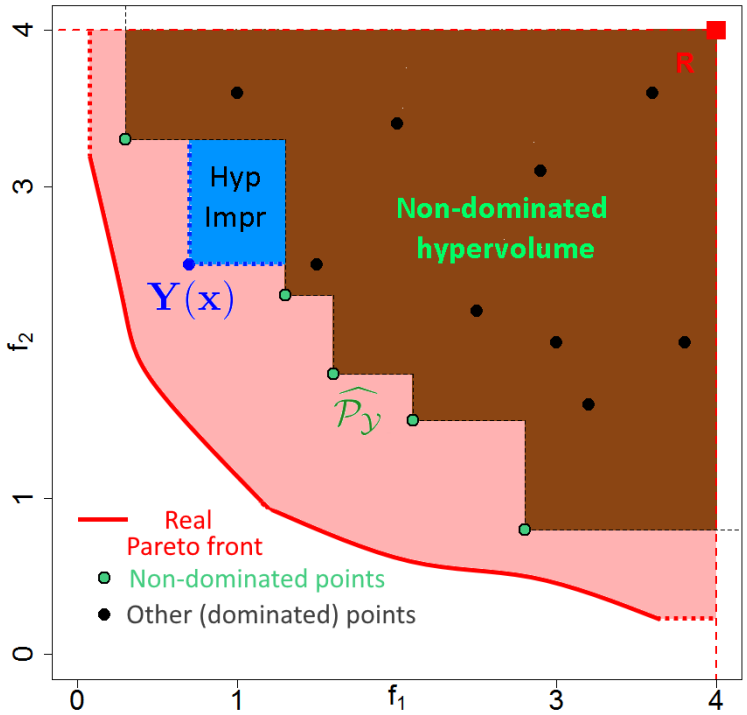


EHI: A Bayesian Multiobjective Infill Criterion

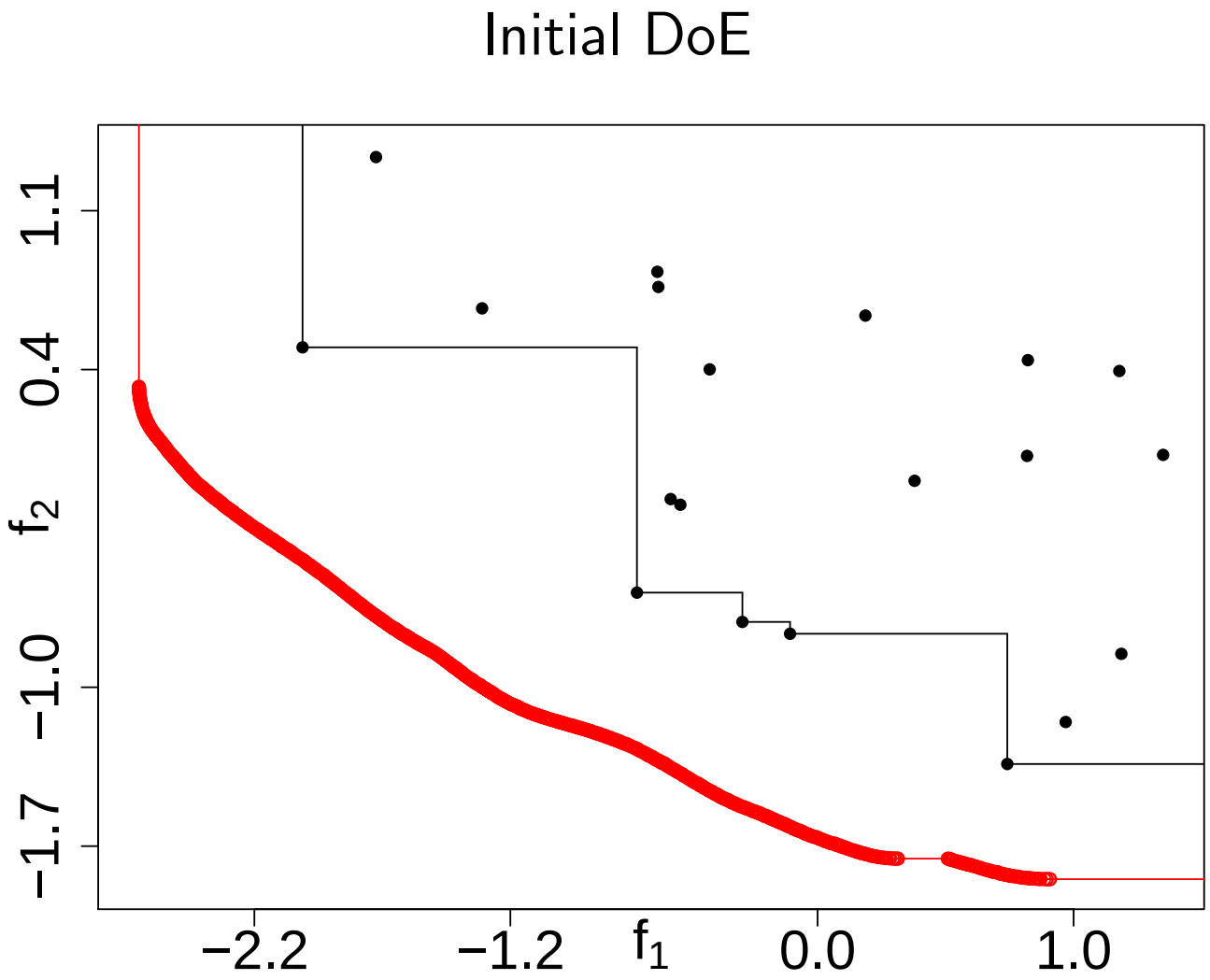
- Expected Hypervolume Improvement (EHI) [Emmerich et al., 2006] aims at uncovering the entire $\mathcal{P}_{\mathcal{Y}}$ in few calls to $\mathbf{f}(\cdot)$.

EHI: A Bayesian Multiobjective Infill Criterion

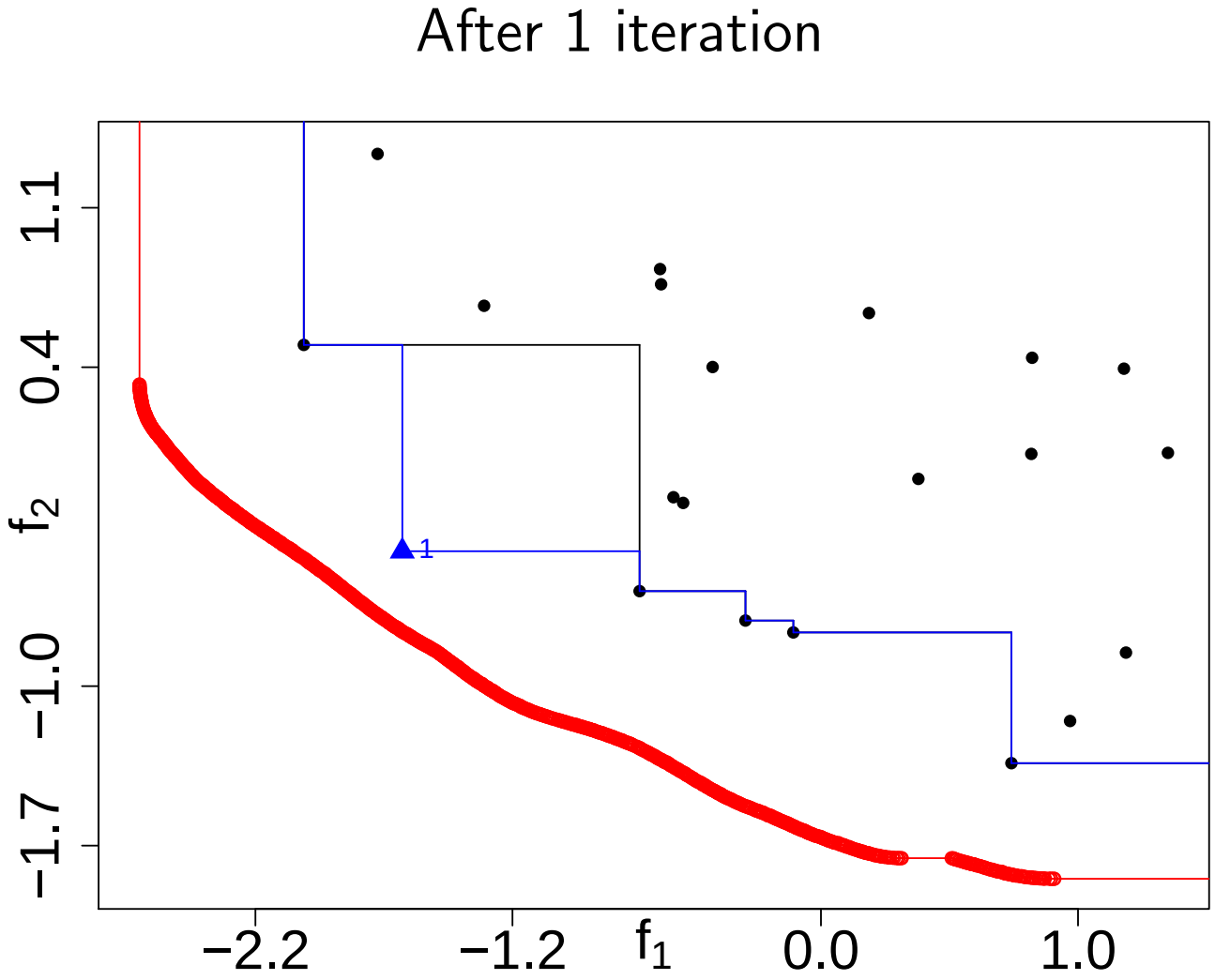
- Expected Hypervolume Improvement (EHI) [Emmerich et al., 2006] aims at uncovering the entire $\mathcal{P}_{\mathcal{Y}}$ in few calls to $\mathbf{f}(\cdot)$.
- $\mathrm{EHI}(\mathbf{x}; \mathbf{R}) = \mathbb{E}[I_H(\widehat{\mathcal{P}_{\mathcal{Y}}} \cup \{\mathbf{Y}(\mathbf{x})\}; \mathbf{R})]$: expected growth of the hypervolume indicator [Zitzler, 1999].



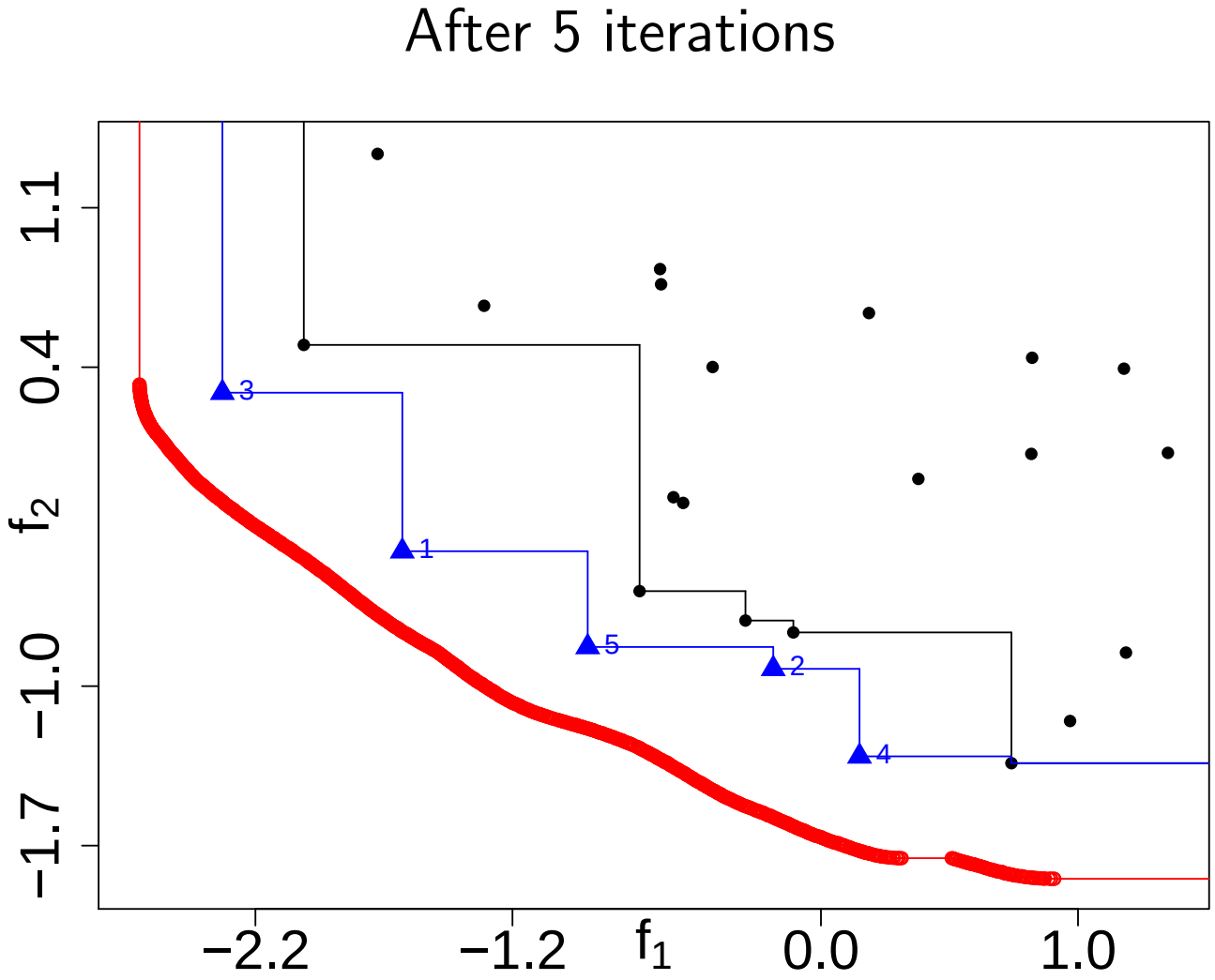
EHI: A Bayesian Multiobjective Infill Criterion



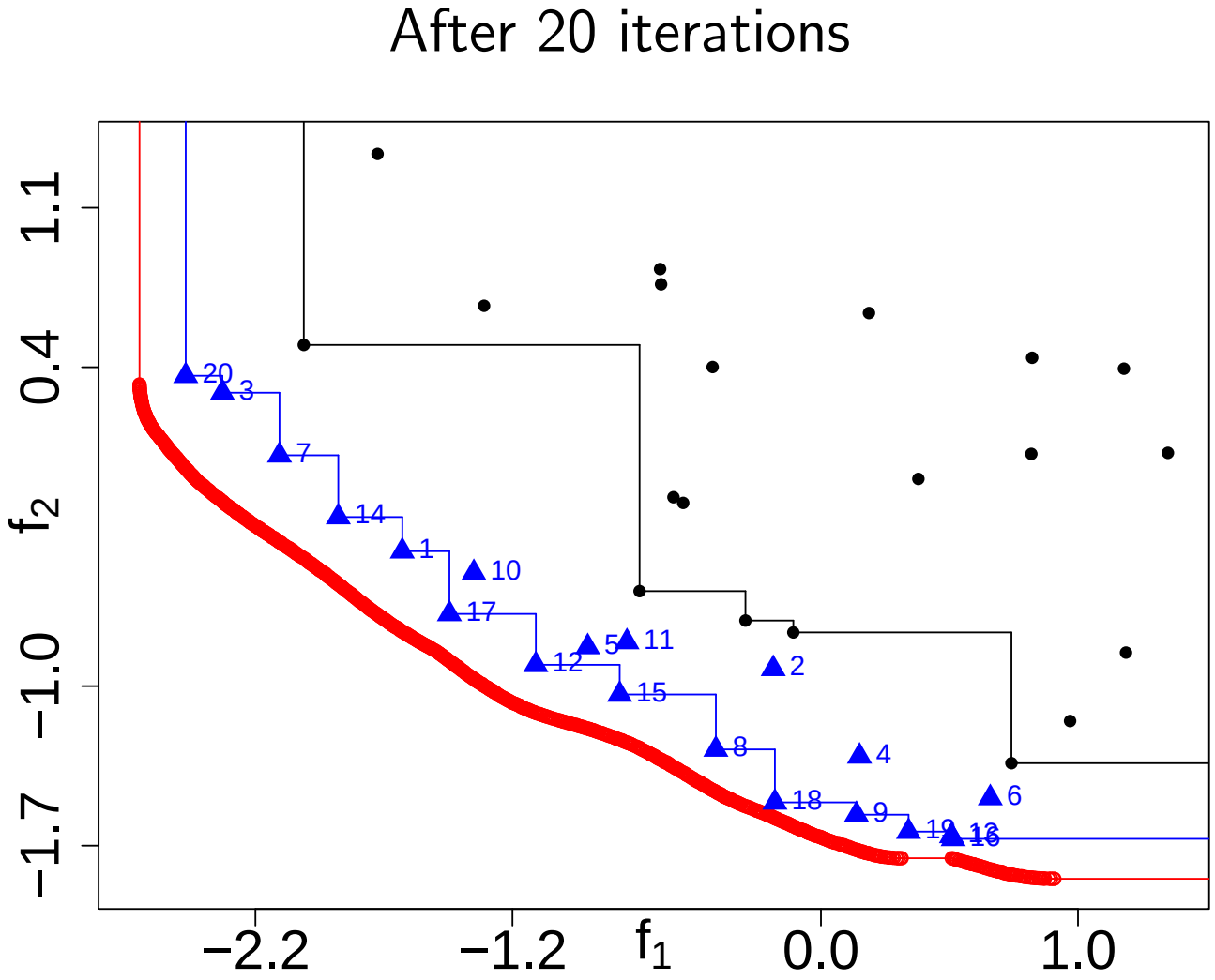
EHI: A Bayesian Multiobjective Infill Criterion



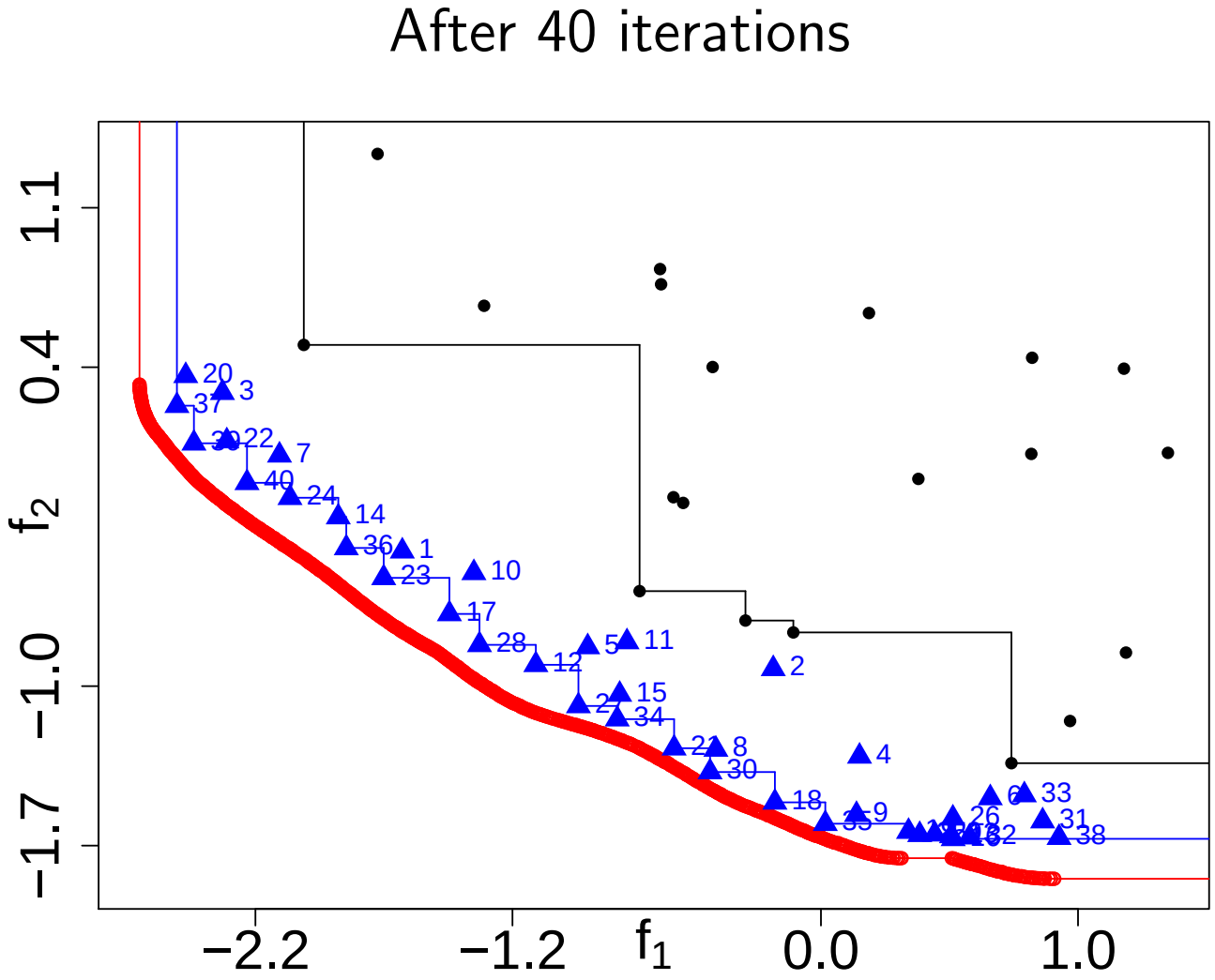
EHI: A Bayesian Multiobjective Infill Criterion



EHI: A Bayesian Multiobjective Infill Criterion



EHI: A Bayesian Multiobjective Infill Criterion



Targeting: Motivation

Good convergence towards $\mathcal{P}_{\mathcal{Y}}$, but

Targeting: Motivation

Good convergence towards $\mathcal{P}_{\mathcal{Y}}$, but

- The size of $\mathcal{P}_{\mathcal{Y}}$ grows exponentially with $m \Rightarrow$ it might not be possible to approximate it accurately under a restricted budget.

Targeting: Motivation

Good convergence towards $\mathcal{P}_{\mathcal{Y}}$, but

- The size of $\mathcal{P}_{\mathcal{Y}}$ grows exponentially with $m \Rightarrow$ it might not be possible to approximate it accurately under a restricted budget.
- All Pareto-optimal solutions may not satisfy the decision maker (especially with large m).

Targeting: Motivation

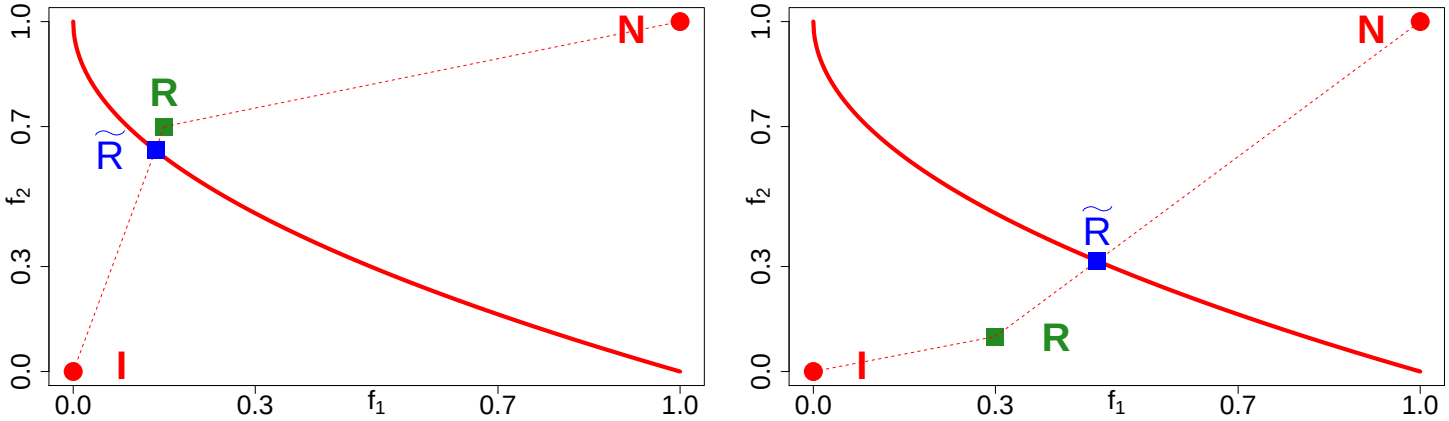
Good convergence towards $\mathcal{P}_{\mathcal{Y}}$, but

- The size of $\mathcal{P}_{\mathcal{Y}}$ grows exponentially with $m \Rightarrow$ it might not be possible to approximate it accurately under a restricted budget.
- All Pareto-optimal solutions may not satisfy the decision maker (especially with large m).

$\Rightarrow$ Prioritize and enhance convergence towards *preferred* parts of the Pareto front.

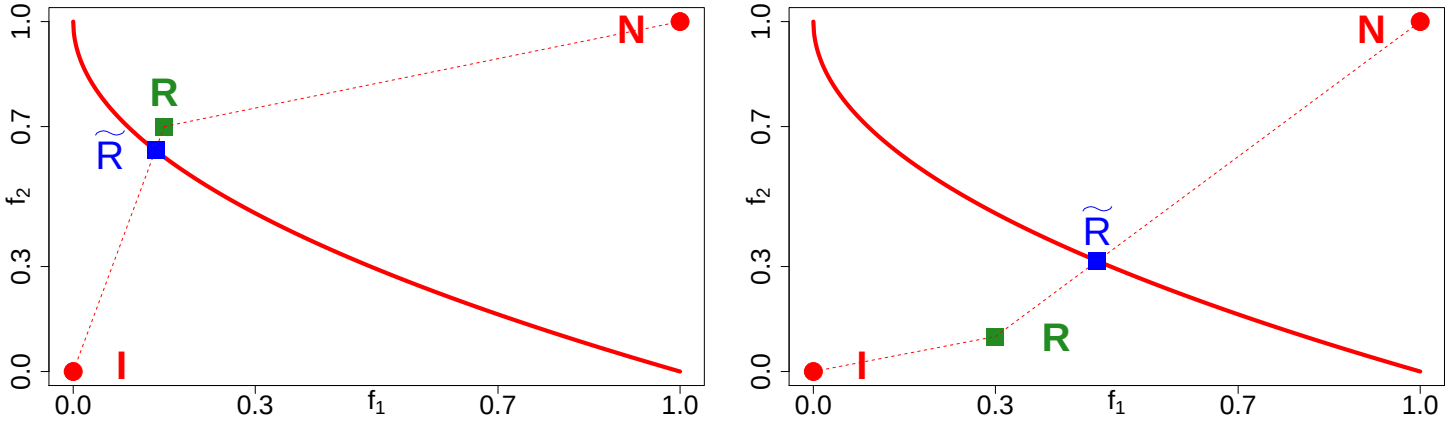
Preferred regions

- Determined through a user-provided reference point $\mathbf{R} \in \mathbb{R}^m$ to be attained/improved.

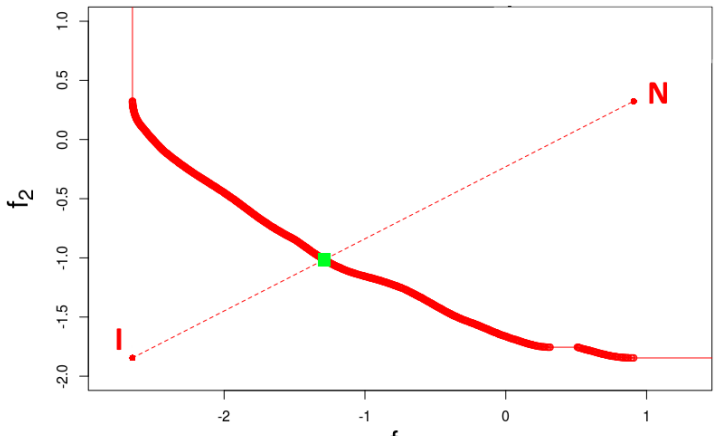


Preferred regions

- Determined through a user-provided reference point $\mathbf{R} \in \mathbb{R}^m$ to be attained/improved.



- No preference expressed $\Rightarrow$ employ the center of the Pareto front, a *well-balanced solution*, as a default preference.



Center of the Pareto front

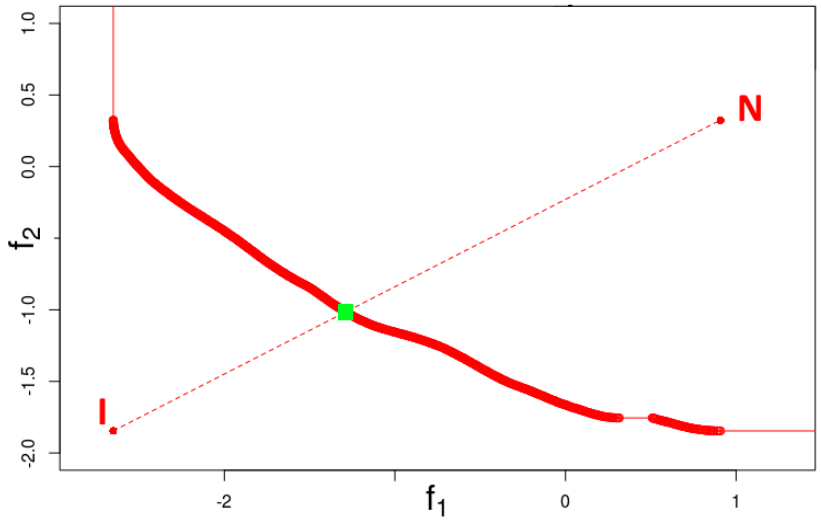
Definition [Gaudrie et al., 2018b]

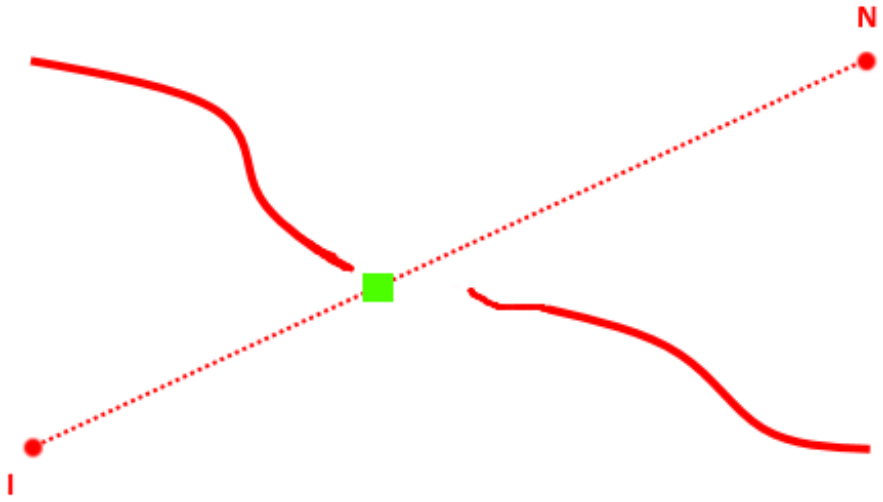
The center of a Pareto front **C** is the closest point in Euclidean distance to $\mathcal{P}_{\mathcal{Y}}$ on the Ideal-Nadir (**IN**) line $\mathcal{L}$.

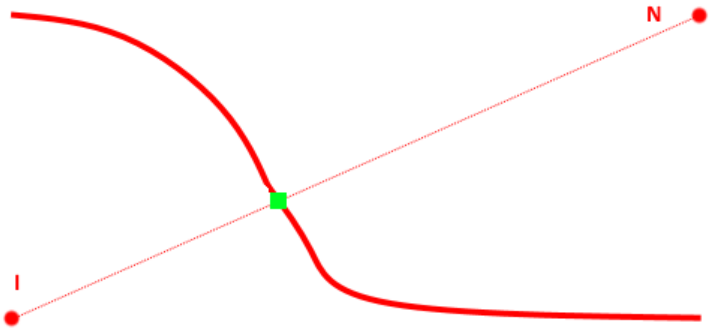
Center of the Pareto front

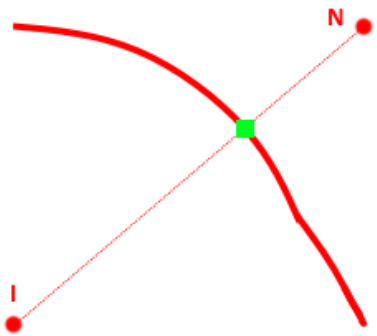
Definition [Gaudrie et al., 2018b]

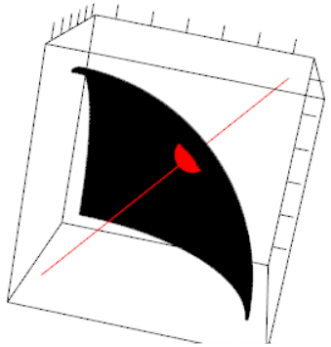
The center of a Pareto front **C** is the closest point in Euclidean distance to $\mathcal{P}_{\mathcal{Y}}$ on the Ideal-Nadir (**IN**) line $\mathcal{L}$.











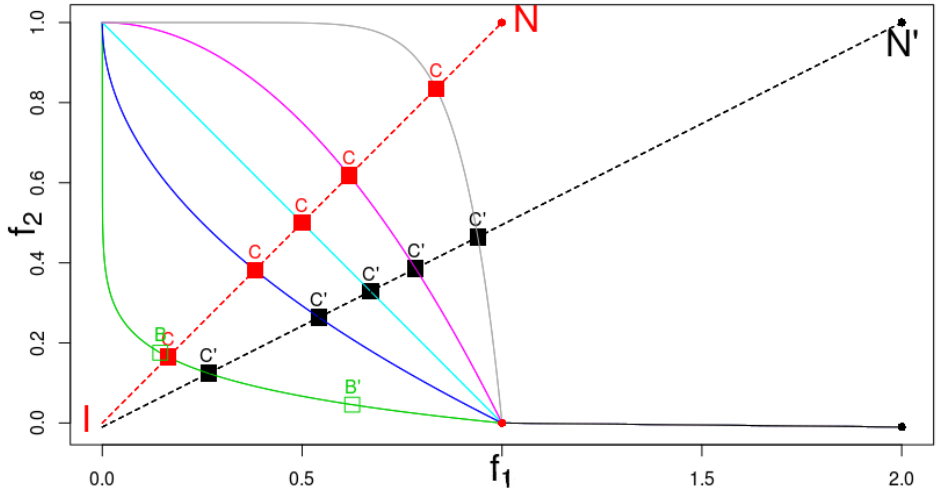
Center of the Pareto front

Properties [Gaudrie et al., 2018a]

- Invariance under an affine transformation of the objective space when $\mathcal{L}$ intersects $\mathcal{P}_{\mathcal{Y}}$ or when $m = 2$.[a]
- Low-sensitivity to the Ideal and to the Nadir point.
- In Game Theory: particular Kalaï-Smorodinsky [Kalai and Smorodinsky, 1975] solution[b] (disagreement point $\equiv$ **N**).

[a] exceptions may occur when $m \geq 3$.

[b] in the case of a convex objective space.



Center of the Pareto front

Empirical front: might lead to weak estimates.

Center of the Pareto front

Empirical front: might lead to weak estimates.

Estimation using Gaussian Processes

Center of the Pareto front

Empirical front: might lead to weak estimates.

Estimation using Gaussian Processes

- Simulate n_{sim} GPs conditioned by $\mathcal{D}_t \Rightarrow$ simulated Pareto fronts $\Rightarrow$ plausible values for **I** and **N**.

Center of the Pareto front

Empirical front: might lead to weak estimates.

Estimation using Gaussian Processes

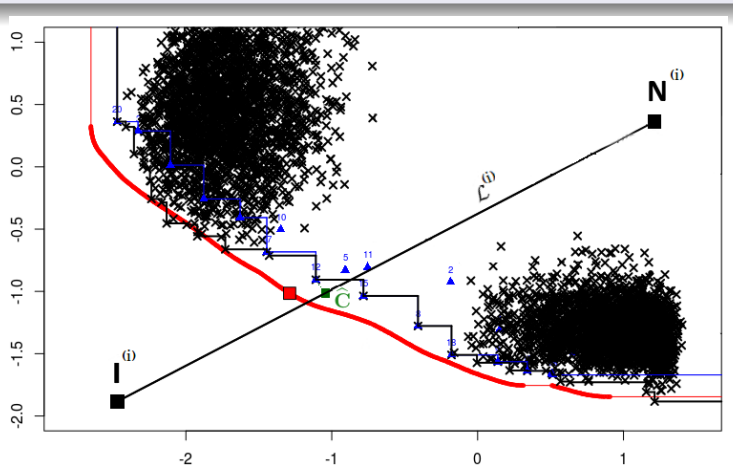
- Simulate n_{sim} GPs conditioned by $\mathcal{D}_t \Rightarrow$ simulated Pareto fronts $\Rightarrow$ plausible values for $\mathbf{I}$ and $\mathbf{N}$.
- Choice of $\mathbf{x}$'s where $\mathbf{Y}(\cdot)$ is simulated is critical!

Center of the Pareto front

Empirical front: might lead to weak estimates.

Estimation using Gaussian Processes

- Simulate n_{sim} GPs conditioned by $\mathcal{D}_t \Rightarrow$ simulated Pareto fronts $\Rightarrow$ plausible values for **I** and **N**.
- Choice of **x**'s where $\mathbf{Y}(\cdot)$ is simulated is critical!
- Choose **x**'s according to probability to lead to **I** or **N** $\Rightarrow$ GP simulations driven towards extreme parts of $\mathcal{P}_{\mathcal{Y}}$.

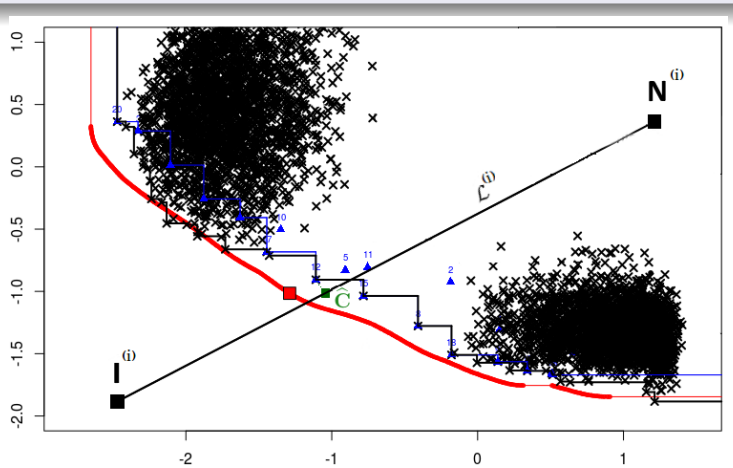


Center of the Pareto front

Empirical front: might lead to weak estimates.

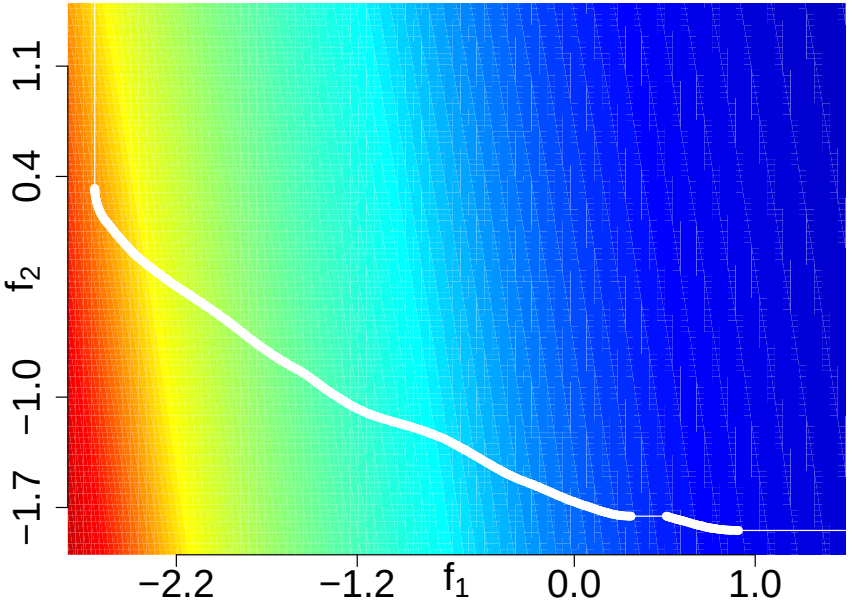
Estimation using Gaussian Processes

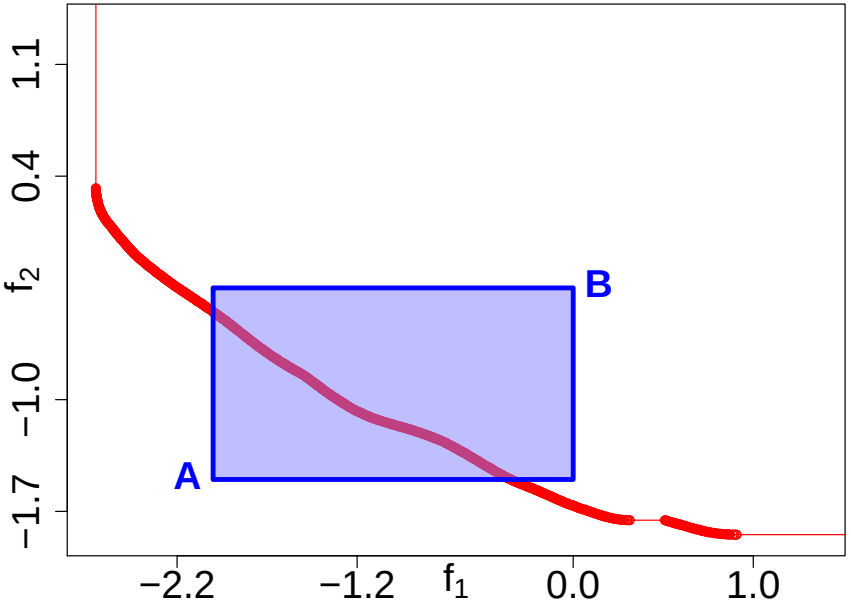
- Simulate n_{sim} GPs conditioned by $\mathcal{D}_t \Rightarrow$ simulated Pareto fronts $\Rightarrow$ plausible values for **I** and **N**.
- Choice of **x**'s where $\mathbf{Y}(\cdot)$ is simulated is critical!
- Choose **x**'s according to probability to lead to **I** or **N** $\Rightarrow$ GP simulations driven towards extreme parts of $\mathcal{P}_{\mathcal{Y}}$.
- Estimated center $\widehat{\mathbf{C}}$: projection of closest point in empirical Pareto front $\widehat{\mathcal{P}_{\mathcal{Y}}}$ on estimated Ideal-Nadir line $\widehat{\mathcal{L}}$.



Targeting in Bayesian Multiobjective Optimization

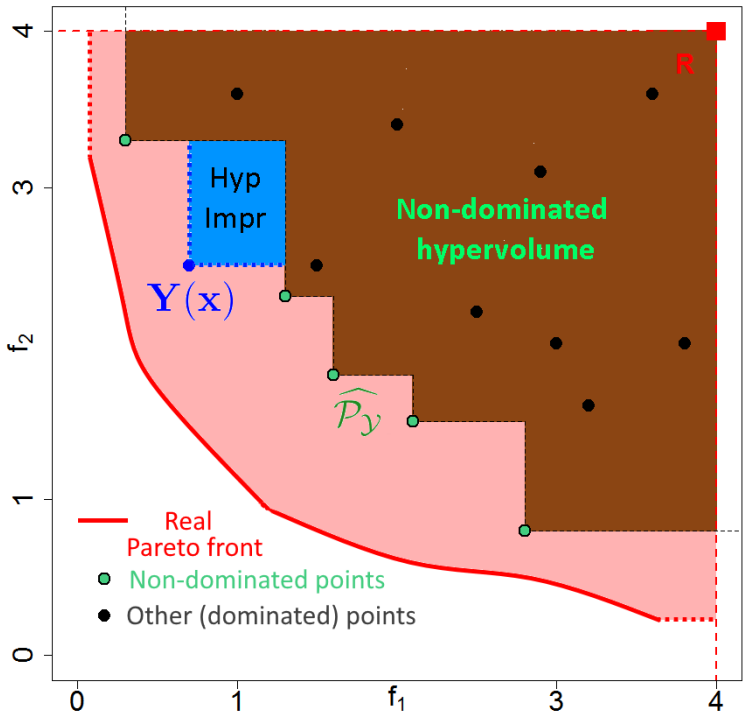
- Weighted EHI [Auger et al., 2009]: externally supplied function $w(\mathbf{y})$ and weighted integration.
- Truncated EHI [Yang et al., 2016]: truncate normal distribution to a user-supplied box $[\mathbf{A}, \mathbf{B}]$.





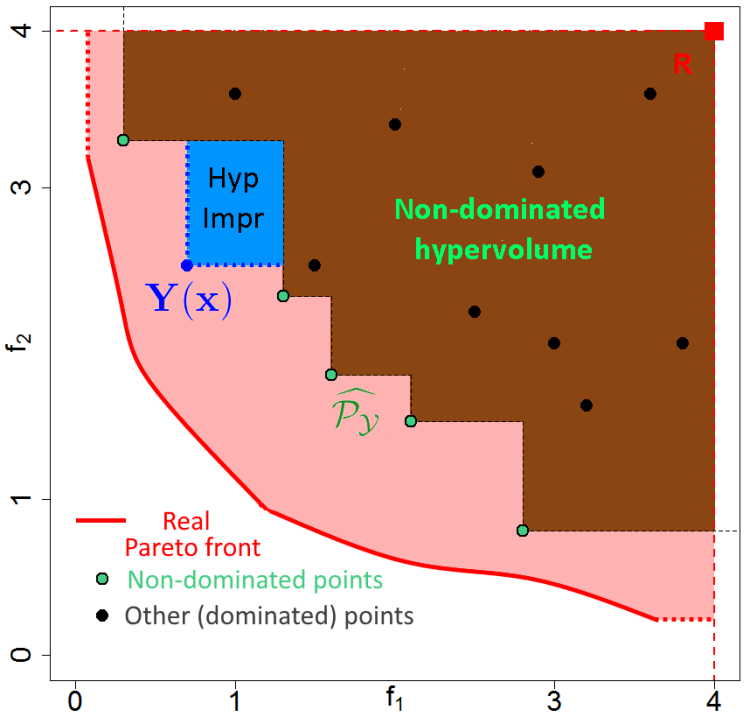
Targeting in Bayesian Multiobjective Optimization

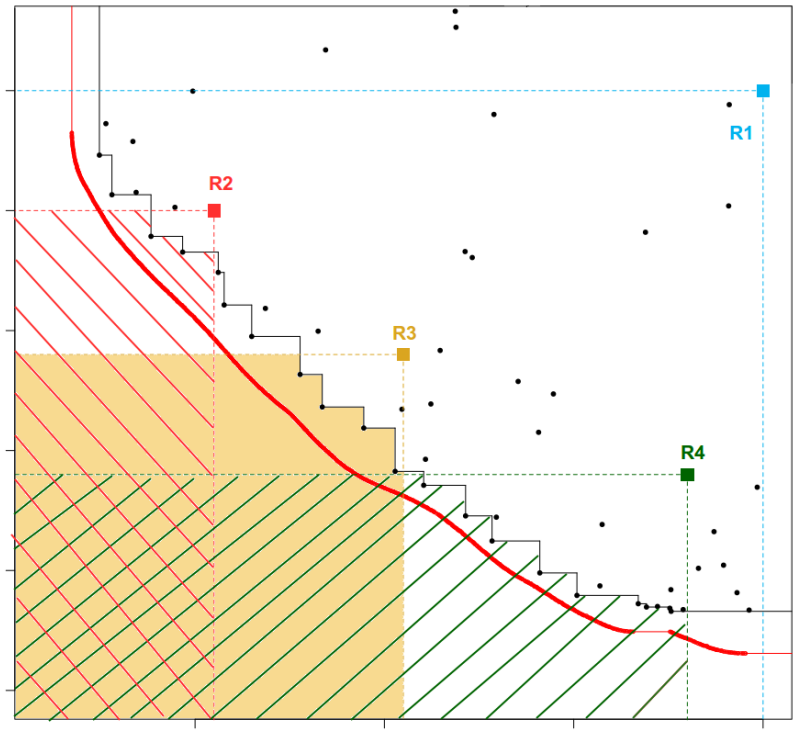
- EHI: hypervolume computed up to $\mathbf{R} \in \mathbb{R}^m$.



Targeting in Bayesian Multiobjective Optimization

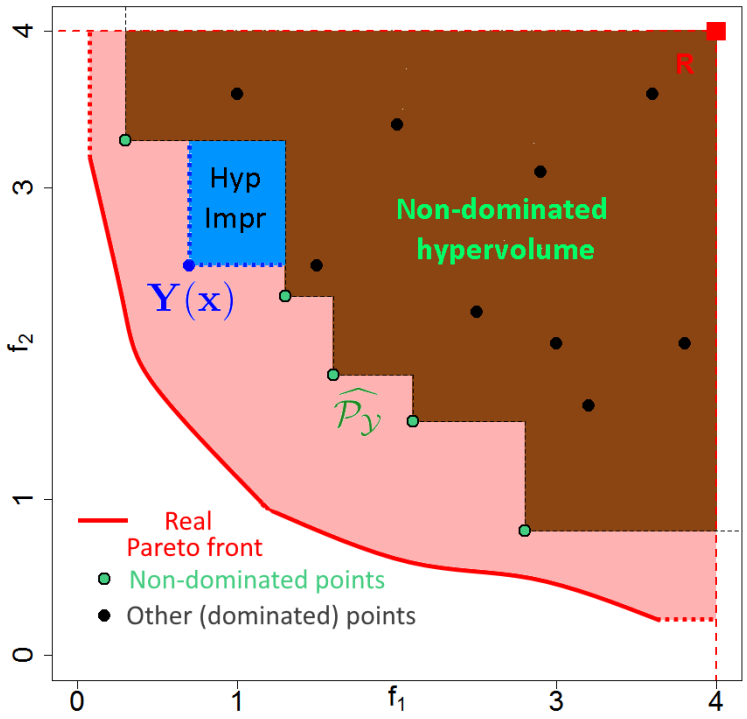
- EHI: hypervolume computed up to $\mathbf{R} \in \mathbb{R}^m$. Defines an improvement region $\mathcal{I}_{\mathbf{R}} := \{\mathbf{z} \in \mathbb{R}^m : \mathbf{z} \preceq \mathbf{R}\}$ where solutions are sought.

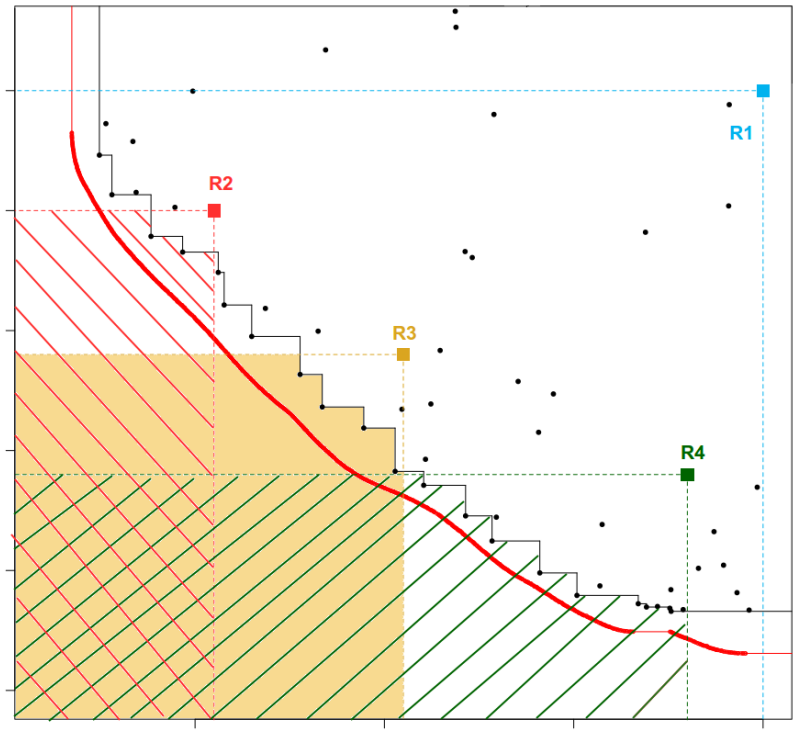




Targeting in Bayesian Multiobjective Optimization

- EHI: hypervolume computed up to $\mathbf{R} \in \mathbb{R}^m$. Defines an improvement region $\mathcal{I}_{\mathbf{R}} := \{\mathbf{z} \in \mathbb{R}^m : \mathbf{z} \preceq \mathbf{R}\}$ where solutions are sought.





$\Rightarrow$ target parts of $\mathcal{P}_{\mathcal{Y}}$ by adjusting $\mathbf{R}$.

The mEI criterion

Proposition [Gaudrie et al., 2018a]

When $\widehat{\mathcal{P}_{\mathcal{Y}}} \not\preceq \mathbf{R}$, $\mathrm{EHI}(\cdot; \mathbf{R}) = \mathrm{mEI}(\cdot; \mathbf{R})$ where $\mathrm{mEI}(\mathbf{x}; \mathbf{R}) = \prod_{j=1}^{m} \mathrm{EI}_j(\mathbf{x}; R_j)$.

The mEI criterion

Proposition [Gaudrie et al., 2018a]

When $\widehat{\mathcal{P}_{\mathcal{Y}}} \not\preceq \mathbf{R}$, $\mathrm{EHI}(\cdot; \mathbf{R}) = \mathrm{mEI}(\cdot; \mathbf{R})$ where $\mathrm{mEI}(\mathbf{x}; \mathbf{R}) = \prod_{j=1}^{m} \mathrm{EI}_j(\mathbf{x}; R_j)$.

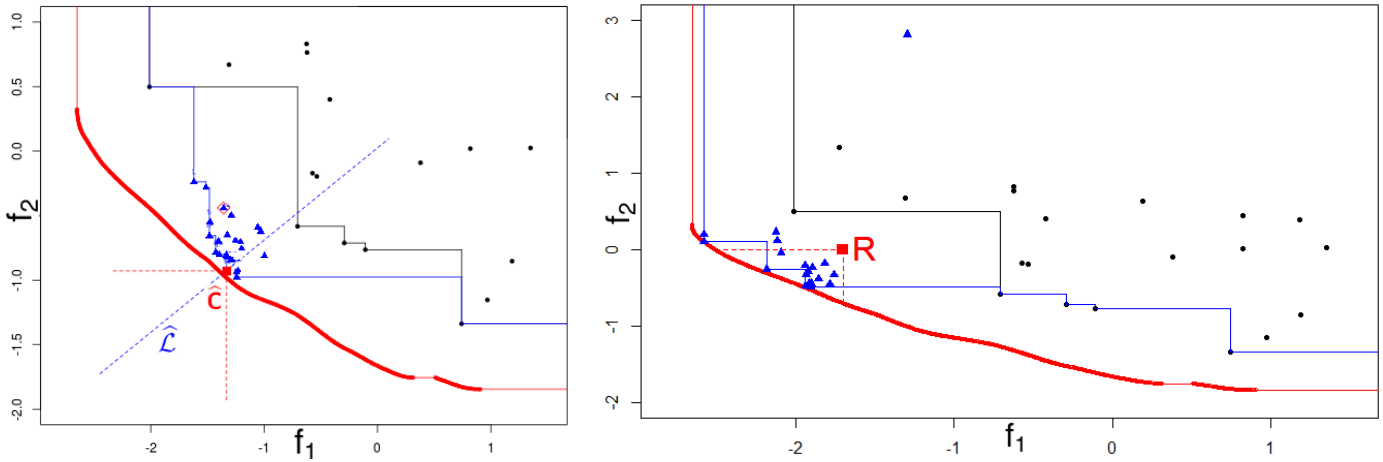
$\Rightarrow$ cheaper criterion, analytical for any m, with computable ∇.

The mEI criterion

Proposition [Gaudrie et al., 2018a]

When $\widehat{\mathcal{P}_{\mathcal{Y}}} \not\preceq \mathbf{R}$, $\mathrm{EHI}(\cdot;\mathbf{R}) = \mathrm{mEI}(\cdot;\mathbf{R})$ where $\mathrm{mEI}(\mathbf{x};\mathbf{R}) = \prod_{j=1}^{m} \mathrm{EI}_j(\mathbf{x};R_j)$.

$\Rightarrow$ cheaper criterion, analytical for any m, with computable ∇.
Evaluate $\mathrm{mEI}(\cdot;\widehat{\mathbf{C}})$'s maximizer $\Rightarrow$ optimization directed towards the center of $\mathcal{P}_{\mathcal{Y}}$.
Evaluate $\mathrm{mEI}(\cdot;\widehat{\mathbf{R}})$'s maximizer $\Rightarrow$ optimization directed towards the user-desired part of $\mathcal{P}_{\mathcal{Y}}$.

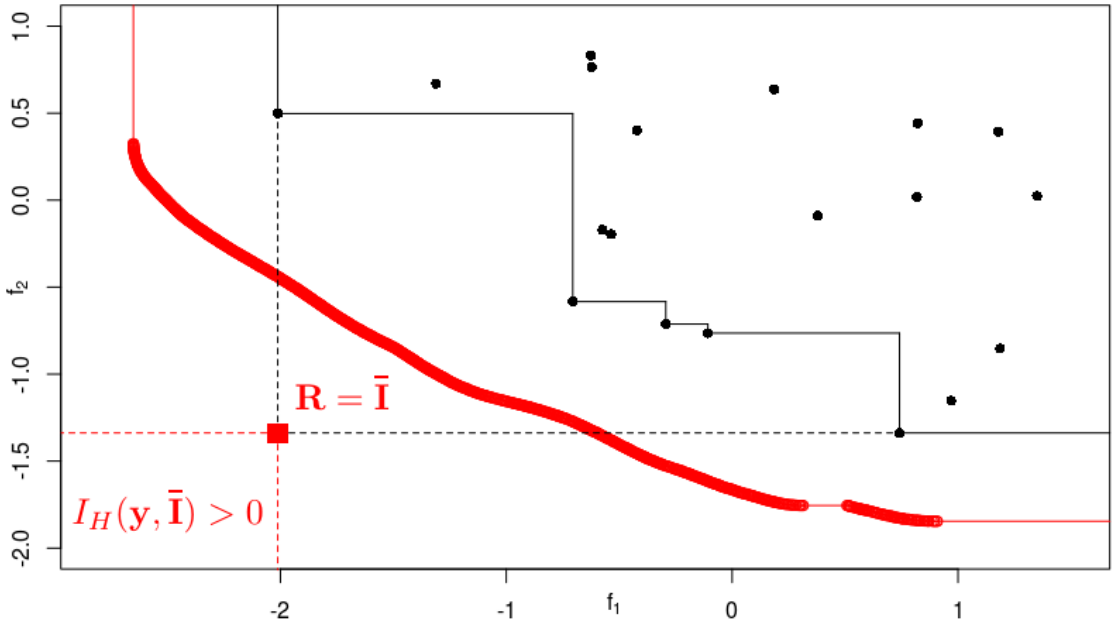


The mEI criterion

- Remark: product of EI's w.r.t $\mathbf{f}_{\min}$?

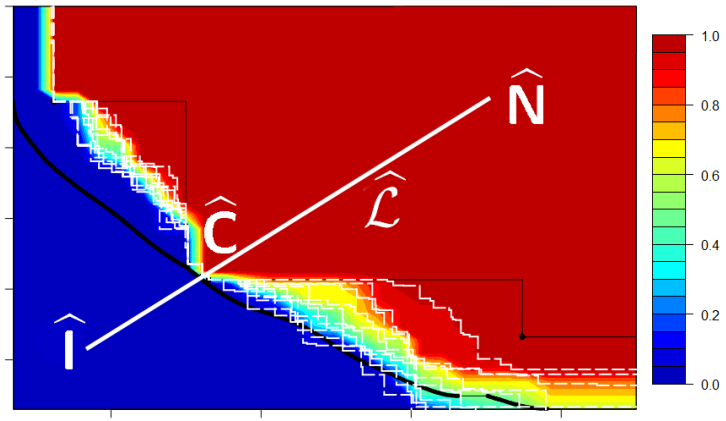
The mEI criterion

- Remark: product of EI's w.r.t $\mathbf{f}_{\min}$?
- $\mathrm{mEI}(\cdot;\bar{\mathbf{I}}) = \mathrm{EHI}(\cdot;\bar{\mathbf{I}})$.

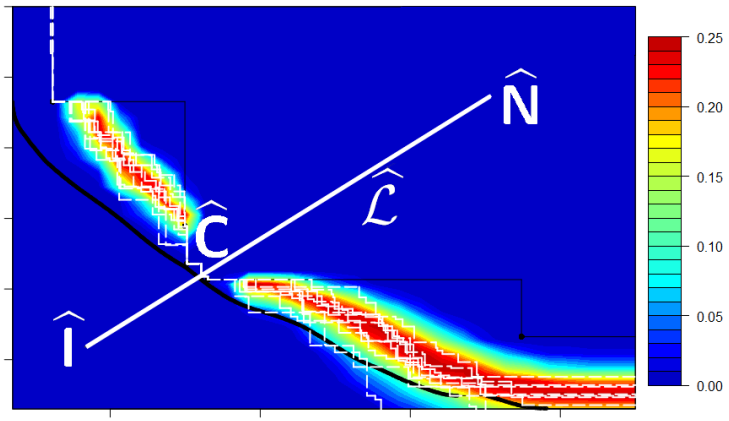


Detecting local convergence to the Pareto front

- $\mathcal{P}_{\mathcal{Y}}$ might be attained before depletion of computational budget $\Rightarrow$ waste of resources.
- Local convergence to $\mathcal{P}_{\mathcal{Y}}$ needs to be verified: stopping criterion.
- Probability of domination $p(\mathbf{y})$: probability that objective vector $\mathbf{y} \in \mathbb{R}^m$ can be dominated by some $(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))^\top$.
- Estimated using simulated fronts $\widetilde{\mathcal{P}_{\mathcal{Y}}}^{(k)}$: $\widehat{p}(\mathbf{y}) = \frac{1}{n_{sim}} \sum_{k=1}^{n_{sim}} \mathbb{1}_{\widetilde{\mathcal{P}_{\mathcal{Y}}}^{(k)} \preceq \mathbf{y}}$.



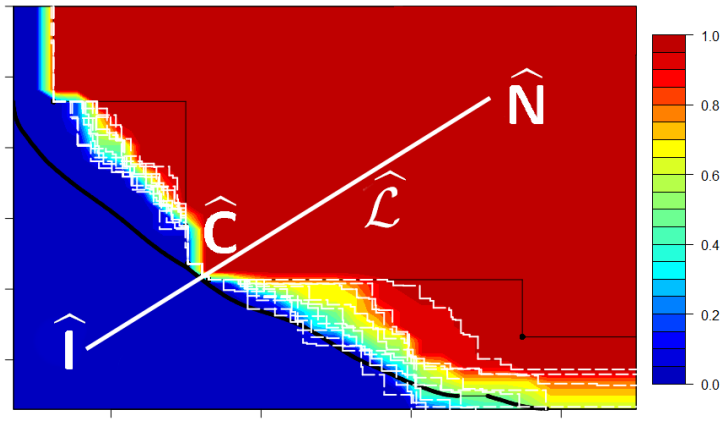
$\widehat{p}(\mathbf{y})$



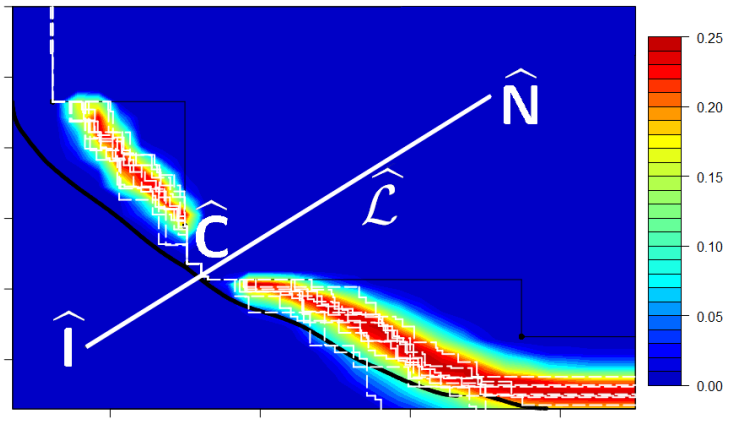
$\widehat{p}(\mathbf{y})(1 - \widehat{p}(\mathbf{y}))$

Detecting local convergence to the Pareto front

- $\mathcal{P}_{\mathcal{Y}}$ might be attained before depletion of computational budget $\Rightarrow$ waste of resources.
- Local convergence to $\mathcal{P}_{\mathcal{Y}}$ needs to be verified: stopping criterion.
- Probability of domination $p(\mathbf{y})$: probability that objective vector $\mathbf{y} \in \mathbb{R}^m$ can be dominated by some $(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))^\top$.
- Estimated using simulated fronts $\widetilde{\mathcal{P}_{\mathcal{Y}}}^{(k)}$: $\widehat{p}(\mathbf{y}) = \frac{1}{n_{sim}} \sum_{k=1}^{n_{sim}} \mathbb{1}_{\widetilde{\mathcal{P}_{\mathcal{Y}}}^{(k)} \preceq \mathbf{y}}$.



$\widehat{p}(\mathbf{y})$

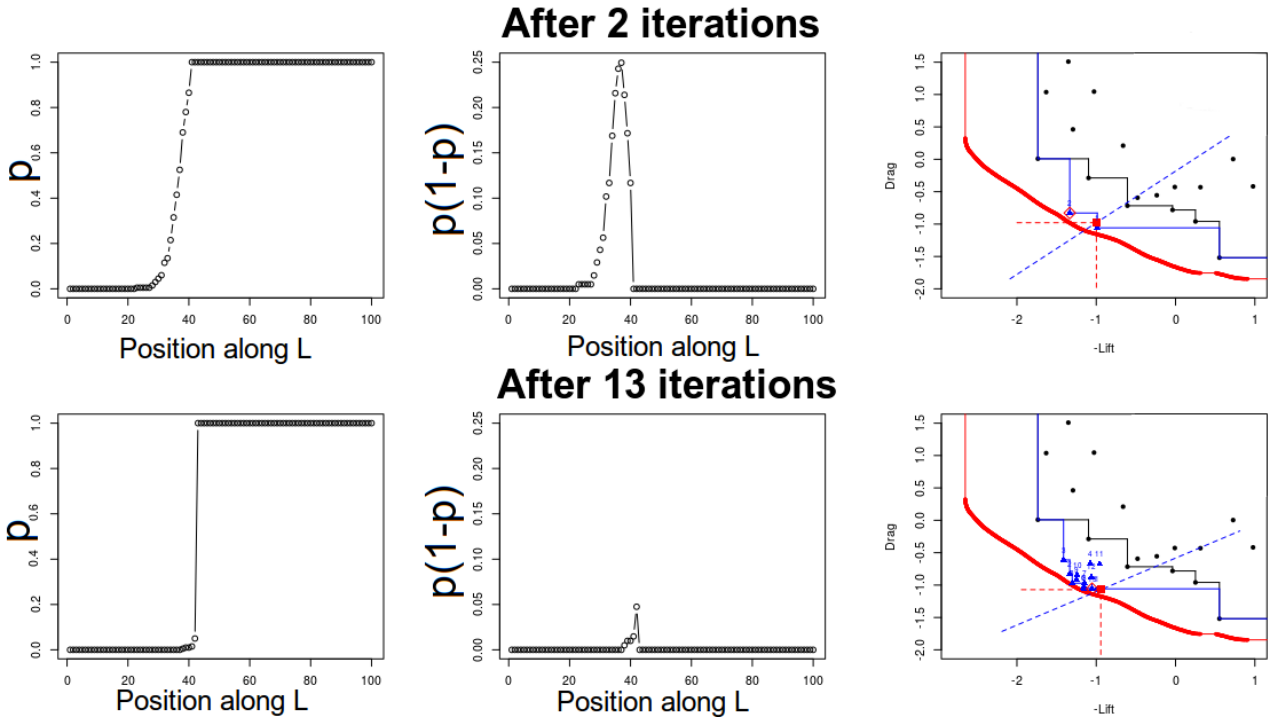


$\widehat{p}(\mathbf{y})(1 - \widehat{p}(\mathbf{y}))$

- Convergence in areas where $\widehat{p}(\cdot)$ goes quickly from 0 to 1 $\Leftrightarrow$ where $\widehat{p}(\cdot)(1 - \widehat{p}(\cdot))$ equals 0.

Detecting local convergence to the Pareto front

- Assume local convergence to $\mathcal{P}_{\mathcal{Y}}$ when the *line-uncertainty* $U(\widehat{\mathcal{L}}) := \frac{1}{|\widehat{\mathcal{L}}|} \int_{\widehat{\mathcal{L}}} \widehat{p}(\mathbf{y})(1 - \widehat{p}(\mathbf{y}))d\mathbf{y} \leq \varepsilon$.

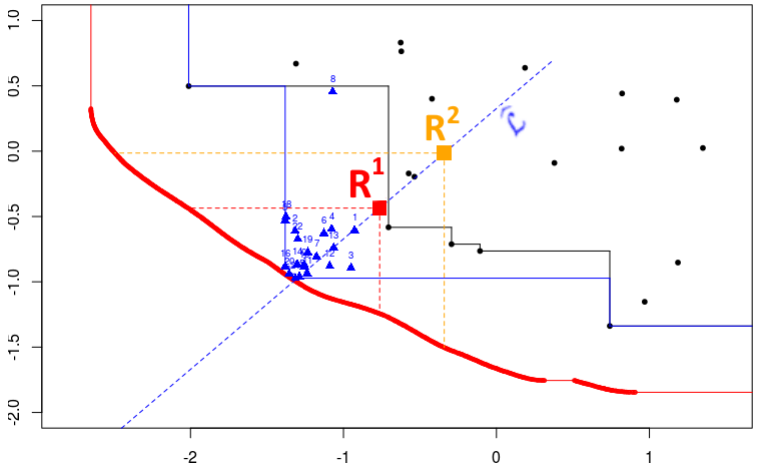


Expansion of the approximation front

- Convergence detected: how to use the remaining budget?

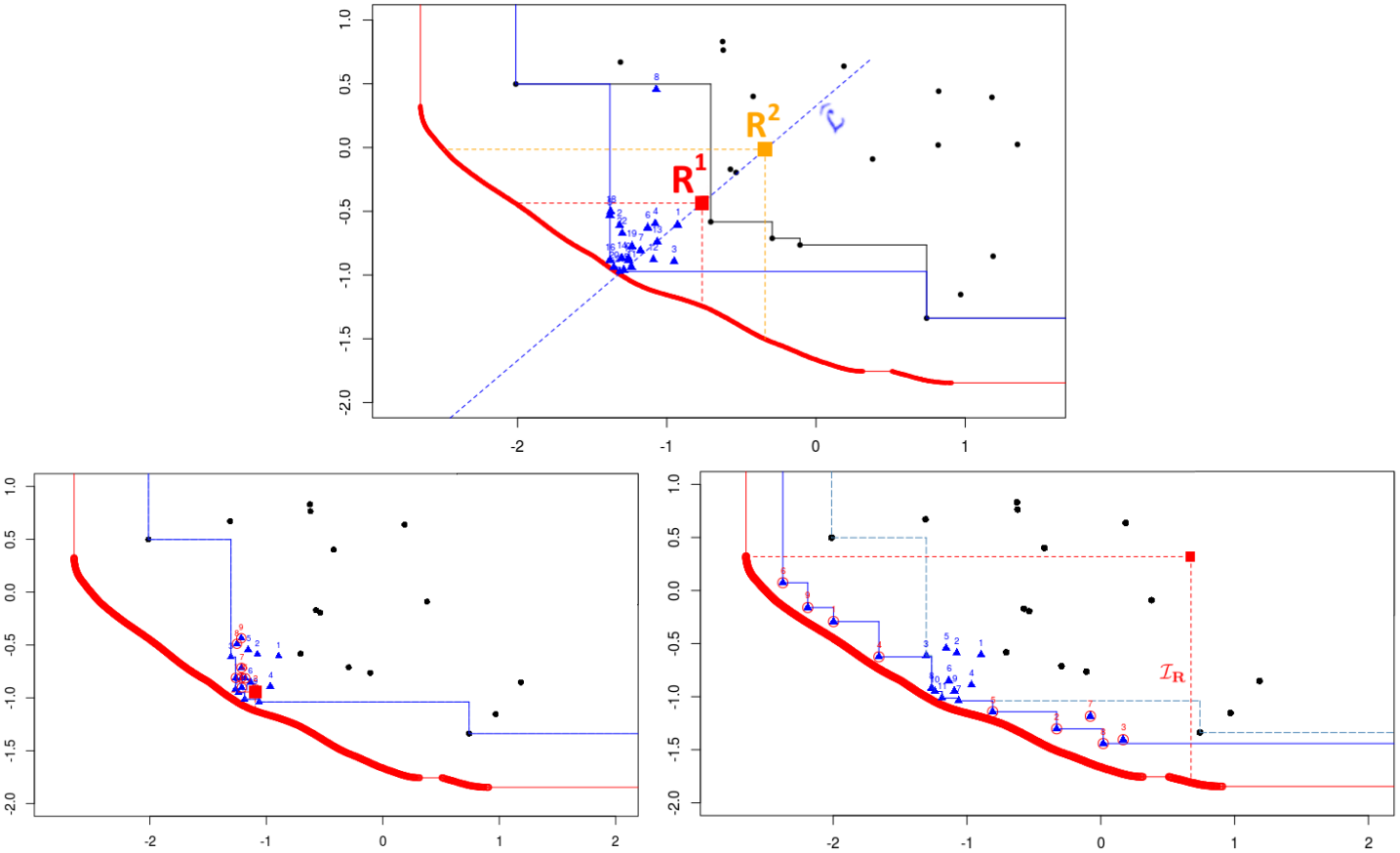
Expansion of the approximation front

- Convergence detected: how to use the remaining budget? Target a *wider* part of $\mathcal{P}_{\mathcal{Y}}$ through $\mathrm{EHI}(\cdot;\mathbf{R})$ with $\mathbf{R}$ shifted backwards on $\widehat{\mathcal{L}}$.



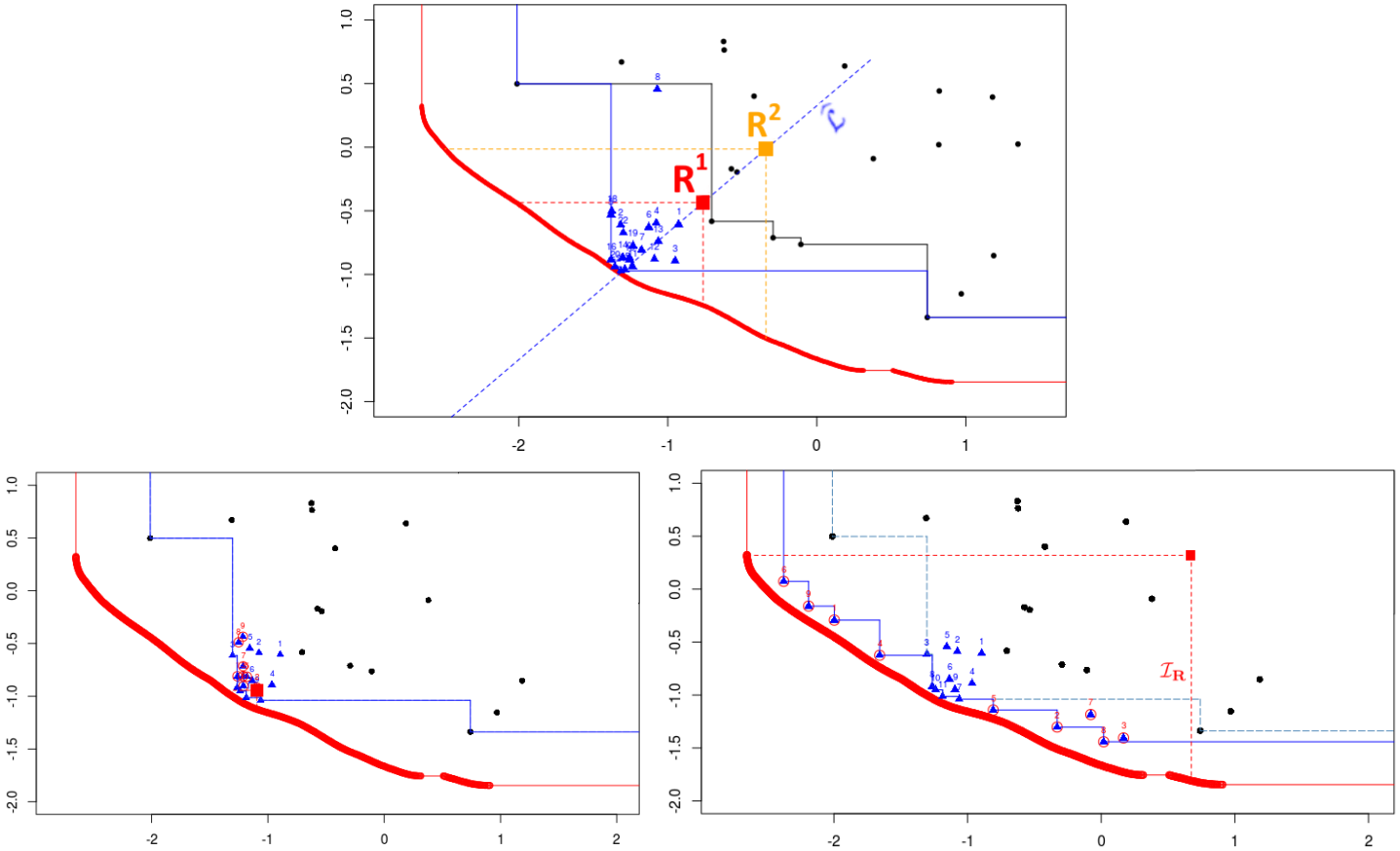
Expansion of the approximation front

- Convergence detected: how to use the remaining budget? Target a *wider* part of $\mathcal{P}_{\mathcal{Y}}$ through $\mathrm{EHI}(\cdot;\mathbf{R})$ with $\mathbf{R}$ shifted backwards on $\widehat{\mathcal{L}}$.



Expansion of the approximation front

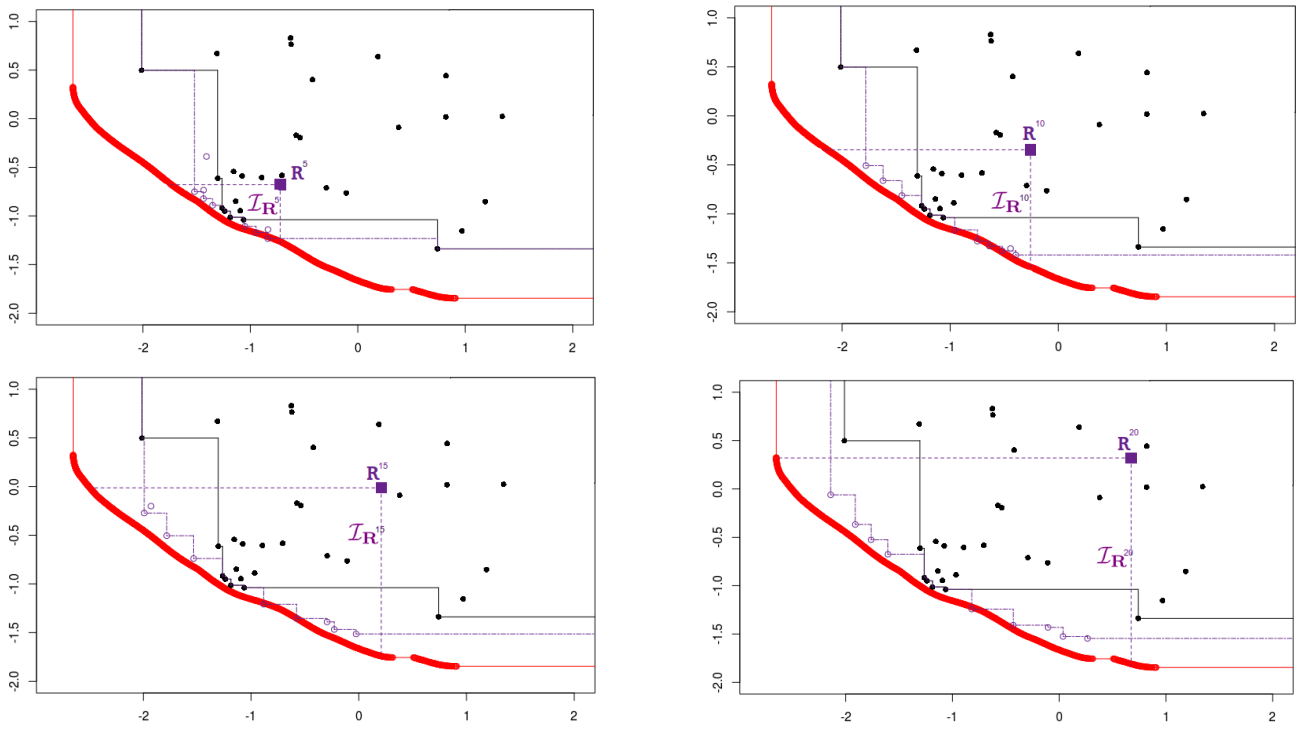
- Convergence detected: how to use the remaining budget? Target a *wider* part of $\mathcal{P}_{\mathcal{Y}}$ through $\mathrm{EHI}(\cdot;\mathbf{R})$ with $\mathbf{R}$ shifted backwards on $\widehat{\mathcal{L}}$.



- Final front depends on $\mathbf{R} \Rightarrow$ anticipate the behavior of the algorithm to determine the widest $\mathcal{I}_{\mathbf{R}}$ with accurate forecasted convergence.

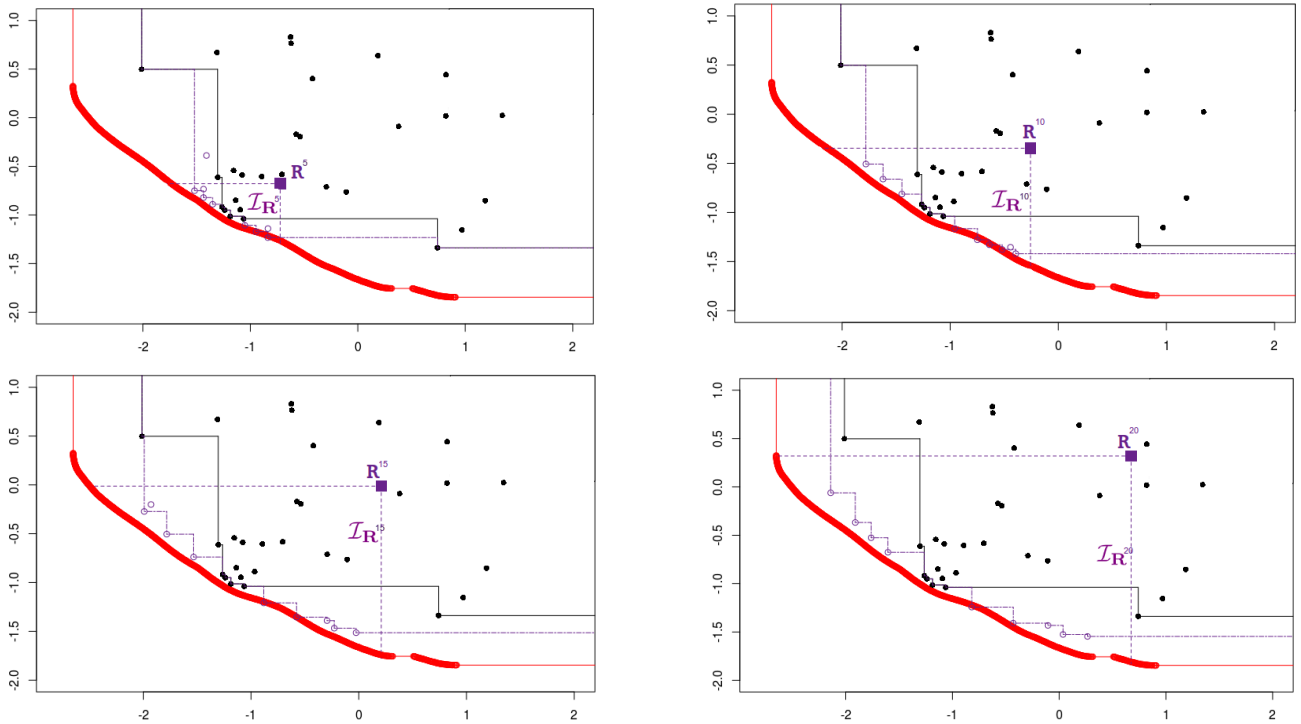
Expansion of the approximation front

- For increasing $\mathbf{R}^c$'s, perform the b last iterations substituting $\mathbf{f}(\cdot)$ by $\widehat{\mathbf{y}}(\cdot) \Rightarrow$ final virtual metamodels $\mathbf{Y}_c^{KB}(\cdot)$ and fronts depending on $\mathbf{R}^c$.

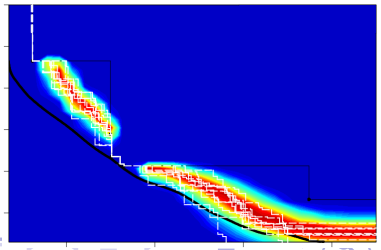


Expansion of the approximation front

- For increasing $\mathbf{R}^c$'s, perform the b last iterations substituting $\mathbf{f}(\cdot)$ by $\widehat{\mathbf{y}}(\cdot) \Rightarrow$ final virtual metamodels $\mathbf{Y}_c^{KB}(\cdot)$ and fronts depending on $\mathbf{R}^c$.

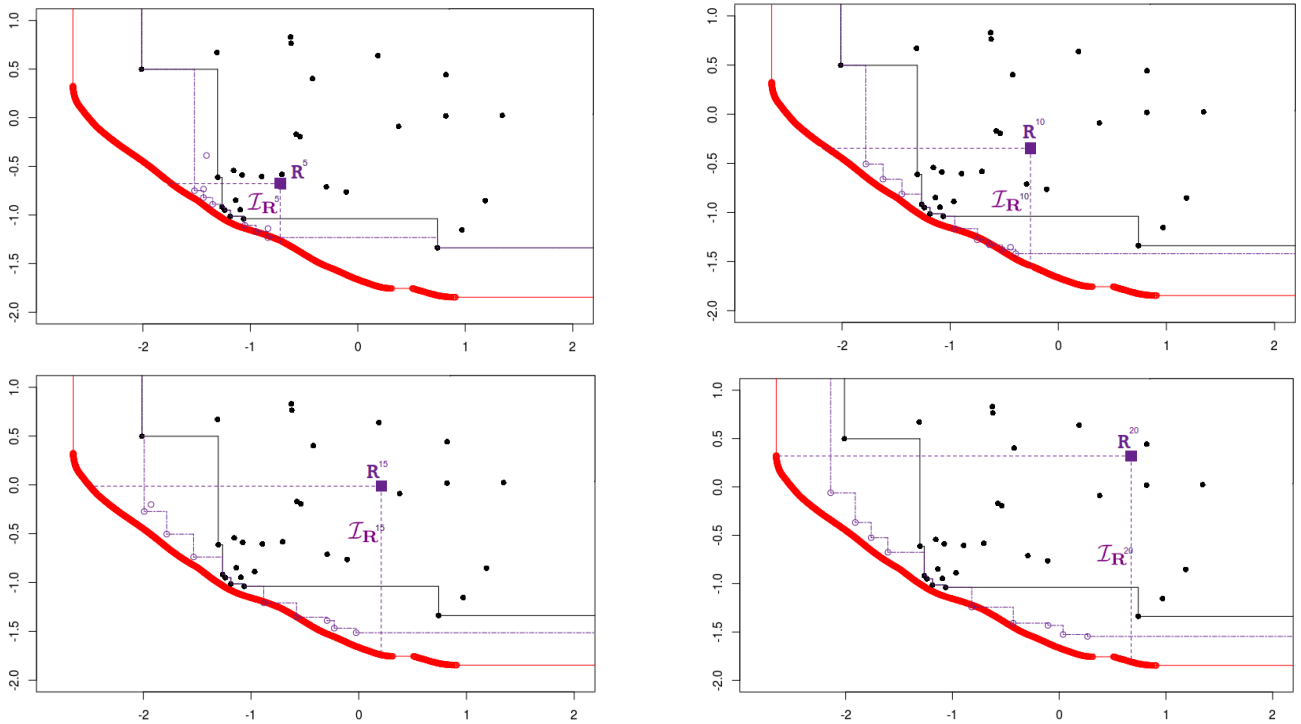


- Quantify uncertainty on each final virtual front through *volume-uncertainty*,
$U(\mathbf{R}; \mathbf{Y}(\cdot)) := \frac{1}{Vol(\mathbf{I},\mathbf{R})} \int_{\mathbf{I} \preceq \mathbf{y} \preceq \mathbf{R}} \widehat{p}(\mathbf{y})(1 - \widehat{p}(\mathbf{y}))d\mathbf{y}$.

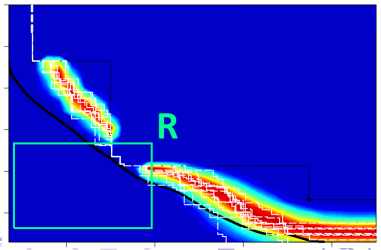


Expansion of the approximation front

- For increasing $\mathbf{R}^c$'s, perform the b last iterations substituting $\mathbf{f}(\cdot)$ by $\widehat{\mathbf{y}}(\cdot) \Rightarrow$ final virtual metamodels $\mathbf{Y}_c^{KB}(\cdot)$ and fronts depending on $\mathbf{R}^c$.

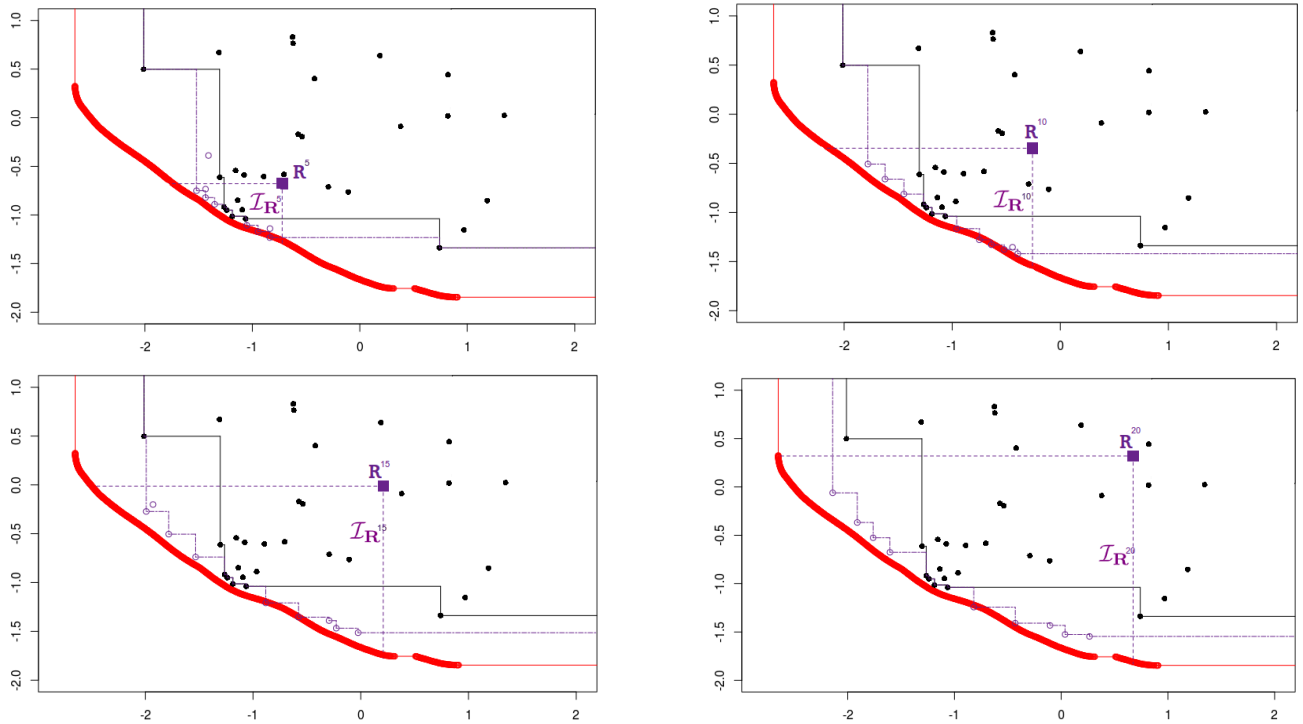


- Quantify uncertainty on each final virtual front through *volume-uncertainty*,
$U(\mathbf{R}; \mathbf{Y}(\cdot)) := \frac{1}{Vol(\mathbf{I},\mathbf{R})} \int_{\mathbf{I} \preceq \mathbf{y} \preceq \mathbf{R}} \widehat{p}(\mathbf{y})(1 - \widehat{p}(\mathbf{y}))d\mathbf{y}$.

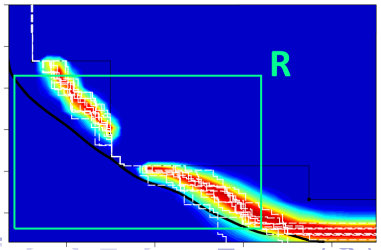


Expansion of the approximation front

- For increasing $\mathbf{R}^c$'s, perform the b last iterations substituting $\mathbf{f}(\cdot)$ by $\widehat{\mathbf{y}}(\cdot) \Rightarrow$ final virtual metamodels $\mathbf{Y}_c^{KB}(\cdot)$ and fronts depending on $\mathbf{R}^c$.

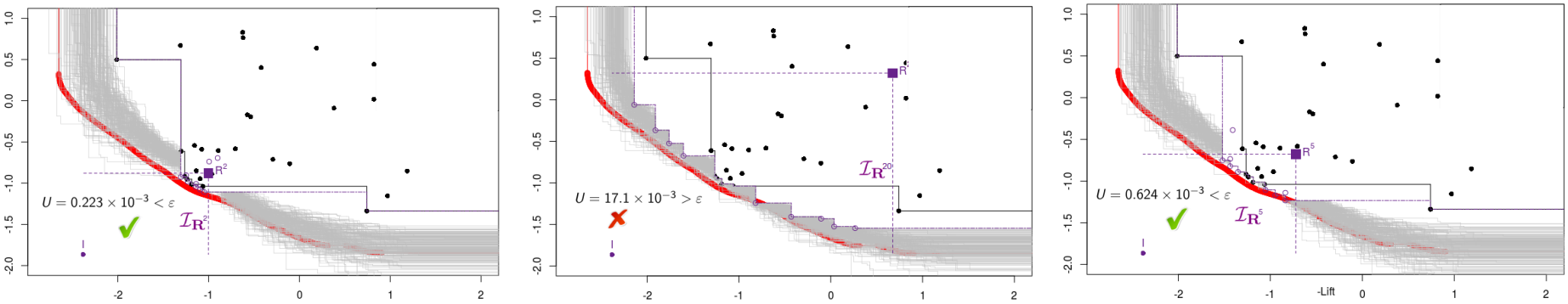


- Quantify uncertainty on each final virtual front through *volume-uncertainty*,
$U(\mathbf{R}; \mathbf{Y}(\cdot)) := \frac{1}{Vol(\mathbf{I},\mathbf{R})} \int_{\mathbf{I} \preceq \mathbf{y} \preceq \mathbf{R}} \widehat{p}(\mathbf{y})(1 - \widehat{p}(\mathbf{y}))d\mathbf{y}$.



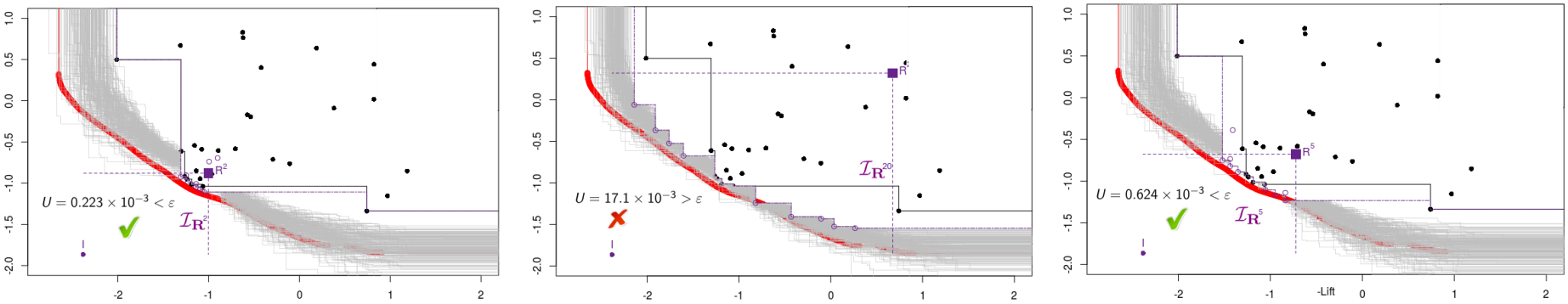
Uncertainty on final fronts

- Determine $\mathbf{R}^* := \arg\max_{\mathbf{R}\in\widehat{\mathcal{L}}} \|\mathbf{IR}\|$ s.t. $U(\mathbf{R}^c; \mathbf{Y}_c^{KB}) < \varepsilon$.

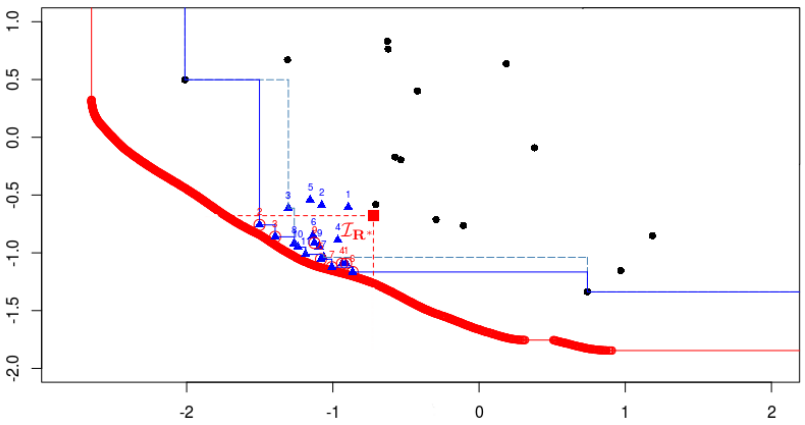


Uncertainty on final fronts

- Determine $\mathbf{R}^* := \arg\max_{\mathbf{R}\in\widehat{\mathcal{L}}} \|\mathbf{IR}\|$ s.t. $U(\mathbf{R}^c; \mathbf{Y}_c^{KB}) < \varepsilon$.

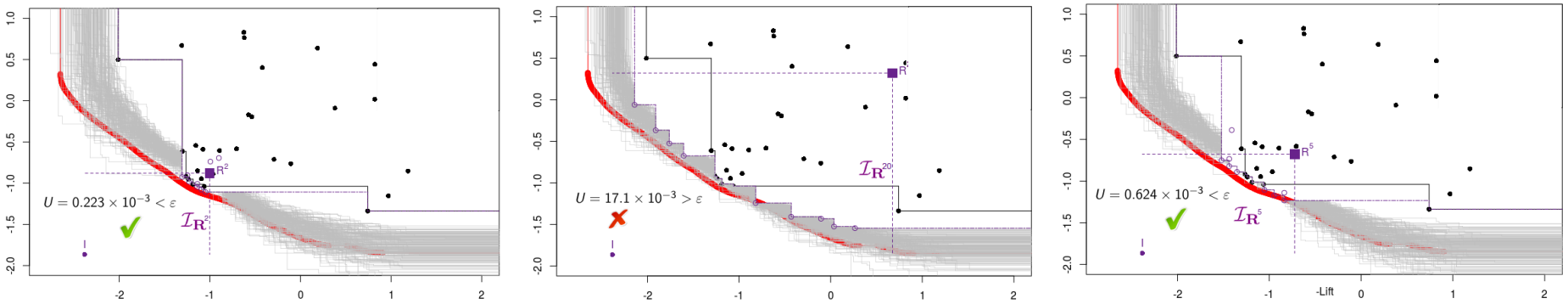


Perform the b remaining iterations with $\mathrm{EHI}(\cdot; \mathbf{R}^*)$.



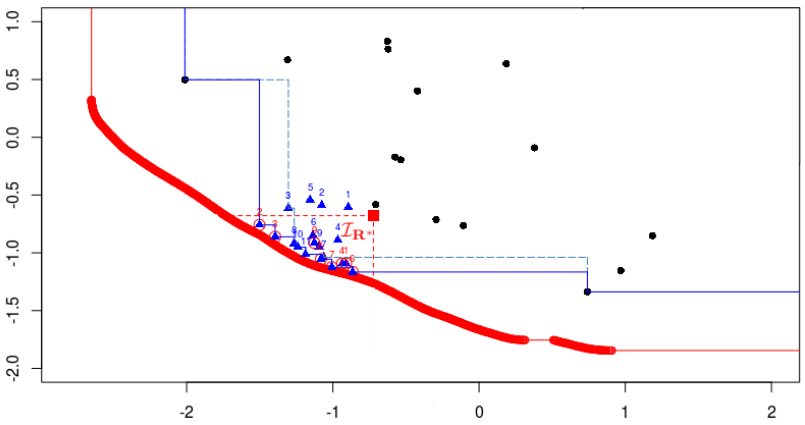
Uncertainty on final fronts

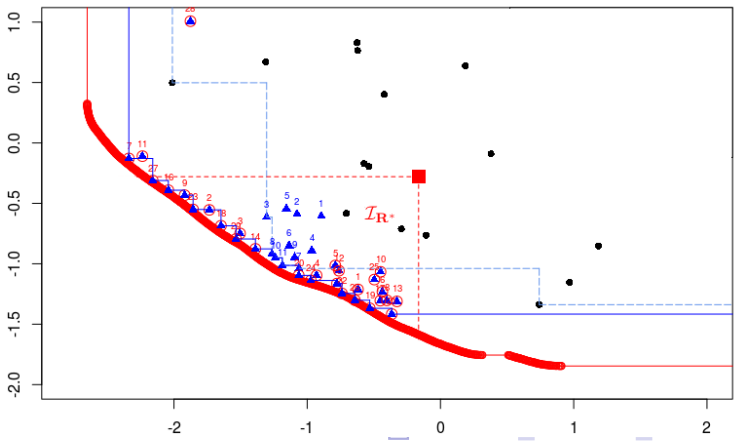
- Determine $\mathbf{R}^* := \arg\max_{\mathbf{R}\in\widehat{\mathcal{L}}} \|\mathbf{IR}\|$ s.t. $U(\mathbf{R}^c; \mathbf{Y}_c^{KB}) < \varepsilon$.



Perform the b remaining iterations with $\mathrm{EHI}(\cdot; \mathbf{R}^*)$.

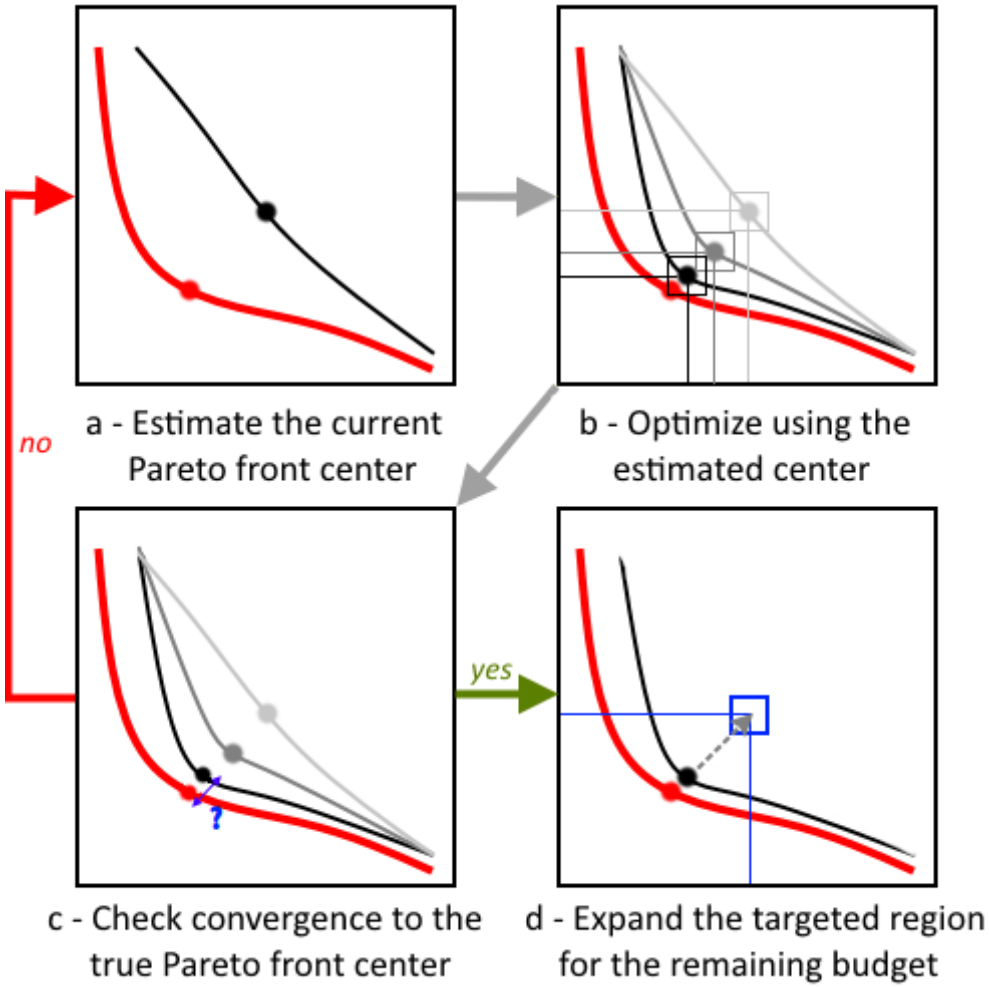
If 29 iterations were available, a broader area would have been targeted.





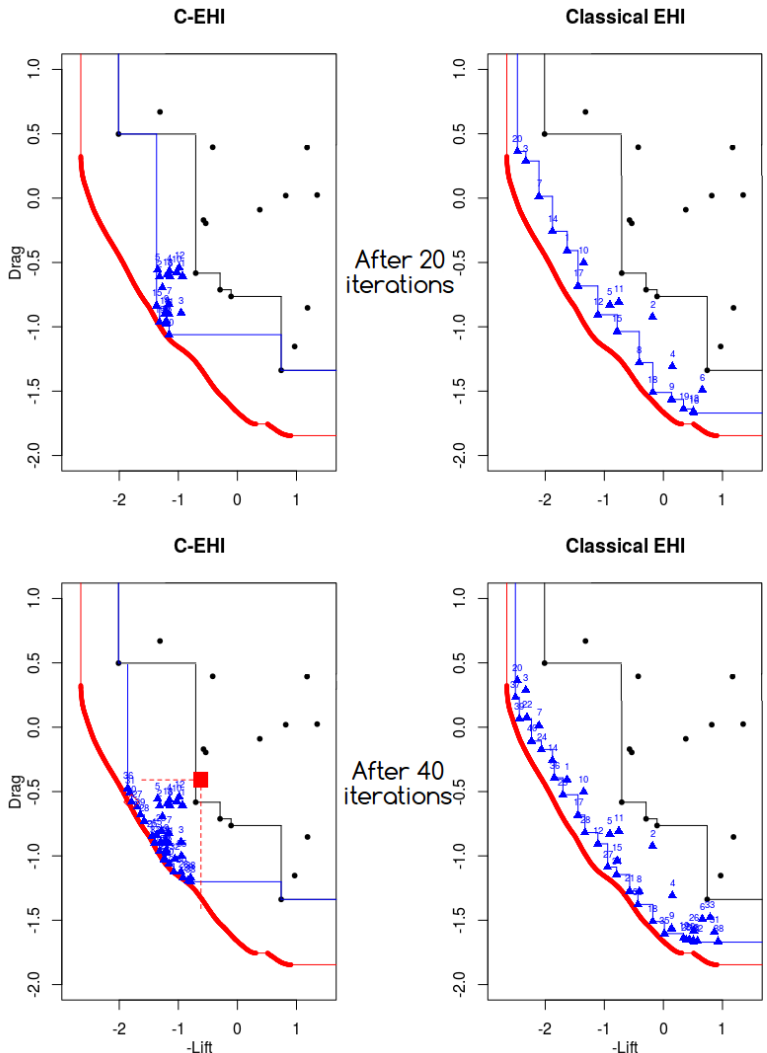
Summary: the C-EHI algorithm

The C-EHI algorithm: an algorithm in two steps to prioritize the center of $\mathcal{P}_{\mathcal{Y}}$, in accordance with resources.



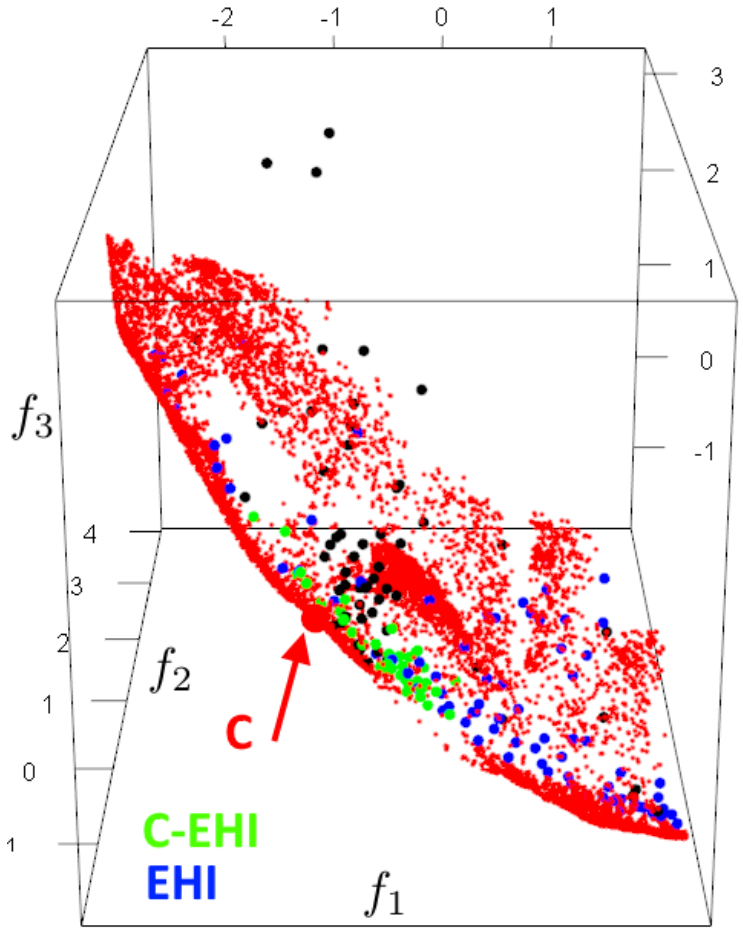
Summary: the C-EHI algorithm

Comparison with the original EHI: better uncovering of the central part of $\mathcal{P}_{\mathcal{Y}}$, at the cost of a narrower covering of the front.



Summary: the C-EHI algorithm

Comparison with the original EHI: better uncovering of the central part of $\mathcal{P}_{\mathcal{Y}}$, at the cost of a narrower covering of the front.



Batch mEI criterion

- Parallel capabilities $\Rightarrow$ return $\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}$.

Batch mEI criterion

- Parallel capabilities $\Rightarrow$ return $\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}$.
- q-mEI [Gaudrie et al., 2019]:

$$\text{q-mEI}(\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}; \mathbf{R}) = \mathbb{E}\left[\max_{i=1,\ldots,q}\left(\prod_{j=1}^{m}(R_j - Y_j(\mathbf{x}^{(t+i)}))_+\right)\right].$$

Batch mEI criterion

- Parallel capabilities $\Rightarrow$ return $\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}$.
- q-mEI [Gaudrie et al., 2019]:

$$\text{q-mEI}(\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}; \mathbf{R}) = \mathbb{E}\left[\max_{i=1,\ldots,q}\left(\prod_{j=1}^{m}(R_j - Y_j(\mathbf{x}^{(t+i)}))_+\right)\right].$$

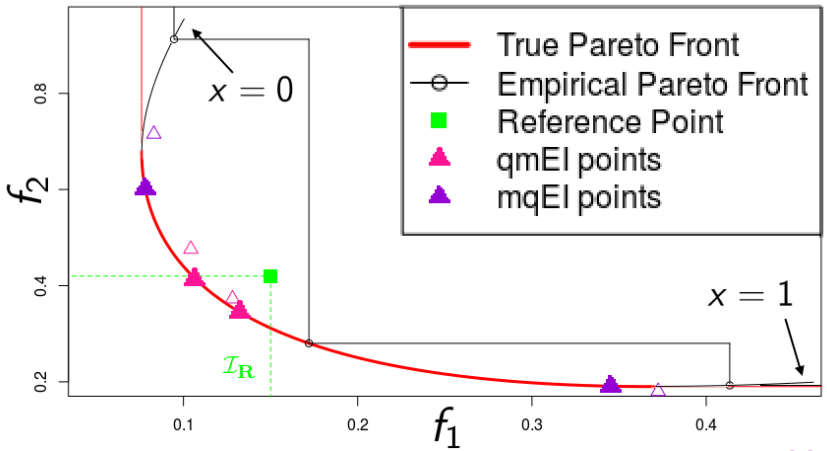
- Remark: product of q-EI's?

$$\text{mq-EI}(\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}; \mathbf{R}) = \prod_{j=1}^{m} \text{q-EI}_j(\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q)}\}; R_j)$$

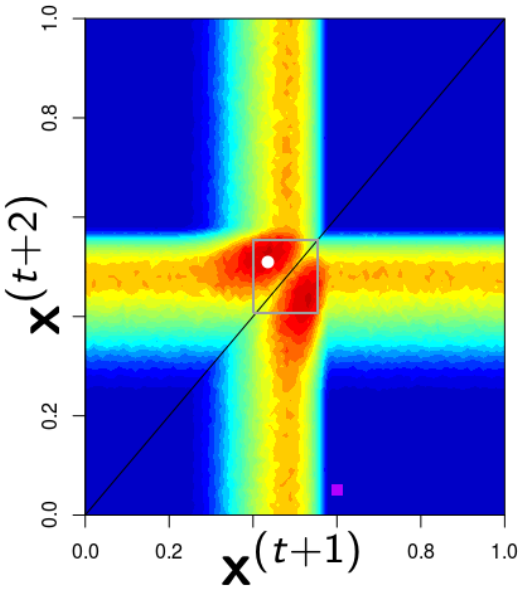
$$= \prod_{j=1}^{m} \mathbb{E}\left[\max_{i=1,\ldots,q}(R_j - Y_j(\mathbf{x}^{(t+i)}))_+\right] = \mathbb{E}\left[\prod_{j=1}^{m} \max_{i=1,\ldots,q}(R_j - Y_j(\mathbf{x}^{(t+i)}))_+\right].$$

Batch mEI criterion

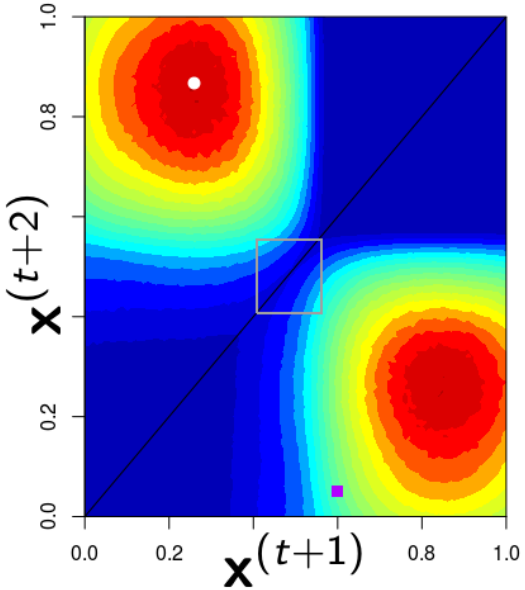
Example: targeting with q-mEI or mq-EI ($x \in [0,1]$, $q = 2$).



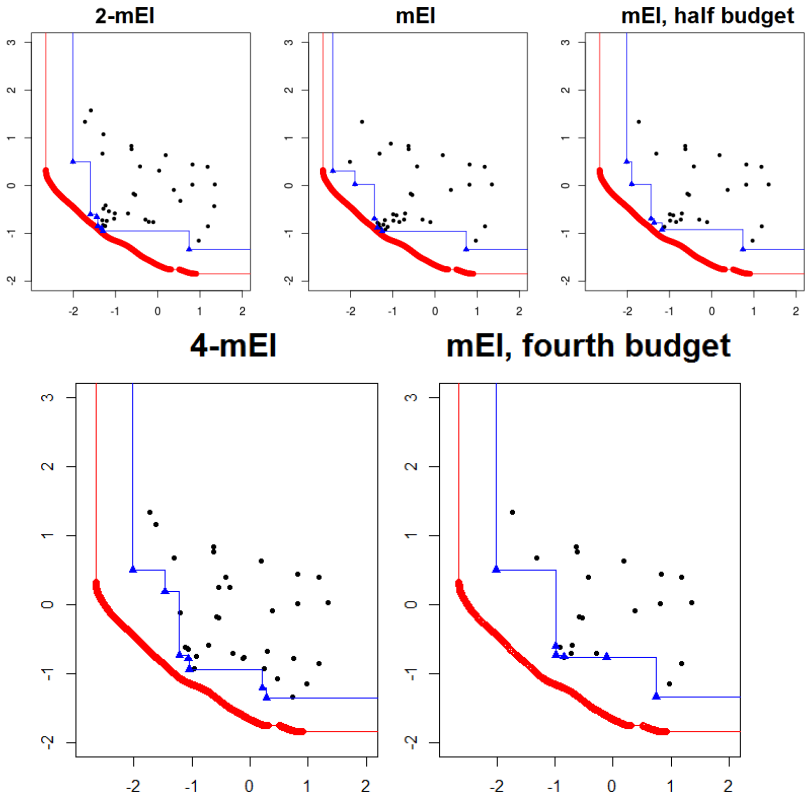
q-mEI



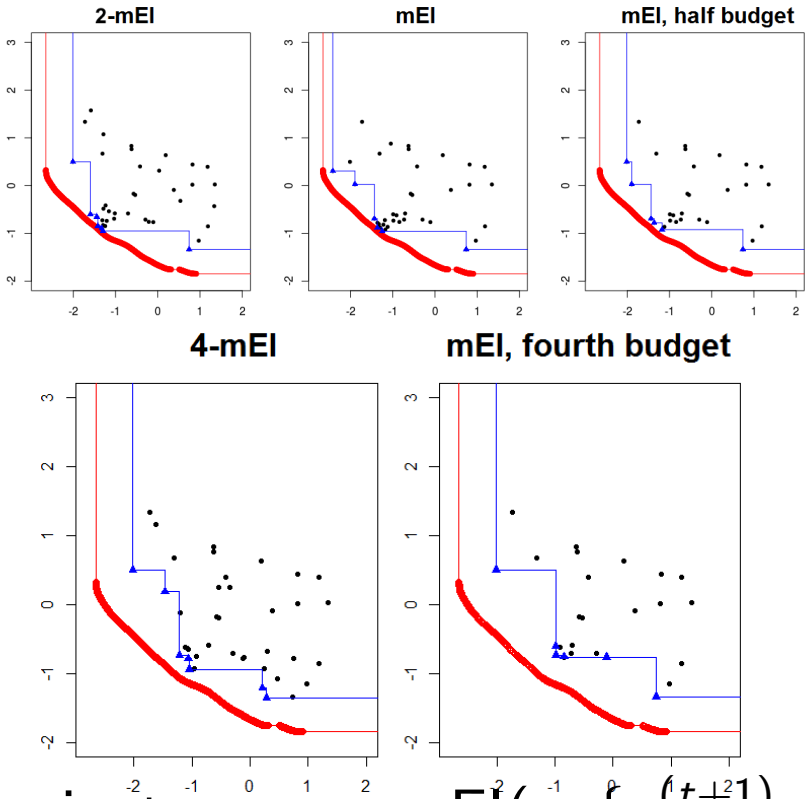
m-qEI



The q-mEI criterion

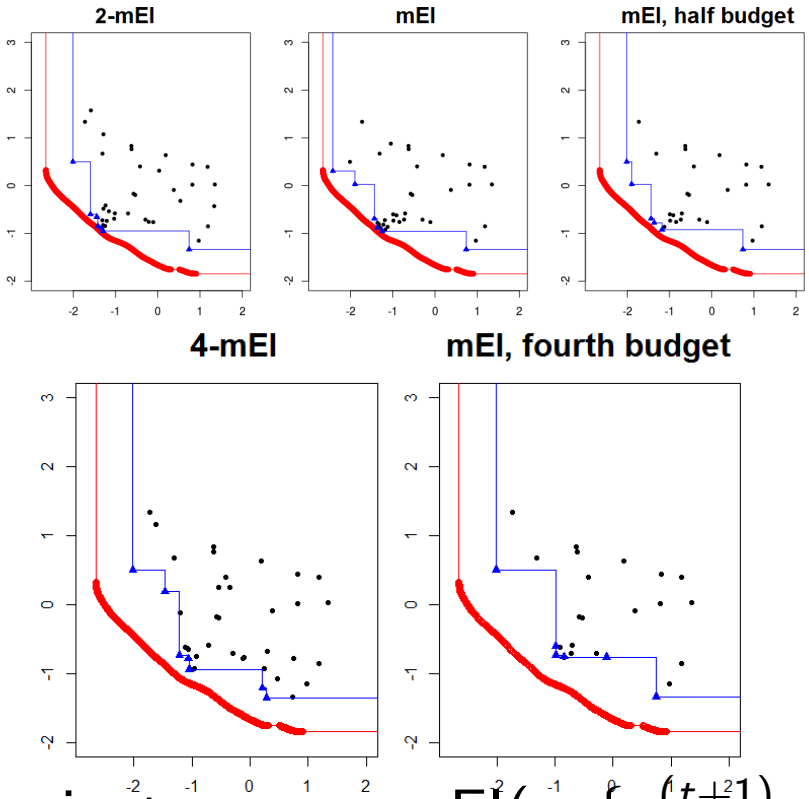


The q-mEI criterion



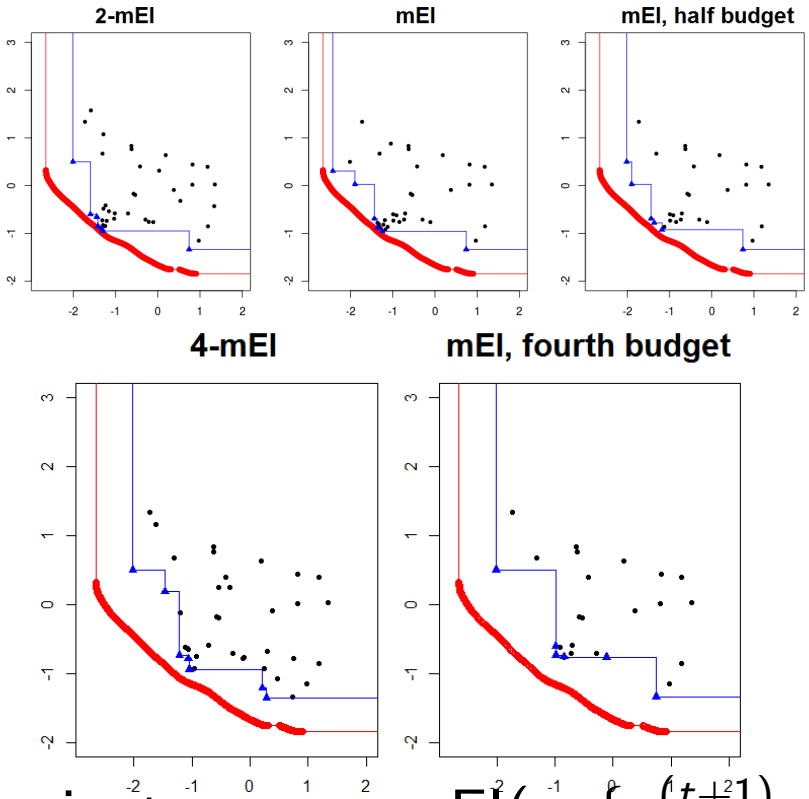
- Asynchronous variant: $\max_{\mathbf{x}\in X}$ q-mEI$(\mathbf{x}; \underbrace{\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q-1)}\}}_{\text{fixed}}; \mathbf{R})$.

The q-mEI criterion



- Asynchronous variant: $\max_{\mathbf{x}\in X}$ q-mEI$(\mathbf{x}; \underbrace{\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q-1)}\}}_{\text{fixed}}; \mathbf{R})$.
- Estimation: Monte Carlo simulation.

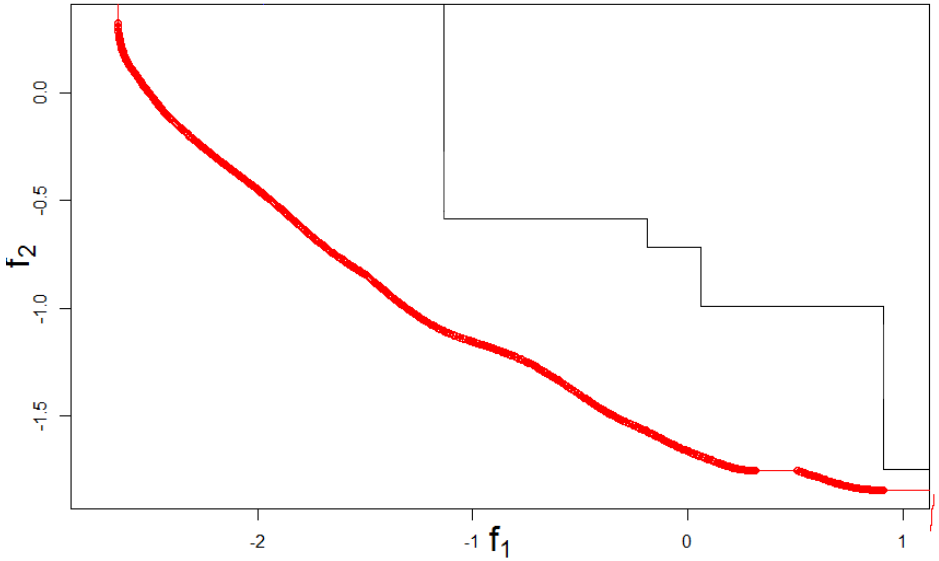
The q-mEI criterion



- Asynchronous variant: $\max_{\mathbf{x}\in X}$ q-mEI$(\mathbf{x}; \underbrace{\{\mathbf{x}^{(t+1)}, \ldots, \mathbf{x}^{(t+q-1)}\}}_{\text{fixed}}; \mathbf{R})$.
- Estimation: Monte Carlo simulation.
- Analytical expression [Chevalier and Ginsbourger, 2013]?
- Proxy for its gradient [Marmin et al., 2016, Wu and Frazier, 2016]?

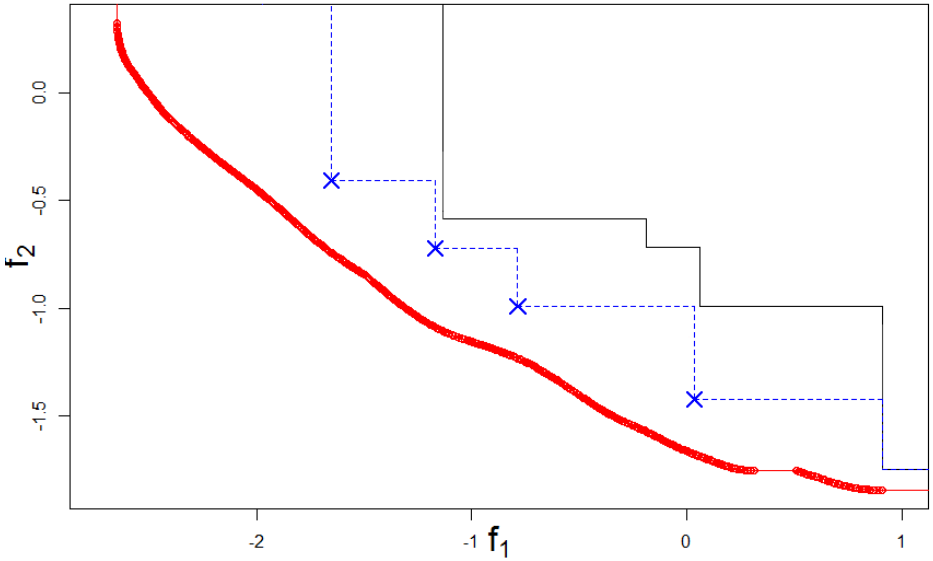
Multi-point EHI: the q-EHI criterion

Multiobjective problems well-suited for batch criteria: collaborative uncovering of $\mathcal{P}_{\mathcal{Y}}$ [Gaudrie, 2019].



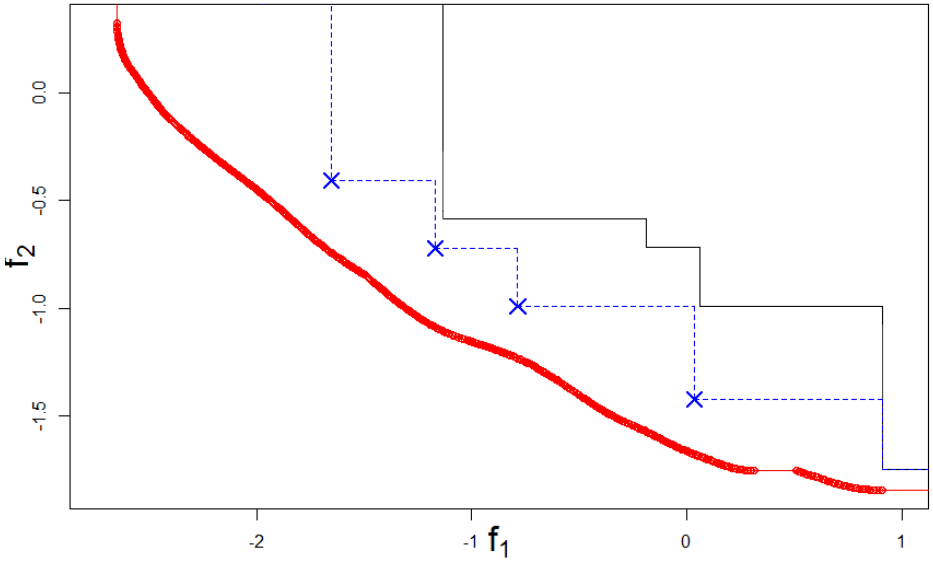
Multi-point EHI: the q-EHI criterion

Multiobjective problems well-suited for batch criteria: collaborative uncovering of $\mathcal{P}_{\mathcal{Y}}$ [Gaudrie, 2019].



Multi-point EHI: the q-EHI criterion

Multiobjective problems well-suited for batch criteria: collaborative uncovering of $\mathcal{P}_{\mathcal{Y}}$ [Gaudrie, 2019].



$$\text{q-EHI}(\{\mathbf{x}^{(t+1)},\ldots,\mathbf{x}^{(t+q)}\};\mathbf{R}) = \mathbb{E}[I_H(\widehat{\mathcal{P}_{\mathcal{Y}}}\cup\{\mathbf{Y}(\mathbf{x}^{(t+1)},\ldots,\mathbf{x}^{(t+q)})\};\mathbf{R})]$$

Multi-point EHI: the q-EHI criterion

- Computationally expensive (Monte Carlo $\Rightarrow$ average of n_{sim} hypervolumes).

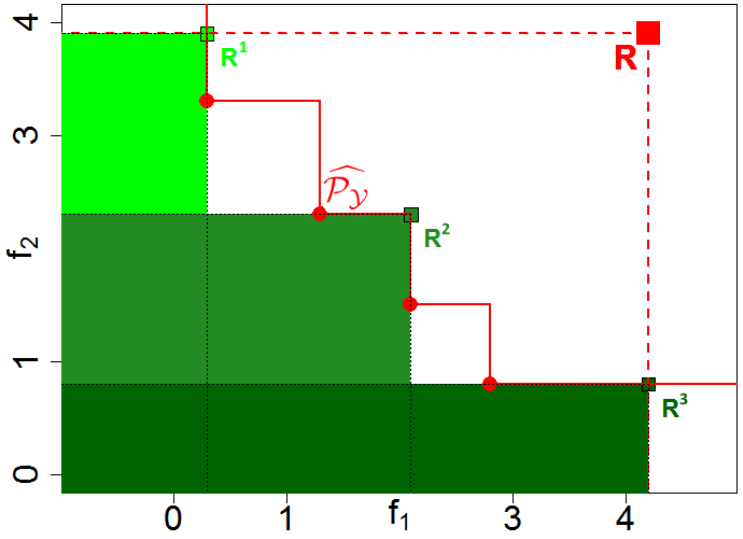
Multi-point EHI: the q-EHI criterion

- Computationally expensive (Monte Carlo $\Rightarrow$ average of n_{sim} hypervolumes).
- Kriging-Believer strategy.

Multi-point EHI: the q-EHI criterion

- Computationally expensive (Monte Carlo $\Rightarrow$ average of n_{sim} hypervolumes).
- Kriging-Believer strategy.
- Batch targeting through $\mathrm{mEI}(\cdot; \mathbf{R}^1), \ldots, \mathrm{mEI}(\cdot; \mathbf{R}^q)$.

$$\mathbf{x}^{(t+1)} \qquad\qquad \mathbf{x}^{(t+q)}$$



Conclusions

- New multiobjective methods that account for the budget: (i) prioritize relevant solutions, (ii) widen the targeted region.
- Batch criteria for parallel computations.
- Adaptations to constraints [Gaudrie, 2019].

References I

Auger, A., Bader, J., Brockhoff, D., and Zitzler, E. (2009).
Articulating user preferences in many-objective problems by sampling the weighted hypervolume.
In Proceedings of the 11th Annual conference on Genetic and evolutionary computation, pages 555–562. ACM.

Chevalier, C. and Ginsbourger, D. (2013).
Fast computation of the multi-points expected improvement with applications in batch selection.
In International Conference on Learning and Intelligent Optimization, pages 59–69. Springer.

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

Gaudrie, D. (2019).
High-dimensional Bayesian Multi-Objective Optimization.
PhD thesis, École Nationale Supérieure des Mines de Saint-Étienne.

Gaudrie, D., Le Riche, R., Picheny, V., Enaux, B., and Herbert, V. (2018a).
Budgeted multi-objective optimization with a focus on the central part of the Pareto front - extended version.
arXiv preprint arXiv:1809.10482.

Gaudrie, D., Le Riche, R., Picheny, V., Enaux, B., and Herbert, V. (2018b).
Targeting well-balanced solutions in multi-objective Bayesian optimization under a restricted budget.
In International Conference on Learning and Intelligent Optimization, pages 175–179. Springer.

Gaudrie, D., Le Riche, R., Picheny, V., Enaux, B., and Herbert, V. (2019).
Targeting solutions in Bayesian multi-objective optimization: Sequential and batch versions.
Annals of Mathematics and Artificial Intelligence.

References II

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

Kalai, E. and Smorodinsky, M. (1975).
Other solutions to Nash's bargaining problem.
Econometrica: Journal of the Econometric Society, pages 513–518.

Marmin, S., Chevalier, C., and Ginsbourger, D. (2016).
Efficient batch-sequential Bayesian optimization with moments of truncated Gaussian vectors.
arXiv preprint arXiv:1609.02700.

Wu, J. and Frazier, P. (2016).
The parallel knowledge gradient method for batch Bayesian optimization.
In Advances in Neural Information Processing Systems, pages 3126–3134.

Yang, K., Li, L., Deutz, A., Back, T., and Emmerich, M. (2016).
Preference-based multiobjective optimization using truncated expected hypervolume improvement.
In Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on, pages 276–281. IEEE.

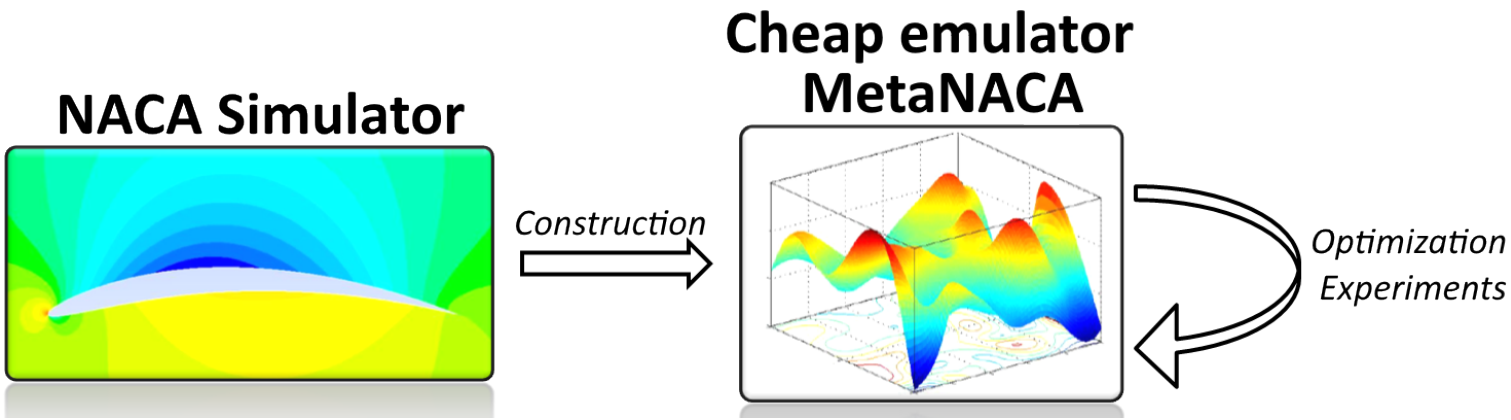
Zitzler, E. (1999).
Evolutionary algorithms for multiobjective optimization: Methods and applications.

Thank you for your attention,

Do you have any question?

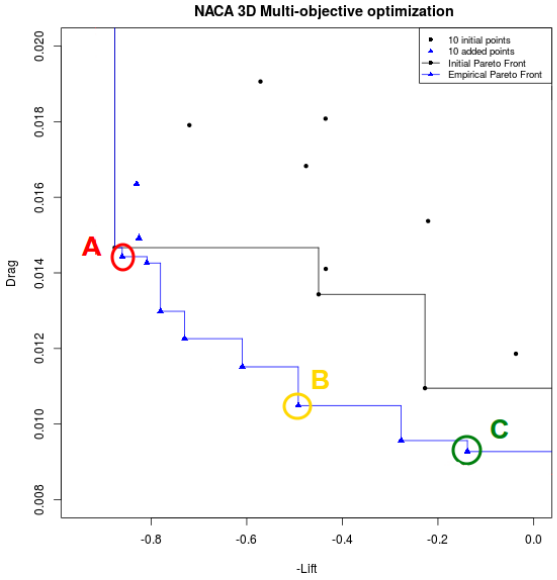
MetaNACA test bed

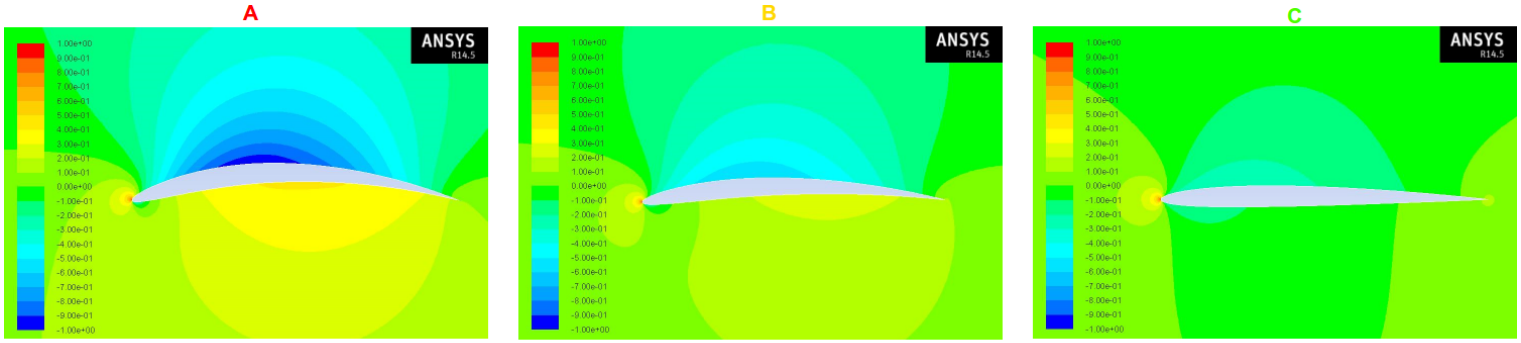
- For each test case (dimension $d = 3$, 8 or 22), surrogate model of an aerodynamic simulation using
  - 1000 points (complete factorial design) in 3D,
  - 1200 points (LHS-maximin design + refinement in areas of compromise) in 8D and 22D.



- Variable number of objectives: $m \in \{2, 3, 4\}$.

Attractive solutions: central solutions





Drag (minimize)

+ ←

Lift (maximize)

Convergence towards preferred parts of $\mathcal{P}_{\mathcal{Y}}$

R-mEI, ZDT3 function

$\#\mathbf{f}(\cdot)$	mEI 20+20	EHI 20+20	$\text{EHI}_{(11,11)}$ 20+20	NSGA-II_1 20+20	NSGA-II_{19} 20+380
Attainment time	24.2 (2.6)	45.3 [7]	103.3 [3]	×	341.5 [7]
Hypervolume	0.634 (0.078)	0.218 (0.353)	0.112 (0.211)	0	0.248 (0.253)
Solutions $\prec$ **R**	4.1 (1.8)	1.1 (1.9)	0.3 (0.5)	0	4.2 (4.1)

R-mEI, P1 function

$\#\mathbf{f}(\cdot)$	mEI 8+12	EHI 8+12	NSGA-II_1 12+12	NSGA-II_{11} 12+132
Attainment time	12.6 (3.5)	25.6 [5]	120 [1]	67.1 [8]
Hypervolume	0.620 (0.165)	0.163 (0.213)	0.043 (0.136)	0.394 (0.295)
Solutions $\prec$ **R**	6.5 (2.5)	0.6 (0.7)	0.2 (0.6)	2.8 (2.4)

C-mEI, MetaNACA

	$d = 8$				$d = 22$			
Criterion	mEI	EHI	NSGA-II_1	NSGA-II_{19}	mEI	EHI	NSGA-II_4	NSGA-I
$\#\mathbf{f}(\cdot)$	20+20	20+20	20+20	20+180	50+50	50+50	20+80	20+48
ainment time	28.4 (5.4)	66.8 [5]	×	261.9 [6]	56.3 (7.2)	71.4 (13.9)	×	191.9 [
ypervolume	0.256 (0.09)	0.025 (0.04)	0	0.044 (0.08)	0.222 (0.12)	0.153 (0.09)	0	0.106 (0

Convergence towards preferred parts of $\mathcal{P}_{\mathcal{Y}}$

C-EHI, P1 function

	Hypervolume			Attainment time		
w	0.05	0.15	0.25	0.05	0.15	0.25
C-EHI	0.185 (0.233)	0.549 (0.263)	0.668 (0.185)	21.6 [7]	13.1 (2.7)	9.5 (1)
EHI	0.155 (0.218)	0.465 (0.179)	0.611 (0.114)	39.4 [4]	13.2 (2.6)	11.4 (2.6)
$\text{EHI}_{\mathcal{P}_{\mathcal{Y}}}$	0.269 (0.260)	0.446 (0.175)	0.636 (0.136)	30.0 [6]	14 (3.2)	11 (2.6)
EHI_{N}	0.130 (0.158)	0.312 (0.223)	0.460 (0.192)	32.4 [5]	16.7 [9]	11.5 (3.5)
EHI_{M}	0.012 (0.039)	0.202 (0.181)	0.389 (0.136)	180 [1]	22.7 [7]	12.6 (4.1)
NSGA-II_b	0	0.052 (0.110)	0.107 (0.183)	×	80 [2]	51.1 [3]
NSGA-II_+	0.188 (0.219)	0.576 (0.109)	0.705 (0.069)	169.6 [5]	50.4 (31.1)	41.3 (31.9)

C-EHI, ZDT1 function

	Hypervolume			Attainment time		
w	0.05	0.15	0.25	0.05	0.15	0.25
C-EHI	0.703 (0.049)	0.895 (0.010)	0.936 (0.006)	26.8 (6.6)	23.4 (2.2)	23.4 (2.2)
EHI	0.065 (0.154)	0.097 (0.204)	0.101 (0.213)	145 [2]	145 [2]	145 [2]
$\text{EHI}_{\mathcal{P}_{\mathcal{Y}}}$	0.611 (0.066)	0.848 (0.029)	0.901 (0.023)	28.7 (2.8)	22.8 (2.3)	21.4 (0.5)
EHI_{N}	0.362 (0.349)	0.650 (0.246)	0.740 (0.206)	48.1 [6]	22.2 (0.4)	22.2 (0.4)
EHI_{M}	0.575 (0.107)	0.845 (0.038)	0.906 (0.022)	24.4 (5.6)	22.2 (0.6)	22.1 (0.3)
$\text{EHI}_{(11,11)}$	0.133 (0.13)	0.327 (0.251)	0.472 (0.218)	120 [2]	120 [2]	59.2 [5]
NSGA-II_b	0	0	0	×	×	×
NSGA-II_+	0.375 (0.161)	0.749 (0.075)	0.842 (0.052)	532.9 (143.4)	331.9 (121)	219.2 (101.5)

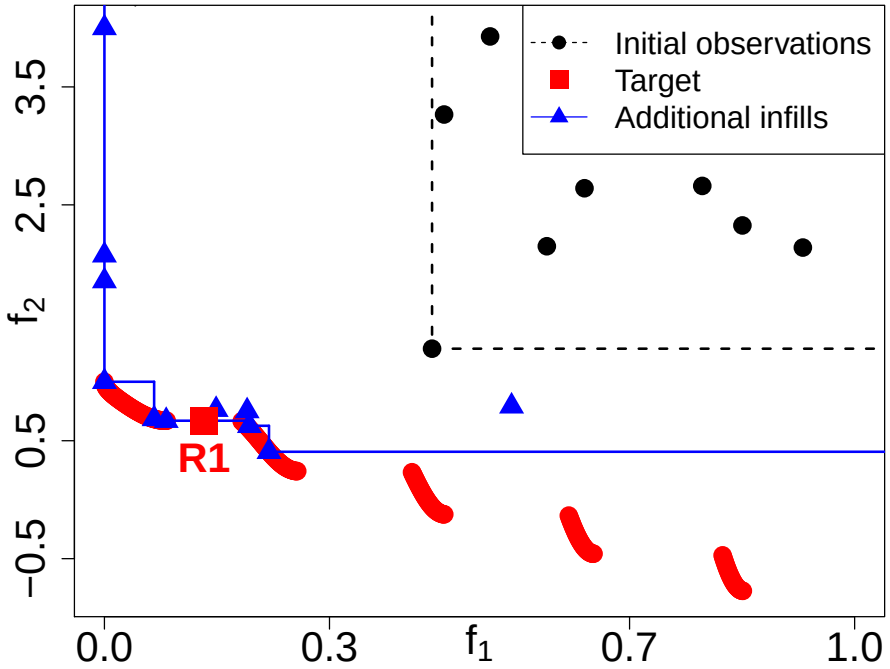
Convergence towards preferred parts of $\mathcal{P}_{\mathcal{Y}}$

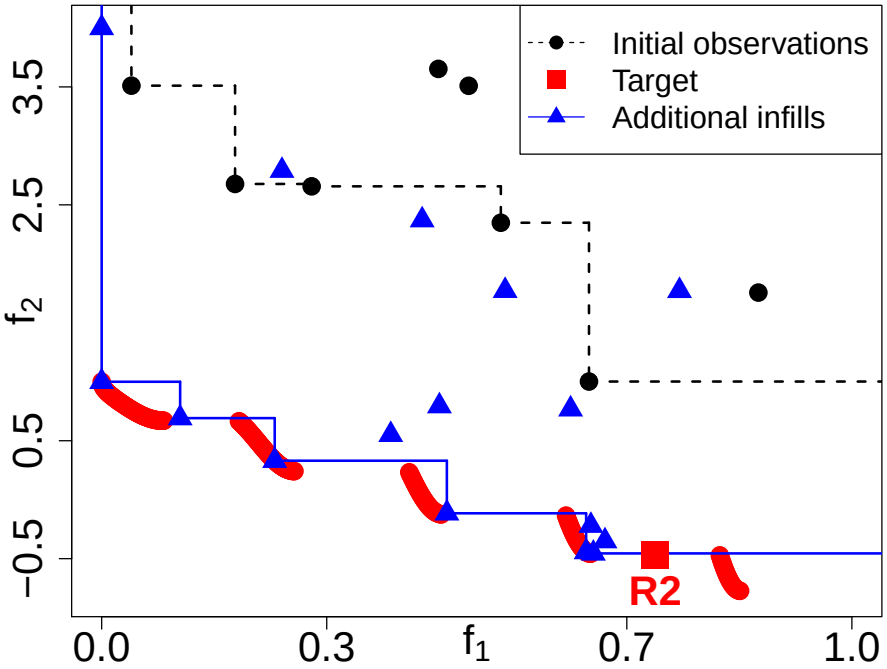
C-EHI, MetaNACA

m	*budget*	$\mathbf{R}^{0.1}$		$\mathbf{R}^{0.2}$		$\mathbf{R}^{0.3}$	
		C-EHI	EHI	C-EHI	EHI	C-EHI	EHI
2	40	0.275 (0.18)	0.025 (0.04)	0.498 (0.17)	0.227 (0.15)	0.581 (0.10)	0.386 (0.19)
	60	0.377 (0.19)	0.096 (0.12)	0.651 (0.11)	0.342 (0.14)	0.719 (0.09)	0.525 (0.12)
	80	0.548 (0.10)	0.118 (0.11)	0.759 (0.05)	0.398 (0.12)	0.821 (0.03)	0.572 (0.11)
	100	0.524 (0.14)	0.153 (0.16)	0.744 (0.08)	0.503 (0.13)	0.831 (0.05)	0.658 (0.08)
3	40	0.013 (0.02)	0 (0)	0.181 (0.09)	0.086 (0.05)	0.319 (0.05)	0.237 (0.07)
	60	0.058 (0.06)	0.010 (0.02)	0.267 (0.08)	0.136 (0.06)	0.394 (0.05)	0.305 (0.04)
	80	0.109 (0.08)	0.012 (0.02)	0.327 (0.14)	0.170 (0.10)	0.447 (0.17)	0.321 (0.13)
	100	0.160 (0.09)	0.016 (0.02)	0.412 (0.07)	0.218 (0.06)	0.546 (0.04)	0.391 (0.06)
4	40	0.113 (0.11)	0.075 (0.10)	0.291 (0.09)	0.240 (0.10)	0.374 (0.06)	0.378 (0.09)
	60	0.187 (0.15)	0.138 (0.09)	0.356 (0.08)	0.340 (0.09)	0.418 (0.05)	0.473 (0.07)
	80	0.312 (0.16)	0.198 (0.08)	0.470 (0.09)	0.413 (0.07)	0.516 (0.09)	0.533 (0.06)
	100	0.519 (0.08)	0.219 (0.07)	0.612 (0.11)	0.464 (0.07)	0.642 (0.12)	0.580 (0.06)

Targeting a hole

ZDT3 function





q-mEI/mEI comparison I

Meta NACA 8

	mEI			q-mEI			
Criterion	mEI	mEI_{half}	$\text{mEI}_{\text{fourth}}$	2mEI	2mEI_t	4mEI	4mEI_t
Budget	20+20	20+20/2	20+20/4	20+2×10	20+2×20	20+4×5	20+4×20
$\mathbf{f}(\cdot)$ to target	28.4 (5.4)	36.9 [7]	72.2 [3]	35.6 [8]	33.1 (5.7)	71.9 [4]	44.5 (17.3
crit to target	8.4 (5.4)	8.4 [7]	5.5 [3]	5.3 [8]	6.6 (2.9)	5.5 [4]	6.1 (4.3)
Hypervolume	0.256 (0.09)	0.134 (0.15)	0.077 (0.13)	0.170 (0.13)	0.280 (0.16)	0.056 (0.09)	0.296 (0.1

	q-mEI-KB					
Criterion	2mEI-KB	2mEI-KB_t	4mEI-KB	4mEI-KB_t	10mEI-KB	10mEI-KB_t
Budget	20+2×10	20+2×20	20+4×5	20+4×20	20+10×2	20+10×8
#$\mathbf{f}(\cdot)$ to target	34.2 [9]	33 (8.9)	57.6 [5]	38.5 (11.3)	49.2 [6]	37.2 (13.7)
#*crit* to target	6.0 [9]	6.5 (2.3)	4.4 [5]	4.6 (11.3)	2.8 [6]	2.5 (1.4)
Hypervolume	0.221 (0.14)	0.361 (0.12)	0.128 (0.16)	0.466 (0.25)	0.040 (0.08)	0.531 (0.17)

Meta NACA 22

	mEI		q-mEI	
Criterion	mEI	mEI_{half}	2mEI	2mEI_t
Budget	50+50	50+50/2	50+2×25	50+2×50
#$\mathbf{f}(\cdot)$ to target	56.3 (7.2)	56.3 (7.2)	71.3 [8]	71.3 [8]
#*crit* to target	6.3 (7.2)	6.3 (7.2)	4.7 [8]	4.7 [8]
Hypervolume	0.222 (0.12)	0.139 (0.10)	0.085 (0.09)	0.119 (0.10)

q-mEI/mEI comparison II

Criterion	q-mEI-KB					
	2mEI-KB	2mEI-KB$_t$	4mEI-KB	4mEI-KB$_t$	10mEI-KB	10mEI-KB$_t$
Budget	50+2×25	50+2×50	50+4×12	50+4×25	50+10×5	50+10×25
#$\mathbf{f}(\cdot)$ to target	73.0 [8]	68.9 (23.0)	63.6 (12.7)	63.6 (12.7)	69.1 [9]	59.7 (7.2)
#*crit* to target	3.3 [8]	5.3 (5.8)	4 (3.1)	4 (3.1)	1.9 [9]	1.7 (0.7)
Hypervolume	0.121 (0.11)	0.260 (0.14)	0.215 (0.14)	0.398 (0.16)	0.100 (0.08)	0.440 (0.26)

ZDT3

Criterion	mEI	q-mEI				q-mEI-KB			
		2mEI	2mEI$_t$	4mEI	4mEI$_t$	2mEI-KB	2mEI-KB$_t$	4mEI-KB	4mEI-KB$_t$
Budget	20+20	20+2×10	20+2×20	20+4×5	20+4×20	20+2×10	20+2×20	20+4×5	20+4×20
#$\mathbf{f}(\cdot)$ to target	24.2 (2.6)	26.3 (4.3)	26.3 (4.3)	32.7 [9]	32.5 (6.6)	24.2 (2.6)	24.2 (2.6)	33.8 [8]	33.2 (15.8)
#*crit* to target	4.2 (2.6)	3.2 (2.2)	3.2 (2.2)	2.6 [9]	3.1 (1.7)	2.4 (1.3)	2.4 (1.3)	2.4 [8]	3.9 (4.0)
Hypervolume	0.634 (0.08)	0.548 (0.20)	0.621 (0.15)	0.424 (0.23)	0.622 (0.09)	0.513 (0.15)	0.513 (0.15)	0.445 (0.25)	0.518 (0.20)
Solutions $\preceq$ **R**	4.1 (1.8)	2.8 (1.0)	3.6 (0.8)	1.5 (1.0)	2.4 (1.0)	2.4 (0.7)	2.4 (0.7)	2.1 (0.9)	2.2 (0.8)

P1

Criterion	mEI	q-mEI				q-mEI-KB			
		2mEI	2mEI$_t$	4mEI	4mEI$_t$	2mEI-KB	2mEI-KB$_t$	4mEI-KB	4mEI-KB$_t$
Budget	8+12	8+2×6	8+2×12	8+4×3	8+4×12	8+2×6	8+2×12	8+4×3	8+4×12
#$\mathbf{f}(\cdot)$ to target	12.6 (3.5)	12.7 (2.6)	12.7 (2.6)	15.1 (3.9)	15.1 (3.9)	12.6 [9]	13 (7.8)	15.5 [8]	14.3 (7.3)
#*crit* to target	4.6 (3.5)	2.4 (1.3)	2.4 (1.3)	1.8 (1.0)	1.8 (1.0)	3.0 [9]	3.4 (2.9)	2.5 [8]	2.4 (1.4)
Hypervolume	0.620 (0.17)	0.624 (0.06)	0.686 (0.04)	0.437 (0.21)	0.718 (0.05)	0.393 (0.21)	0.451 (0.17)	0.205 (0.19)	0.540 (0.18)
Solutions $\preceq$ **R**	6.5 (2.5)	5.6 (1)	9.7 (1.1)	2.6 (1.6)	11.4 (2.1)	1.8 (0.9)	2.3 (1.5)	1.2 (0.8)	4.1 (3.1)

q-EHI/EHI comparison I

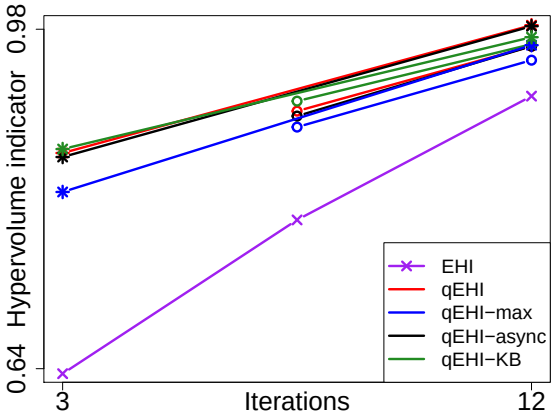
Criterion	EHI		
$\#\mathbf{f}(\cdot)$	12	6	3
P1	0.913 *(0.029)*	0.789 *(0.068)*	0.635 *(0.107)*
ZDT1	0.939 *(0.003)*	0.885 *(0.007)*	0.786 *(0.017)*
NACA	0.748 *(0.067)*	0.644 *(0.094)*	0.587 *(0.085)*

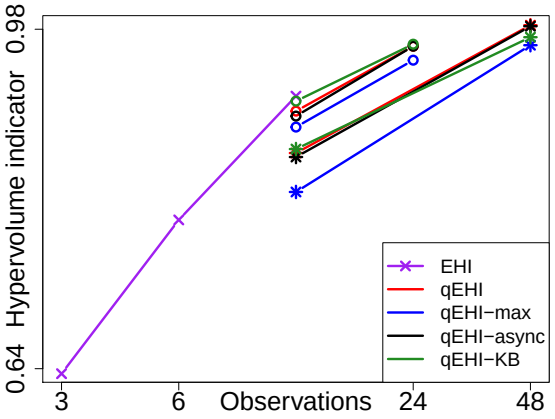
Criterion	2-EHI		4-EHI	
$\#\mathbf{f}(\cdot)$	2×12	2×6	2×12	2×6
P1	0.963 *(0.007)*	0.898 *(0.036)*	0.984 *(0.002)*	0.856 *(0.054)*
ZDT1	0.970 *(0.001)*	0.939 *(0.002)*	0.960 *(0.004)*	0.883 *(0.012)*
NACA	0.831 *(0.046)*	0.718 *(0.093)*	0.845 *(0.018)*	0.712 *(0.053)*

Criterion	2-EHI$_{\text{async}}$		4-EHI$_{\text{async}}$	
$\#\mathbf{f}(\cdot)$	2×12	2×6	4×12	4×3
P1	0.963 *(0.009)*	0.893 *(0.048)*	0.983 *(0.002)*	0.852 *(0.061)*
ZDT1	0.970 *(0.001)*	0.940 *(0.003)*	0.983 *(0.001)*	0.933 *(0.003)*
NACA	0.841 *(0.038)*	0.744 *(0.064)*	0.887 *(0.025)*	0.714 *(0.049)*

q-EHI/EHI comparison II

Criterion	2-EHI-KB		4-EHI-KB	
$\#\mathbf{f}(\cdot)$	2×12	2×6	4×12	4×3
P1	0.965 *(0.006)*	0.908 *(0.038)*	0.972 *(0.005)*	0.860 *(0.059)*
ZDT1	0.970 *(0.001)*	0.939 *(0.003)*	0.985 *(0)*	0.941 *(0.004)*
NACA	0.830 *(0.031)*	0.715 *(0.097)*	0.897 *(0.028)*	0.703 *(0.093)*





q-EHI/EHI comparison III