Multiobjective Bayesian Optimization under uncertainties

Victor Trappler, Céline Helbert, Christophette Blanchet-Scalliet, Rodolphe Le Riche

(Multi)objective Bayesian Optimization under uncertainties

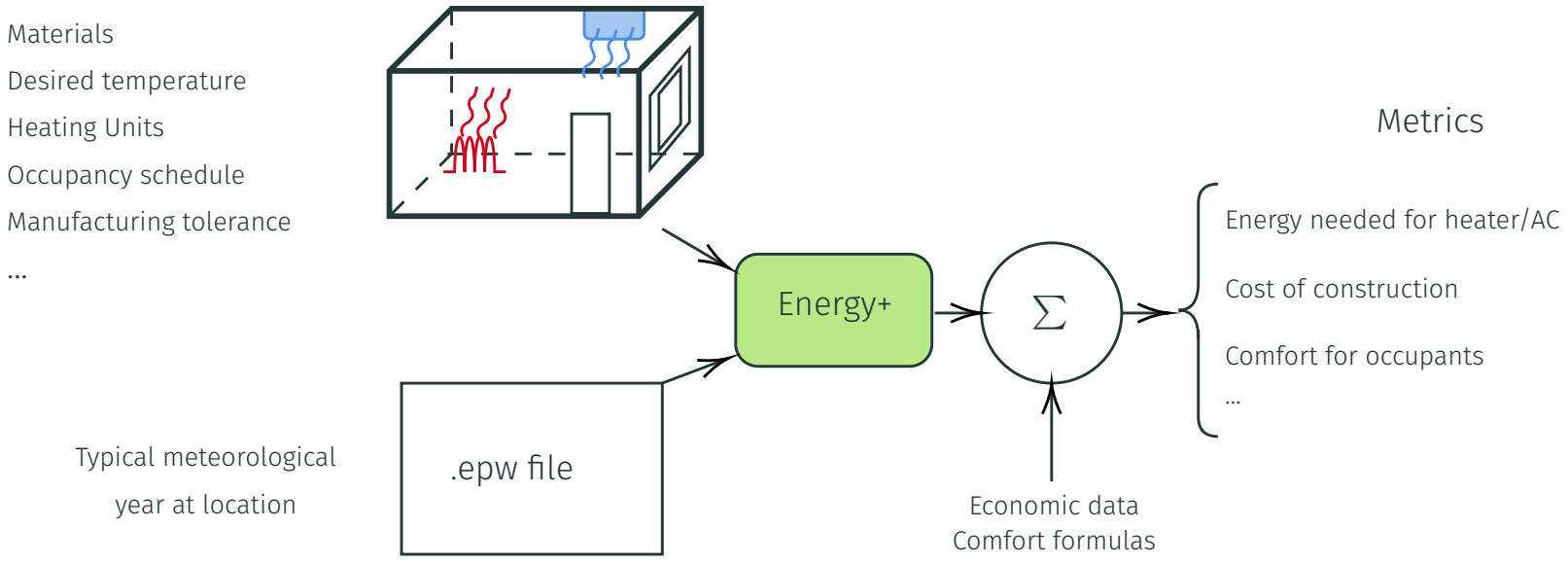
Victor Trappler[1]
Céline Helbert[1] Christophette Blanchet-Scalliet[1] Rodolphe Le Riche[2]
Rencontres Annuelles du réseau MEXICO

December 6, 2024

[1]Institut Camille Jordan
[2]LIMOS

vtrappler.github.io

How to design buildings: EnergyPlus



- $x \in \mathcal{X} \to$ Materials (type and quantity), Heating/AC, Rate of air change...

Desirable building:

- Low cost of construction
- Low energy consumption
- High comfort for occupants

Sadly, no design optimizes simultaneously all objectives.

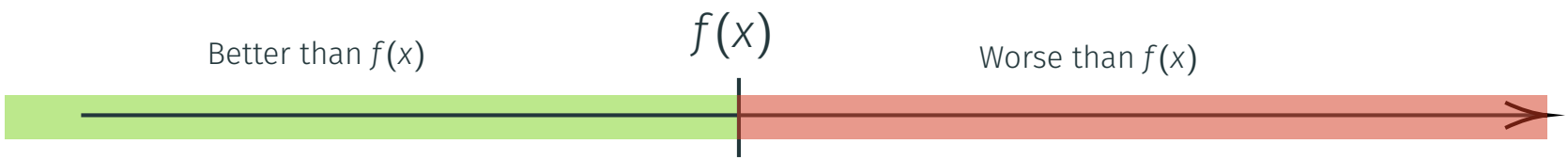
Back to basics

How to compare if x is better than x' ?

Order relation on $\mathbb{R}$: $\leq$

Let $a, b \in \mathbb{R}$. a is not worse than b if $a \leq b$

$f(x')$ *not worse* than $f(x)$ if $f(x') \leq f(x)$



$$\min_{x \in \mathcal{X}} f(x) = y \text{ verifies } y \leq f(x), \quad \forall x \in \mathcal{X}$$

What if f is very expensive to evaluate ?

Bayesian Optimization

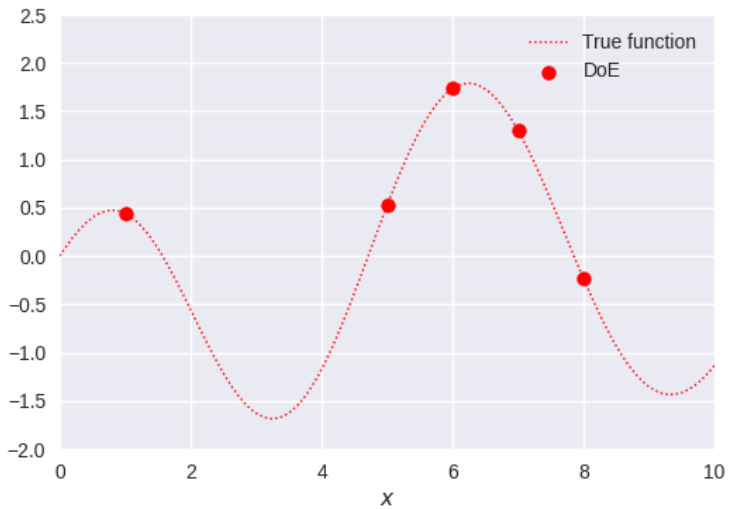
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



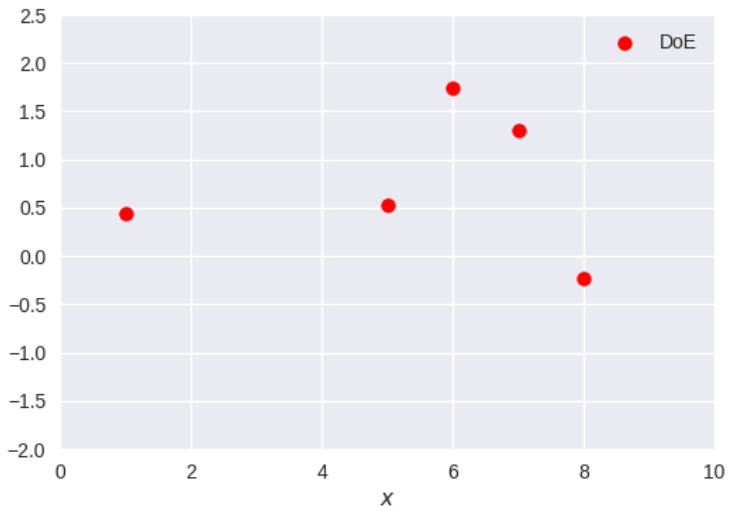
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



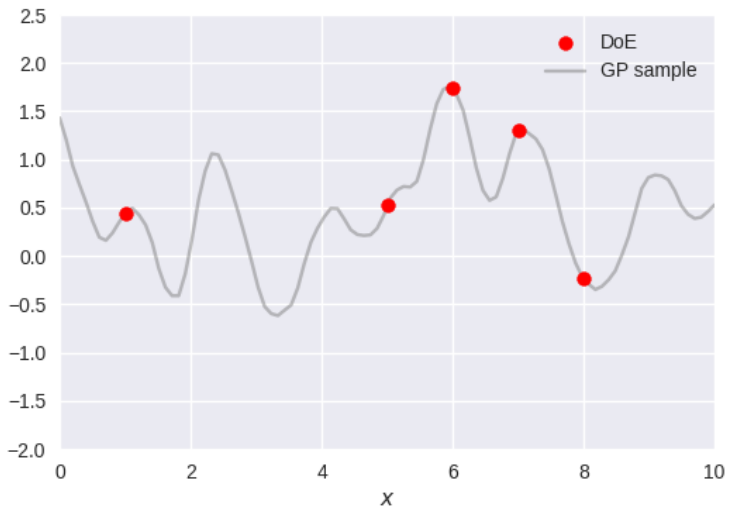
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



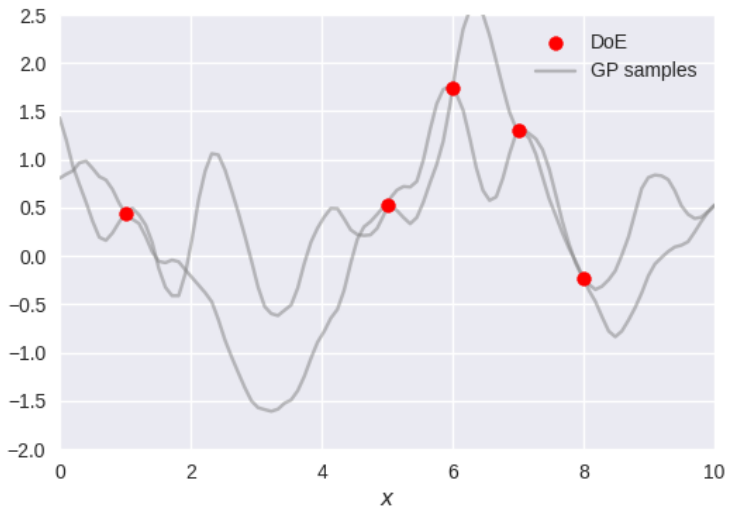
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



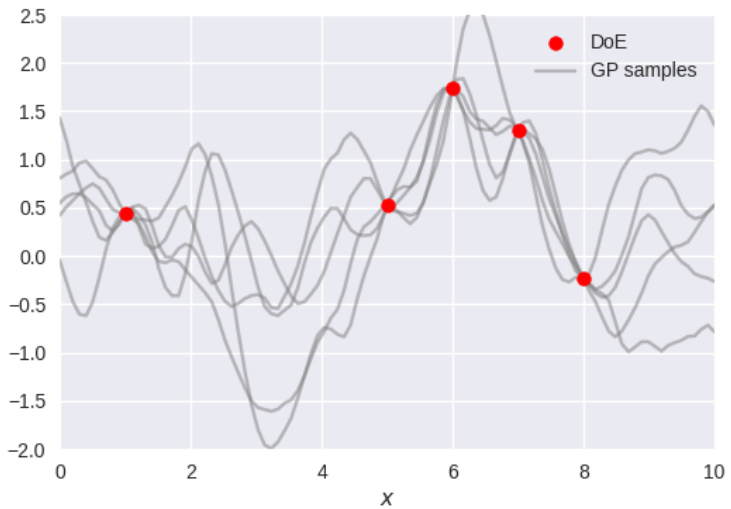
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



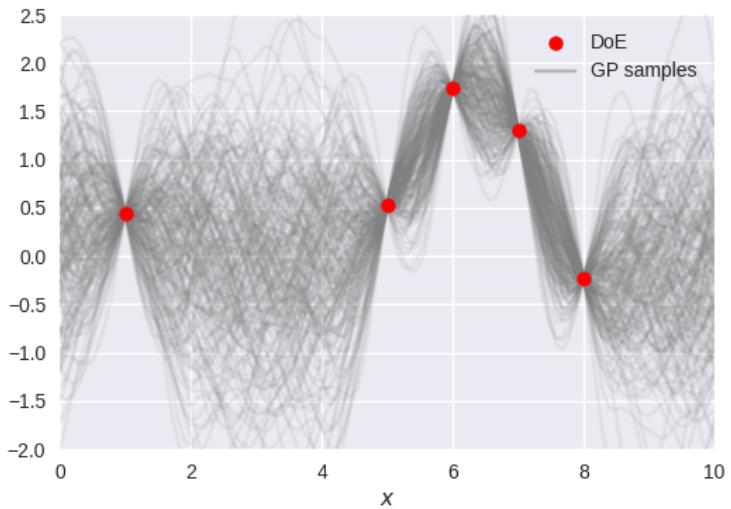
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



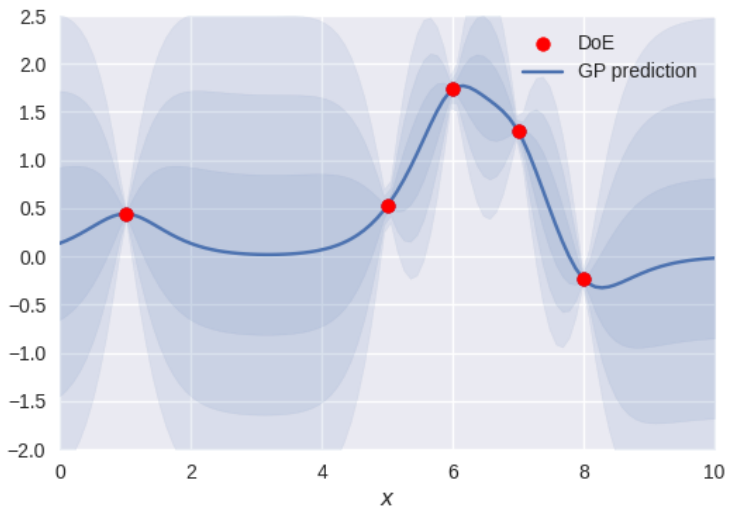
GP modelling

BO relies on Gaussian Process Regression:

GP modelling in very short

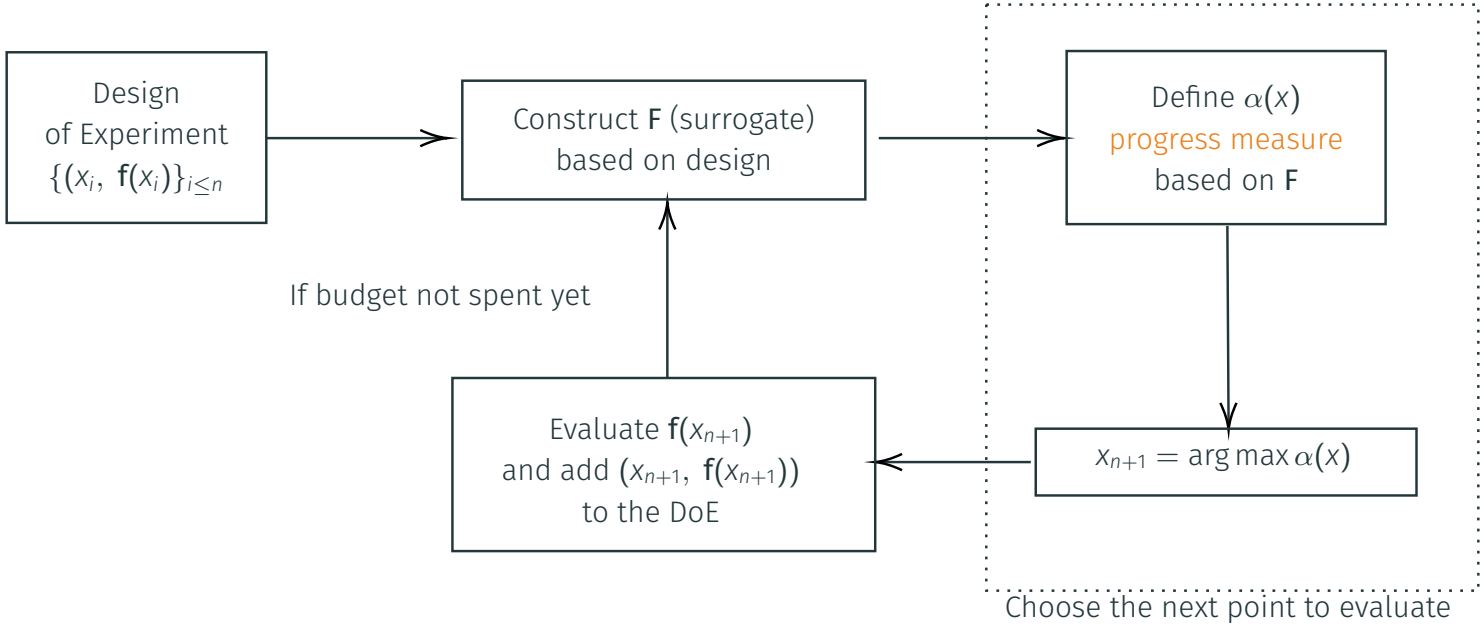
Gaussian Process F: Surrogate model of (the expensive to evaluate) f, conditioned on a Design of Experiment

$$F(x) \sim \mathcal{N}(\underbrace{m_F(x)}_{\text{prediction}}, \underbrace{\sigma_F^2(x)}_{\text{uncertainty}})$$



Works well for noisy evaluations of the function !

Bayesian Optimization Loop: Active Learning



- Different Acquisition methods for different objectives
  - EGO, EFI(SUR), LCB, Entropy-based acquisitions...

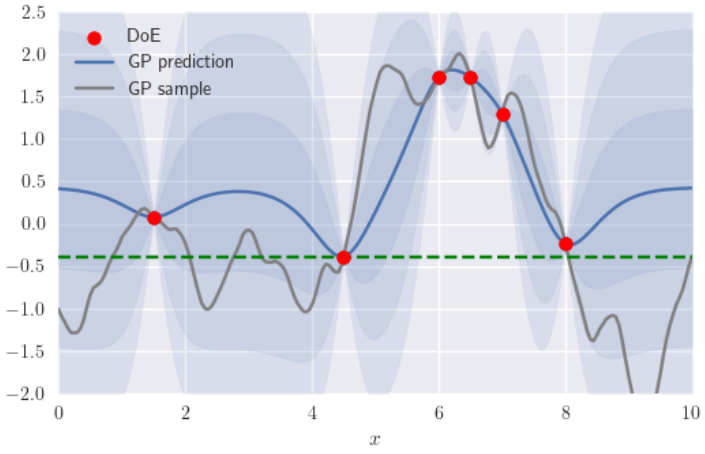
Efficient Global Optimization [JS98]

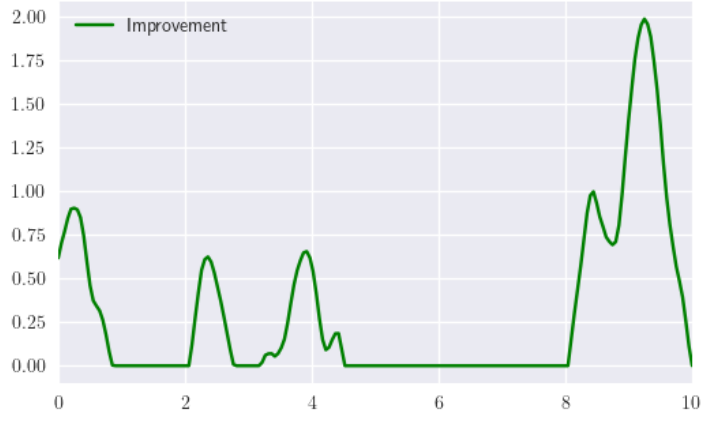
Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$

r.v.





Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

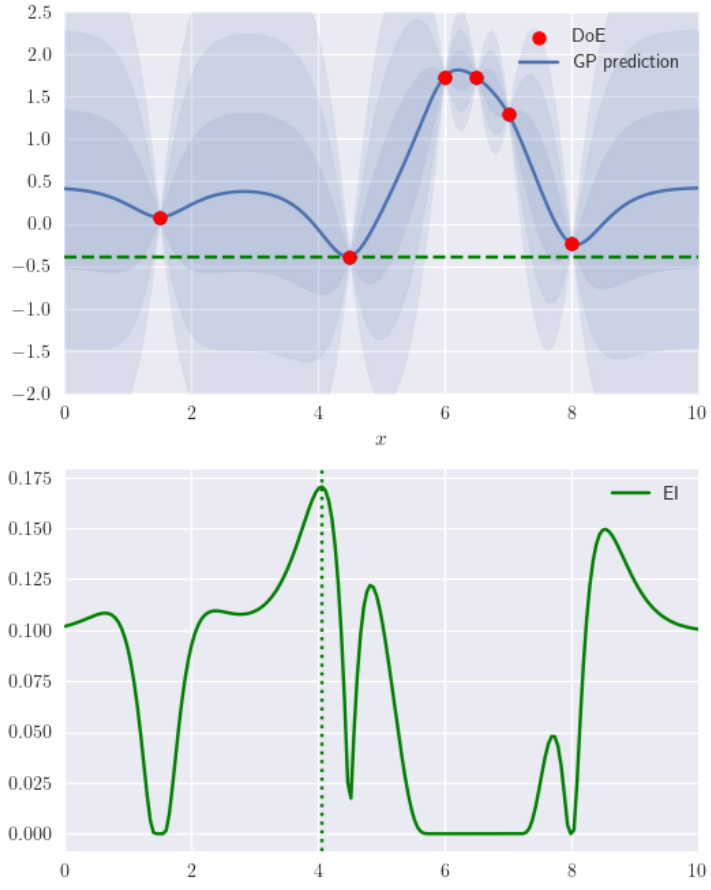
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

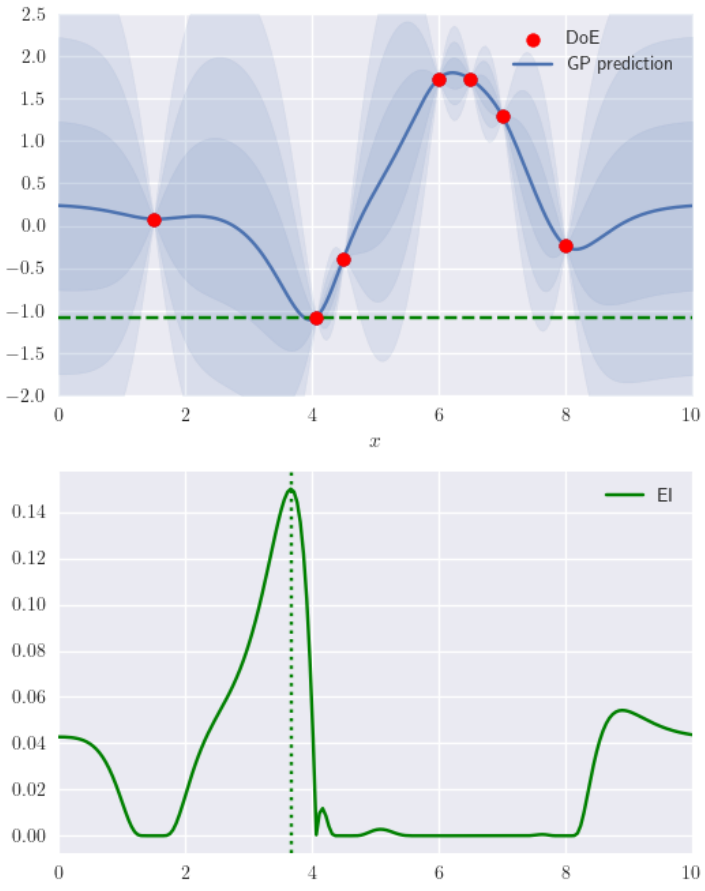
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

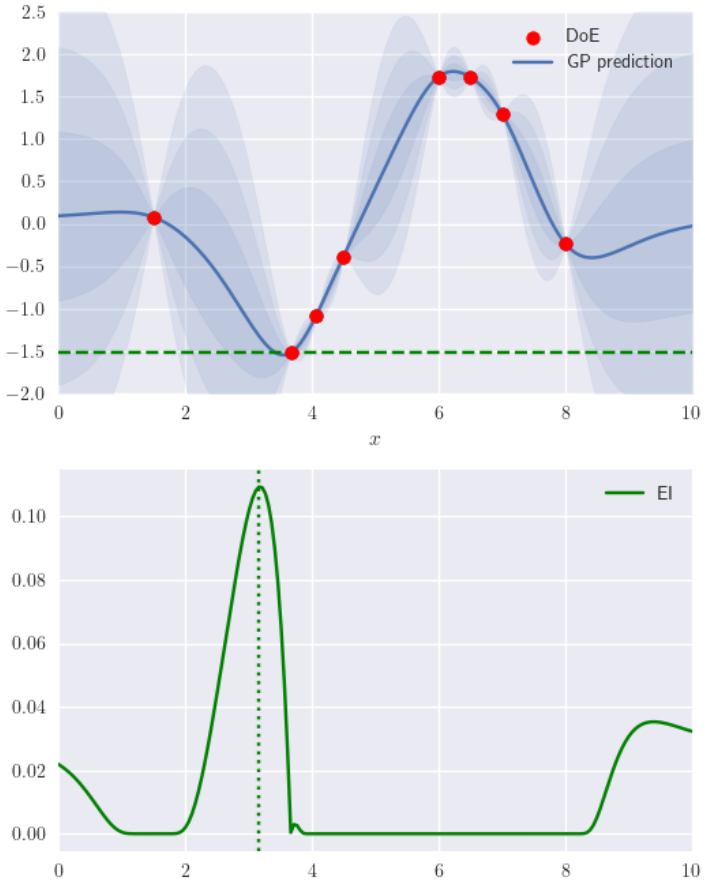
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

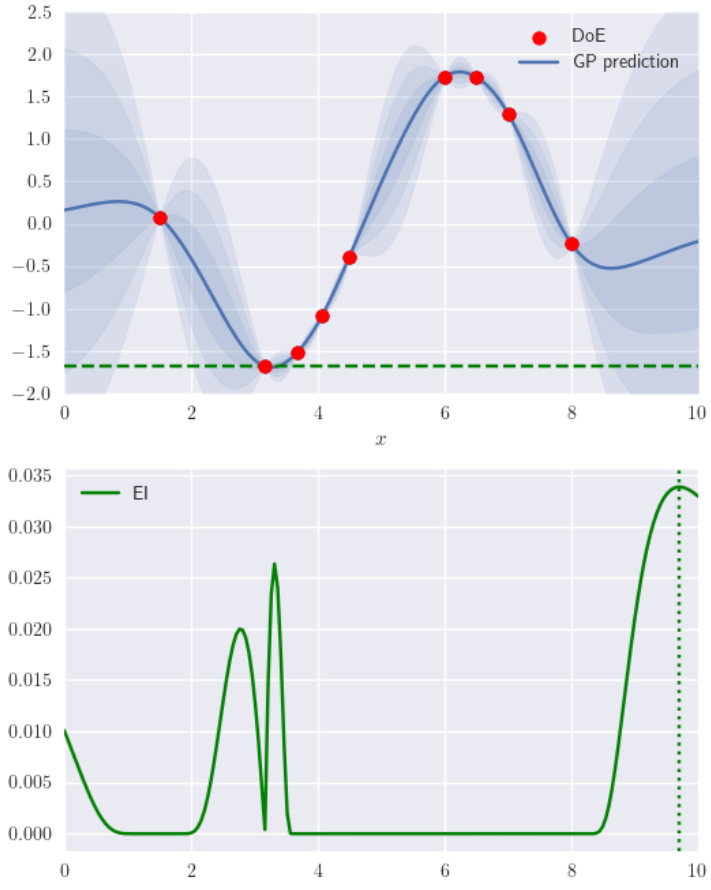
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

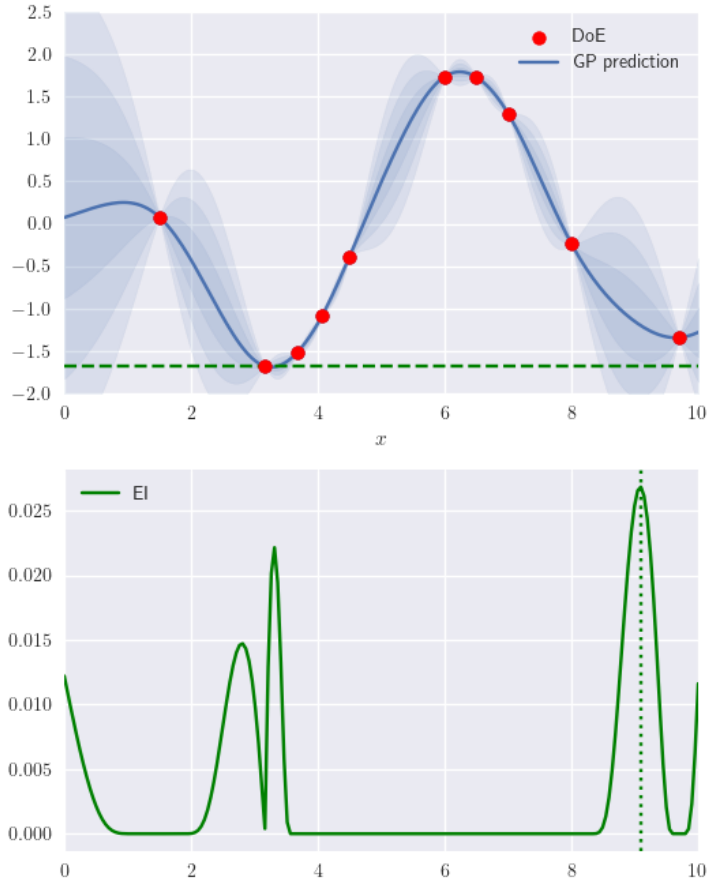
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

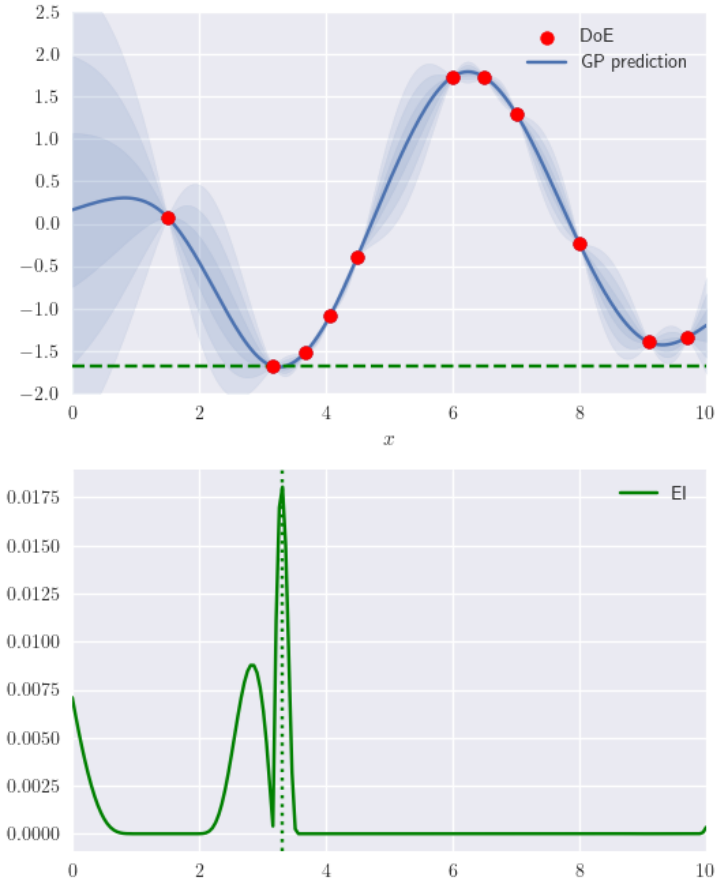
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Efficient Global Optimization [JS98]

Improvement

How much does $F(x)$ improves the current best value $f_{\min}$?

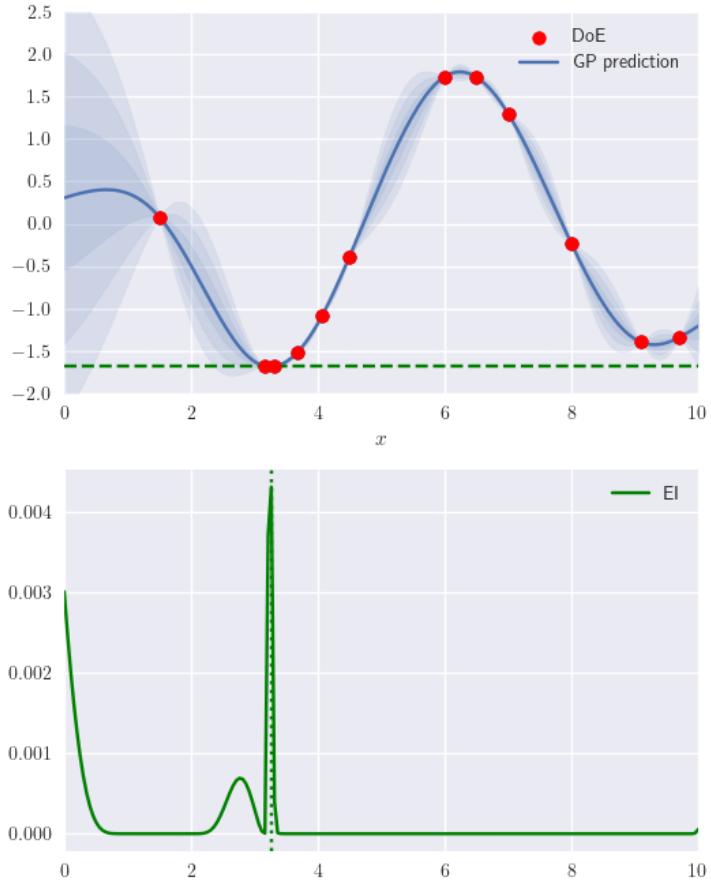
$$I(x) = \begin{cases} f_{\min} - F(x) & \text{if } f_{\min} \geq F(x) \\ 0 & \text{elsewhere} \end{cases}$$ r.v.

Progress Measure: Expected Improvement

$$\alpha(x) = \mathbb{E}\left[I(x)\right] = \mathbb{E}\left[(f_{\min} - F(x))_+\right]$$

$$x_{n+1} = \arg\max_x \alpha(x)$$

Analytical expression exists for EI (and its gradient) since F is a GP !



$\rightarrow$ Tradeoff exploration vs exploitation

Multiobjective Optimization

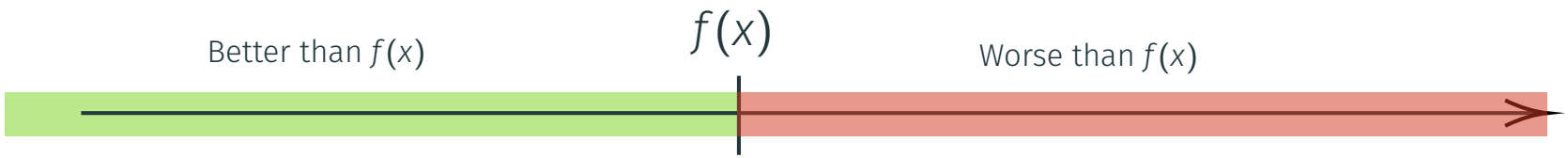
Back to "Back to basics"

How to compare if x is better than x' ?

Order relation on $\mathbb{R}$: $\leq$

Let $a, b \in \mathbb{R}$. a is not worse than b if $a \leq b$

$f(x')$ *not worse* than $f(x)$ if $f(x') \leq f(x)$



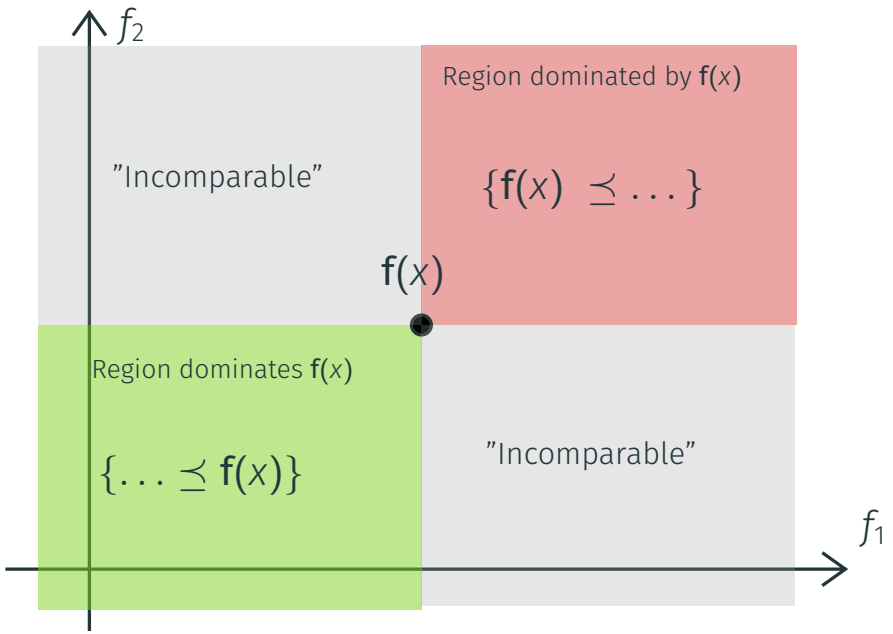
$$\min_{x \in \mathcal{X}} f(x) = y \text{ verifies } y \leq f(x), \quad \forall x \in \mathcal{X}$$

Pareto Domination relation

Partial order relation on $\mathbb{R}^2$: $\preceq$

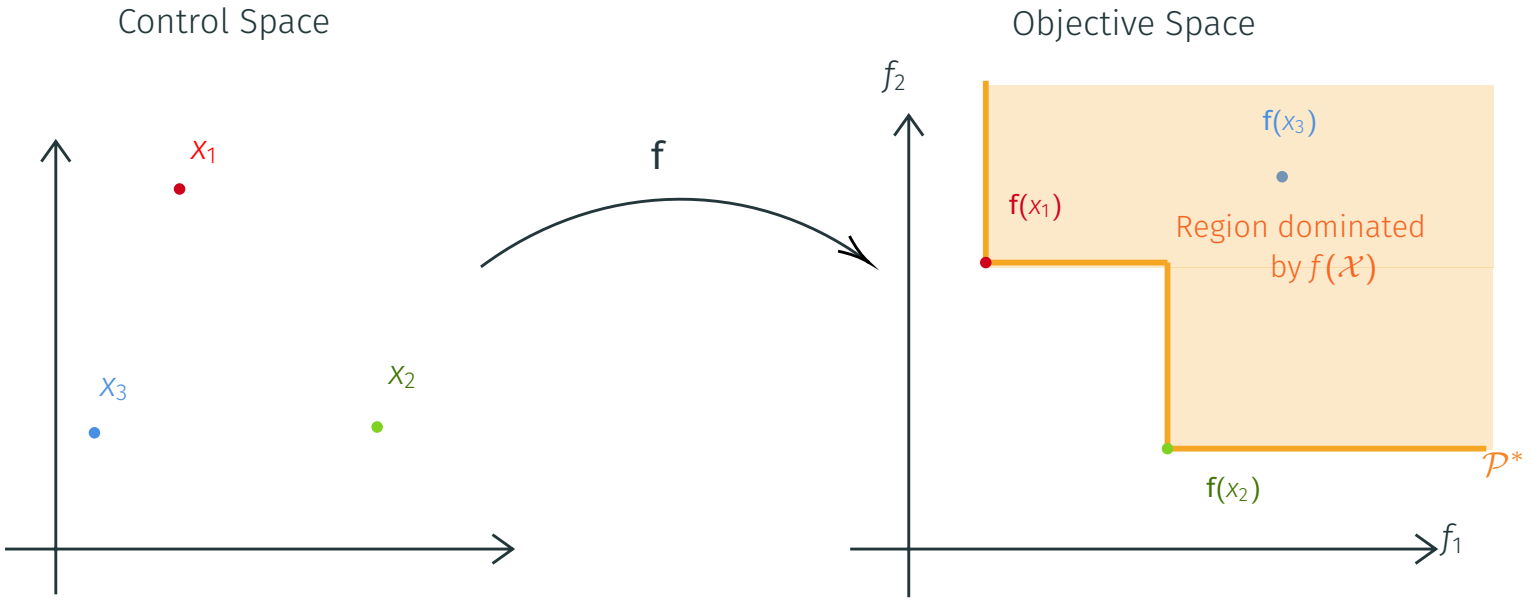
Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$.

$\mathbf{a} \preceq \mathbf{b} \iff a_i \leq b_i \quad \forall i$: all components of $\mathbf{a}$ are *not worse* than those of $\mathbf{b}$



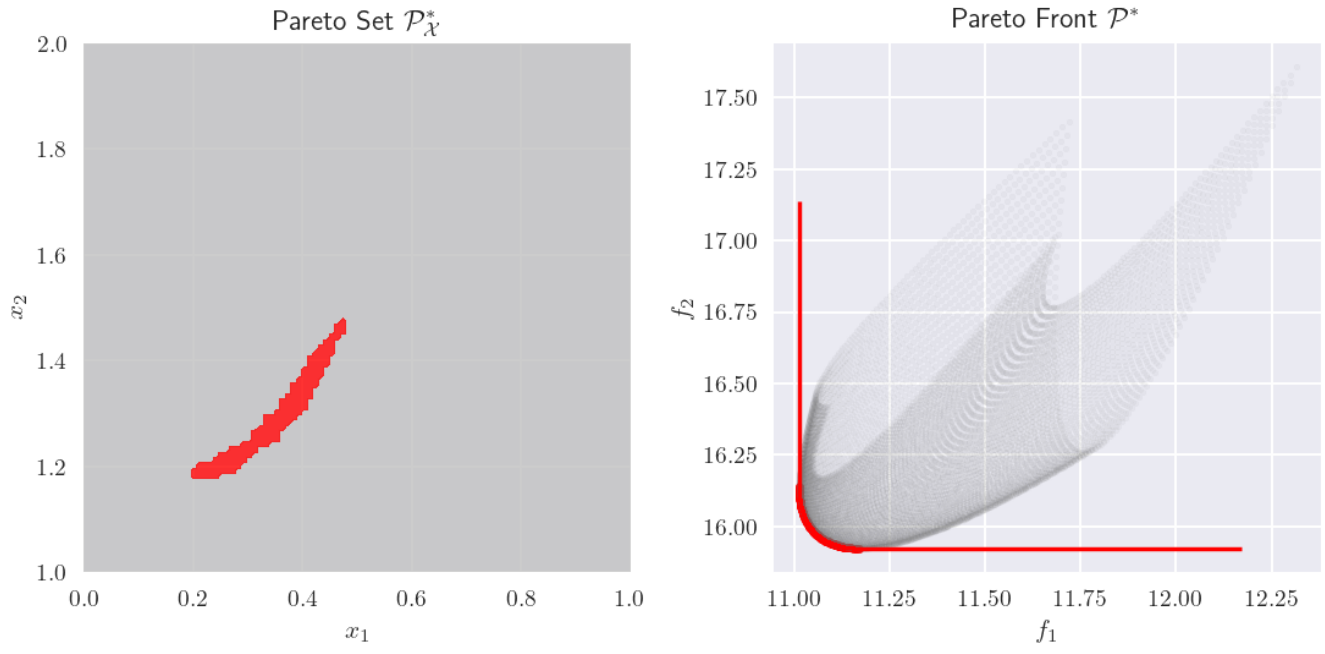
"Incomparable"
$\rightarrow$ some components are better, some worse

Pareto Front, Pareto Set



- Pareto Front $\mathcal{P}^*$: Set of **non-dominated** point of $\mathbf{f}(\mathcal{X})$. Subset of objective space
- Pareto set $\mathcal{P}^*_{\mathcal{X}}$: Preimage of Pareto front. Subset of control space

Pareto Front, Pareto Set



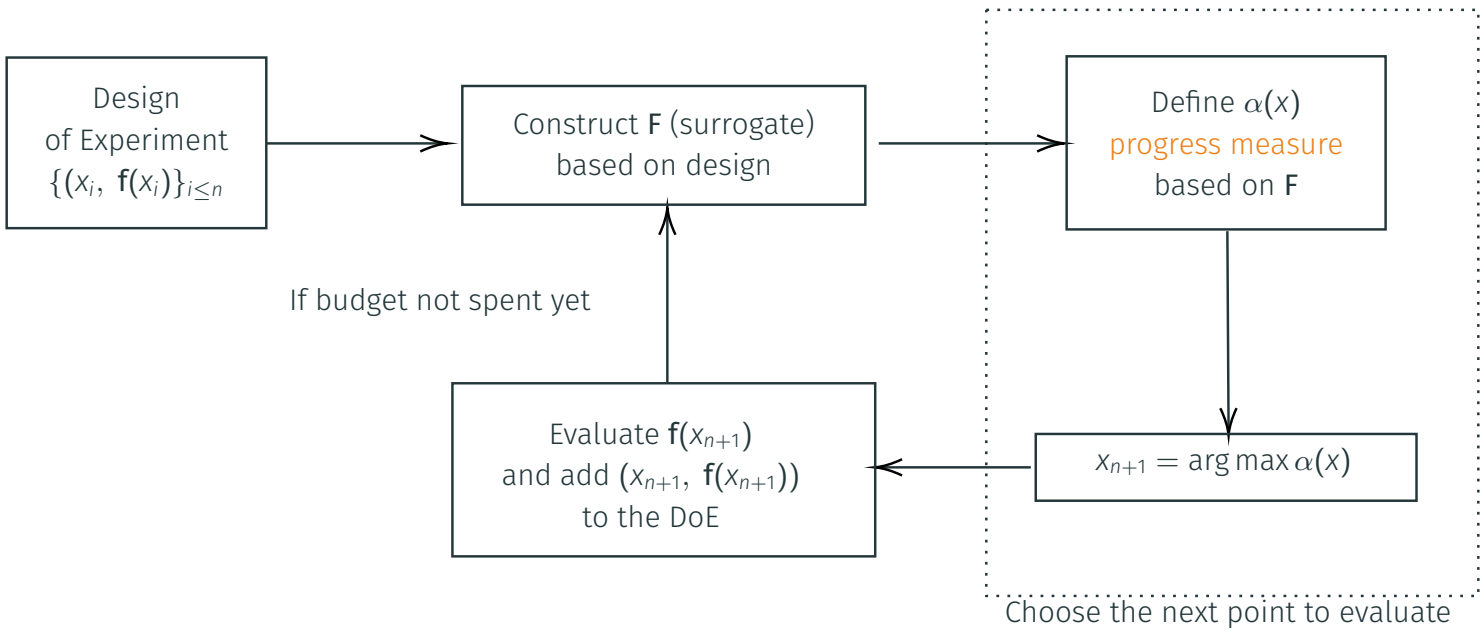
- Pareto Front $\mathcal{P}^*$: Set of **non-dominated** point of $\mathbf{f}(\mathcal{X})$. Subset of objective space
- Pareto set $\mathcal{P}^*_{\mathcal{X}}$: Preimage of Pareto front. Subset of control space

Typical Bayesian Optimization Loop

Multidim GP

$$\mathbf{F} \sim \mathrm{GP}(\mathbf{m}, \mathbf{k})$$

$$\mathbf{F}(x,u) \sim \mathcal{N}\left(\begin{bmatrix} m_1(x,u) \\ m_2(x,u) \end{bmatrix}, \begin{bmatrix} \sigma_1(x,u) & 0 \\ 0 & \sigma_2(x,u) \end{bmatrix}\right)$$



Multiobjective Bayesian Optimisation

Bayesian Optimisation methods exist to find non dominated points !

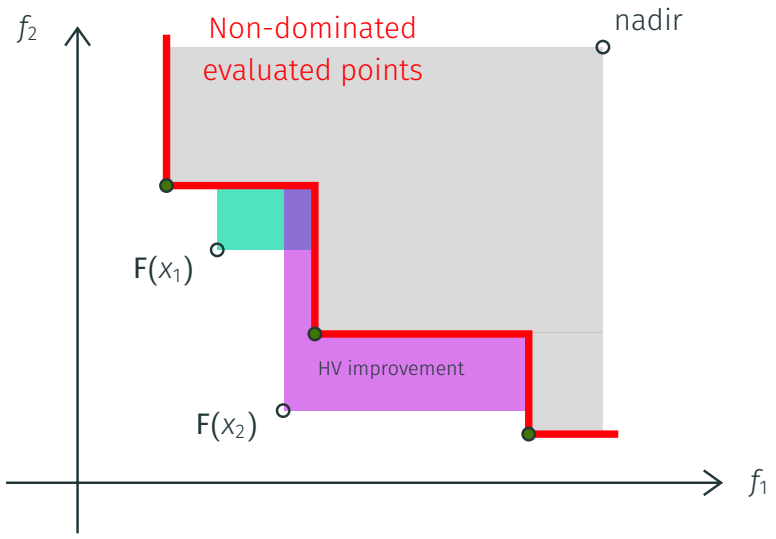
EHVI, SUR, $\star$-ESMO . . . [Pic13, HHSA16]

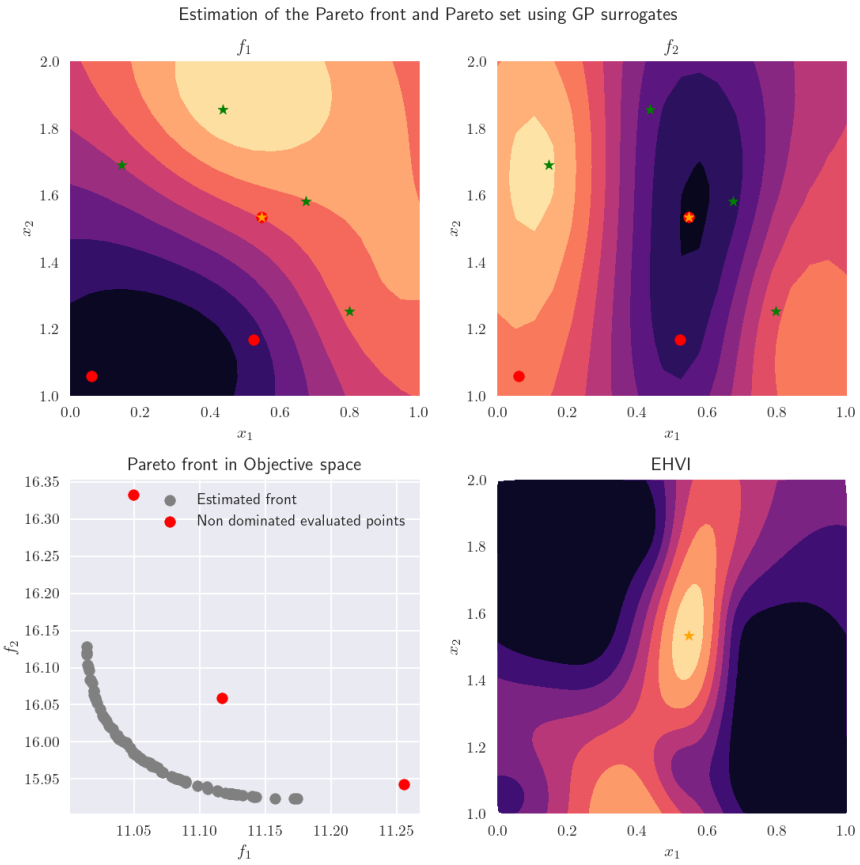
Hypervolume Improvement

$$\mathrm{HVI}(x_1) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\mathrm{HVI}(x_2) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\alpha(x) = \mathbb{E}\left[HVI(x)\right]$$





Multiobjective Bayesian Optimisation

Bayesian Optimisation methods exist to find non dominated points !

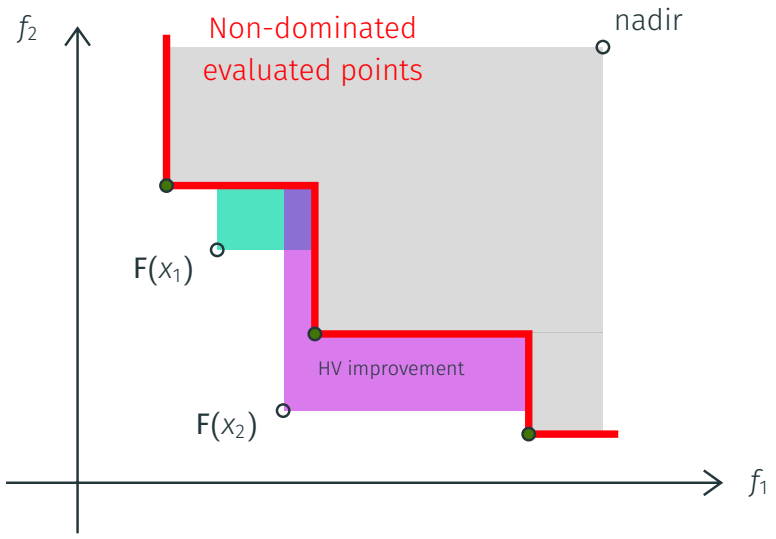
EHVI, SUR, $\star$-ESMO . . . [Pic13, HHSA16]

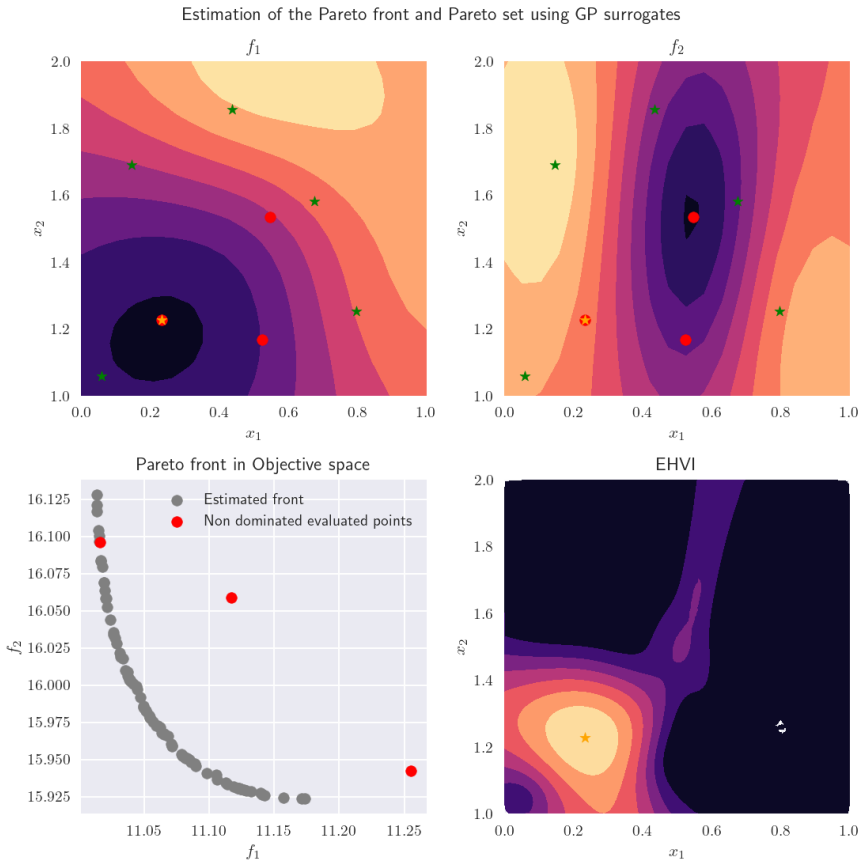
Hypervolume Improvement

$$\mathrm{HVI}(x_1) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\mathrm{HVI}(x_2) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\alpha(x) = \mathbb{E}\left[HVI(x)\right]$$





Multiobjective Bayesian Optimisation

Bayesian Optimisation methods exist to find non dominated points !

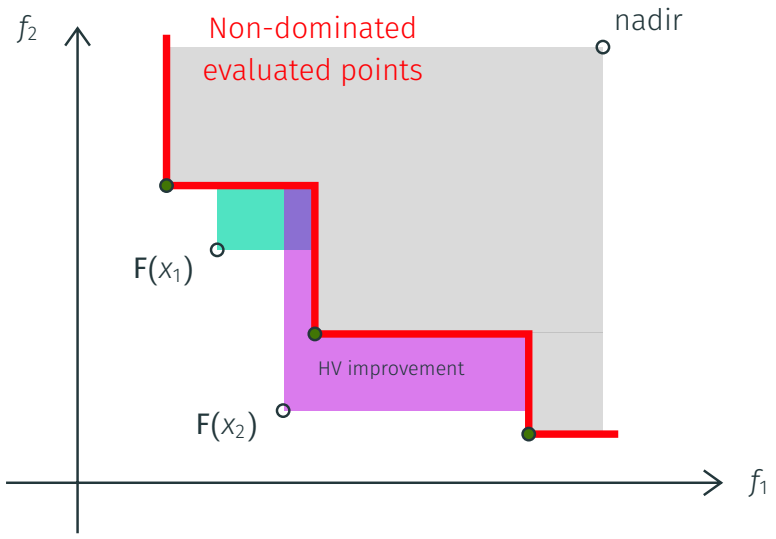
EHVI, SUR, $\star$-ESMO . . . [Pic13, HHSA16]

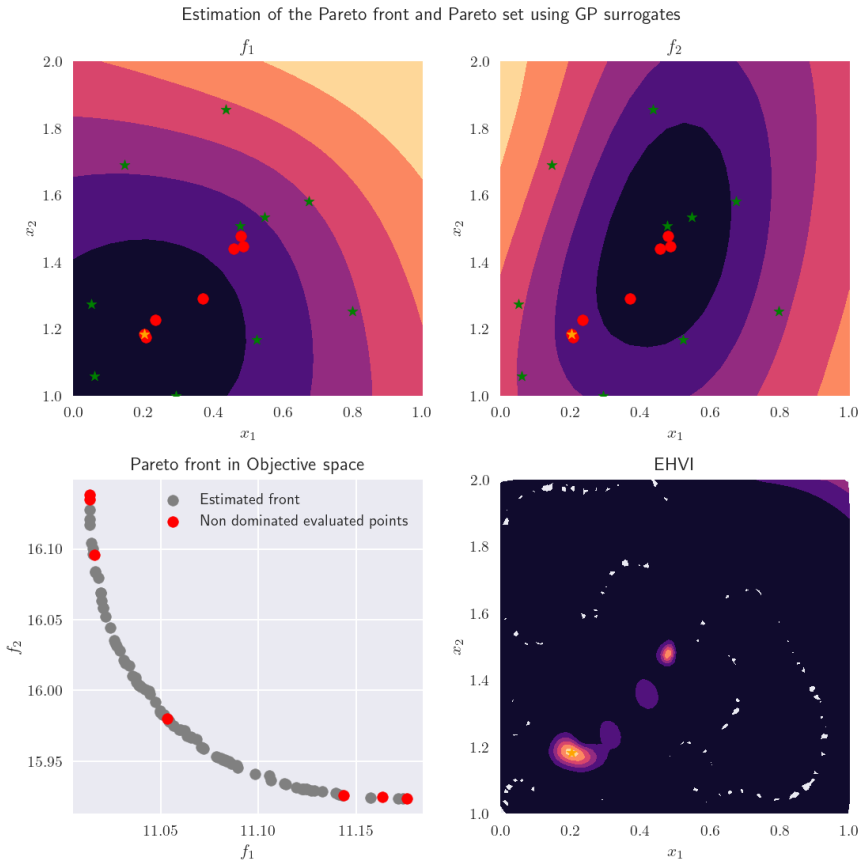
Hypervolume Improvement

$$\mathrm{HVI}(x_1) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\mathrm{HVI}(x_2) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\alpha(x) = \mathbb{E}\left[HVI(x)\right]$$





Multiobjective Bayesian Optimisation

Bayesian Optimisation methods exist to find non dominated points !

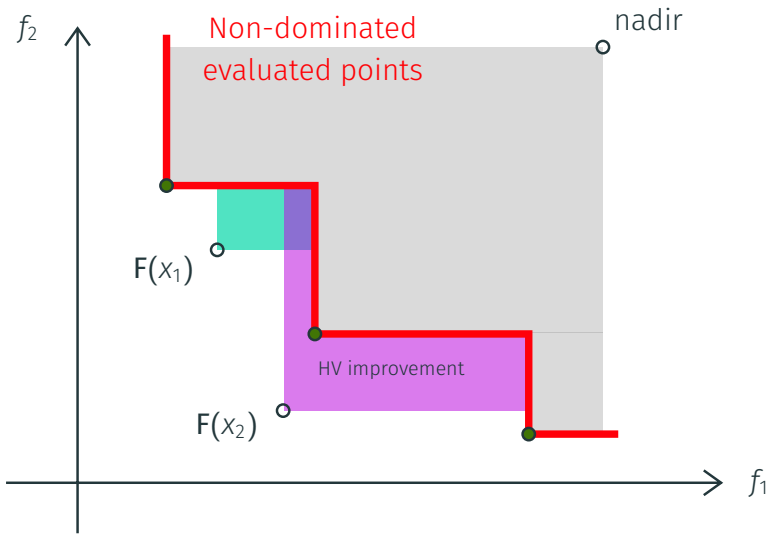
EHVI, SUR, $\star$-ESMO . . . [Pic13, HHSA16]

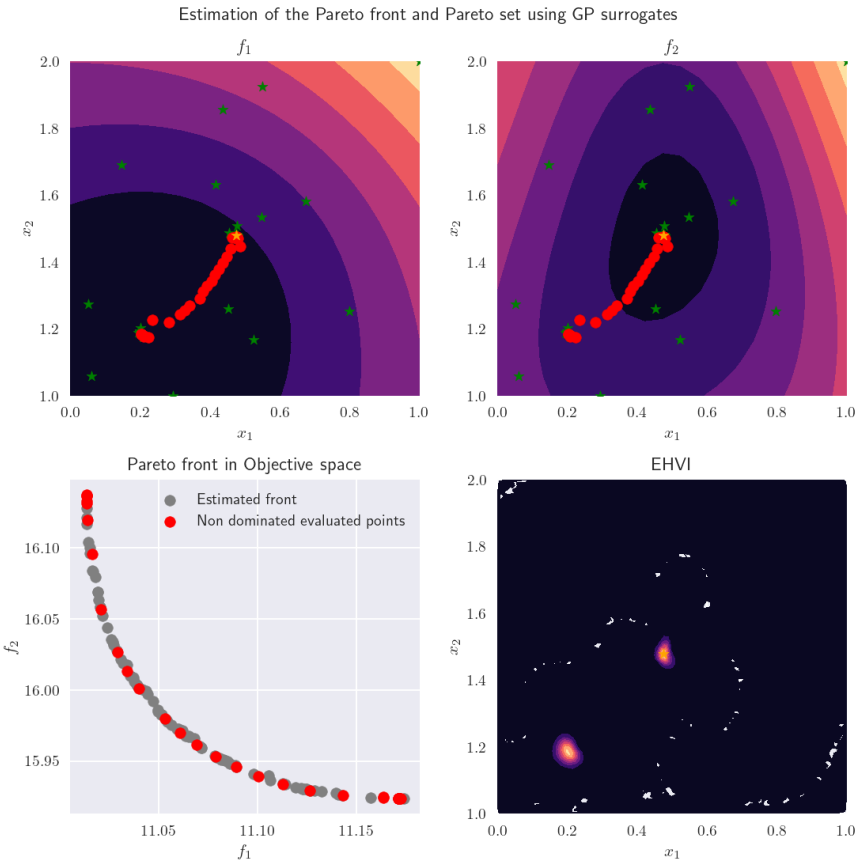
Hypervolume Improvement

$$\mathrm{HVI}(x_1) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\mathrm{HVI}(x_2) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\alpha(x) = \mathbb{E}\left[HVI(x)\right]$$





Multiobjective Bayesian Optimisation

Bayesian Optimisation methods exist to find non dominated points !

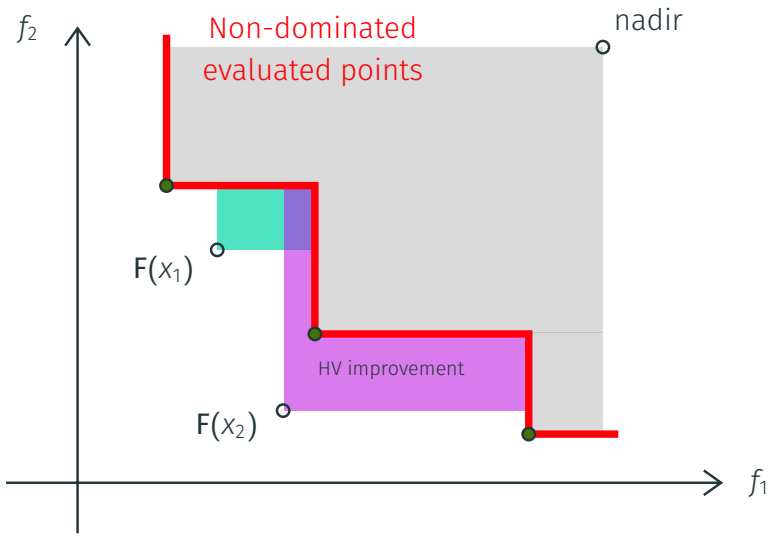
EHVI, SUR, $\star$-ESMO . . . [Pic13, HHSA16]

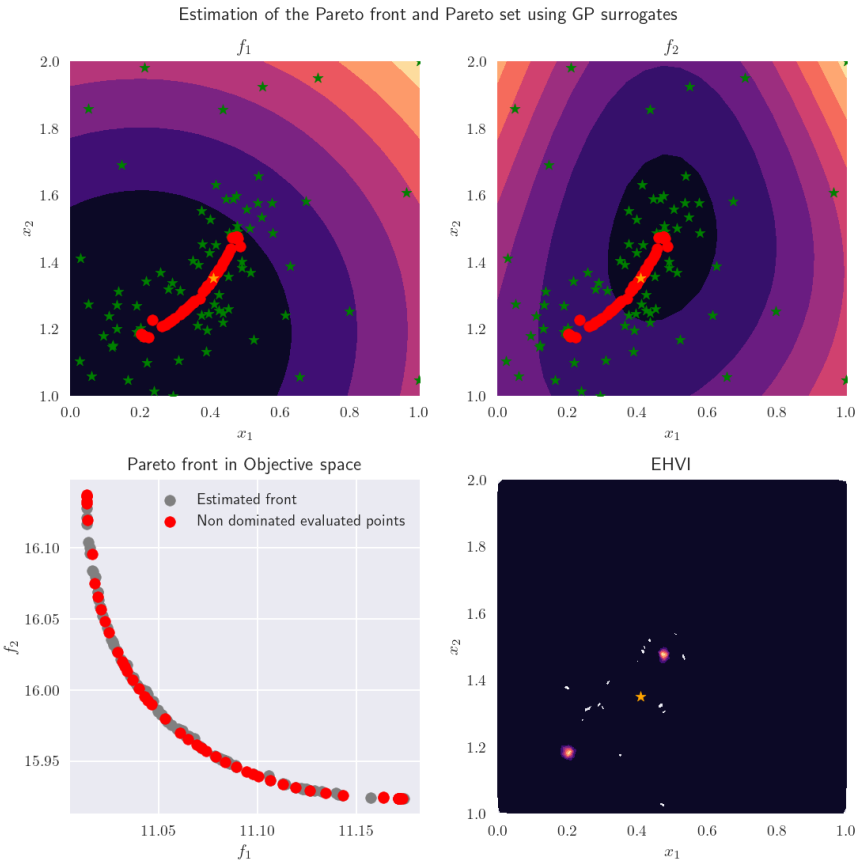
Hypervolume Improvement

$$\mathrm{HVI}(x_1) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\mathrm{HVI}(x_2) = \mathrm{Vol}\left(\square \cup \square\right) - \mathrm{Vol}\left(\square\right)$$

$$\alpha(x) = \mathbb{E}\left[HVI(x)\right]$$





Optimal Conditions today ?

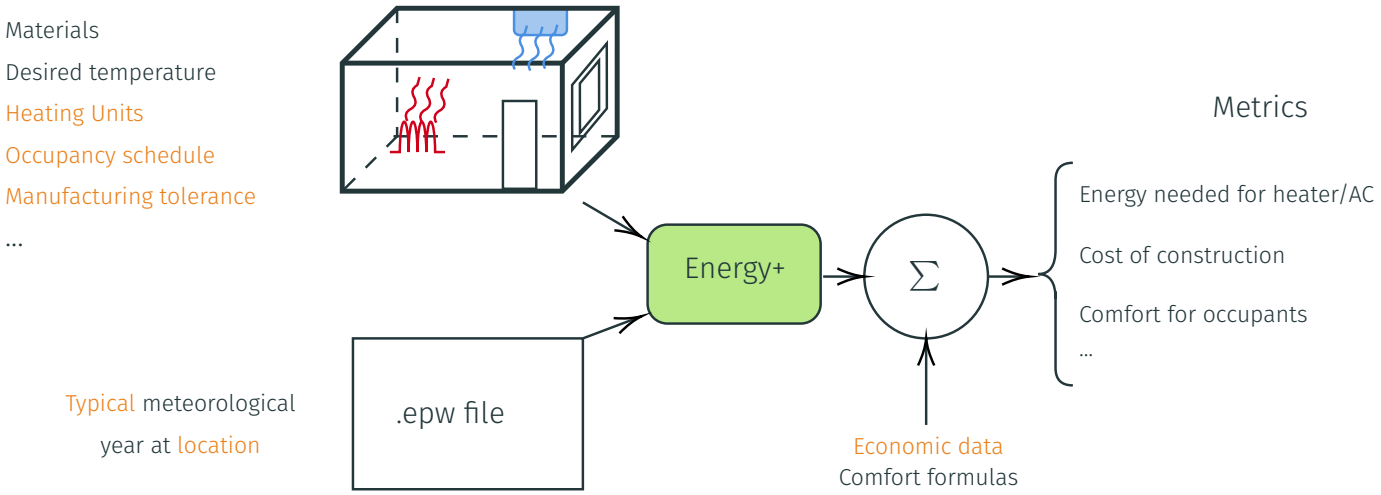
Buildings have a lifespan > 20 years.

If the building is ”optimal” today, how about in 5 years, 10 years or 20+ years ?

$\Rightarrow$ Many uncertainties can be taken into account for this modelling !

Multiobjective Optimization under uncertainties

Introducing uncertainties



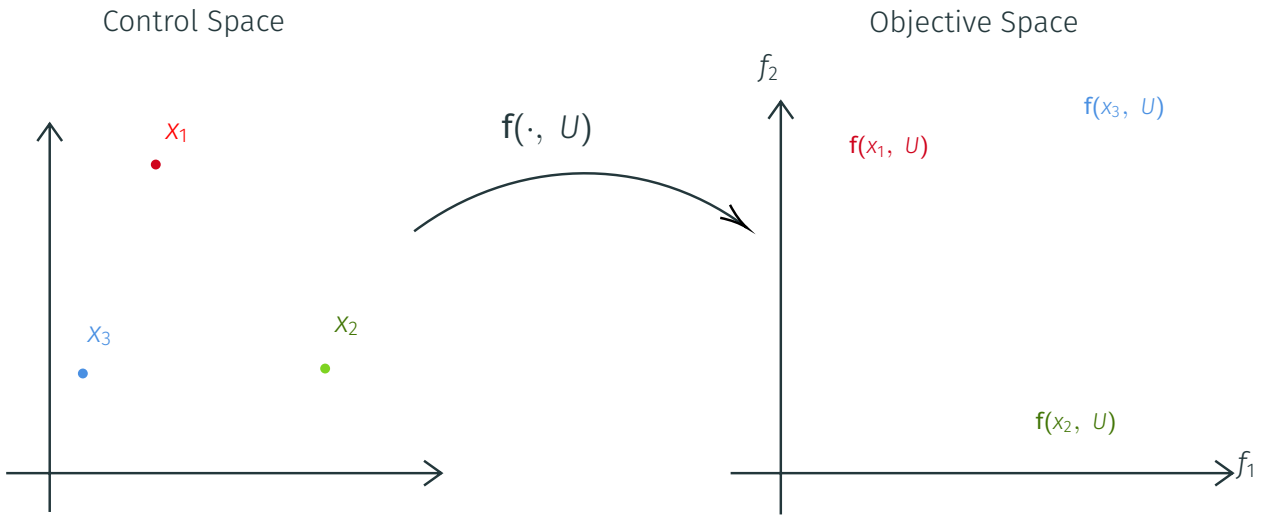
Let $u \sim U$ with support $\mathcal{U}$ which models

- Nominal vs real physical properties
- Energy prices / Environmental impact
- Climate Change
- Time variability of materials properties
- Occupancy schedules...

$\mathbf{f}(x, U)$: Multivariate random variable to *optimize* ?

Comparing sets ?

$\mathbf{f}(x, U)$: Multivariate random variable

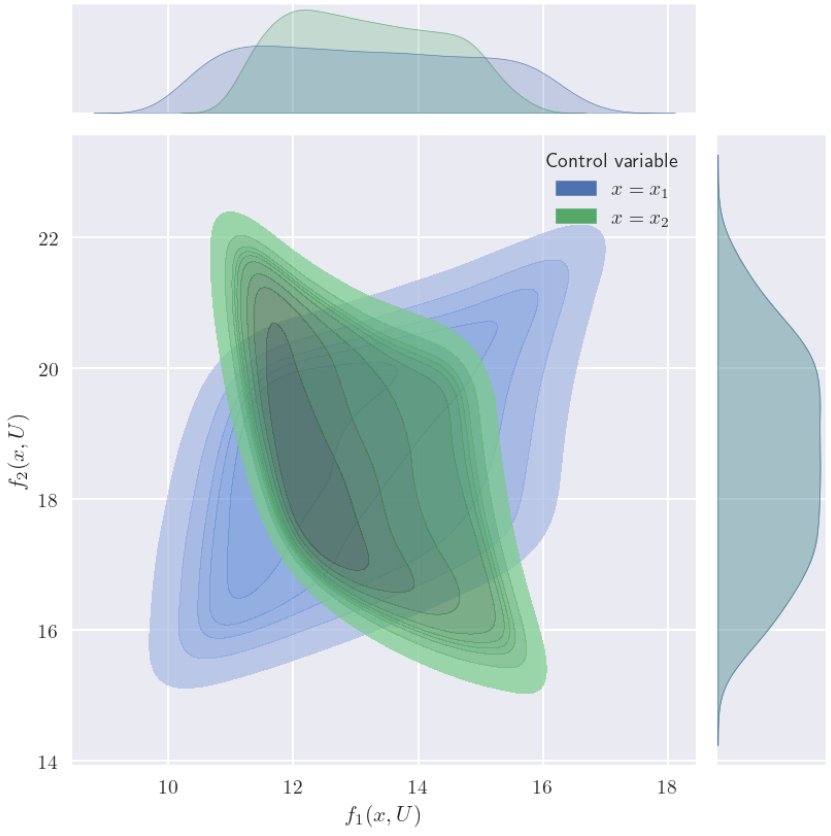


What does it mean to be better/not worse in this context ?

How to deal with U ?

Comparing sets ?

$\mathbf{f}(x_1, U)$ vs $\mathbf{f}(x_2, U)$



What does it mean to be better/not worse in this context ?

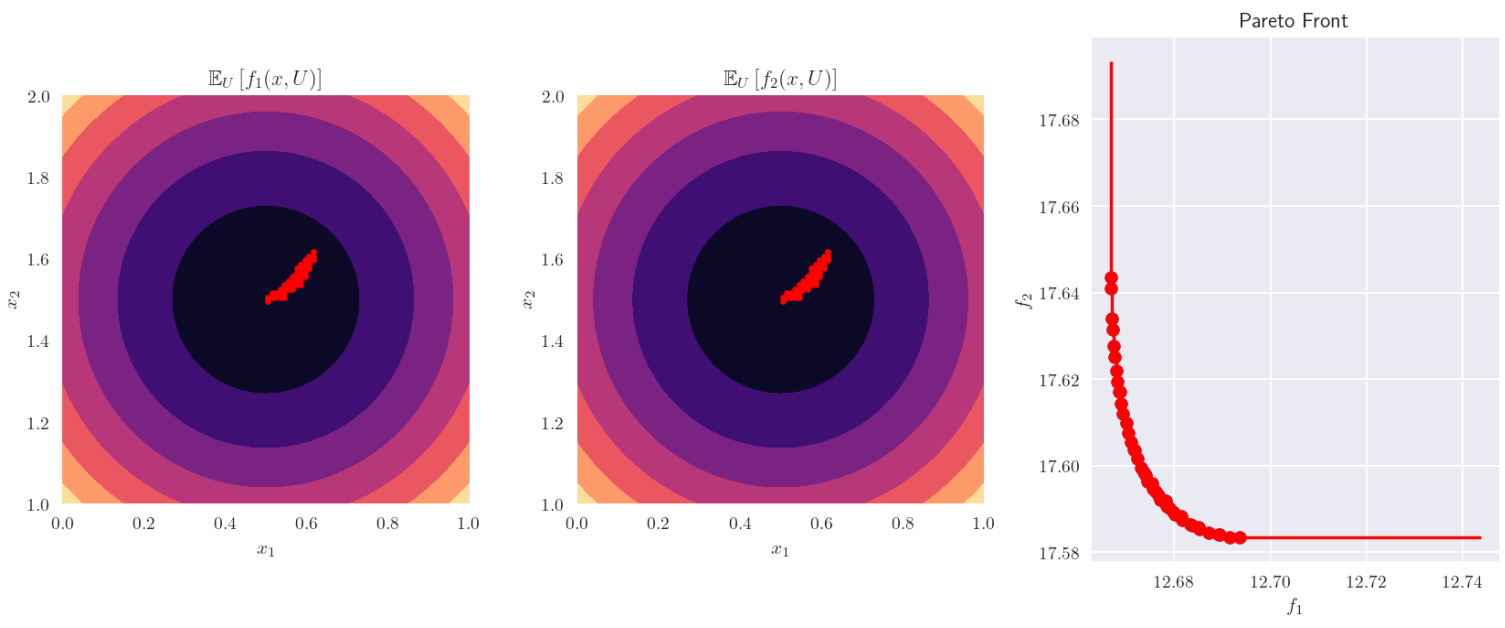
How to deal with U ?

Mean Objective Optimization

We can look for statistics of $\mathbf{f}(x, U)$.

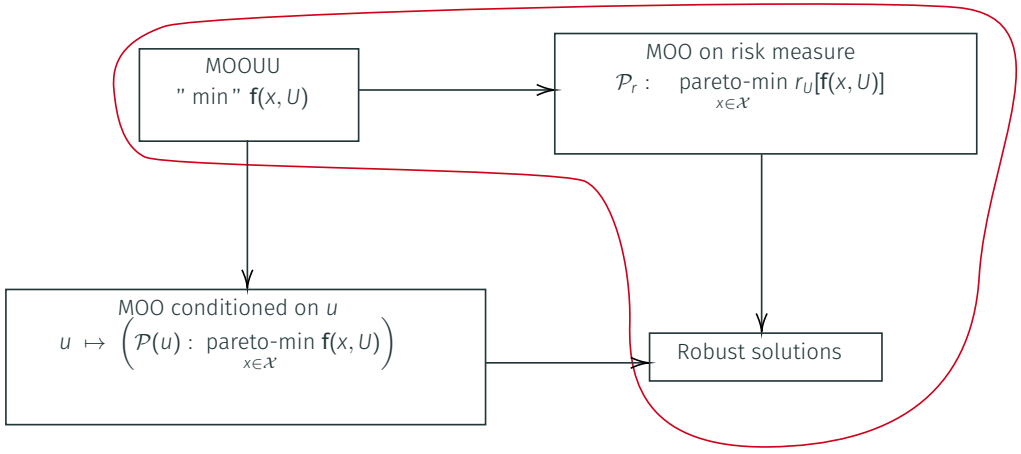
Mean Objective Optimization

Non dominated points of $\mathbb{E}_U[\mathbf{f}(x, U)] = (\mathbb{E}_U[f_1(x, U)], \ldots, \mathbb{E}_U[f_m(x, U)])$



$\Rightarrow$ We can construct an acquisition function (like in EFISUR, or Noisy EHVI)

Robustify or Solve first ?



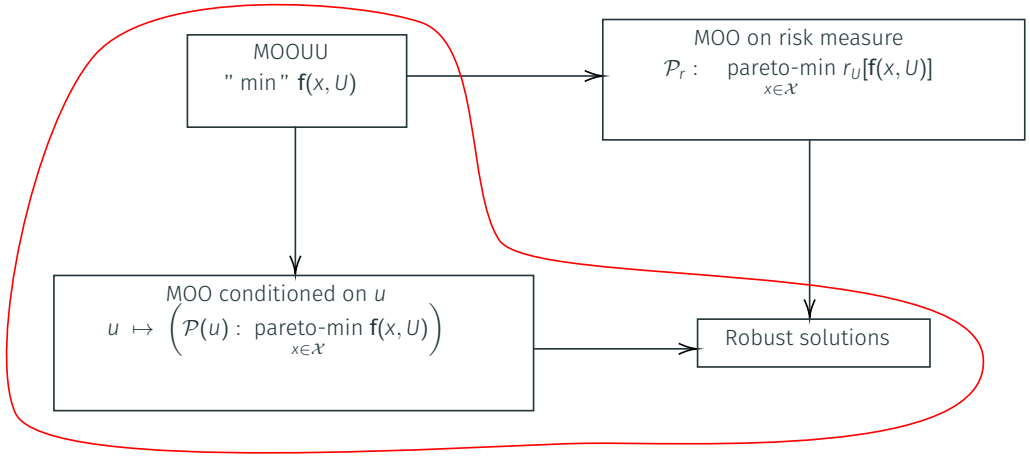
"Robustify then Solve"

$$r_U[\mathbf{f}(x,U)] = \begin{cases} \mathbb{E}_U\left[\mathbf{f}(x,U)\right] & \text{[QCD23, DG05]} \\ \text{Multivariate VaR} & \text{[TKLS24, DCB}^+\text{22]} \\ \text{Bounding boxes approaches} & \text{[RC22]} \end{cases}$$

"Solve then Robustify"

- For a given u, MOO problem $\rightarrow$ Conditional Pareto front $\mathcal{P}^*(u)$
- Do some things using $\mathcal{P}^*(U)$ and $\mathcal{P}^*_{\mathcal{X}}(U)$?

Robustify or Solve first ?



"Robustify then Solve"

$$r_U[\mathbf{f}(x,U)] = \begin{cases} \mathbb{E}_U\left[\mathbf{f}(x,U)\right] & [\text{QCD23, DG05}] \\ \text{Multivariate VaR} & [\text{TKLS24, DCB}^+\text{22}] \\ \text{Bounding boxes approaches} & [\text{RC22}] \end{cases}$$

"Solve then Robustify"

- For a given u, MOO problem $\rightarrow$ Conditional Pareto front $\mathcal{P}^*(u)$
- Do some things using $\mathcal{P}^*(U)$ and $\mathcal{P}^*_{\mathcal{X}}(U)$?

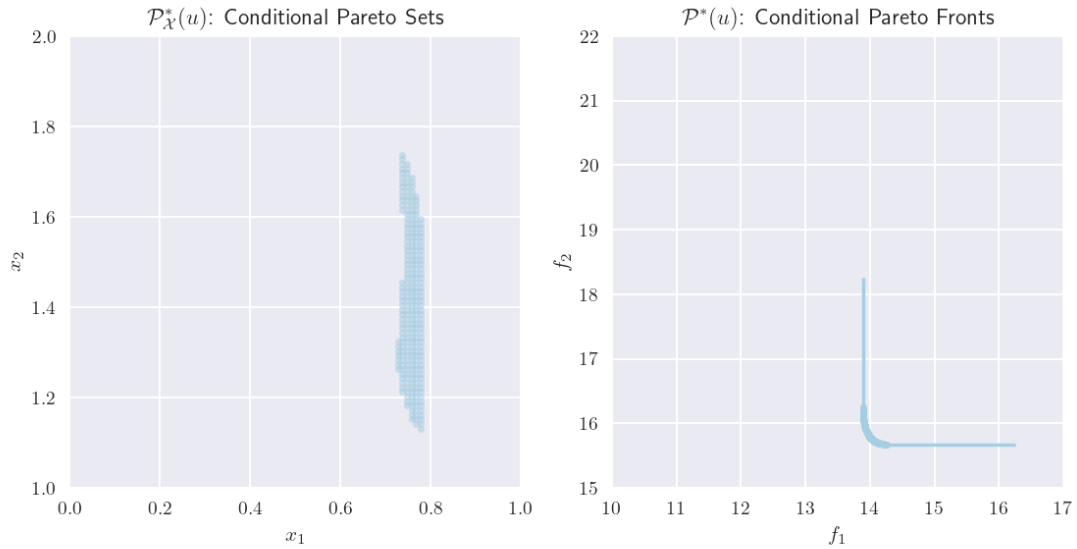
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



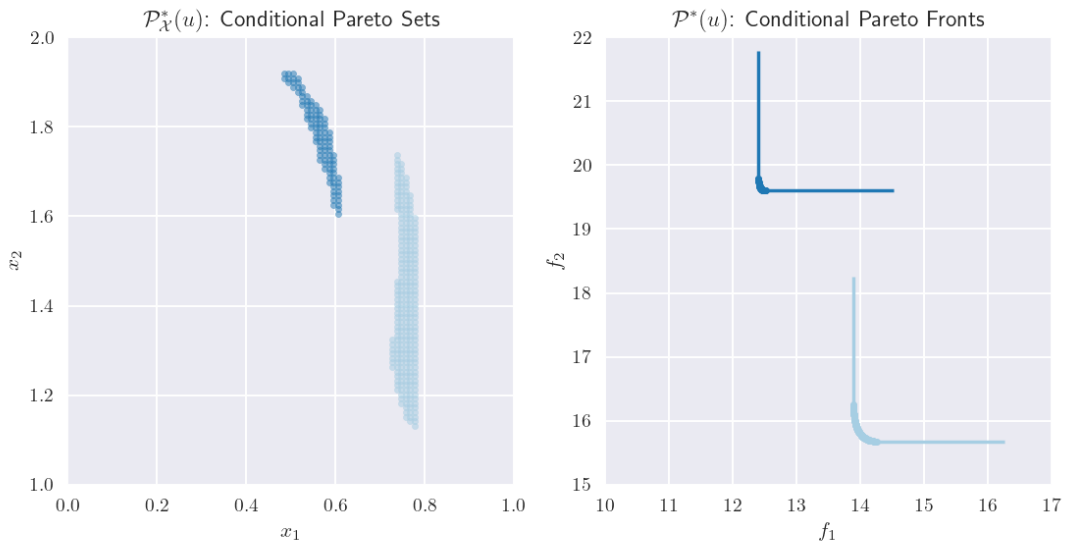
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



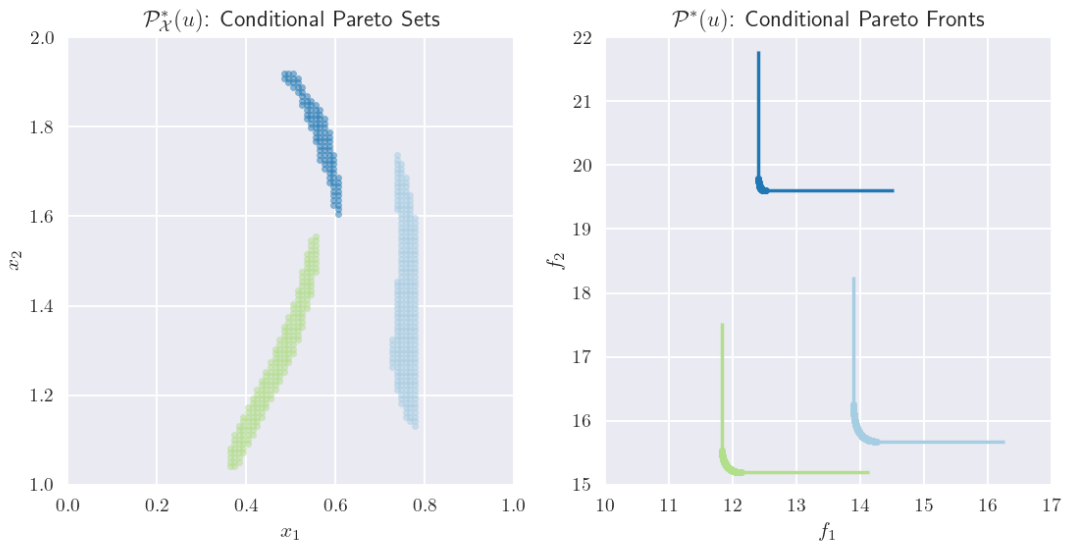
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



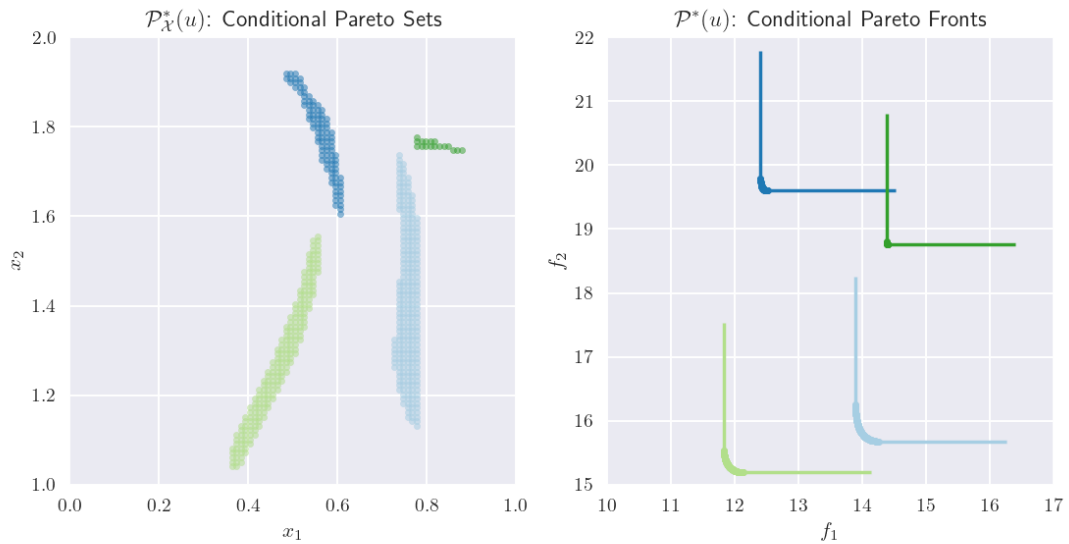
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



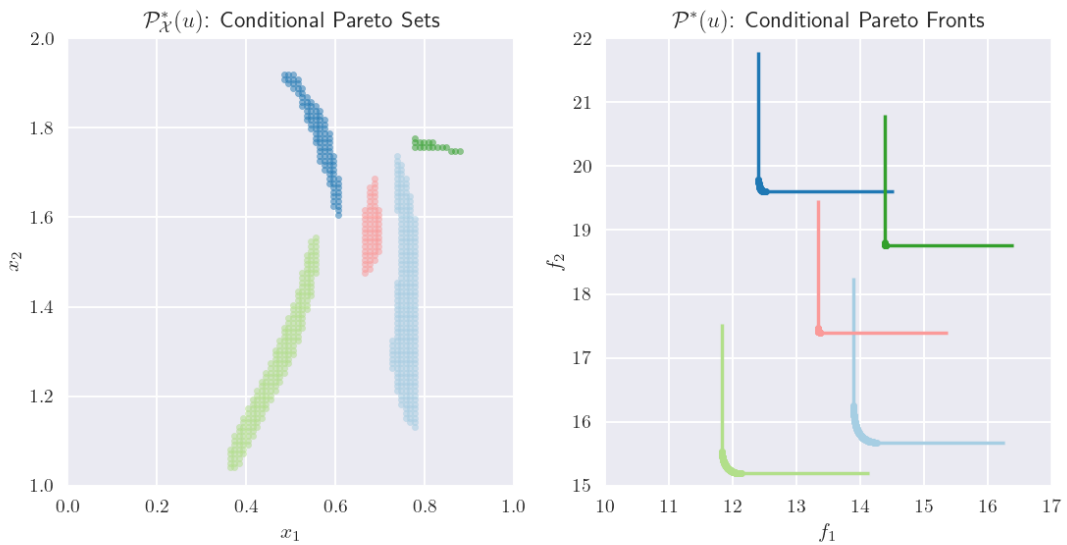
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



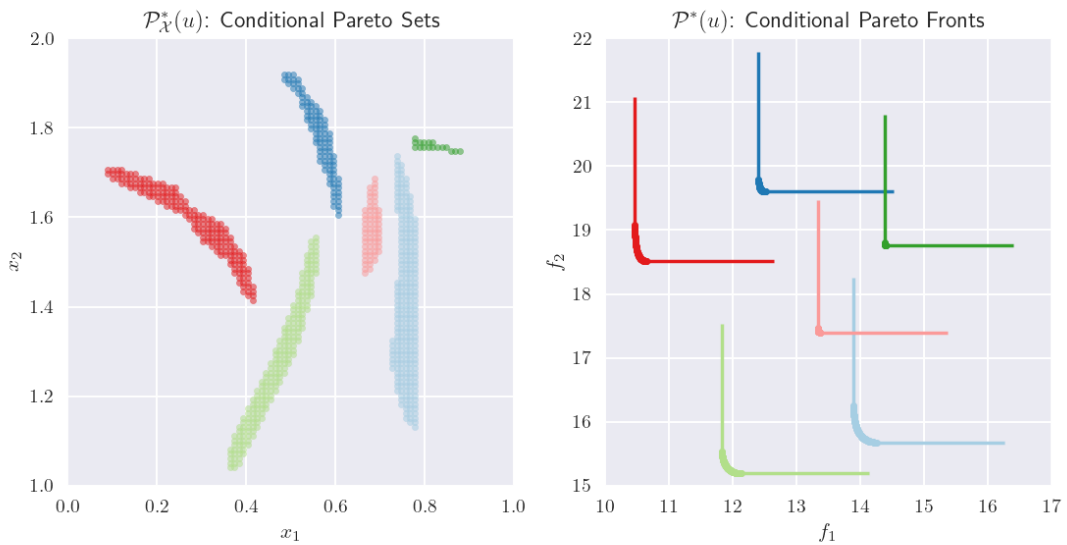
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



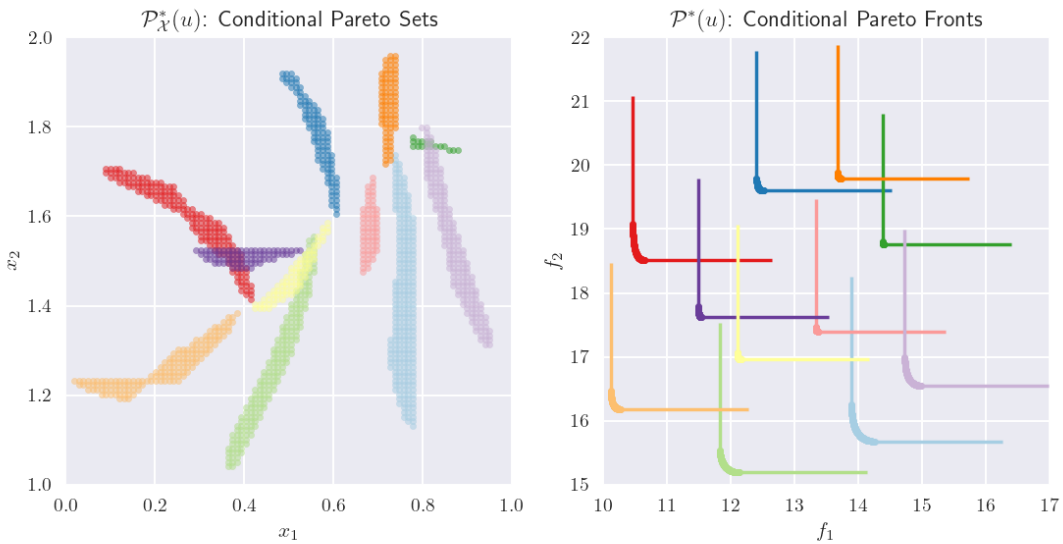
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



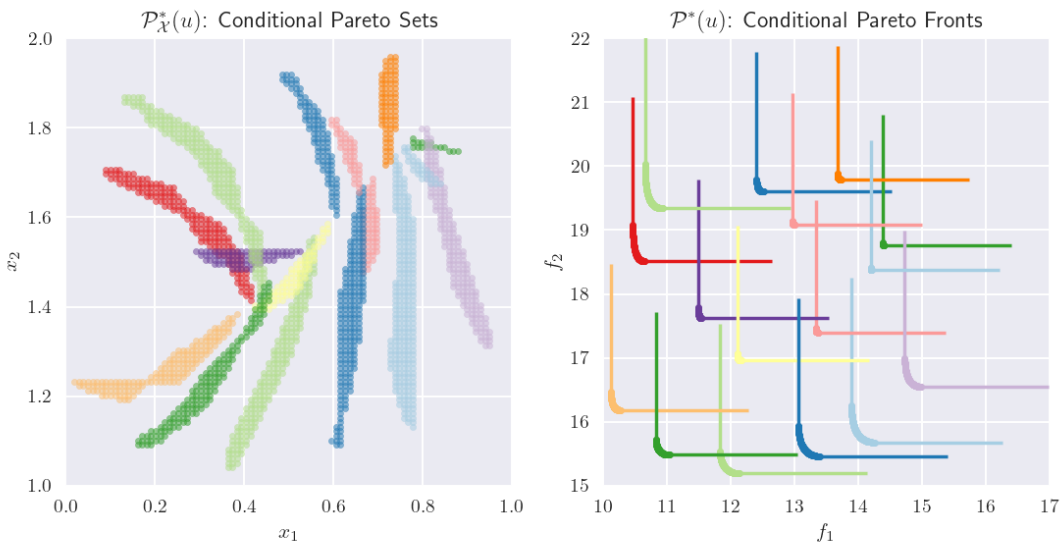
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



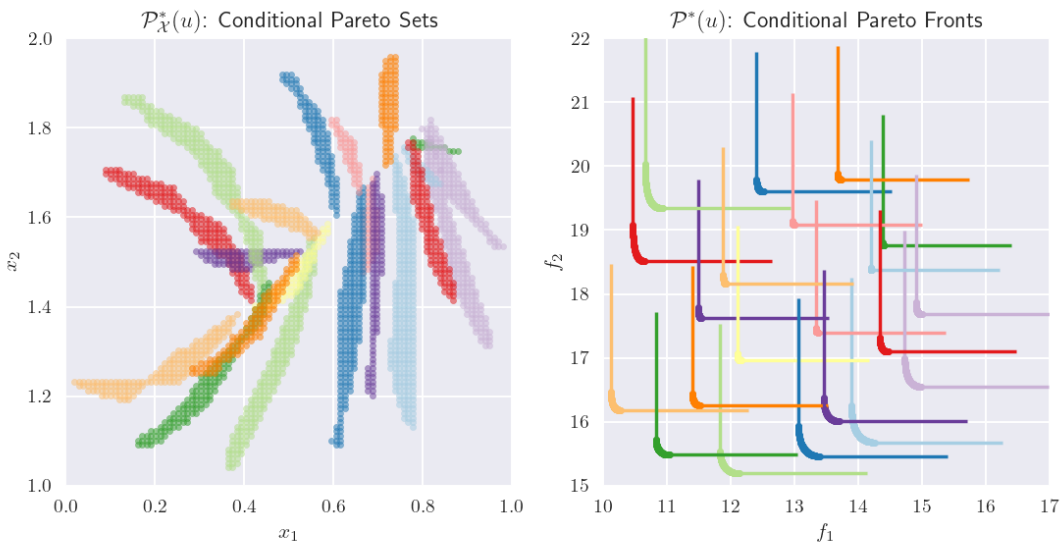
Conditional Pareto Front/Set – Contextual Optimization

Given $u \in \mathcal{U}$, we have a deterministic MOO problem:

Conditional Pareto Front/Set

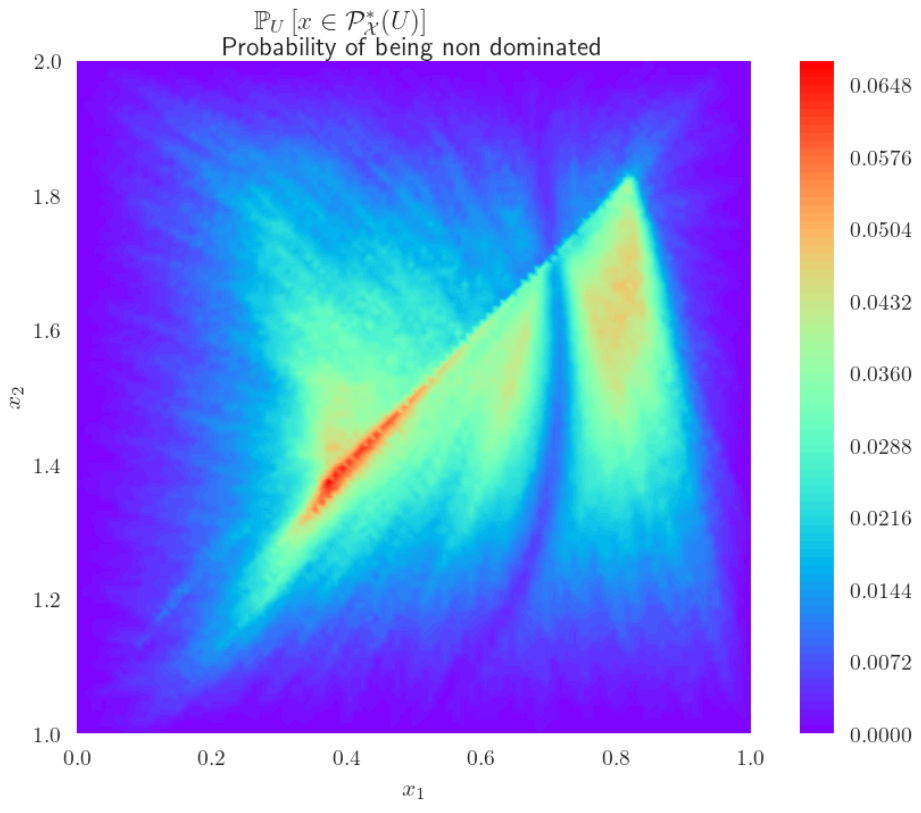
$$\mathcal{P}^*(u) = \text{non-dom}\,\{\mathbf{f}(x, u) \text{ for } x \in \mathcal{X}\}$$

$$\mathcal{P}^*_{\mathcal{X}}(u) = \{x \in \mathcal{X} \text{ s.t. } \mathbf{f}(x, u) \text{ non-dominated}\}$$



Probability of being non-dominated

$\rightarrow \mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$ the probability of being "Pareto-optimal"

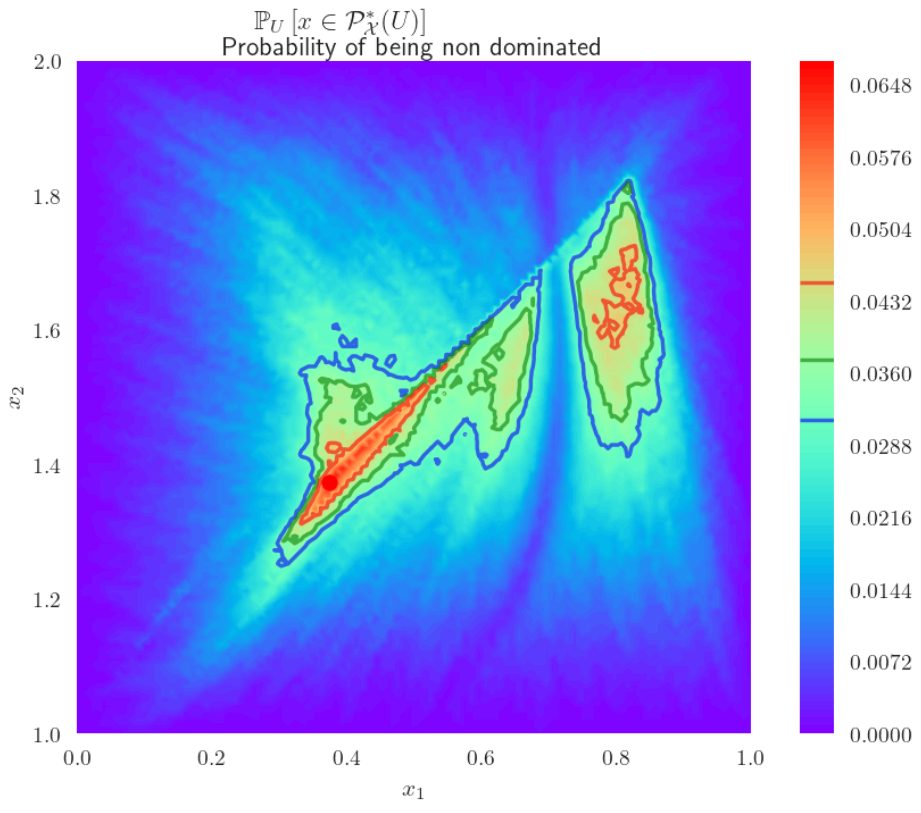


- $\max \mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$
- Get subset of $\mathcal{X}$ which corresponds to the top k% of values

How to estimate $u \mapsto \mathcal{P}^*_{\mathcal{X}}(u)$?

Probability of being non-dominated

$\rightarrow \mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$ the probability of being "Pareto-optimal"



- $\max \mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$
- Get subset of $\mathcal{X}$ which corresponds to the top k% of values

How to estimate $u \mapsto \mathcal{P}^*_{\mathcal{X}}(u)$?

GP framework

$\mathsf{f}(x, u)$ is expensive to evaluate once, it is even worse when taking into account the random nature of $u \sim U$

We consider the joint input space $\mathcal{X} \times \mathcal{U}$:

$$\mathsf{F}(x, u) \sim \mathcal{N}(\mathsf{m}(x, u), \Sigma(x, u)) \tag{1}$$

Since F is a GP, it works well with linear operators, ie $\mathsf{Z}(x) = \mathbb{E}[\mathsf{F}(x, U)]$ still a GP !

Modify the EHVI for CPF estimation

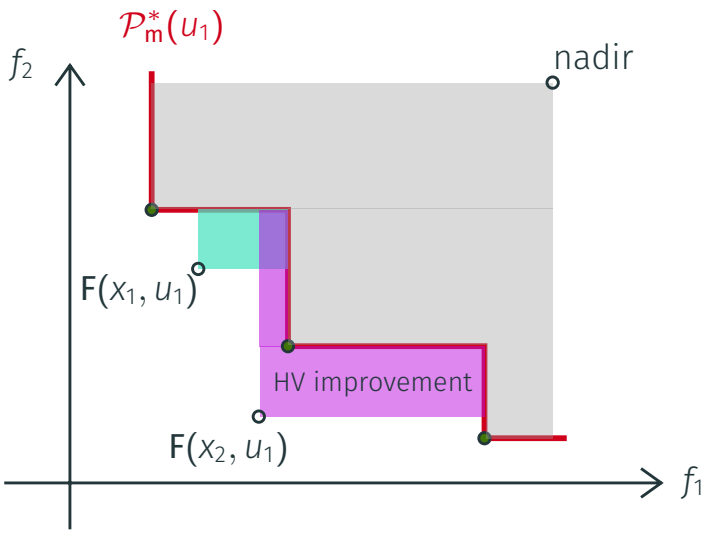
We want to estimate $u \mapsto \mathcal{P}^*(u), \mathcal{P}^*_{\mathcal{X}}(u)$ using the GP

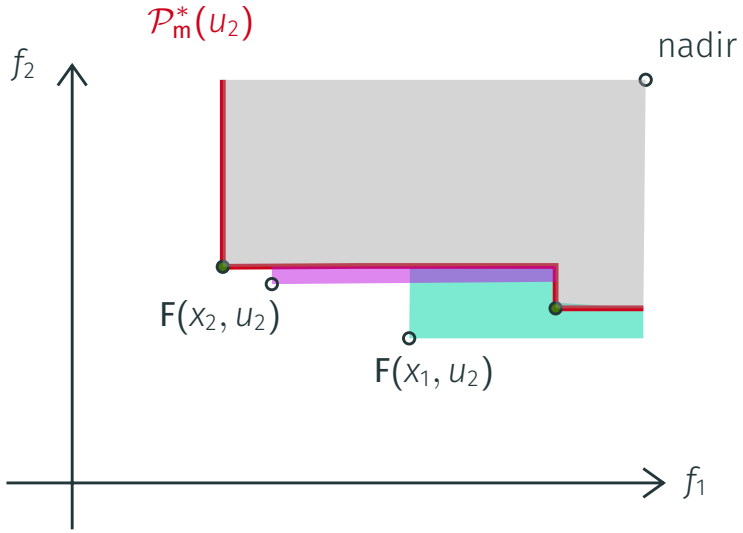
Profile-EHVI

Similar to the PEI defined in [GBC+14], acquisition defined on $\mathcal{X} \times \mathcal{U}$

$$\mathrm{PEHVI}(x, u) = \mathbb{E}_{\mathsf{F}}\left[\underbrace{\mathrm{HV}(\mathcal{P}^*_{\mathsf{m}}(u) \cup \mathsf{F}(x, u)) - \mathrm{HV}(\mathcal{P}^*_{\mathsf{m}}(u))}_{\text{HVI of } \mathcal{P}^*_{\mathsf{m}}(u)}\right]$$

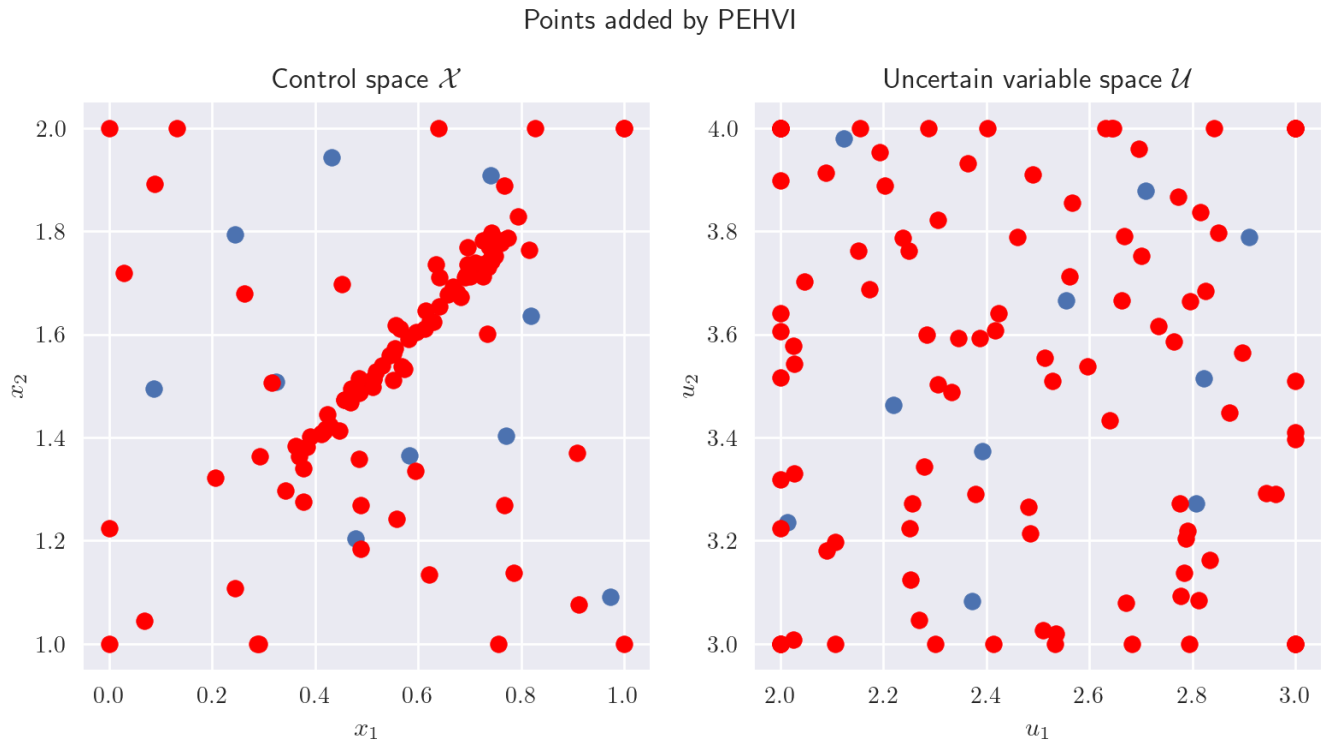
Estimation of $\mathcal{P}^*_{\mathsf{m}}(u)$ based on GP means





Design of Experiment using PEHVI

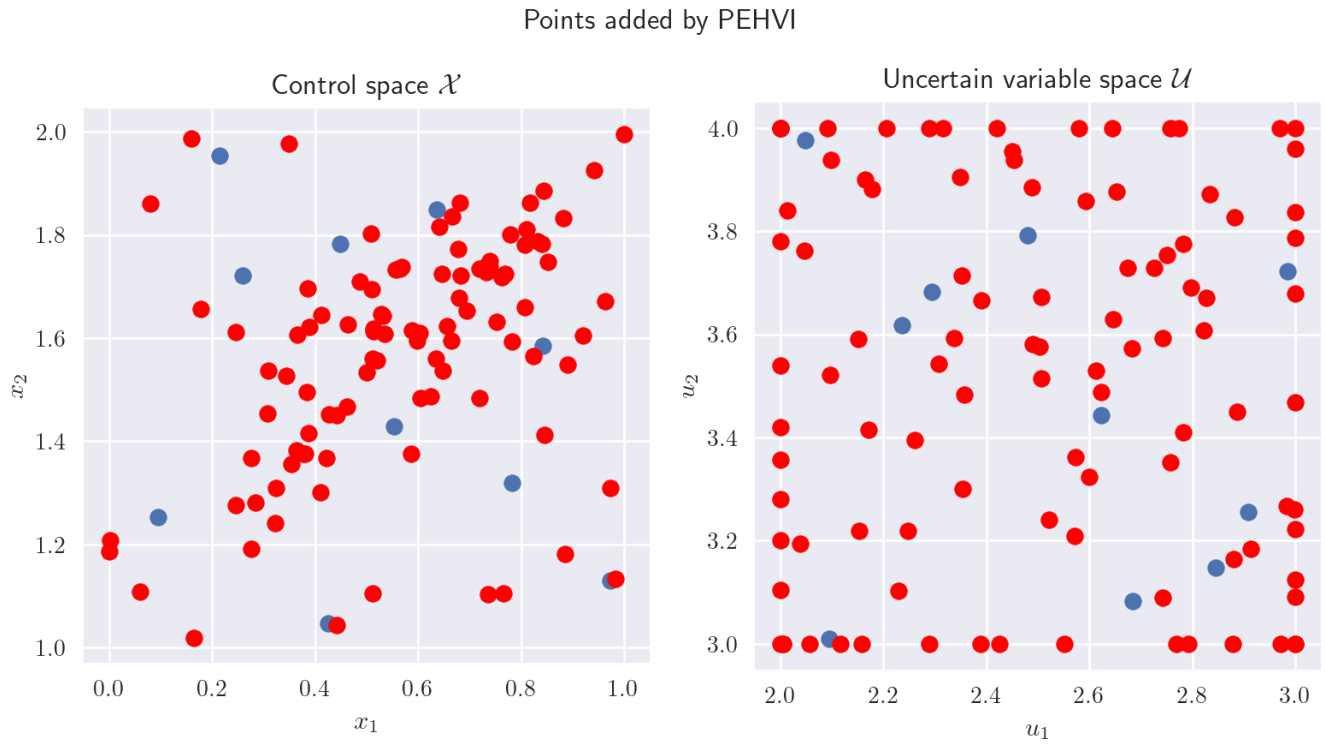
- PEHVI relies on a single optimization in the joint space $\mathcal{X} \times \mathcal{U}$
- This acquisition function leads to an intensification around the diagonal on $\mathcal{X}$, while exploring $\mathcal{U}$.
- Depending on the estimation of $\mathcal{P}^*_{\text{m}}(u)$, we can tend toward more exploration $\Rightarrow$ Better quality metrics



Use the metamodel to estimate $\mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$

Design of Experiment using PEHVI

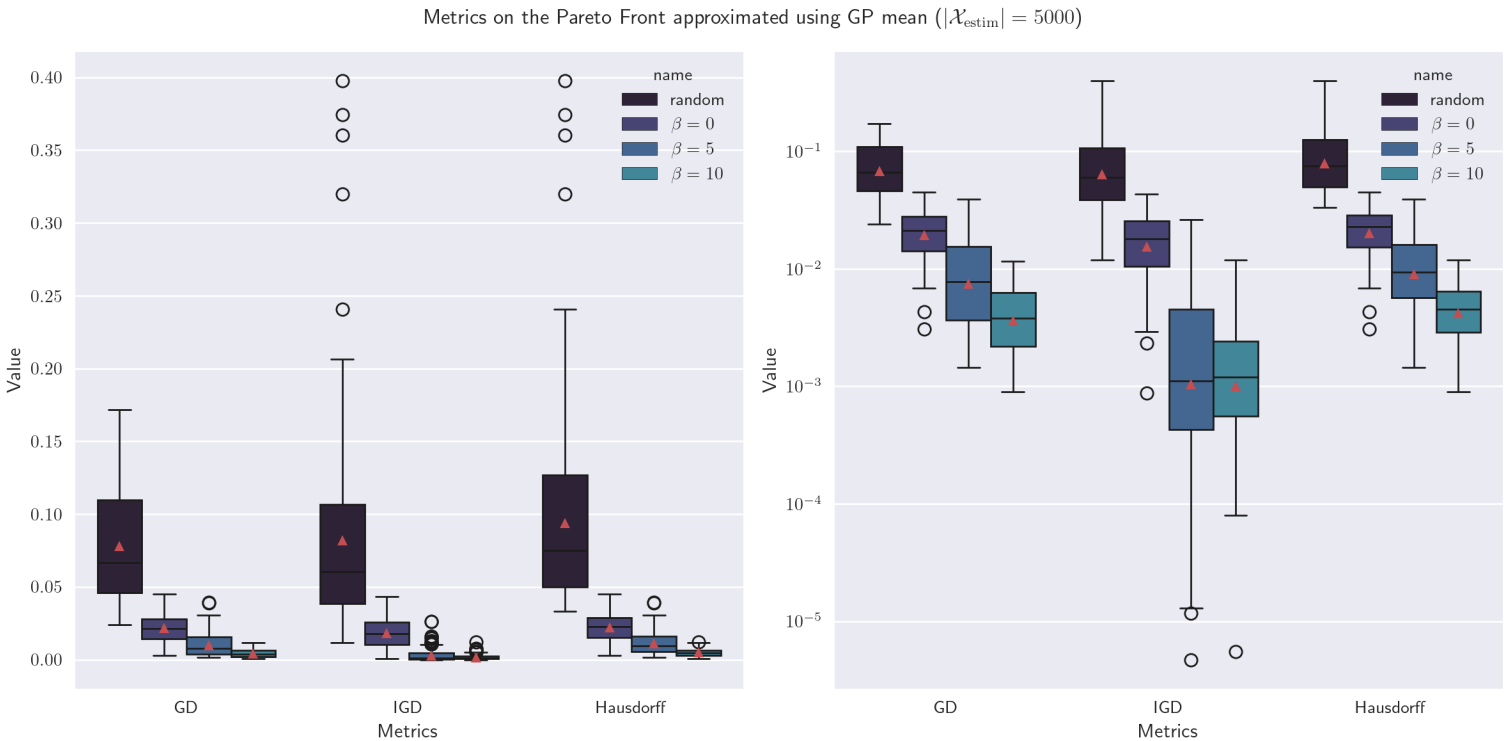
- PEHVI relies on a single optimization in the joint space $\mathcal{X} \times \mathcal{U}$
- This acquisition function leads to an intensification around the diagonal on $\mathcal{X}$, while exploring $\mathcal{U}$.
- Depending on the estimation of $\mathcal{P}^*_{\mathsf{m}}(u)$, we can tend toward more exploration $\Rightarrow$ Better quality metrics



Use the metamodel to estimate $\mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$

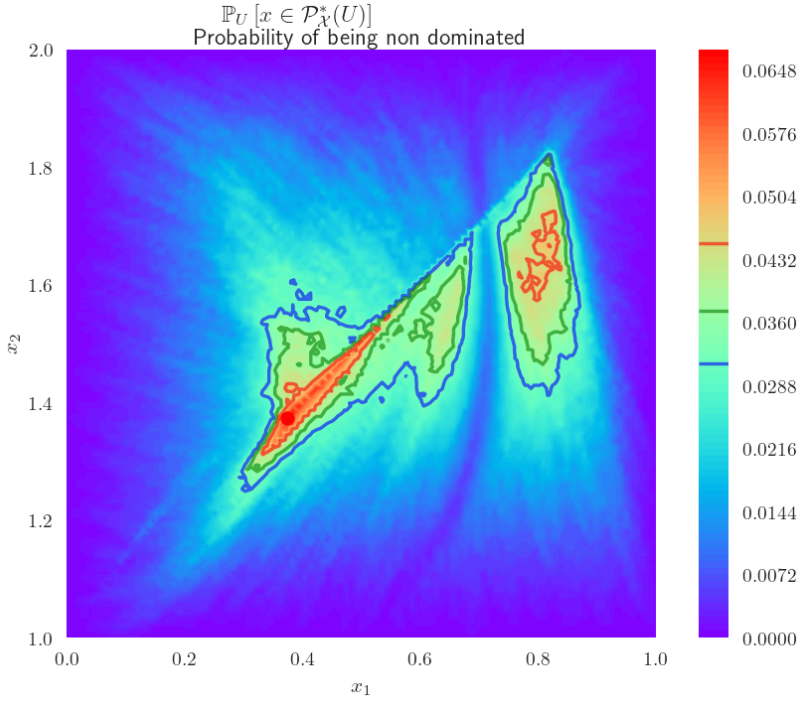
Design of Experiment using PEHVI

- PEHVI relies on a single optimization in the joint space $\mathcal{X} \times \mathcal{U}$
- This acquisition function leads to an intensification around the diagonal on $\mathcal{X}$, while exploring $\mathcal{U}$.
- Depending on the estimation of $\mathcal{P}^*_{\mathsf{m}}(u)$, we can tend toward more exploration $\Rightarrow$ Better quality metrics



Use the metamodel to estimate $\mathbb{P}_U\left[x \in \mathcal{P}^*_{\mathcal{X}}(U)\right]$

Limitations of CPF



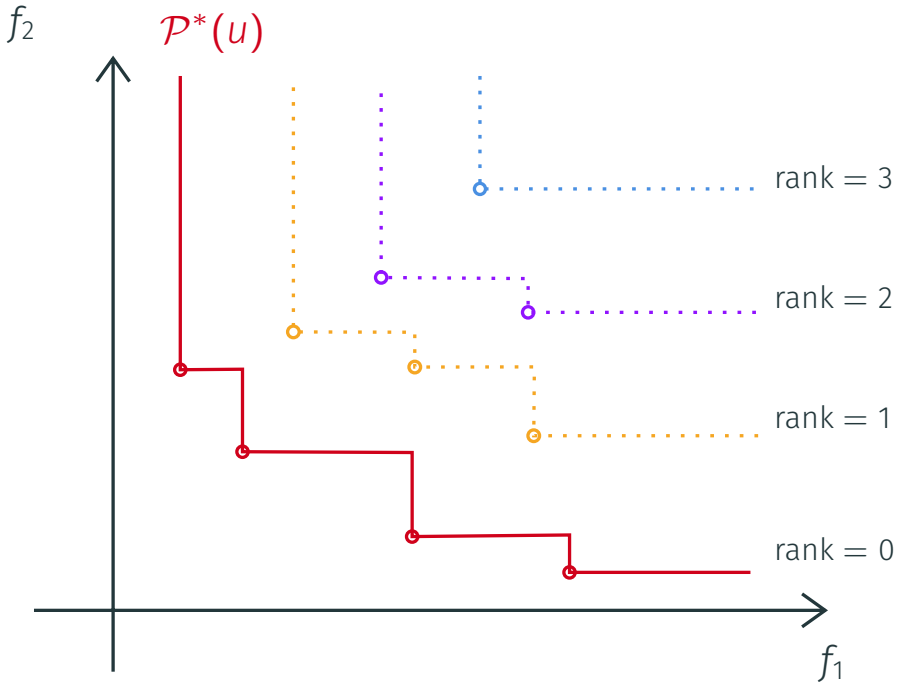
But

- No control on the probability of optimality: what if no x has a "high" probability of being optimal
- If $x \notin \mathcal{P}^*_{\mathcal{X}}(u)$, is it still a reasonable solution or not ?

How to quantify the proximity to the CPF, while keeping some notion of "incommensurability" (don't compare apples to oranges) ?

Non-dominated Sorting

Let $u \in \mathcal{U}$, let us say that we evaluated $\mathbf{f}(\mathfrak{X}, u)$. At the heart of NSGA-II ([DPAM02])



- $r(x, u) =$ rank of $\mathbf{f}(x, u)$ among $\mathbf{f}(\mathfrak{X} \cup \{x\}, u)$
- $r(x, u) = 0 \iff x \in \mathcal{P}^*_{\mathcal{X}}(u)$
- Look for points which have low $\mathbb{E}_U[r(x, U)]$?

Proximity indicator ρ

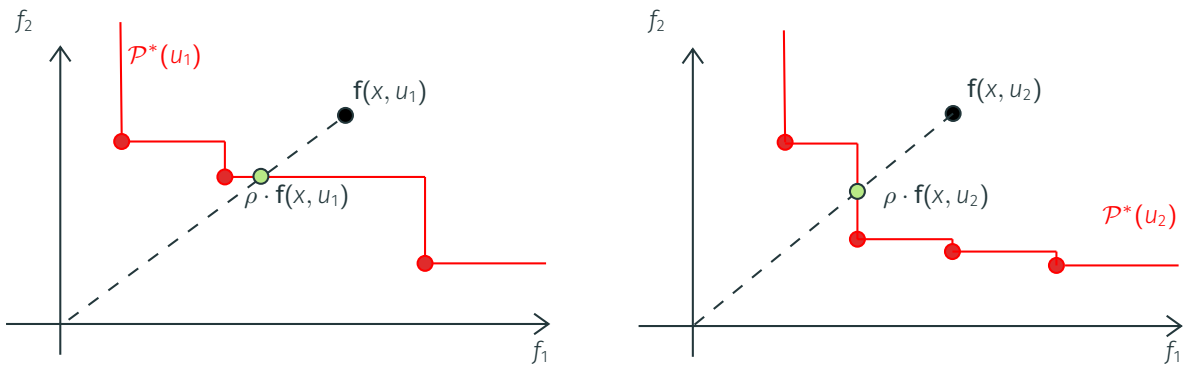
Let us assume that $\forall x, u$, $\mathbf{f}(x, u) > 0$, and that $\mathbf{0}$ is a "nice" ideal point

Definition of ρ

Let $x \in \mathcal{X}, u \in \mathcal{U}$

$$\text{largest } \rho \text{ such that } \rho \cdot \mathbf{f}(x, u) \preceq \mathcal{P}^*(u)$$

- $0 < \rho \leq 1$
- $\rho = 1$ means $\mathbf{f}(x, u)$ not dominated thus $x \in \mathcal{P}^*_{\mathcal{X}}(u)$
- Invariant to linear scaling of one or more objectives



We can define $(x, u) \mapsto \rho(x, u)$.

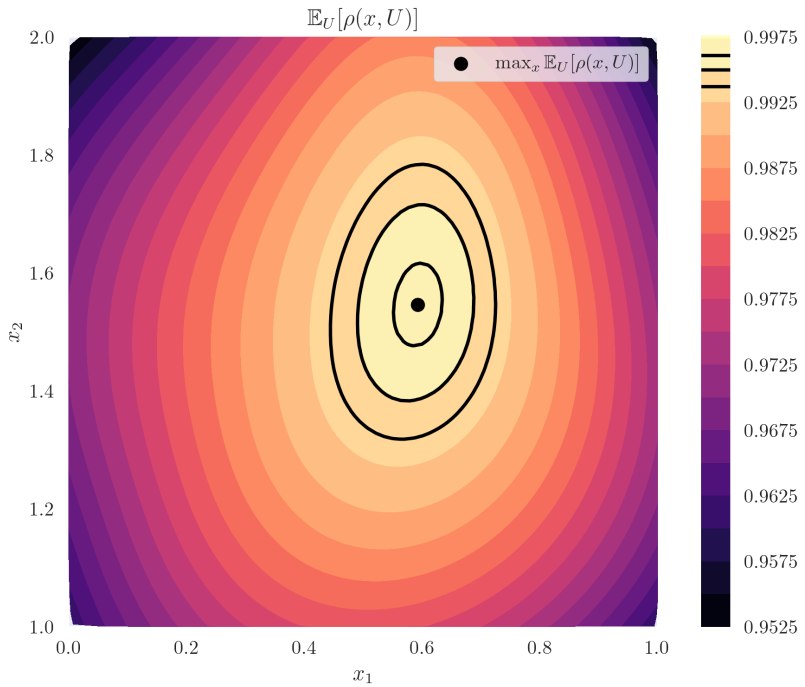
$\rightarrow$ Maximize average proximity ?

ρ-based designs

Definition of ρ

$$\rho(x, u) = \sup_{\rho} \{\rho \cdot \mathbf{f}(x, u) \preceq \mathcal{P}^*(u)\} = \ell(\mathbf{f}(x, u), \mathcal{P}^*(u))$$

We are interested in high values of $\mathbb{E}_U[\rho(x, U)]$



- $\max \mathbb{E}_U[\rho(x, U)]$
- get top k% of values

Estimation of ρ using GP

Definition of ρ

$$\rho(x,u) = \sup_{\rho} \{\rho \cdot \mathbf{f}(x,u) \preceq \mathcal{P}^*(u)\} = \ell(\mathbf{f}(x,u), \mathcal{P}^*(u))$$

ρ can not be modelled directly as a GP using $\mathbf{F}$.

Estimation of $\rho(x,u)$ using GP samples:

- Samples of the GP at $\underbrace{\mathbf{F}_\omega(x,u)}_{\text{for } \mathbf{f}(x,u)}, \underbrace{\mathbf{F}_\omega(\xi_1,u),\ldots,\mathbf{F}_\omega(\xi_N,u)}_{\text{for } \mathcal{P}^*_\omega(u)}$ to get $\rho_\omega(x,u)$

Estimation of $\mathbb{E}_U[\rho(x,U)]$ using GP samples and samples of U

- Using a finite set of samples $\{u^{(i)}\}_i$, we can approach $\mathbb{E}_U[\rho_\omega(x,U)]$ as $\hat{\rho}_\omega(x)$

$\Rightarrow$ using enough GP samples, and doing this (hopefully) in parallel, we can have a prediction + measure of uncertainty $\rightarrow$ Bayesian Optimization

Optimization of ρ ?

Definition of ρ

$$\rho(x,u) = \sup_{\rho} \{\rho \cdot \mathbf{f}(x,u) \preceq \mathcal{P}^*(u)\} = \ell(\mathbf{f}(x,u), \underbrace{\mathcal{P}^*(u)}_{\text{does not depend on } x})$$

Uncertainty on $U \neq$ Uncertainty ω from the GP

To optimize $\mathbb{E}_U[\rho(x,U)]$ using GP: we decouple choice of x and u

- Construct and optimize UCB acquisition function using GP samples, to get x_{target} (where we aim at reducing the uncertainty)
- ... but lookahead methods in the joint space on $\mathbb{V}\text{ar}_\omega[\mathbb{E}_U[\rho_\omega(x,U)]]$ is too expensive
- ... reducing uncertainty at x_{target} does not necessarily means evaluating at some $\mathbf{f}(x_{\text{target}}, \tilde{u})$

  $\rightarrow$ uncertainty might come from the estimation of $\mathcal{P}^*(u)$.

Comparison between MOOUU methods

$\mathbb{E}_U[\mathbf{f}(x, U)]$

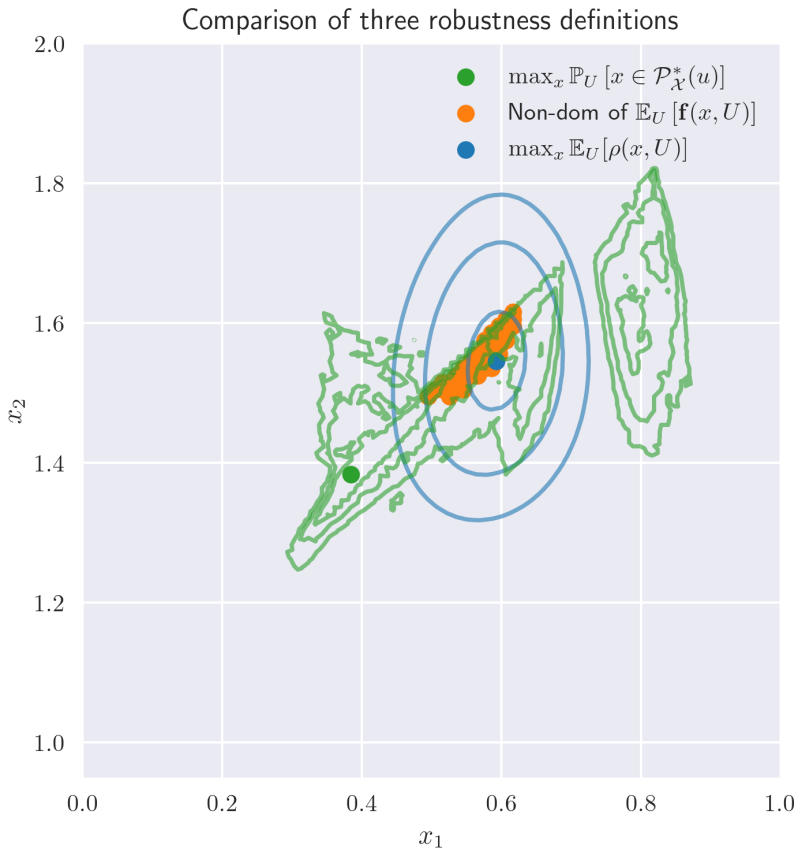
- Mean performances
- Ease of implementation
- Information on marginals, not on dependence

$\mathbb{P}_U[x \in \mathcal{P}^*_{\mathcal{X}}(u)]$

- Probability of being non-dominated
- PEHVI
- No control on performance outside of this optimality

$\mathbb{E}_U[\rho(x, U)]$ / Average NDS Rank

- Average "proximity" with $\mathcal{P}^*(u)$
- More expensive to estimate
- Tricky acquisition procedure



Wrapping-up/Take-Home messages

What does it mean to be robust in a MO setting ?

- What does it mean to be robust in my problem ?
- Comparing multivariate random variables is not trivial
- Conditional Pareto Front: Solve then Robustify strategy
- Profile-EHVI helps improve the estimation of $u \mapsto \mathcal{P}^*(u)$
- ρ is more challenging, and acquisition function is not easy to construct

What is next ?

- Continue numerical experiments
- Metrics to compare acquisitions functions
- Application to EnergyPlus

Three steps optimization

- Identify the x_{target} where we want to reduce the variance of the estimation (UCB)

$$x_{\text{target}} = \arg\max_{x \in \mathcal{X}} \hat{\rho}(x) + \beta \hat{v}(x)^{1/2}$$

- Identify the env. var. which brings the most uncertainty at x_{target}

$$u_{\text{target}} = \arg\max_{u \in \mathcal{U}} \mathbb{V}\text{ar}_{\omega} \left[\rho_{\omega}(x_{\text{target}}, u) \right]$$

  Find which MOO problem (dependent on u) carries the most uncertainty on the estimation of $\hat{\rho}(x_{\text{target}})$

- Minimize lookahead variance

$$x_{n+1} = \arg\min_{x \in \mathcal{X}} \mathbb{V}\text{ar}_{\omega} \left[\rho^{(n+1)}_{\omega,(x,u_{\text{target}})}(x_{\text{target}}, u_{\text{target}}) \right]$$

- Finally, the next point to evaluate is $(x_{n+1}, u_{n+1}) \leftarrow (x_{n+1}, u_{\text{target}})$

References i

Samuel Daulton, Sait Cakmak, Maximilian Balandat, Michael A. Osborne, Enlu Zhou, and Eytan Bakshy.
Robust Multi-Objective Bayesian Optimization Under Input Noise, June 2022.

Kalyanmoy Deb and Himanshu Gupta.
Searching for Robust Pareto-Optimal Solutions in Multi-objective Optimization.
In Carlos A. Coello Coello, Arturo Hernández Aguirre, and Eckart Zitzler, editors, *Evolutionary Multi-Criterion Optimization*, pages 150–164, Berlin, Heidelberg, 2005. Springer.

K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan.
A fast and elitist multiobjective genetic algorithm: NSGA-II.
Trans. Evol. Comp, 6(2):182–197, April 2002.

References ii


Reda El Amri, Rodolphe Le Riche, Céline Helbert, Christophette Blanchet-Scalliet, and Sébastien Da Veiga.
A sampling criterion for constrained Bayesian optimization with uncertainties, December 2023.


David Ginsbourger, Jean Baccou, Clément Chevalier, Frédéric Perales, Nicolas Garland, and Yann Monerie.
Bayesian Adaptive Reconstruction of Profile Optima and Optimizers.
SIAM/ASA Journal on Uncertainty Quantification, 2(1):490–510, January 2014.


Daniel Hernández-Lobato, José Miguel Hernández-Lobato, Amar Shah, and Ryan P. Adams.
Predictive Entropy Search for Multi-objective Bayesian Optimization, February 2016.


Donald R Jones and Matthias Schonlau.
Efficient Global Optimization of Expensive Black-Box Functions.
1998.

References iii

Victor Picheny.
Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction, October 2013.

Jixiang Qing, Ivo Couckuyt, and Tom Dhaene.
A robust multi-objective Bayesian optimization framework considering input uncertainty.
Journal of Global Optimization, 86(3):693–711, July 2023.

M. Rivier and P.M. Congedo.
Surrogate-Assisted Bounding-Box approach applied to constrained multi-objective optimisation under uncertainty.
Reliability Engineering & System Safety, 217:108039, January 2022.

Ben Tu, Nikolas Kantas, Robert M. Lee, and Behrang Shafei.
Scalarisation-based risk concepts for robust multi-objective optimisation, May 2024.