

Restricting ambitions in global multi-objective optimization of costly functions

David Gaudrie^{1,2}, Rodolphe Le Riche¹, Victor Picheny³,
Benoît Enaux², Vincent Herbert²

¹ CNRS LIMOS at Mines de Saint-Etienne, France
² PSA ³ Prowler.io

December 2018
ONERA
Châtillon, France

Context: metamodels for the optimization of costly functions (1)

- Multi-objective optimization of costly functions

$$\min_{\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^d} (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) \quad (1)$$

$f_i()$'s are costly functions: results of PDE solvers (e.g., finite element/volume models, boundary elements), Monte Carlo simulations, results of real experiments.

- Issue: cannot call more than (say) 100 times each $f_i()$ \Rightarrow cannot optimize, identify, propagate uncertainties.

Context: metamodels for the optimization of costly functions (2)

- Issue: cannot call more than (say) 100 times each $f_i()$ \Rightarrow cannot optimize, identify, propagate uncertainties.
- Solution (?): work with a metamodel i.e., a statistical model of already obtained $(\mathbf{x}^i, \mathbf{f}(\mathbf{x}^i))$, $i = 1, n$.
- Metamodels = neural nets, support vector machines, regression, GAMs, kriging, random forests, splines A lot of contributions, e.g., [17, 1, 13, 7, 3, 15, 14].
- With the metamodels, you can perform multiobjective, robust, global optimization?

Context: metamodels for the optimization of costly functions (3)

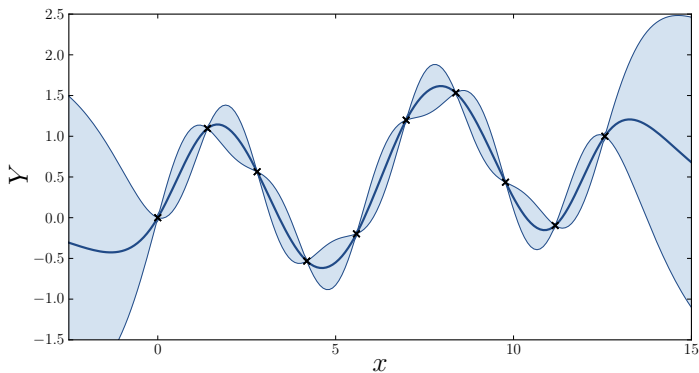
- With the metamodels, you can perform multiobjective, robust, global optimization?
- Not that simple: the curse of dimensionality is still hidden under the bed. You need *many* points to learn an accurate metamodel.
- So, back to the start? Not really:

metamodels as a way to proportionate ambitions and experimental budget.

Background: conditional Gaussian processes (1)

Conditional Gaussian processes (GPs) will be our metamodels here as they come with an uncertainty measure,

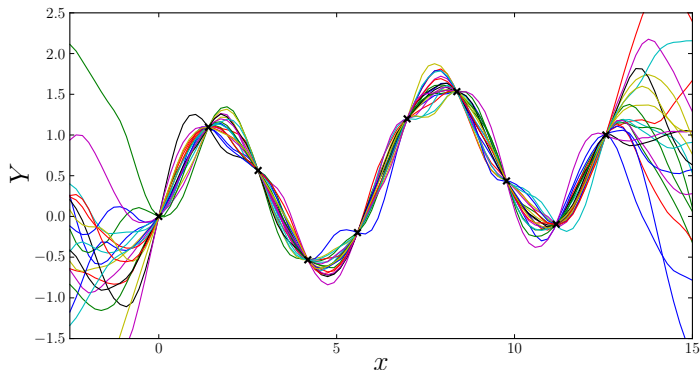
$$Y(\mathbf{x}) := F(\mathbf{x}) \mid f(\mathbf{x}^1) = f^1, \dots, f(\mathbf{x}^n) = f^n \sim \mathcal{N}(\mu(\mathbf{x}), C(\mathbf{x}, \mathbf{x}'))$$



$\mu(\cdot)$ and $C(\cdot, \cdot)$ known analytically, \mathbf{x} d -dimensional.

Background: conditional Gaussian processes (2)

Trajectories (possible functions) can be sampled, allowing Monte Carlo simulations



A lot more to be said about GPs, cf. [18, 4].

Background: multi-objective optimization (1)

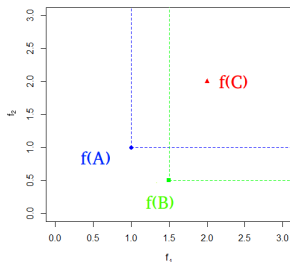
In general there are many (even an infinite number of) trade-off solutions to

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

called the Pareto set (in \mathcal{X}) or front (in \mathcal{Y}).

It is composed of Non-Dominated points,

$$\{\mathbf{x} \in \mathcal{X} : \nexists \mathbf{x}' \neq \mathbf{x} \in \mathcal{X}, \forall i f_i(\mathbf{x}') \leq f_i(\mathbf{x}) \ \& \ \exists j f_j(\mathbf{x}') < f_j(\mathbf{x})\}.$$



C is dominated,
A and **B** non-dominated

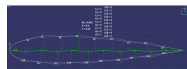
Notation:

$$\mathbf{A} \prec \mathbf{C}, \mathbf{B} \prec \mathbf{C}$$

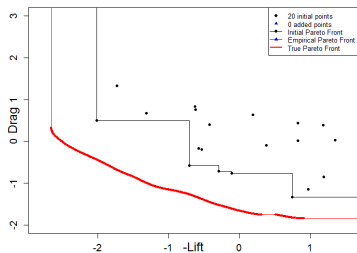
Background: multi-objective optimization (2)

True Pareto front vs. empirical Pareto front.

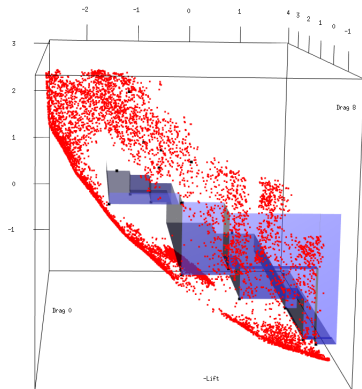
Examples from the metaNACA test bed [8], $\mathbf{x} :=$



$m = 2$ objectives



$m = 3$ objectives

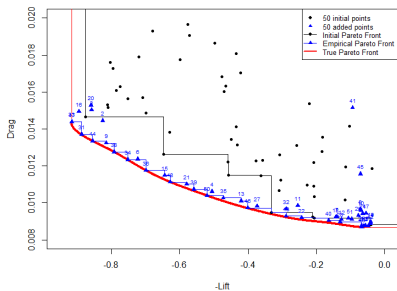


The Pareto fronts can have holes.

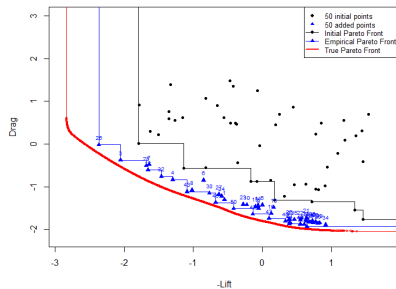
Curse of dimensionality: number of variables

At a given budget, optimization performance degrades with the number of variables:

$d = 3$



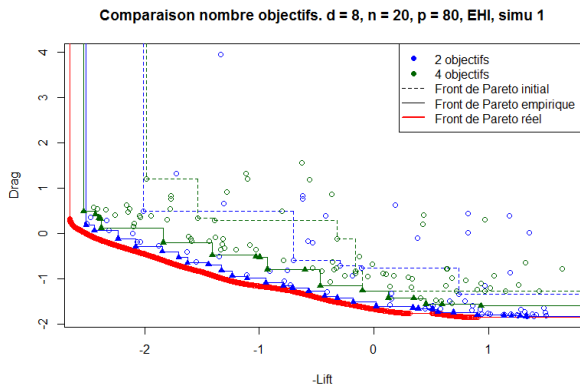
$d = 22$



(optimization algorithm: EHI with GPareto, [5, 2])

Curse of dimensionality: number of objectives (1)

At a given budget, optimization performance degrades with the number of objectives:



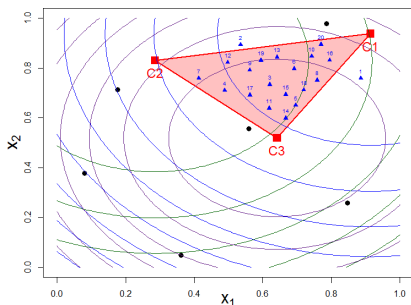
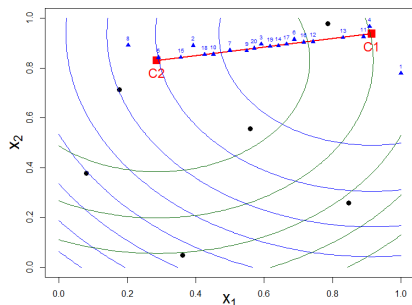
(optimization algorithm: EHI with GPareto, [5, 2])

Curse of dimensionality: number of objectives (2)

As the number of objectives increases, a larger part of \mathcal{X} becomes Pareto optimal:

$m = 2$

$m = 3$



Ex: sphere functions centered on C_1 , C_2 , C_3 . Pareto sets (in red) are all convex combinations of the C 's. Blue triangles: points sampled by MO Bayesian algorithm (GPareto). With 4 objectives at the corners of \mathcal{X} , every point could be a Pareto solution. As the Pareto set becomes larger, the optimization algorithm degenerates in a space filling algorithm. Give up the utopian search for all of the Pareto set.

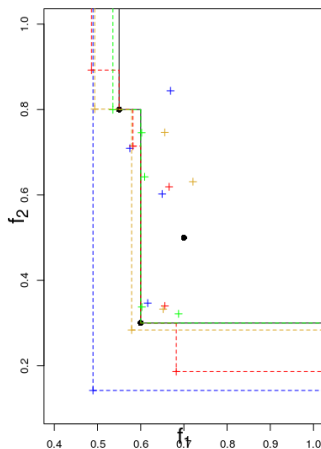
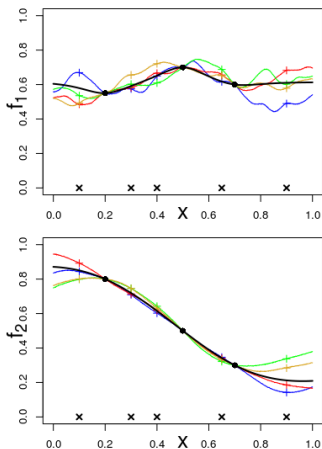
Content of the talk

Recently, we have explored ways to proportionate ambitions to search budget:

- 1 **Today's talk:** how to focus on specific regions of the Pareto front. The C-EHI algorithm explained step by step. Details and proofs in [8].
- 2 **See also:** how to select variables on which optimization is performed ($m = 1$). Work with Adrien Spagnol and Sébastien Da Veiga, Safran Tech, cf. [19].

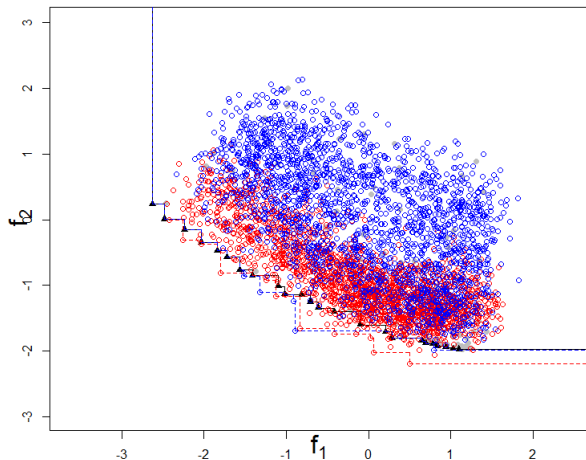
Bayesian multi-objective optimization (1)

Equipped with observations of the true functions and GPs, we can simulate possible Pareto fronts at given \mathbf{x} 's:



Simulations points for the Pareto front (1)

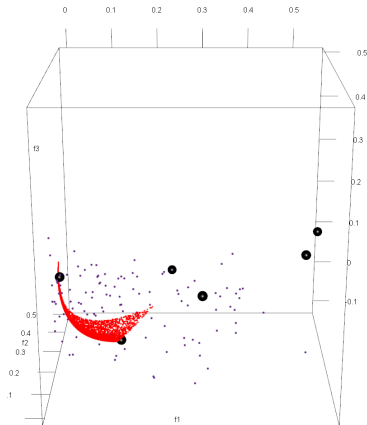
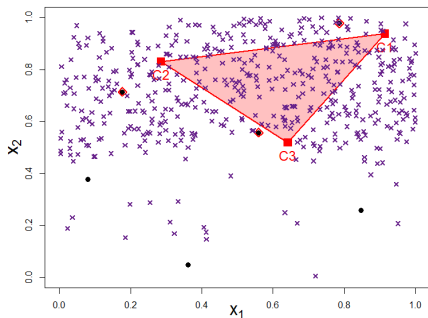
The choice of the \mathbf{x} 's where the simulations are performed matters. Below, blue points are random, red points selected proportionally to their probability of being non dominated by the empirical front $\widehat{\mathcal{P}}_y$:



($d = 22$)

Simulations points for the Pareto front (2)

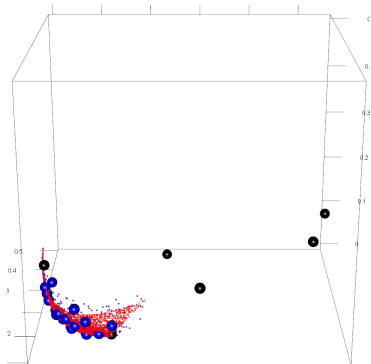
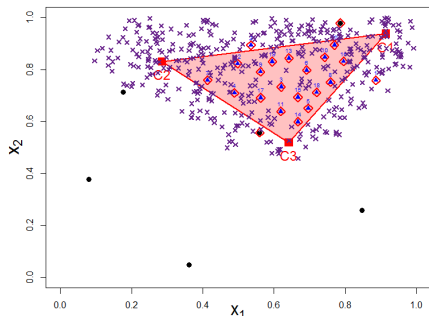
Choose \mathbf{x} 's with a probability proportional to $\mathbb{P}(\widehat{\mathcal{P}}_Y \not\subseteq \mathbf{Y}(\mathbf{x}))$.
In the quadratic case, $d = 2$, $m = 3$, initial DoE



DoE \cdot & \bullet , Non Dominated \diamond , selected point \times , sampled \mathbf{Y} ND \cdot

Simulations points for the Pareto front (3)

Choose \mathbf{x} 's with a probability proportional to $\mathbb{P}(\widehat{\mathcal{P}}_Y \not\subseteq \mathbf{Y}(\mathbf{x}))$.
In the quadratic case, $d = 2$, $m = 3$, after 20 iterations

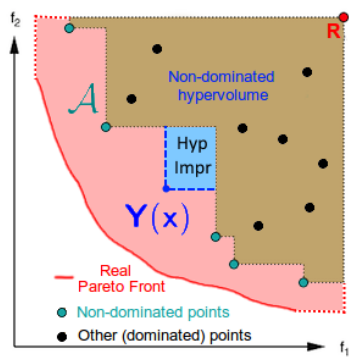


DoE \cdot & \bullet & \bullet , Non Dominated \diamond , selected point \times , sampled \mathbf{Y} ND \cdot .

Simulation points are uniformly distributed near the Pareto set.

Bayesian multi-objective optimization (2)

Where to put the next point, \mathbf{x}^{n+1} , where to call the costly \mathbf{f} ? At the point that maximizes, on the average of the $\mathbf{Y}(\mathbf{x})$ samples, the *Hypercube Improvement* (over the empirical Pareto front $\widehat{\mathcal{P}}_{\mathbf{y}}$):



EHI implementation from [16]

$$H(\mathcal{A}; \mathbf{R}) = \bigcup_{\mathbf{y} \in \mathcal{A}} \int_{\mathbf{y} \preceq \mathbf{z} \preceq \mathbf{R}} dz$$

$$I_H(\mathbf{f}; \mathbf{R}) = H(\widehat{\mathcal{P}}_{\mathbf{y}} \cup \{\mathbf{f}\}; \mathbf{R}) - H(\widehat{\mathcal{P}}_{\mathbf{y}}; \mathbf{R})$$

$$\text{EHI}(\mathbf{x}; \mathbf{R}) = \mathbb{E}(I_H(\mathbf{Y}(\mathbf{x}); \mathbf{R}))$$

EHI favors $\mathbf{Y}(\mathbf{x})$ dominating the empirical Pareto front and far from already observed $\mathbf{f}(\mathbf{x}^i)$ s.

Notice: the omnipresence of \mathbf{R} ; $\forall \mathbf{R} \succeq \widehat{\mathcal{P}}_{\mathbf{y}}$, EHI generalizes the EI criterion of EGO [11].

Bayesian multi-objective optimization (3)

Algorithm 1 Multi-objective EHI Bayesian optimizer

Require: DoE = $\{(\mathbf{x}^1, \mathbf{f}(\mathbf{x}^1)), \dots, (\mathbf{x}^n, \mathbf{f}(\mathbf{x}^n))\}$, \mathbf{R} , n^{\max}

- 1: **while** $n < n^{\max}$ **do**
 - 2: Build m independent GPs, $\mathbf{Y}() = (Y_1(), \dots, Y_m())$, from current DoE
 - 3: Find next iterate by solving $\mathbf{x}^{n+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \text{EHI}(\mathbf{x}; \mathbf{R})$
{internal optimization problem, no call to $\mathbf{f}()$ }
 - 4: Calculate $\mathbf{f}(\mathbf{x}^{n+1})$
 - 5: $n \leftarrow n + 1$
 - 6: **end while**
-

There are other acquisition criteria in addition to EHI [20] or EHI which includes constraints [6].

Targeting improvement regions with EHI

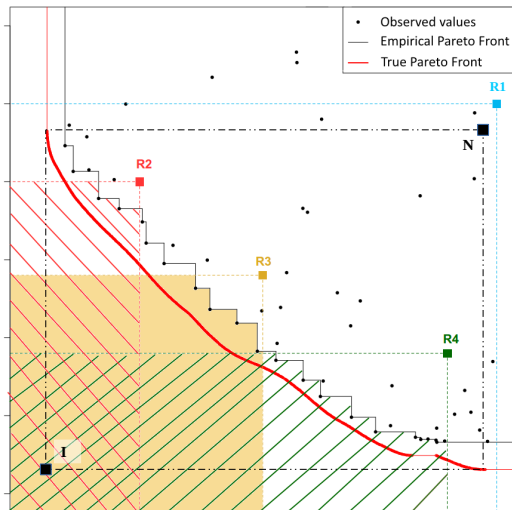
To find the entire Pareto front, \mathbf{R} must be dominated by the Nadir point, \mathbf{N} : **R1** is the default in the literature.

But the entire Pareto front is *i)* too large to be described *ii)* not interesting in general (e.g., extreme solutions).

\Rightarrow move \mathbf{R} and control the *improvement region*,

$$\mathcal{I}_{\mathbf{R}} := \{\mathbf{y} \in \mathcal{Y} : \mathbf{y} \preceq \mathbf{R}\}$$

(keeps the Pareto rank for non comparable functions)



mEl, a computationally efficient proxy to EHI

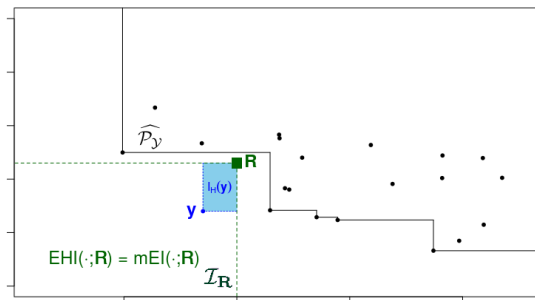
Once \mathbf{R} is freed from $\widehat{\mathcal{P}}_y$, a new acquisition criterion is possible.

Definition:

$$\text{mEl}(\mathbf{x}; \mathbf{R}) := \prod_{j=1}^m \text{El}_j(\mathbf{x}; R_j) = \prod_{j=1}^m \mathbb{E}(\max(0, R_j - Y_j(\mathbf{x})))$$

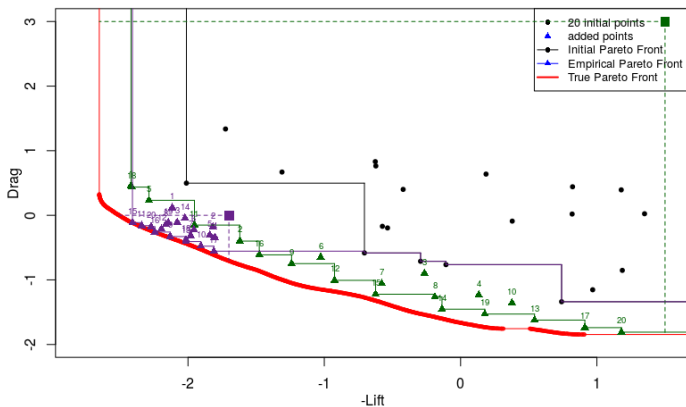
Property:

If $\widehat{\mathcal{P}}_y \not\subseteq \mathbf{R}$, $\text{EHI}(\cdot; \mathbf{R}) = \text{mEl}(\cdot; \mathbf{R})$.



$\text{mEl}(\mathbf{x}; \mathbf{R})$ is analytical in $m_i(\mathbf{x})$ and $C_i(\mathbf{x}, \mathbf{x})$, computationally much more efficient than EHI which involves Monte Carlo simulations when $m > 2$ (*ms vs min*).

Example: targeted mEI versus EHI



Violet: mEI convergence with \mathbf{R} at ■.

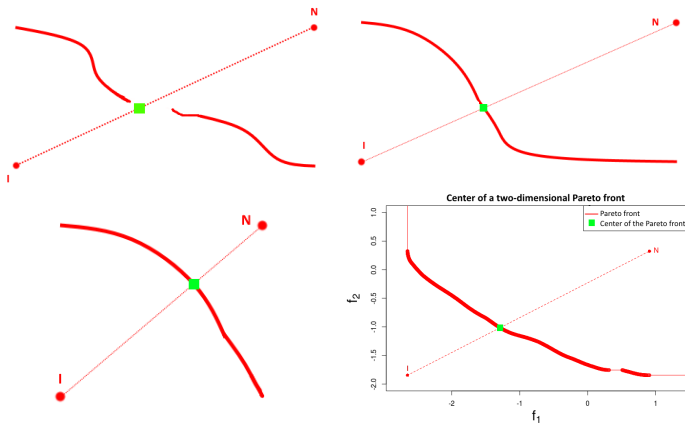
Green: EHI convergence with \mathbf{R} at ■.

Note the more local and accurate convergence with mEI. ($d = 8$)

The Pareto front center

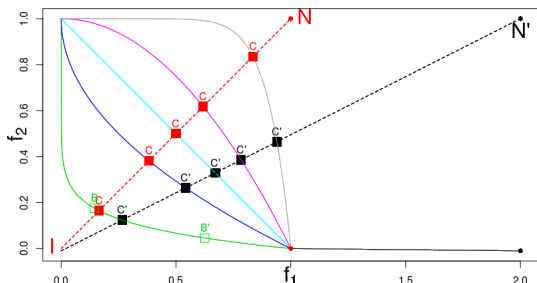
Which point should be targeted through \mathbf{R} ? By default, the point where objectives are “balanced”. Need a definition.

Definition: The center \mathbf{C} is the point of the Ideal-Nadir line the closest in Euclidean distance to the Pareto front.



Properties of the Pareto front center

- The Pareto front center is equivalent, in game theory, to the Kalai-Smordinsky solution with a disagreement point at the Nadir [12].
- The Pareto front center is invariant w.r.t. a linear scaling of the objectives either when the Pareto front intersects the Ideal-Nadir line, or when $m = 2$ (not true in general though).
- The Pareto front center is stable w.r.t. perturbations in Ideal and Nadir: $\|\Delta \mathbf{C}\|_2 < \|\Delta \mathbf{N}\|_2$ and $\|\Delta \mathbf{C}\|_2 < \|\Delta \mathbf{I}\|_2$.



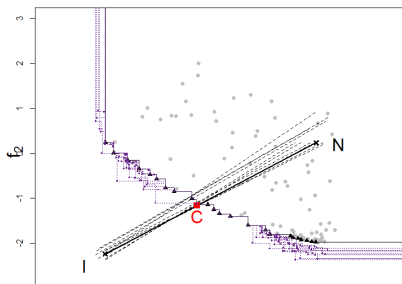
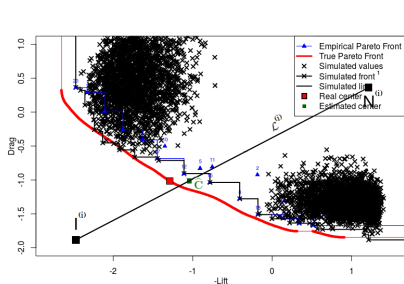
Estimating the Pareto front center

Crude estimators:

$$\hat{\mathbf{I}} = \left(\min_{\mathbf{y} \in \text{DoE}} (y_1), \dots, \min_{\mathbf{y} \in \text{DoE}} (y_m) \right),$$

$$\hat{\mathbf{N}} = \left(\max_{\mathbf{y} \in \widehat{\mathcal{P}}_y} (y_1), \dots, \max_{\mathbf{y} \in \widehat{\mathcal{P}}_y} (y_m) \right),$$

but they may be misleading early in the search. Take advantage of the GPs uncertainties \Rightarrow estimate them from Pareto front simulations (at carefully selected \mathbf{x} 's, see next slides) and take their median.



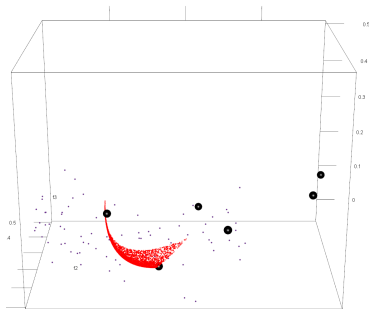
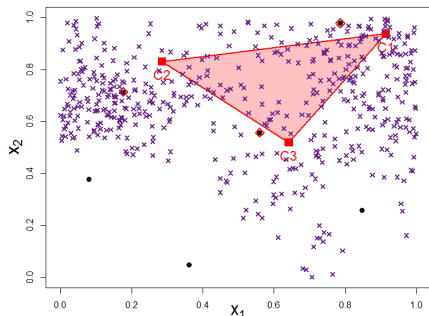
Simulation points for the Ideal and the Nadir (1)

- (For the Pareto front, choose \mathbf{x} 's with a probability proportional to $\mathbb{P}\left(\mathbf{Y}(\mathbf{x}) \preceq \widehat{\mathcal{P}}_y\right)$.) \leftarrow see earlier
- For the Ideal, choose \mathbf{x} 's with a probability proportional to $\mathbb{P}\left(\mathbf{Y}(\mathbf{x}) \leq \min f_i^j\right)$, $j = 1, n$, $i = 1, m$ (analytical).
- For the Nadir, choose \mathbf{x} 's with a probability proportional to $\mathbb{P}\left(Y_i(\mathbf{x}) > \widehat{\mathbf{N}}_i, \mathbf{Y}(\mathbf{x}) \text{ non dominated}\right) + \mathbb{P}\left(\mathbf{Y}(\mathbf{x}) \preceq \arg \widehat{\mathbf{N}}_i\right)$, $i = 1, m$

More details in [8]

Simulation points for the Ideal and the Nadir (2)

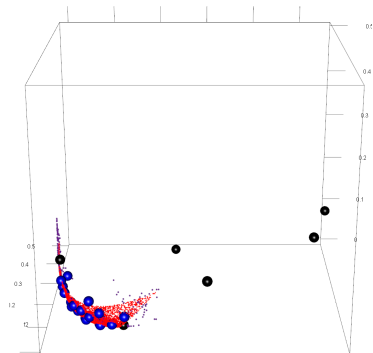
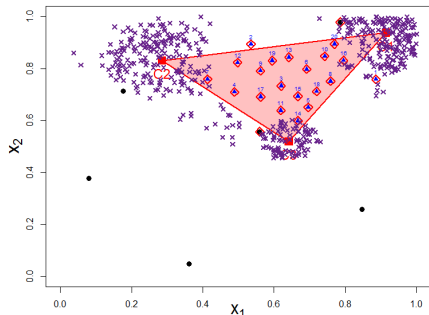
In the quadratic case, $d = 2$, $m = 3$, initial DoE



DoE \bullet , Non Dominated \blacklozenge , selected point \times , sampled \mathbf{Y} ND \bullet ,

Simulation points for the Ideal and the Nadir (3)

In the quadratic case, $d = 2$, $m = 3$, after 20 iterations



DoE . & ● & ● , Non Dominated ◇ , selected point × , sampled Y ND .

Simulation points are grouped around the centers which make the Ideal and Nadir.

Partial progress report

Require: DoE = $\{(\mathbf{x}^1, \mathbf{f}(\mathbf{x}^1)), \dots, (\mathbf{x}^n, \mathbf{f}(\mathbf{x}^n))\}$, n^{\max}

- 1: Build the m independent GPs;
- 2: **repeat**
- 3: estimate $\hat{\mathbf{I}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{C}}$;
- 4: $\mathbf{x}^{n+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \text{mEl}(\mathbf{x}; \hat{\mathbf{C}})$;
- 5: evaluate $\mathbf{f}(\mathbf{x}^{n+1})$ and update the GPs;
- 6: $n \leftarrow n+1$;
- 7: **until** $n > n^{\max}$

but often $\hat{\mathbf{C}}$ has converged to the true \mathbf{C} before the end. Cannot be further improved. Waste of computation. Need a better stopping criterion.

Uncertainty in center location (1)

Need a stopping criterion. mEI and EHI are too unstable: depend on f_i 's scales and \mathbf{R} .

Define the domination probability,

$$p(\mathbf{y}) := \mathbb{P}(\exists \mathbf{x} \in \mathcal{X} : \mathbf{Y}(\mathbf{x}) \preceq \mathbf{y})$$

Estimation: simulate n_{sim} Pareto fronts (at well-chosen \mathbf{x} 's), $\widetilde{\mathcal{P}}_y^{(i)}$, and

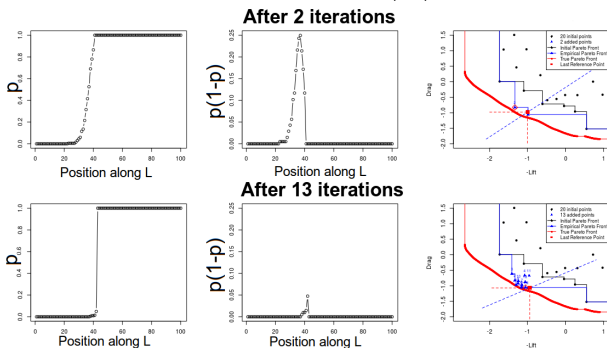
$$\widehat{p}(\mathbf{y}) = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \mathbb{1}(\widetilde{\mathcal{P}}_y^{(i)} \preceq \mathbf{y})$$

Uncertainty in center location (2)

If $\widehat{p(\mathbf{y})}$ is near 1 or 0, we are quite sure that \mathbf{y} is dominated or not. The uncertainty is $p(\mathbf{y})(1 - p(\mathbf{y}))$, the variance of the Bernoulli variable $D(\mathbf{y}) = \mathbb{1}(\mathcal{P}_{\mathbf{Y}(\cdot)} \preceq \mathbf{y})$.

Define the uncertainty in center location as

$$U(\widehat{\mathcal{L}}) := \frac{1}{|\widehat{\mathcal{L}}|} \int_{\widehat{\mathcal{L}}} p(\mathbf{y})(1 - p(\mathbf{y})) d\mathbf{y}.$$



($d = 8$)

Targeted MO Bayesian optimization

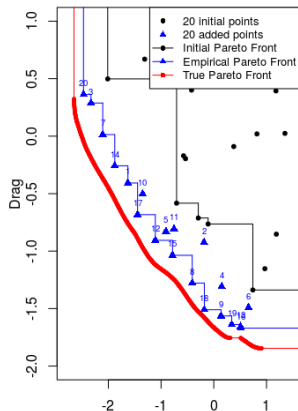
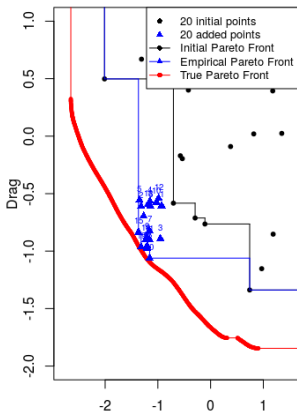
Algorithm 2 First phase of the C-EHI algorithm

Require: DoE = $\{(\mathbf{x}^1, \mathbf{f}(\mathbf{x}^1)), \dots, (\mathbf{x}^n, \mathbf{f}(\mathbf{x}^n))\}$, ε_1 , n^{\max}

- 1: Build the m independent GPs;
 - 2: **repeat**
 - 3: estimate $\hat{\mathbf{I}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{C}}$;
 - 4: $\mathbf{x}^{n+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \text{mEl}(\mathbf{x}; \hat{\mathbf{C}})$;
 - 5: evaluate $\mathbf{f}(\mathbf{x}^{n+1})$ and update the GPs;
 - 6: compute $U(\hat{\mathcal{L}})$;
 - 7: $n \leftarrow n+1$;
 - 8: **until** $U(\hat{\mathcal{L}}) \leq \varepsilon_1$ **or** $n > n^{\max}$
-

Example of targeted MO Bayesian opt. vs EHI

MetaNACA, $n^{\max} = 40$



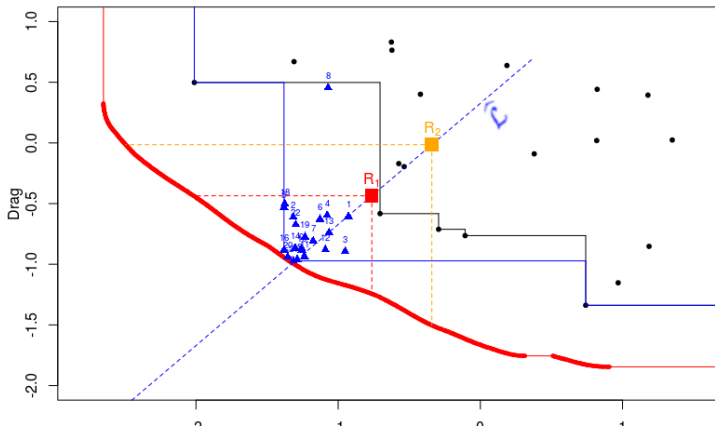
(Statistically significant results can be found in [8])

($d = 8$)

Remaining budget

What if convergence to the center occurs before n^{\max} ?

⇒ widen the search around the center by moving \mathbf{R} along $\hat{\mathcal{L}}$ towards the Nadir by a distance that is *compatible with the remaining budget*, $n^{\max} - n$.



Optimal final search region

- For a given \mathbf{R} , anticipate the future space filling of the algorithm by virtual iterates (Kriging Believer, [10]) $\Rightarrow \mathbf{Y}^{KB}(\mathbf{x})$ built from $\{(\mathbf{x}^1, \mathbf{f}(\mathbf{x}^1)), \dots, (\mathbf{x}^n, \mathbf{f}(\mathbf{x}^n))\} \cup \{(\mathbf{x}^{n+1}, \boldsymbol{\mu}(\mathbf{x}^{n+1})), \dots, (\mathbf{x}^{n^{\max}}, \boldsymbol{\mu}(\mathbf{x}^{n^{\max}}))\}$
- Measure the remaining uncertainty in Pareto domination

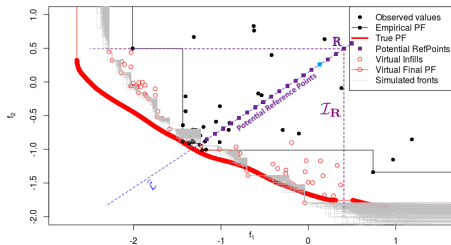
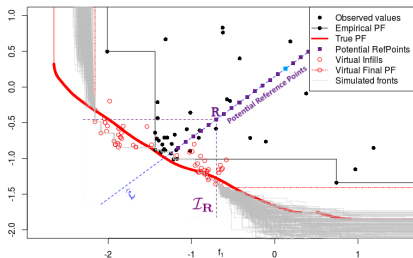
$$U(\mathbf{R}; \mathbf{Y}) := \frac{1}{\text{Vol}(\mathbf{I}, \mathbf{R})} \int_{\mathbf{I} \preceq \mathbf{y} \preceq \mathbf{R}} p(\mathbf{y})(1 - p(\mathbf{y})) d\mathbf{y} .$$

- Optimal \mathbf{R} defined through

$$\mathbf{R}^* = \arg \min_{\mathbf{R} \in \hat{\mathcal{L}}} \|\mathbf{R} - \hat{\mathbf{N}}\| \quad \text{such that} \quad U(\mathbf{R}; \mathbf{Y}^{KB}) \leq \varepsilon_2 \quad (2)$$

by enumeration.

Optimal final search region: illustration



The remaining uncertainty in Pareto domination can be seen by the sampled fronts roaming (in grey). It is small enough on the left, too large on the right.

R^* is in blue. $d = 8$.

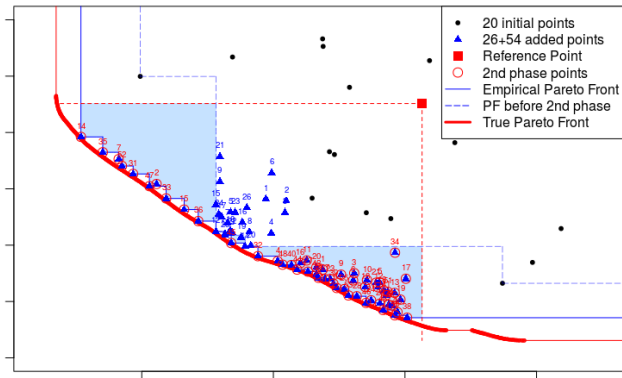
Budgeted and Targeted MO Bayesian Optimization

Algorithm 3 The C-EHI algorithm

Require: DoE = $\{(\mathbf{x}^1, \mathbf{f}(\mathbf{x}^1)), \dots, (\mathbf{x}^n, \mathbf{f}(\mathbf{x}^n))\}$, ε_1 , ε_2 , n^{\max}

- 1: Build the m independent GPs;
 - 2: **repeat**
 - 3: estimate $\hat{\mathbf{I}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{C}}$;
 - 4: $\mathbf{x}^{n+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \text{mEl}(\mathbf{x}; \hat{\mathbf{C}})$;
 - 5: evaluate $\mathbf{f}(\mathbf{x}^{n+1})$ and update the GPs;
 - 6: compute $U(\hat{\mathcal{L}})$;
 - 7: $n \leftarrow n+1$;
 - 8: **until** $U(\hat{\mathcal{L}}) \leq \varepsilon_1$ **or** $n > n^{\max}$
 - 9: **if** $n < n^{\max}$ **then**
 - 10: Calculate \mathbf{R}^* solution of Eq. (2); # needs ε_2
 - 11: **end if**
 - 12: **while** $n < n^{\max}$ **do**
 - 13: $\mathbf{x}^{n+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \text{EHI}(\mathbf{x}; \mathbf{R}^*)$;
 - 14: evaluate $f_i(\mathbf{x}^{(t+1)})$ and update the GPs;
 - 15: $n = n + 1$;
 - 16: **end while**
 - 17: **return** final DoE, final GPs, and approximation front $\hat{\mathcal{P}}_y$
-

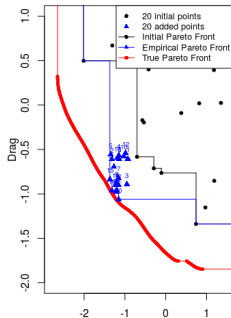
C-EHI: illustration of the 2nd phase



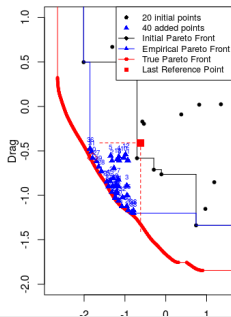
The objective values added during the 2nd phase are circled in red. Compared to the initial front obtained when searching for the center, the last approximation front is expanded as highlighted by the blue hypervolumes. $d = 8$.

C-EHI vs. EHI: illustration $m = 2$

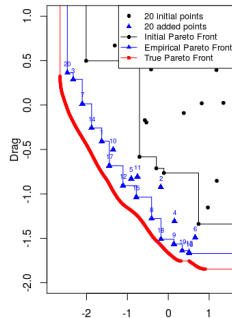
C-EHI (left) vs. EHI (right), top after 20 iterations, bottom after 40 iterations. C-EHI local convergence has occurred at 22 iterations, a wider optimal improvement region (under the red square) is targeted for the 18 remaining iterations. Compared to the standard EHI, C-EHI searches in a smaller balanced part of the objective space, at the advantage of a better convergence. $d = 8$



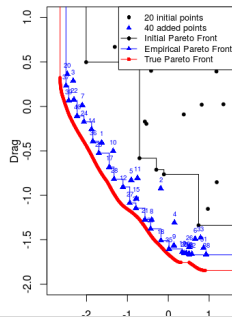
C-EHI



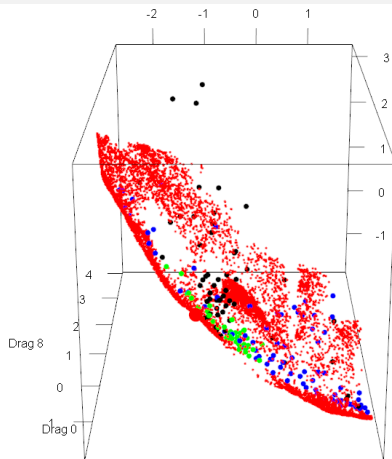
Classical EHI



Classical EHI



C-EHI vs. EHI: illustration $m = 3$



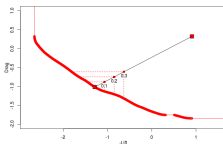
-Lift

($d = 8$)

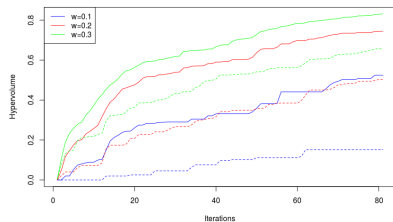
green, C-EHI; blue, EHI; black, initial front; red, true front, ●, true center.

C-EHI vs. EHI: tests

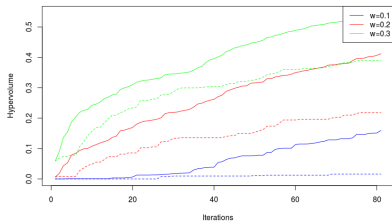
Hypervolumes of the C-EHI (continuous line) and EHI (dashed) averaged over 10 runs. Initial DoE of size 20, 80 iterations. Blue, red and green correspond to the improvement regions $\mathcal{I}_{0.1}$, $\mathcal{I}_{0.2}$ and $\mathcal{I}_{0.3}$, respectively. $d = 8$.



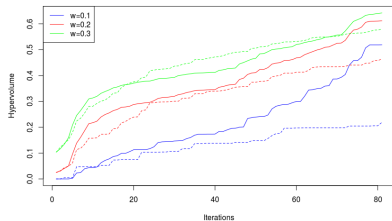
$m = 2$



$m = 3$



$m = 4$



C-EHI $>$ EHI, except when $m = 4$ and $\mathbf{R}_{0.3}$ because it is a large region.

Conclusions

Summary We have seen with the C-EHI Bayesian multi-objective algorithm how to

- target a specific region of improvement,
- target the center of the Pareto front,
- search for a part of the Pareto front adapted to the budget.

Perspectives

- Make it parallel. First steps in [9].
- Account for couplings between the objectives.
- Other ways to match ambitions to computational constraints: variables selection *for optimization*, first step with Hilbert-Schmidt sensitivity indices in [19].

References I

- [1] Brian M Adams, WJ Bohnhoff, KR Dalbey, JP Eddy, MS Eldred, DM Gay, K Haskell, Patricia D Hough, and Laura P Swiler.
Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: version 5.0 user's manual. *Sandia National Laboratories, Tech. Rep. SAND2010-2183*, 2009.
- [2] Mickael Binois and Victor Picheny.
GPareto: An R package for Gaussian-process based multi-objective optimization and analysis.
- [3] Mickaël Binois.
Uncertainty quantification on Pareto fronts and high-dimensional strategies in Bayesian optimization, with applications in multi-objective automotive design.
PhD thesis, École Nationale Supérieure des Mines de Saint-Etienne, 2015.
- [4] Nicolas Durrande and Rodolphe Le Riche.
Introduction to Gaussian Process Surrogate Models.
Lecture at 4th MDIS form@ter workshop, Clermont-Fd, France, October 2017.
HAL report cel-01618068.

References II

- [5] Michael Emmerich, André Deutz, and Jan Willem Klinkenberg.
Hypervolume-based expected improvement: Monotonicity properties and exact computation.
In Evolutionary Computation (CEC), 2011 IEEE Congress on, pages 2147–2154. IEEE, 2011.
- [6] Paul Feliot, Julien Bect, and Emmanuel Vazquez.
A Bayesian approach to constrained single- and multi-objective optimization.
Journal of Global Optimization, 67(1):97–133, January 2017.
- [7] Alexander I. J. Forrester, Andras Sobester, and Andy J. Keane.
Engineering Design via Surrogate Modelling - A Practical Guide.
Wiley, 2008.
- [8] David Gaudrie, Rodolphe Le Riche, Victor Picheny, Benoit Enaux, and Vincent Herbert.
Budgeted multi-objective optimization with a focus on the central part of the pareto front – extended version.
preprint arXiv:1809.10482, 2018.
- [9] David Gaudrie, Rodolphe Le Riche, Victor Picheny, Benoit Enaux, and Vincent Herbert.
Targeting solutions in bayesian multi-objective optimization: Sequential and parallel versions.
preprint arXiv:1811.03862, 2018.

References III

- [10] David Ginsbourger, Rodolphe Le Riche, and Laurent Carraro.
Kriging is well-suited to parallelize optimization.
In *Computational Intelligence in Expensive Optimization Problems*, pages 131–162.
Springer, 2010.
- [11] Donald R Jones, Matthias Schonlau, and William J Welch.
Efficient Global Optimization of expensive black-box functions.
Journal of Global optimization, 13(4):455–492, 1998.
- [12] Ehud Kalai and Meir Smorodinsky.
Other solutions to Nash’s bargaining problem.
Econometrica: Journal of the Econometric Society, pages 513–518, 1975.
- [13] Jack P.C. Kleijnen.
Kriging metamodeling in simulation: A review.
European Journal of Operational Research, 192(3):707 – 716, 2009.
- [14] Rodolphe Le Riche.
Surrogates and (mono-objective) optimization: a long-term relationship.
Lecture at the SAMCO workshop, Lorentz Center, Leiden, NL, February 2016.
HAL report cel-01281650.

References IV

- [15] Ilya Loshchilov, Marc Schoenauer, and Michèle Sebag.
Comparison-based optimizers need comparison-based surrogates.
In Robert Schaefer, Carlos Cotta, Joanna Kołodziej, and Günter Rudolph, editors, *Parallel Problem Solving from Nature, PPSN XI*, pages 364–373, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [16] Olaf Mersmann.
EMOA: Evolutionary multiobjective optimization algorithms.
R package version 0.5-0, 2012.
- [17] Nestor V Queipo, Raphael T Haftka, Wei Shyy, Tushar Goel, Rajkumar Vaidyanathan, and P Kevin Tucker.
Surrogate-based analysis and optimization.
Progress in aerospace sciences, 41(1):1–28, 2005.
- [18] Carl Edward Rasmussen and Christopher KI Williams.
Gaussian Processes for Machine Learning.
The MIT Press, 2006.
- [19] Adrien Spagnol, Rodolphe Le Riche, and Sebastien Da Veiga.
Global sensitivity analysis for optimization with variable selection.
preprint arXiv:1811.04646, 2018.

References V

- [20] Tobias Wagner, Michael Emmerich, André Deutz, and Wolfgang Ponweiser.
On expected-improvement criteria for model-based multi-objective optimization.
In *International Conference on Parallel Problem Solving from Nature*, pages 718–727.
Springer, 2010.

Comparisons of estimations for the Pareto front center

Example, $d = 8$:

Ideal-Nadir line of the empirical PF: a less robust estimator for the center of the Pareto front

