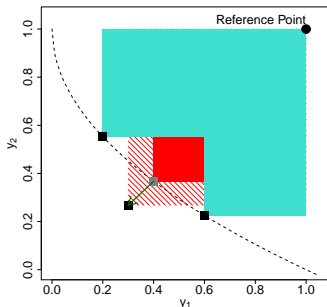
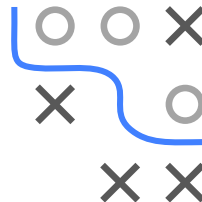
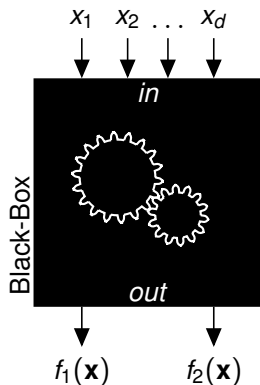


Multicriteria Bayesian Optimization



- Multicriteria Optimization
- Taxonomy
- ParEGO, SMS-EGO, EHI

MULTICRITERIA BAYESIAN OPTIMIZATION



$$f : \mathcal{S} \rightarrow \mathbb{R}^m$$

$$\min_{\mathbf{x} \in \mathcal{S}} \quad f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$$

- A configuration \mathbf{x} **dominates** (\prec) $\tilde{\mathbf{x}}$ if

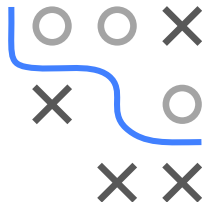
$$\forall i \in \{1, \dots, m\} : \quad f_i(\mathbf{x}) \leq f_i(\tilde{\mathbf{x}})$$

$$\text{and } \exists j \in \{1, \dots, m\} : \quad f_j(\mathbf{x}) < f_j(\tilde{\mathbf{x}})$$

- Set of non-dominated solutions:

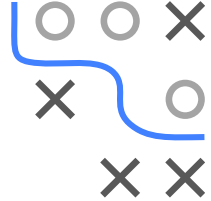
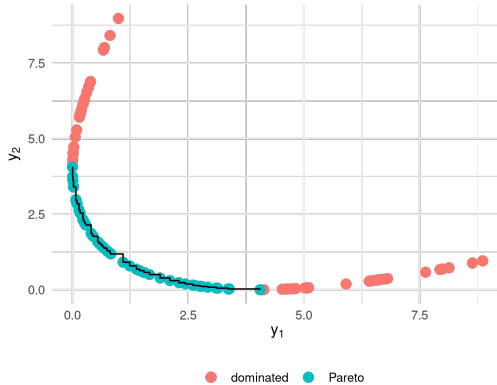
$$\mathcal{P} := \{\mathbf{x} \in \mathcal{S} \mid \nexists \tilde{\mathbf{x}} \in \mathcal{S} : \tilde{\mathbf{x}} \prec \mathbf{x}\}$$

- Pareto set \mathcal{P} , Pareto front $\mathcal{F} = f(\mathcal{P})$
- Goal: Find $\hat{\mathcal{P}}$ of non-dominated points that estimates the true Pareto set \mathcal{P}



MULTICRITERIA BAYESIAN OPTIMIZATION / 2

Example Pareto front:



MULTICRITERIA BAYESIAN OPTIMIZATION / 3

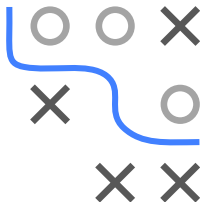
The most popular quality indicator is the hypervolume indicator (also called dominated hypervolume or \mathcal{S} -metric).

The hypervolume, HV, of an approximation of the Pareto front $\hat{\mathcal{F}} = f(\hat{\mathcal{P}})$ can be defined as the combined volume of the dominated hypercubes domHC_r of all solution points $\mathbf{x} \in \hat{\mathcal{P}}$ regarding a reference point \mathbf{r} , i.e.,

$$\text{HV}_r(\hat{\mathcal{P}}) := \mu \left(\bigcup_{\mathbf{x} \in \hat{\mathcal{P}}} \text{domHC}_r(\mathbf{x}) \right)$$

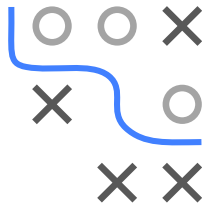
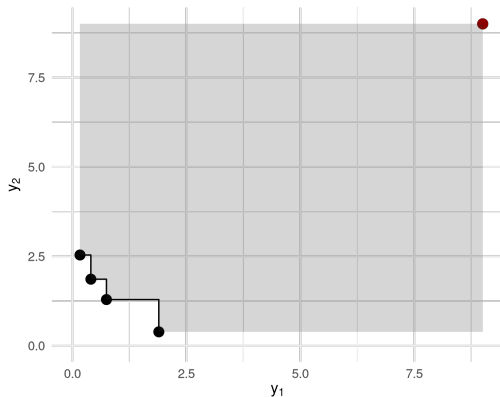
where μ is the Lebesgue measure and the dominated hypercube is given as:

$$\text{domHC}_r(\mathbf{x}) := \{ \mathbf{u} \in \mathbb{R}^m \mid f_i(\mathbf{x}) \leq u_i \leq r_i \ \forall i \in \{1, \dots, m\} \}$$



MULTICRITERIA BAYESIAN OPTIMIZATION / 4

Hypervolume example:



Reference point \mathbf{r} in red, estimated Pareto front $\hat{\mathcal{F}}$ in black,
corresponding $\text{HV}_{\mathbf{r}}(\hat{\mathcal{F}})$ is given by the grey area

```
graph TD
    subgraph Initial_design [Initial design]
        A[Arbitrary DOE  
(e. g. random,  
Sobol, grid)]
        B[Latin Hypercube  
Sampling]
    end

    subgraph Model_fitting [Model fitting]
        C[Individual models for  
each objective]
        D[Scalarize objectives]
        E[Model of scalarization]
    end

    subgraph Candidate_generation [Candidate generation]
        F[Multi-obj.  
optimization of  
infill criterion  
on each model]
        G[Multiple single-  
obj. optim. of  
(aggregating)  
infill criterion]
        H[Single-obj.  
optimization of  
(aggregating)  
infill criterion]
    end

    subgraph Candidate_selection [Candidate selection]
        I[Select design  
point(s) with  
regard to  
other criterion]
        J[Select design  
point(s) with  
regard to  
infill criterion]
    end

    subgraph Stopping_decision [Stopping decision]
        K[Decide  
termination  
based on  
approximation]
        L[Decide  
termination  
based on  
total budget]
    end

    A --> C
    A --> D
    B --> E
    C --> F
    D --> G
    E --> H
    F --> I
    G --> I
    G --> J
    H --> J
    I --> K
    J --> L
    K --> L
    L --> Update[Update design]
    Update --> A
```

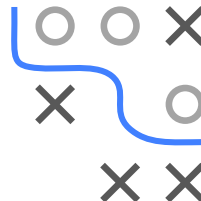
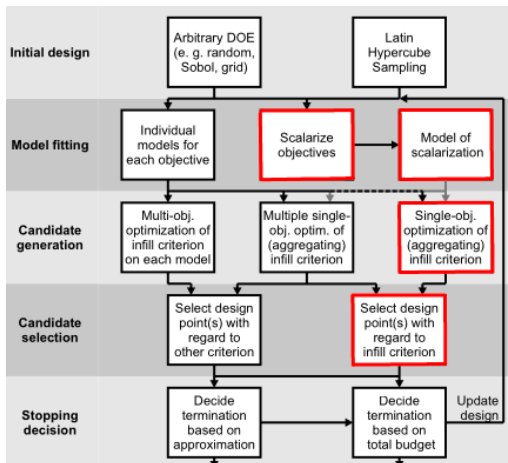
The flowchart illustrates the sequential design process for multi-objective optimization, organized into five main stages:

- Initial design:** Starts with two options: "Arbitrary DOE (e. g. random, Sobol, grid)" and "Latin Hypercube Sampling".
- Model fitting:** The process branches into three parallel paths:
 - "Individual models for each objective" (from Arbitrary DOE)
 - "Scalarize objectives" (from Arbitrary DOE)
 - "Model of scalarization" (from Latin Hypercube Sampling)
- Candidate generation:** Each model fitting path leads to a candidate generation step:
 - "Multi-obj. optimization of infill criterion on each model" (from Individual models)
 - "Multiple single-obj. optim. of (aggregating) infill criterion" (from Scalarize objectives)
 - "Single-obj. optimization of (aggregating) infill criterion" (from Model of scalarization)
- Candidate selection:** The candidate generation steps lead to selection:
 - "Select design point(s) with regard to other criterion" (from Multi-obj. optimization)
 - "Select design point(s) with regard to infill criterion" (from both Multiple single-obj. optim. and Single-obj. optimization)
- Stopping decision:** The selection steps lead to a decision point:
 - "Decide termination based on approximation" (from Select design point(s) with regard to other criterion)
 - "Decide termination based on total budget" (from Select design point(s) with regard to infill criterion)

Both stopping decisions lead to an "Update design" step, which loops back to the "Initial design" stage.

A 3x3 grid with a blue path starting at the top-left cell (0,0) and ending at the bottom-right cell (2,2). The path consists of the following cells: (0,0), (0,1), (1,1), (1,2), and (2,2). The cells (0,2), (1,0), and (2,0) are marked with 'X', while the other cells are empty.

PAREGO



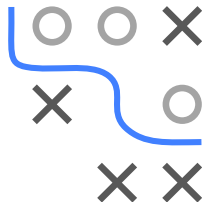
PAREGO / 2

- 1 Scalarize standardized objectives using the augmented Tchebycheff norm

$$\max_{i \in \{1, \dots, m\}} w_i f_i(\mathbf{x}) + \rho \sum_{i=1}^m w_i f_i(\mathbf{x})$$

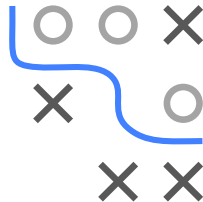
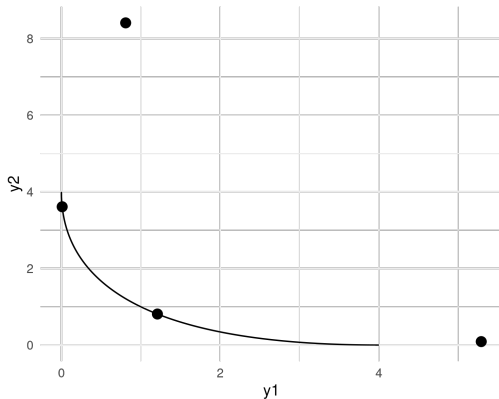
with weight vector \mathbf{w} drawn uniformly from the set of evenly distributed weight vectors \mathcal{W}

- 2 Fit SM on the scalarized objective function
- 3 Proceed to use any standard single-objective acquisition function (EI, PI, LCB, ...)



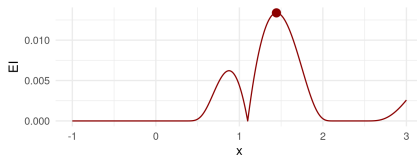
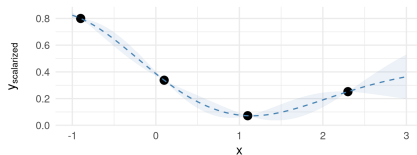
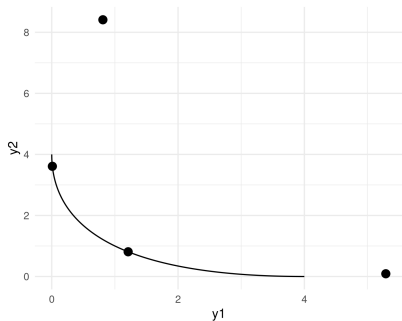
PAREGO / 3

ParEGO Example, initial design and true Pareto front in black ...



PAREGO / 4

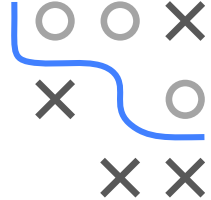
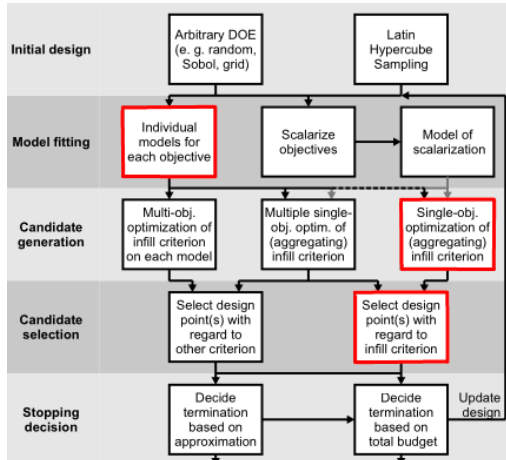
... standardize objectives, obtain scalarized objective via augmented Tchebycheff norm, fit SM and optimize EI ...





Optimization in Machine Learning – 10 / 13

SMS-EGO

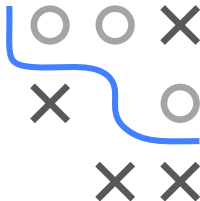
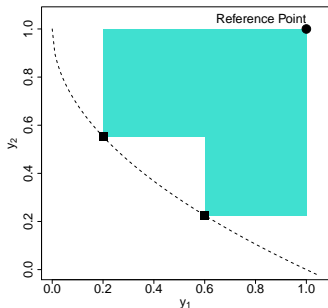


SMS-EGO

Individual models for each objective f_i

Single-objective optimization of aggregating acquisition function:
Calculate contribution of the confidence bound of candidate to the current front approximation

- Calculate LCB for each objective
- Measure contribution with regard to the hypervolume improvement
- For ε -dominated (\prec_ε) solutions, a penalty is added

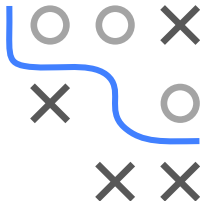
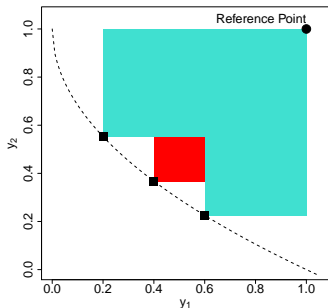


SMS-EGO

Individual models for each objective f_i

Single-objective optimization of aggregating acquisition function:
Calculate contribution of the confidence bound of candidate to the current front approximation

- Calculate LCB for each objective
- Measure contribution with regard to the hypervolume improvement
- For ε -dominated (\prec_ε) solutions, a penalty is added

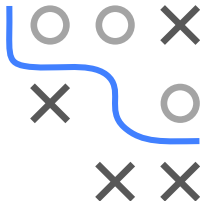
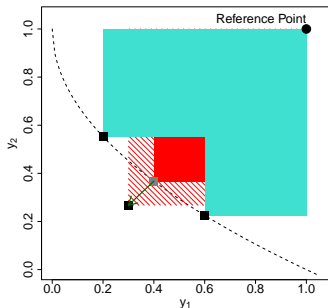


SMS-EGO

Individual models for each objective f_j

Single-objective optimization of aggregating acquisition function:
Calculate contribution of the confidence bound of candidate to the
current front approximation

- Calculate LCB for each objective
- Measure contribution with regard to the hypervolume improvement
- For ε -dominated (\prec_ε) solutions, a penalty is added



OUTLOOK

Many more options exist:

- Expected Hypervolume improvement
- Multi-Ego
- Entropy based: PESMO, MESMO
- ...

