Removing nonoptimal points for D-optimal exact designs

Radoslav Harman and Samuel Rosa

Comenius University Bratislava, Slovakia

mODa14, June 2026

Removing points that cannot support an optimal design: only for *approximate* designs:

- D-optimality: Harman and Pronzato (2007)[1]
- Φ_p-optimality: Pronzato (2013)[2]
- E-optimality: Harman and Rosa (2019)[3]
- c-, A-optimality: Pronzato and Sagnol (2021)[4]

Now: for D-optimal *exact* designs

- Harman and Rosa (2026)[5]

---

[1] Harman and Pronzato (2007). Improvements on removing nonoptimal support points in D-optimum design algorithms. *Statist. Probab. Lett.*

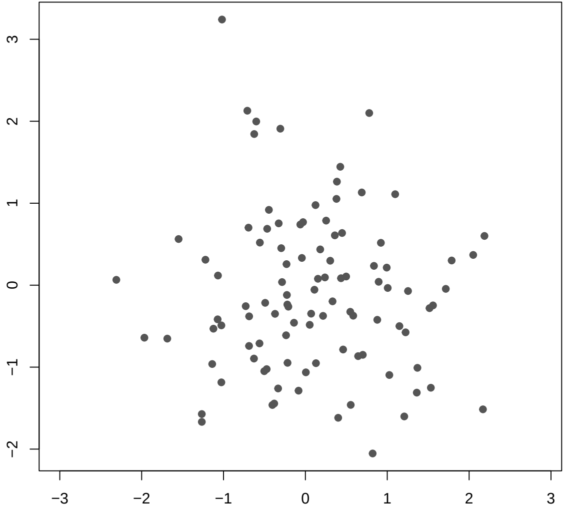
[2] Pronzato (2013). A delimitation of the support of optimal designs for Kiefers ϕ_p-class of criteria, *Statist. Probab. Lett.*

[3] Harman and Rosa (2019). Removal of the points that do not support an E-optimal experimental design, *Statist. Probab. Lett.*

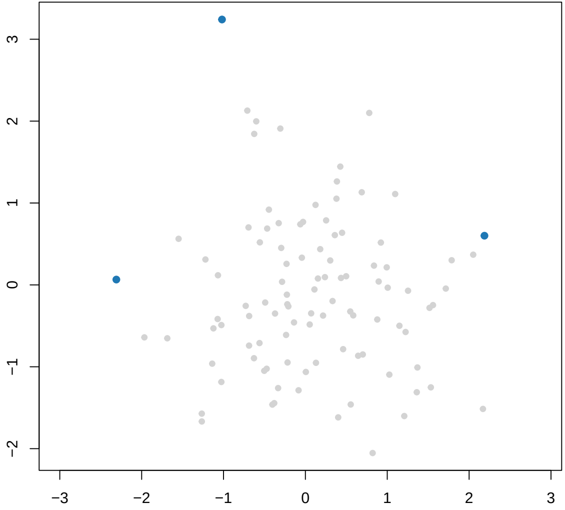
[4] Pronzato and Sagnol (2021). Removing inessential points in c-and A-optimal design, *JSPI*

[5] Harman and Rosa (2026). Removal of redundant candidate points for the exact D-optimal design problem, *Statist. and Comput.*

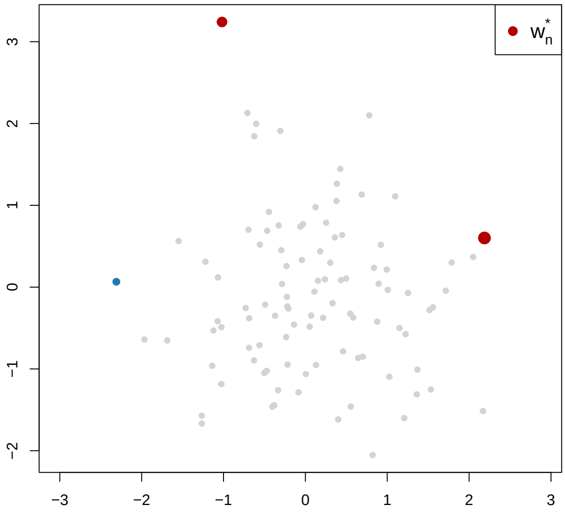
$$y = \mathbf{f}^\top(\mathbf{x})\beta + \varepsilon, \quad n = 5 \text{ trials}$$



$$y = \mathbf{f}^{\top}(\mathbf{x})\beta + \varepsilon, \quad n = 5 \text{ trials}$$



$$y = \mathbf{f}^\top(\mathbf{x})\beta + \varepsilon, \quad n = 5 \text{ trials}$$



Design

- $\mathbf{x} \in \mathfrak{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ ... design space
- Exact design for n trials:

$$\mathbf{w} = (w_1, \ldots, w_N)^T, \quad \sum_{i=1}^{N} w_i = 1, \quad nw_i \in \mathbb{N}_0 \, \forall i$$

  nw_i: replications
- Example: $N = 7, n = 5$

$$\mathbf{w} = (\frac{1}{5}, 0, 0, 0, \frac{2}{5}, \frac{1}{5}, 0)$$

  replications: $1, 0, 0, 0, 2, 1, 0$

- Exact design for n trials:

$$\mathbf{w} = (w_1, \ldots, w_N)^T, \quad \sum_{i=1}^{N} w_i = 1, \quad n w_i \in \mathbb{N}_0 \, \forall i$$

$n w_i$: replications

- Approximate design:

$$\mathbf{w} = (w_1, \ldots, w_N)^T, \quad \sum_{i=1}^{N} w_i = 1, \quad w_i \geq 0 \, \forall i$$

$n w_i$: approximate replications

- Design space $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$
- Known vectors $\{\mathbf{f}_1, \ldots, \mathbf{f}_N\} \subseteq \mathbb{R}^m$
- Information matrix:

$$\mathbf{M}(\mathbf{w}) := \sum_{i=1}^{N} w_i \mathbf{f}_i \mathbf{f}_i^\top$$

- Linear model $y = \mathbf{f}^\top(\mathbf{x})\beta + \varepsilon$: $\quad \mathbf{f}_i = \mathbf{f}(\mathbf{x}_i)$
- D-optimality: $\Phi(\mathbf{M}(\mathbf{w})) = (\det(\mathbf{M}(\mathbf{w})))^{1/m} \to \max$
- Set of all exact designs: Ξ_n
- Set of all D-optimal exact designs: Ξ_n^*

- Support of $\mathbf{w}$: $\mathcal{S}(\mathbf{w}) = \{i \,|\, w_i > 0\}$

Objective:

Find conditions that must be satisfied for any $\ell \in \{1, \ldots, N\}$ that supports any D-optimal exact design:

$$\ell \in \mathcal{S}_n^* := \cup_{\mathbf{w}_n^* \in \Xi_n^*} \mathcal{S}(\mathbf{w}_n^*)$$

Augmentation condition

Idea:

- $\mathbf{w}_n^+$: an exact design

$\Rightarrow$ $\ell \in \mathcal{S}_n^*$ must satisfy

$$\max_{\mathbf{w} \in \Xi_n, w_\ell \geq 1/n} \Phi(\mathbf{M}(\mathbf{w})) \geq \Phi(\mathbf{M}(\mathbf{w}_n^+)) \tag{1}$$

Augmentation condition

Idea:

- $\mathbf{w}_n^+$: an exact design

$\Rightarrow$ $\ell \in \mathcal{S}_n^*$ must satisfy

$$\max_{\mathbf{w} \in \Xi_n, w_\ell \geq 1/n} \Phi(\mathbf{M}(\mathbf{w})) \geq \Phi(\mathbf{M}(\mathbf{w}_n^+)) \tag{1}$$

- Difficult $\Rightarrow$ upper bounds on LHS of (1)
- Algebraic methods in Branch and Bound (weak for $w_\ell \geq 1/n$)
- Convex relaxation of LHS (slow)

Augmentation condition

Inputs:

- any exact design $\mathbf{w}_n^+$ with nonsingular $\mathbf{M}(\mathbf{w}_n^+)$
  ... exact design heuristics
- a D-optimal *approximate* design with information matrix $\mathbf{M}_*$
  ... approximate design algorithms

Augmentation condition

Inputs:

- any exact design $\mathbf{w}_n^+$ with nonsingular $\mathbf{M}(\mathbf{w}_n^+)$
  ... exact design heuristics
- a D-optimal *approximate* design with information matrix $\mathbf{M}_*$
  ... approximate design algorithms

Theorem 1. If $\ell \in \mathcal{S}_n^*$, then

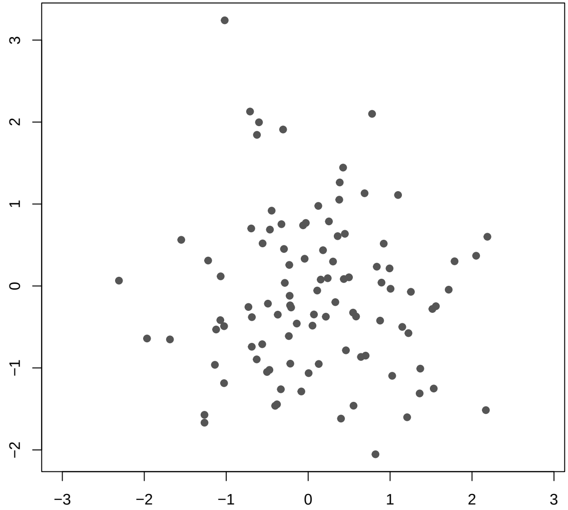
$$\mathbf{f}_\ell^\top \mathbf{M}_*^{-1} \mathbf{f}_\ell \geq mn \left(\mathrm{eff}_\infty(\mathbf{w}_n^+) - \frac{n-1}{n} \right),$$

where

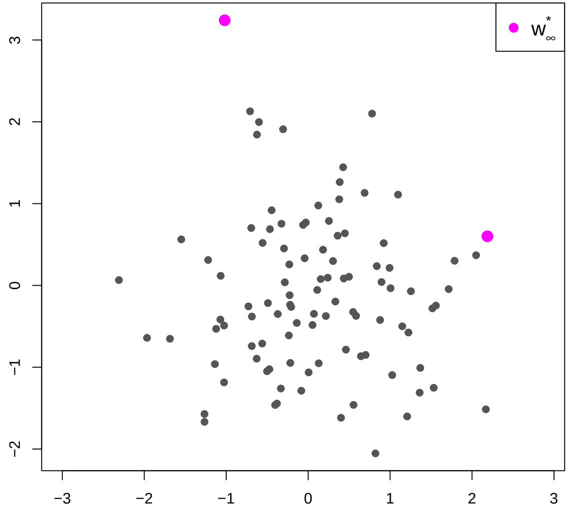
$$\mathrm{eff}_\infty(\mathbf{w}) := \frac{\Phi(\mathbf{M}(\mathbf{w}))}{\Phi(\mathbf{M}_*)}.$$

Possible use

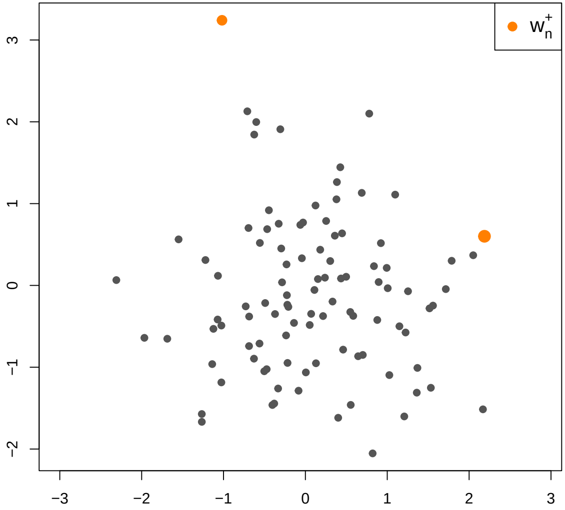
1. Compute optimal approximate design $\mathbf{w}^*_\infty$ (e.g. `od_REX` from `OptimalDesign` package)
2. Compute efficient/optimal exact design $\mathbf{w}^+_n$ *on the support of* $\mathbf{w}^*_\infty$ (e.g. `od_KL`, `od_MISOCP`)
3. Remove points using Theorem 1
4. Compute efficient/optimal exact design on the remaining points



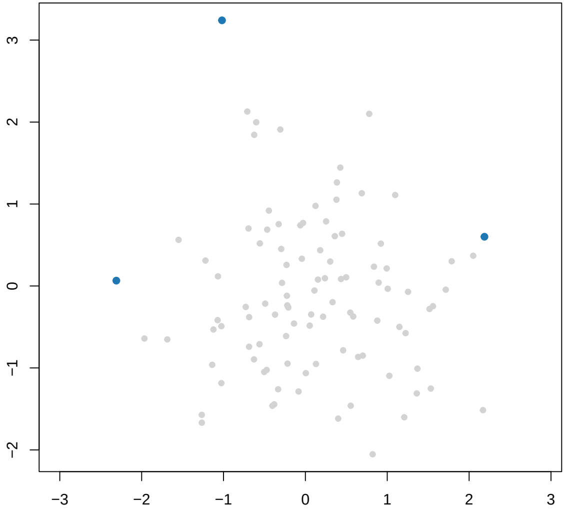
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3

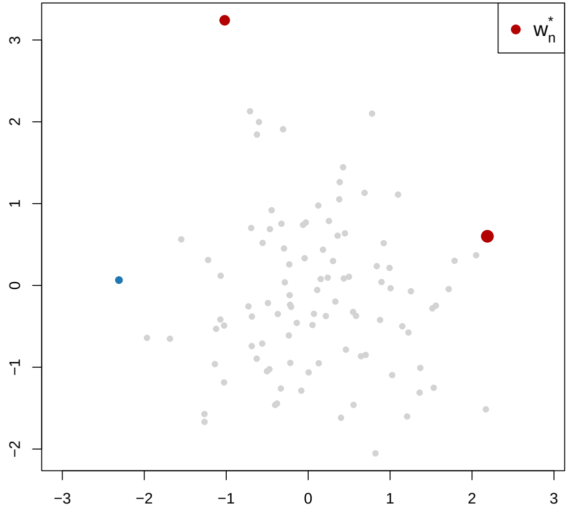


w^*_∞
−3
−2
−1
0
1
2
3



W_n^+
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3


3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3



W_n^*
−3
−2
−1
0
1
2
3

Corollary.
There exists a natural number n_0 such that $\mathcal{S}_n^* \subseteq \{i : \mathbf{f}_i^\top \mathbf{M}_*^{-1} \mathbf{f}_i = m\}$ for all $n \geq n_0$.

Exchange condition

- **Idea:** if $\ell \in \mathcal{S}(\mathbf{w}_n^*)$, then no exchange of $\mathbf{x}_\ell$ for another $\mathbf{x}_i$ can improve Φ:

$$\Phi\left(\mathbf{M}(\mathbf{w}_n^*) - \frac{1}{n}\mathbf{f}_\ell\mathbf{f}_\ell^\top + \frac{1}{n}\mathbf{f}_i\mathbf{f}_i^\top\right) \leq \Phi(\mathbf{M}(\mathbf{w}_n^*)) \quad \forall i$$

- Reformulation and bounds ...

Inputs:

- any exact design $\mathbf{w}_n^+$ with nonsingular $\mathbf{M}(\mathbf{w}_n^+)$
- a D-optimal *approximate* design with information matrix $\mathbf{M}_*$

Theorem 2. If $\ell \in \mathcal{S}_n^*$, then $d_+^{1/m} \leq \frac{1}{m} t_\ell$ and

$$\frac{\Delta(\mathbf{s}_i, \mathbf{s}_\ell)}{\overline{g}(1, t_\ell, d_+)} - \frac{\Delta(\mathbf{s}_\ell, \mathbf{s}_i)}{\underline{g}(1, t_\ell, d_+)} \leq \frac{\|\mathbf{s}_i\|^2 \|\mathbf{s}_\ell\|^2 - (\mathbf{s}_i^\top \mathbf{s}_\ell)^2}{n \underline{g}^2(2, t_\ell, d_+)} \quad \forall i,$$

Inputs:

- any exact design $\mathbf{w}_n^+$ with nonsingular $\mathbf{M}(\mathbf{w}_n^+)$
- a D-optimal *approximate* design with information matrix $\mathbf{M}_*$

Theorem 2. If $\ell \in \mathcal{S}_n^*$, then $d_+^{1/m} \leq \frac{1}{m} t_\ell$ and

$$\frac{\Delta(\mathbf{s}_i, \mathbf{s}_\ell)}{\overline{g}(1, t_\ell, d_+)} - \frac{\Delta(\mathbf{s}_\ell, \mathbf{s}_i)}{\underline{g}(1, t_\ell, d_+)} \leq \frac{\|\mathbf{s}_i\|^2 \|\mathbf{s}_\ell\|^2 - (\mathbf{s}_i^\top \mathbf{s}_\ell)^2}{n \underline{g}^2(2, t_\ell, d_+)} \quad \forall i,$$

where

$$\mathbf{s}_i = \mathbf{M}_*^{-1/2} \mathbf{f}_i \, \forall i, \quad d_+ := (\mathrm{eff}_\infty(\mathbf{w}_n^+))^m, \quad t_\ell := \frac{n-1}{n} m + \frac{1}{n} \mathbf{f}_\ell^\top \mathbf{M}_*^{-1} \mathbf{f}_\ell$$

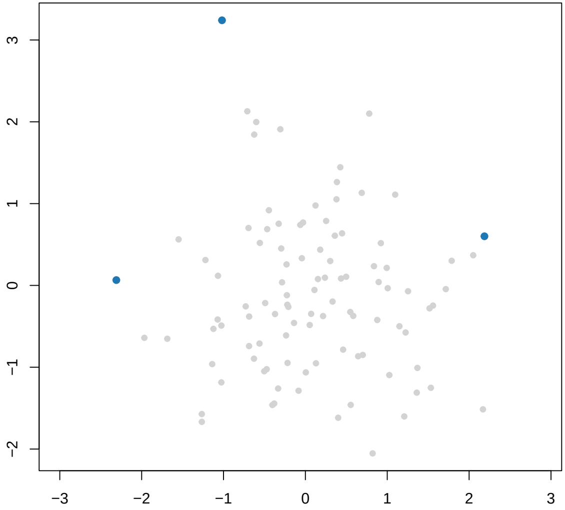
$$\Delta(\mathbf{v}, \mathbf{z}) := \frac{1}{2} \left(\|\mathbf{v} + \mathbf{z}\| \|\mathbf{v} - \mathbf{z}\| + \|\mathbf{v}\|^2 - \|\mathbf{z}\|^2 \right),$$

$\underline{g}(k, t_\ell, d_+) \in [0, \frac{t_\ell}{m}]$ and $\overline{g}(k, t_\ell, d_+) \in [\frac{t_\ell}{m}, \frac{t_\ell}{k}]$ are the solutions of $R_{k, t_\ell}(g) = d_+^{1/m}$ for

$$R_{k, t_\ell}(g) := \left[g^k \left(\frac{t_\ell - kg}{m-k} \right)^{m-k} \right]^{1/m}.$$

Possible use

1. Compute optimal approximate design $\mathbf{w}^*_\infty$ (e.g. `od_REX` from `OptimalDesign` package)
2. Compute efficient/optimal exact design *on the support of* $\mathbf{w}^*_\infty$ (e.g. `od_KL`, `od_MISOCP`)
3. Remove points using Theorem 1
4. **Among the remaining points, remove points using Theorem 2**
5. Compute efficient/optimal exact design on the remaining points



3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3

Deletion conditions depend on $\mathbf{w}_n^+$ only through $\mathrm{eff}_\infty(\mathbf{w}_n^+)$:

- Theorem 1:

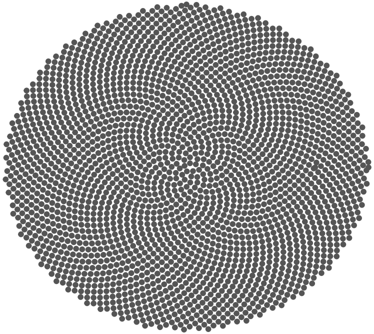
$$\mathbf{f}_\ell^\top \mathbf{M}_*^{-1} \mathbf{f}_\ell \geq mn \left(\mathrm{eff}_\infty(\mathbf{w}_n^+) - \frac{n-1}{n} \right),$$

- Theorem 2:

$$\frac{\Delta(\mathbf{s}_i, \mathbf{s}_\ell)}{\overline{g}(1, t_\ell, d_+)} - \frac{\Delta(\mathbf{s}_\ell, \mathbf{s}_i)}{\underline{g}(1, t_\ell, d_+)} \leq \frac{\|\mathbf{s}_i\|^2 \|\mathbf{s}_\ell\|^2 - (\mathbf{s}_i^\top \mathbf{s}_\ell)^2}{n \underline{g}^2(2, t_\ell, d_+)} \quad \forall i,$$

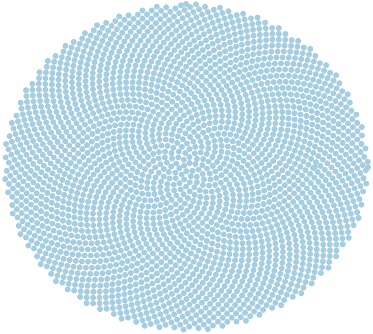
Toy example 2:

- $m = 2$, $n = 9$, $N = 2500$
- assume $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.8$



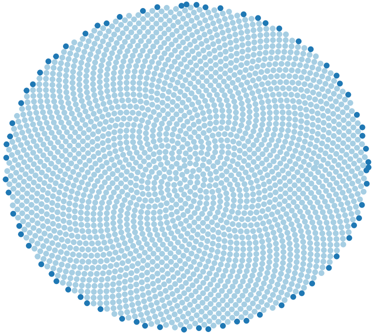
Toy example 2:

- $m = 2$, $n = 9$, $N = 2500$
- assume $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.8$
- light blue: remaining after Augmentation removal



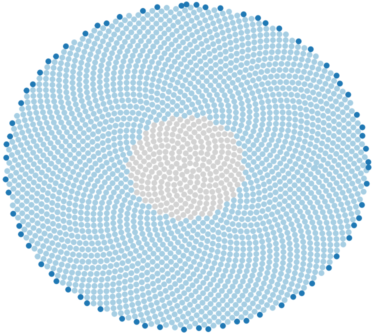
Toy example 2:

- $m = 2,\ n = 9,\ N = 2500$
- assume $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.8$
- light blue: remaining after Augmentation removal
- dark blue: remaining after Exchange removal



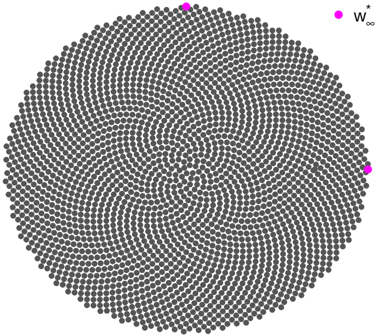
Toy example 2:

- $m = 2,\ n = 9,\ N = 2500$
- assume $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.9$
- light blue: remaining after Augmentation removal
- dark blue: remaining after Exchange removal



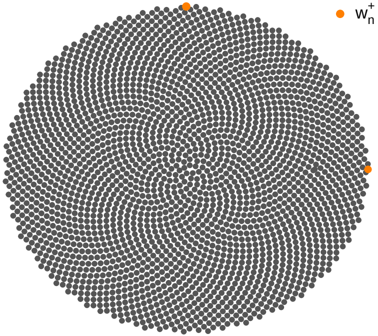
Toy example 2:

- $m = 2,\ n = 9,\ N = 2500$



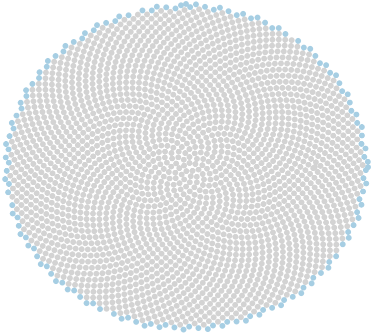
Toy example 2:

- $m = 2,\ n = 9,\ N = 2500$
- $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.994$ (D-optimal design on the support of $\mathbf{w}_\infty^*$)



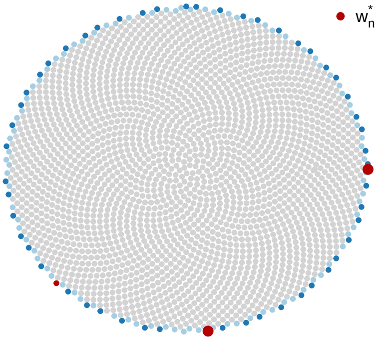
Toy example 2:

- $m = 2$, $n = 9$, $N = 2500$
- $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.994$ (D-optimal design on the support of $\mathbf{w}_\infty^*$)
- light blue: remaining after Augmentation removal



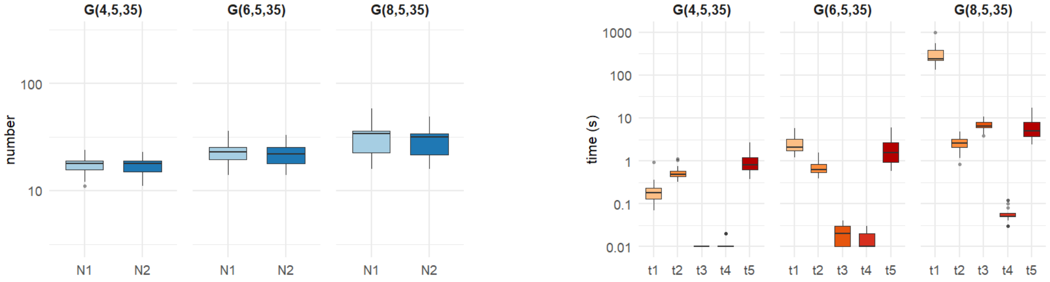
Toy example 2:

- $m = 2, \ n = 9, \ N = 2500$
- $\mathrm{eff}_\infty(\mathbf{w}_9^+) = 0.994$ (D-optimal design on the support of $\mathbf{w}_\infty^*$)
- light blue: remaining after Augmentation removal
- dark blue: remaining after Exchange removal



Random regressors

20 times: $N = 10^c$ generated vectors $\mathbf{f}_i \sim \mathcal{N}(\mathbf{0}_m, \mathbf{I}_m)$, n trials … $G(c, m, n)$



- N_1, N_2: number of points remaining after Theorem 1, Theorem 2
- Steps: **1.** $\mathbf{w}^*_\infty$ (`od_REX`), **2.** $\mathbf{w}^+_n$ (`od_MISOCP` on $\mathcal{S}(\mathbf{w}^*_\infty)$), **3.** Theorem 1, **4.** Theorem 2, **5.** $\mathbf{w}^*_n$ (`od_MISOCP`)

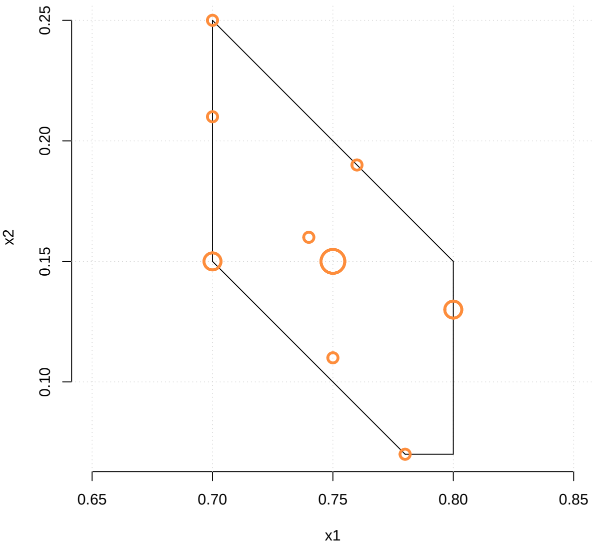
Mixture model

Quadratic Scheffé model used in Paucar-Menacho et al. (2025)[6]:

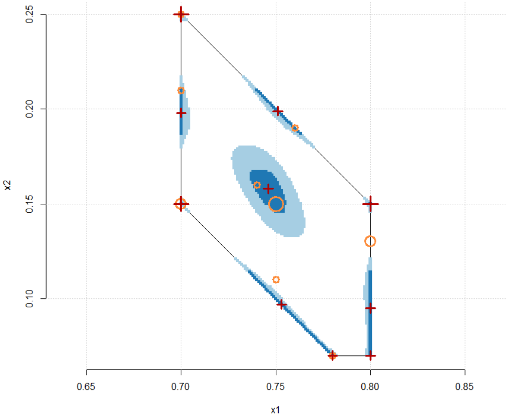
$$\mathfrak{X} = \{(x_1, x_2, x_3)^\top \in \mathbb{R}^3 \,|\, x_1 + x_2 + x_3 = 1, \\ x_1 \in [0.7, 0.8], x_2 \in [0.07, 0.25], x_3 \in [0.05, 0.15]\},$$

- each x_i discretized to three decimal places $\Rightarrow N = 9991 \approx 10^4$
- $\mathbf{f}_i = (x_{i1}, x_{i2}, x_{i3}, x_{i1}x_{i2}, x_{i1}x_{i3}, x_{i2}x_{i3})^\top \in \mathbb{R}^6$
- $n = 13$

[6] Paucar-Menacho L. M. et al. (2025). Optimization of a craft ale-type beer enriched with cañihua malt (Chenopodium pallidicaule) and banana passionfruit juice (Passiflora tripartita var. mollisima). *Heliyon 11*


0.25
0.20
0.15
0.10
x2
0.65
0.70
0.75
0.80
0.85
x1

- $N \approx 10^4$
- Augmentation removal: $N_1 = 1\,644$ remaining (light blue), 1.55 s
- Exchange removal: $N_2 = 390$ remaining (dark blue), 0.49 s
- red: final design (MISOCP for 60 s)



- each x_i discretized to four decimal places $\Rightarrow N = 981901 \approx 10^6$
- t_2: time of Exchange removal; eff: $\mathrm{eff}_\infty(\mathbf{w}_n^+)$