Optimal experimental design under linear constraints

A column generation framework for Kiefer's criteria

Selin D. Ahipaşaoğlu

mODa14, Drongen Abbey, Belgium 2026

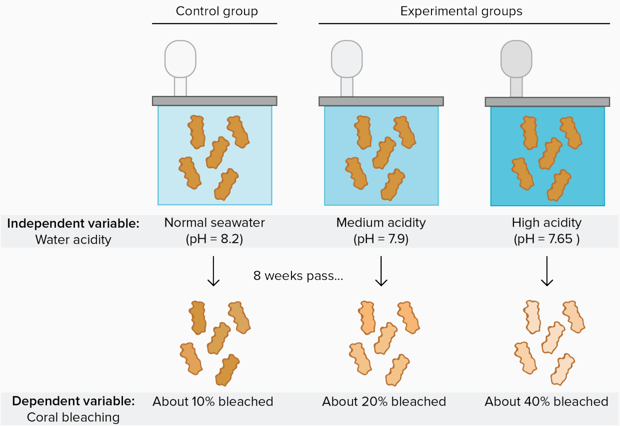
Mathematical Sciences, University of Southampton

Outline

- Optimal Experimental Design
  - Motivation and problem definition
  - Formulations (exact and approximate optimal designs, side constraints)
- Continuous relaxation
  - Duality, optimality conditions, geometric interpretation
  - Column generation
- Recent and on-going work

Motivation: Why Optimal Experimental Design?

Goal: Optimal Experimental Design (OED) aims to maximize information gained from samples by choosing them carefully, often before the output is observed.



Motivation: Why Optimal Experimental Design?

Goal: Maximize information gained from samples by choosing them carefully.

- Experiments are costly (clinical trials, engineering, A/B testing, simulation optimization). OED can decrease the number of experiments or maximize the information gain from a fixed number of experiments.
- OED can identify an optimal subsample (with or without repetitions) from a large dataset after experimentation.
- Often there are constraints such as budget, interactions, group designs
- Nonlinear models dominate (dose–response, logistic regression, growth curves).
- Linearized models play a key part in finding nonlinear designs.

Optimal Experimental Design

Model: $y \sim f(x;\theta)$ where $x, \theta \in \mathbb{R}^n$.

$$\xi = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ n_1 & n_2 & \cdots & n_m \end{bmatrix}$$

- x_i: design points (experimental setting)
- n_i: number of replicates (or allocation weight) at x_i
- $\sum_{i=1}^{m} n_i = N$ (total sample size)

Fisher information under design ξ

$$M(\xi;\theta) = \sum_{i=1}^{m} \frac{n_i}{N} M(x_i, \theta)$$

Choose ξ to maximize information about θ, i.e., to create a “large" information matrix.

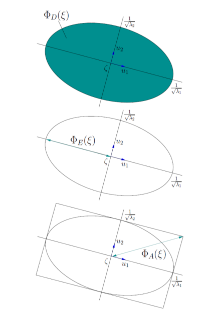
Kiefer's ϕ_p Criteria

$$\phi_p(M) = \begin{cases} \left(\frac{1}{n}\text{Trace } M^p\right)^{1/p} & \text{for} \quad p \neq 0, \pm\infty; \\ (\det M)^{1/n} & \text{for} \quad p = 0; \\ \lambda_{\min} & \text{for} \quad p = -\infty. \end{cases}$$

$\phi_p(M)$ is an information function for all $p \leq 1$. Special cases:

- **D-optimality**: maximize $\det(M)$ (minimize generalized variance).
- **A-optimality**: minimize $\mathrm{tr}(M^{-1})$ (average variance).
- **E-optimality**: maximize smallest eigenvalue (worst-case control).

Geometric view - D: volume, A: average axis length, E: shortest axis.



Exact Optimal Experimental Design

Let $u_i := \frac{n_i}{N}$ and $M(u) := \sum_i u_i x_i x_i^T$.

The exact optimal design problem can be formulated as:

$$\begin{array}{rrcll} \max & g_p(u) & := & \ln \phi_p(M(u)) & \\ s.t. & \sum_{i=1}^m u_i & = & 1, & \\ & u_i & \geq & 0, & \forall i \in [m], \\ & u_i N & \in & \mathbb{Z}, & \forall i \in [m]. \end{array}$$

NP-hard for D-optimality (Welch; 1982) and A-optimality (Nikolov et al.; 2019)

Cardinality constraints:

- No repeated points: $u_i N \in \{0, 1\}$
- Repetition allowed: $|\mathrm{supp}(u)| \leq \kappa$

Optimal Experimental Design

Practical experiments often impose additional linear restrictions:

- **Budget / capacity :** $c^\top u \leq \bar{c}$
- **Stratified allocation:**
  $\sum_{i \in G_k} u_i = \pi_k$
- **Lower and upper bounds:**
  $\alpha_i \leq u_{i_0} \leq \beta_i$
  (min/max obs, B&B
  subproblems)
- **Symmetry:** $u_i = u_j$

Optimal Experimental Design

Practical experiments often impose additional linear restrictions:

- **Budget / capacity :** $c^\top u \le \bar{c}$
- **Stratified allocation:** $\sum_{i \in G_k} u_i = \pi_k$
- **Lower and upper bounds:** $\alpha_i \le u_{i_0} \le \beta_i$ (min/max obs, B&B subproblems)
- **Symmetry:** $u_i = u_j$

We collect these into the polytope:

$$\mathcal{P} = \{u \ge 0 : \mathbf{1}^\top u = 1,\ Au \le b,\ Cu = d\},$$

$$A \in \mathbf{R}^{r \times m},\ \ C \in \mathbf{R}^{\ell \times m}.$$

Prior work
Special cases treated in isolation.
A unified approach for arbitrary $\mathcal{P}$ and Φ_p family for $p \le 1$ can be developed.

Exact Optimal Designs with Polyhedral Constraints

$$
\begin{array}{lrcll}
\max & g_p(u) & := & \ln \phi_p(M(u)) & \\
s.t. & \sum_{i=1}^{m} u_i & = & 1, & \\
 & u_i & \geq & 0, \quad \forall i \in [m], & \qquad (D_p^{\text{Exact}}) \\
 & Au & \leq & b & \\
 & Cu & = & d & \\
 & u_i N & \in & \mathbb{Z}, \quad \forall i \in [m]. &
\end{array}
$$

Sagnol (2013) showed that ϕ_p has conic representation for rational $p \leq 1$.

Example:

$$\Phi_A(M) := \left(\operatorname{trace} M^{-1}\right)^{-1}$$

Semidefinite representation of Φ_A:

$$\Phi_A(M) \geq t \iff \exists\, Y \in \mathbb{S}_m : \ \operatorname{trace} Y \leq t \quad \text{and} \quad \begin{bmatrix} Y & tI \\ tI & M \end{bmatrix} \succeq 0.$$

SDP representation

A-optimality SDP:

$$\begin{aligned}
\max_{\mathbf{u},t,Y} \quad & t \\
\text{s.t.} \quad & \operatorname{trace} Y \leq t \\
& \begin{bmatrix} Y & tI \\ tI & M(\mathbf{u}) \end{bmatrix} \succeq 0 \\
& \sum_i u_i = 1, \quad \mathbf{u} \geq \mathbf{0}, \\
& Au \leq b, \quad Cu = d, \\
& u_i N \in \mathbb{Z}, \quad \forall i \in [m].
\end{aligned}$$

Solving Exact Optimal Design Problem - Literature

Exact methods:

1. **Custom Branch-and-Bound Algorithms:**
   1.1 Welch (1982): simple BaB, very small instances (no side constraints)
   1.2 Ahipasaoglu (2021): customized BaB with FW (box constraints)
   1.3 Hendrych et al. (2023): BaB with FW for general purpose nonlinear problems (box constraints)
   1.4 Ponte et al.(2025): customized BaB with FW, variable tightening, and spectral bounds (box constraints, D-opt only)
2. **Commercial Solvers:**
   2.1 Sagnol and Harman (2015): Mixed Integer Second Order Conic Program reformulations by for $p = 0$ and $p = -1$. (polyhedral constraints)

Custom BaB is typically much faster for moderately sized instances.

They all fail for large m!

Solving Exact Optimal Design Problem - Literature

Heuristics and approximations:

1. **Local search:** Rounding to integer solutions based on the convex relaxation or 2-exchange algorithms.
   1.1 Federov (1972), Pukelsheim (1993), etc.
   1.2 Wang et al. (2016); Allen-Zhu et al. (2017); Singh and Xie (2018); Nikolov et al. (2019); Madan et al. (2019); Ahipasaoglu et al. (2025)
2. **Quadratic Approximations:** Quadratic Integer programming to solve an approximation for $p = 0$ and $p = -1$. (Harman and Filova, 2014; Filova and Harman, 2019)

Convex Subproblems: The Primal Problem (D_p)

$$\max_{u \in \mathbb{R}^m} \; g_p(u) := \ln \Phi_p(M(u)) \quad \text{s.t.} \quad \mathbf{1}^\top u = 1, \quad u \geq 0, \quad Au \leq b, \quad Cu = d. \quad (D_p)$$

Assumption (Relaxed Slater condition)
$\exists\, \bar{u} \geq 0$ with $\mathbf{1}^\top \bar{u} = 1$, $A\bar{u} \leq b$, $C\bar{u} = d$, and $M(\bar{u}) \succ 0$.

(D_p) can be solved as an SDP for rational p (Sagnol, 2013), size of SDP varies with p.

Challenge: In applications, m can reach 10^7: beyond SDP solvers.

Key observation: There is a sparse optimal solution (independent of m).

The Dual Problem $(D_p^\star)$

Let $f_q(H) := -\ln \Phi_q(H)$, $\omega_i(H) := x_i^\top H x_i$, and $\frac{1}{p} + \frac{1}{q} = 1$.

$$\min_{H \succeq 0,\, \lambda \geq 0,\, \nu \geq 0,\, \eta} \quad f_q(H) \quad \text{s.t.} \quad \omega_i(H) + \lambda_i - (A^\top \nu)_i - (C^\top \eta)_i = n - \nu^\top b - \eta^\top d, \quad i \in [m]. \tag{$D_p^\star$}$$

Dual variables:

- $H \in \mathbb{S}^n$: dual matrix $(n \times n)$
- $\lambda_i \geq 0$: multiplier for $u_i \geq 0$
- $\nu \geq 0$: multiplier for $Au \leq b$
- η free: multiplier for $Cu = d$

- $O(n^2 + r + \ell)$ free variables.
- $|\operatorname{supp}(u^\star)| \leq \frac{n(n+1)}{2} + 1 + \ell + r$ (Carathéodory).
- $\max(D_p) = \min(D_p^\star)$ (Strong duality)

Set $n^\star := n - {\nu^\star}^\top b - {\eta^\star}^\top d$.

Theorem (Necessary and sufficient conditions)
$u^\star$ and $(H^\star, \lambda^\star, \nu^\star, \eta^\star)$ are simultaneously optimal if and only if:

(i) **Hölder equality:** $H^\star = \dfrac{M(u^\star)^{p-1}}{\Phi_p(M(u^\star))^p}$.

Optimality Conditions

Set $n^\star := n - {\nu^\star}^\top b - {\eta^\star}^\top d$.

Theorem (Necessary and sufficient conditions)
$u^\star$ and $(H^\star, \lambda^\star, \nu^\star, \eta^\star)$ are simultaneously optimal if and only if:

(i) **Hölder equality:** $\quad H^\star = \dfrac{M(u^\star)^{p-1}}{\Phi_p(M(u^\star))^p}.$

(ii) **Complementary slackness:** $\quad \lambda_i^\star u_i^\star = 0 \ \forall\, i, \quad {\nu^\star}^\top (b - Au^\star) = 0.$

Optimality Conditions

Set $n^\star := n - {\nu^\star}^\top b - {\eta^\star}^\top d$.

Theorem (Necessary and sufficient conditions)
$u^\star$ and $(H^\star, \lambda^\star, \nu^\star, \eta^\star)$ are simultaneously optimal if and only if:

(i) **Hölder equality:** $H^\star = \dfrac{M(u^\star)^{p-1}}{\Phi_p(M(u^\star))^p}$.

(ii) **Complementary slackness:** $\lambda_i^\star u_i^\star = 0 \; \forall\, i, \quad {\nu^\star}^\top (b - Au^\star) = 0$.

(iii) **Point-shifted containment cuts:** $\forall\, i \in [m]$,

$$\omega_i(H^\star) \;\leq\; n^\star + (A^\top \nu^\star)_i + (C^\top \eta^\star)_i,$$

with equality when $i \in \operatorname{supp}(u^\star)$.

Generalization of **Kiefer's equivalence theorem** for polyhedral constraints.

Geometric Interpretation: Point-Shifted Ellipsoids

For each candidate x_i, define the point-shifted ellipsoid

$$\mathcal{E}_i^\star \; := \; \Big\{ x \in \mathbf{R}^n : x^\top H^\star x \; \le \; n^\star + (A^\top \nu^\star)_i + (C^\top \eta^\star)_i \Big\}.$$

Optimality conditions:

- Every x_i lies inside or on $\mathcal{E}_i^\star$.
- Support points $x_i \in \operatorname{supp}(u^\star)$ lie on $\partial \mathcal{E}_i^\star$.

No additional constraints:
When $\nu^\star = \eta^\star = 0$: all ellipsoids collapse to

$$\mathcal{E}^\star = \{ x : x^\top H^\star x \le n \}.$$

At $p = 0$: exactly the **MVEE characterisation** of D-optimality.

Running Example: Stratified Linear Regression

Model. $y_j = \theta_1 + \theta_2 t_{i_j} + \varepsilon_j$; $x(t) = (1, t)^\top \in \mathbf{R}^2$, $n = 2$.

Candidates. $m = 4$ points at $t \in \{-1, -\frac{1}{2}, \frac{1}{2}, 1\}$:

$$x_1 = (1, -1)^\top, \quad x_2 = (1, -\tfrac{1}{2})^\top, \quad x_3 = (1, \tfrac{1}{2})^\top, \quad x_4 = (1, 1)^\top.$$

Constraint. Strata $G_1 = \{x_1, x_2\}$ (left, $t \leq 0$) and $G_2 = \{x_3, x_4\}$ (right, $t > 0$) with prescribed allocations $u_1 + u_2 = 0.4$, $u_3 + u_4 = 0.6$. Encoded as $Cu = d$ with $C = (1, 1, 0, 0)$, $d = 0.4$; no inequalities.

Criterion. D-optimality ($p = 0$): maximise $\frac{1}{2} \ln \det M(u)$.

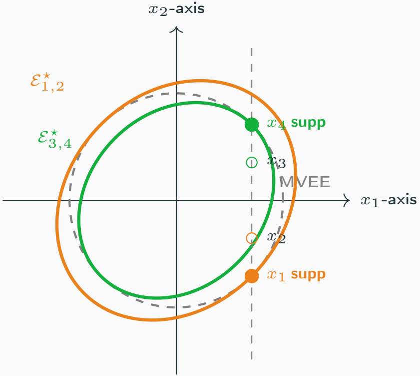
Cost of stratification

Unconstrained D-optimal: $u^\star = (0.5, 0, 0, 0.5)$, $\det M = 1$.

Stratification forces $u_1 \leq 0.4$, yielding $\det M^\star = 0.96$ at the constrained optimum.

Running Example: Point-Shifted Ellipsoids

Constrained D-optimum: $u^\star = (0.4, 0, 0, 0.6), \quad \eta^\star = \frac{5}{6}, \quad n^\star = \frac{5}{3}.$



Left stratum G_1 (brown):

$(C^\top \eta^\star)_i = +\frac{5}{6}$, threshold $= \frac{5}{3} + \frac{5}{6} = 2.5.$

$\omega_1(H^\star) = 2.5$ on $\partial\mathcal{E}^\star_1$ ✓

$\omega_2(H^\star) \approx 1.51$ inside ✓

Right stratum G_2 (green):

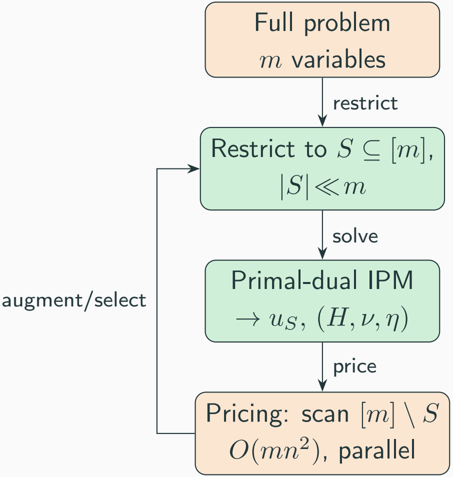
$(C^\top \eta^\star)_i = 0$, threshold $= \frac{5}{3}$.

$\omega_4(H^\star) = \frac{5}{3}$ on $\partial\mathcal{E}^\star_4$ ✓

$\omega_3(H^\star) \approx 1.09$ inside ✓

The dashed circle is the unconstrained MVEE. The equality multiplier $\eta^\star > 0$ shifts the left-stratum ellipsoid outward.

Column Generation: Exploiting Sparsity



1. Maintain active set $S \subseteq [m]$, $|S| \ll m$.
2. Solve the restricted primal-dual pair on S via a primal-dual IPM for u_S and (H, ν, η) simultaneously.
3. Price $i \notin S$; identify violated dual constraints.
4. **If** no violations: optimality certificate. **Else** augment S (multi-cut) and repeat.

Benefit

IPM runtime is independent of m and in our control. Pricing is $O(m(n^2 + r + l))$ for all, but can be parallelized or done partially.

Restricted Masters and the Separation Oracle

Restrict the m dual feasibility constraints to $i \in S$: the restricted master $(D^\star_{p,S})$.

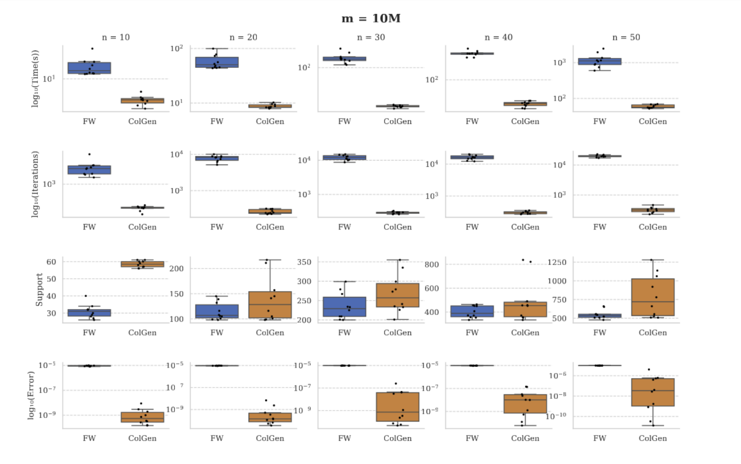
Lemma (Separation oracle)
*Let $(H, \lambda_S, \nu, \eta)$ be optimal for $(D^\star_{p,S})$. Define the **violation***

$$\boxed{\delta_i(H,\nu,\eta) \ := \ \omega_i(H) - (A^\top \nu)_i - (C^\top \eta)_i - n^\star}, \qquad n^\star := n - \nu^\top b - \eta^\top d.$$

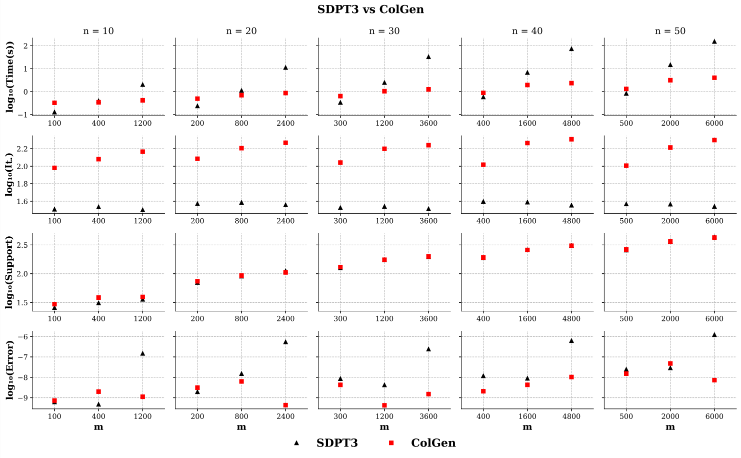
- *If $\delta_i \leq 0$ for all $i \notin S$: set $\lambda_i := -\delta_i \geq 0$; this extends to an optimal solution of $(D^\star_p)$.*
- *Otherwise: any i with $\delta_i > 0$ can be added to S.*

Pricing. Factor $H = LL^\top$; set $y_i = L^\top x_i$. Then $\omega_i(H) = \|y_i\|^2$, computed in $O(n^2)$.

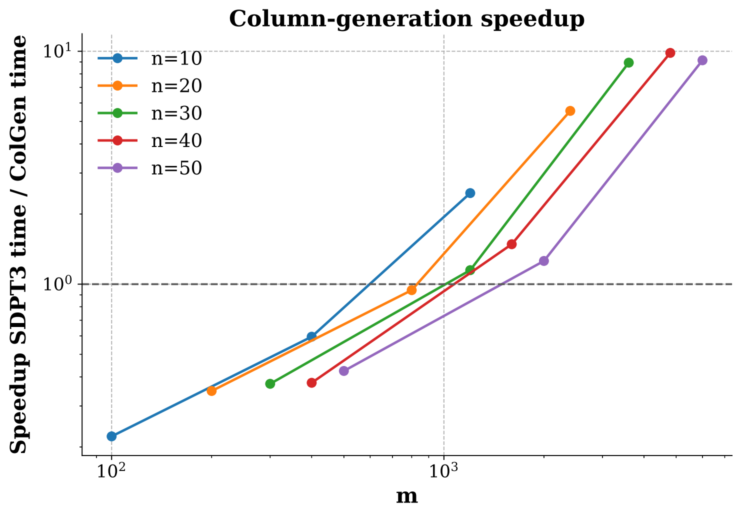
Experiments I - no extra constraints



Experiments I - no extra constraints



Experiments II - with inequality constraints



Special case: D-optimality

S. D. Ahipaşaoğlu, S. Cipolla, J. Gondzio. A column generation approach to exact experimental design, Math. Prog. Part C (2026, in press). https://arxiv.org/abs/2507.03210

- Exact D-optimal design with no further constraints
- Column Generation and Interior-Point SDP solvers for relaxation
- Restricted Local Search with Provable Bounds
- Extensive Numerical Validation on Synthetic and Real-World Data.
- Similar methods can be applied to other optimal design criteria, especially A-optimal exact design.

Remark: Similar work by Rosa and Harman, 2022; Harman and Rosa, 2026.

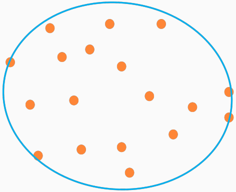
Future/on-going work

- Testing the CG framework for several criteria and polyhedral constraints for problems with large m
- Incorporating the CG framework to branch-cut-price algorithms to solve exact problems
- General design spaces with large or uncountable number of elements (instead of $\{x_1, ...x_m\}$): pricing problem can be used to 'generate' columns (design vectors), similar to the "cutting stock problem"

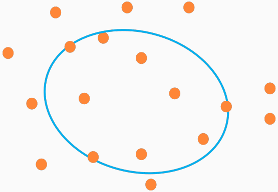
Remark: See also recent work by Filová, Somogyi, Harman, 2026; Pilai et.al, 2025.

Related problems in robust statistics

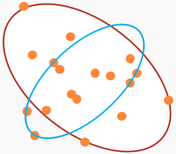
MVEE



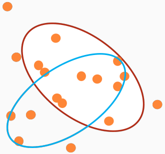
MVE estimator



k-*MVEEs*



k-*MVEEs-outliers*



THANK YOU!

References

- S. D. Ahipaşaoğlu, A branch-and-bound algorithm for the exact optimal experimental design problem, Statistics and Computing, 31, 1–11 (2021)
- S. D. Ahipaşaoğlu, A First-Order Algorithm for the A-Optimal Experimental Design Problem: A Mathematical Programming Approach, Statistics and Computing, 25, 1113–1127 (2015)
- S. D. Ahipaşaoğlu and P. Sun and M. J. Todd, Linear convergence of a Modified Frank-Wolfe algorithm for computing minimum-volume enclosing ellipsoids, Optimization Methods and Software, 23, 5–19 (2008)
- V. V. Fedorov, Theory of Optimal Experiments. Academic Press, New York (1972)
- L. Filova, P. Somogyi, and R. Harman, Optimal Exact Designs of Multiresponse Experiments Under Linear and Sparsity Constraints, 2026 (Arxiv)

References

- L. Filova and R. Harman, Ascent with Quadratic Assistance for the Construction of Exact Experimental Designs, Computational Statistics 35 (2), 775-801 (2020)
- R. Harman and L. Filova, Computing efficient exact designs of experiments using integer quadratic programming, Computational Statistics & Data Analysis, 71, pp. 1159–1167 (2014)
- R. Harman and S. Rosa, Removal of redundant candidate points for the exact D-optimal design problem, Statistics and Computing 36 (1), 56(2026)
- R. Harman and S. Rosa, Mixed-integer linear programming for computing optimal experimental designs, J. of Statistical Planning and Inference 234, 106200 (2025)
- G. Ponte, M. Fampa, and J. Lee, Branch-and-bound for integer D-optimality with fast local search and variable-bound tightening, Mathematical Programming, 2025.

References

- A. Pillai, G. Ponte, M. Fampa, J. Lee, M. Singh, and W. Xie, Computing Experiment-Constrained D-Optimal Designs, 2025 Proceedings of the Conference on Applied and Computational Discrete Algorithms (ACDA).
- F. Pukelsheim, Optimal Design of Experiments. John Wiley and Sons, NY (1993)
- S. Rosa and R. Harman, Computing minimum-volume enclosing ellipsoids for large datasets, Computational Statistics Data Analysis 171, 107452 (2022)
- G. Sagnol, R. Harman R, Computing exact D-optimal designs by mixed integer second-order cone programming, The Annals of Statistics, 43, 2198–2224 (2015)
- W. J. Welch, Algorithmic Complexity: Three NP-hard problems in computational statistics, Journal of Statistical Computation and Simulation, 15, 17–25 (1982)
- W. J. Welch, Branch-and-Bound search for experimental designs based on D optimality and other criteria, Technometrics, 24, 41–48 (1982)