# *Advances in Orthogonal Minimally Aliased Response Surface (OMARS) Designs*

José Núnez Ares and Peter Goos

KU LEUVEN

EFFEX

APEX workshop 2023
4/10/23

# *Summary of this talk*

OMARS designs

How to choose an experimental design

- Example 1: choosing an optimization design
- Example 2: choosing a screening design
- Example 3: choosing a screening design in blocks

Break

EFFEX: software interactive demo

Discussion

# *Response surface designs*

- Experimental plans used in product and process optimization.
- Involves the study of several quantitative factors
- The estimation of a complete second-order response surface is often the goal:
    - Main effects
    - Interaction effects
    - Quadratic effects
- Best-known designs are:
    - (Small) Central Composite Design
    - Box-Behnken design

# *3-level screening designs*

- Experimental plans used in product and process optimization.
- Involves the study of several quantitative factors
- The estimation of a partial second-order response surface is often the goal:

  Main effects

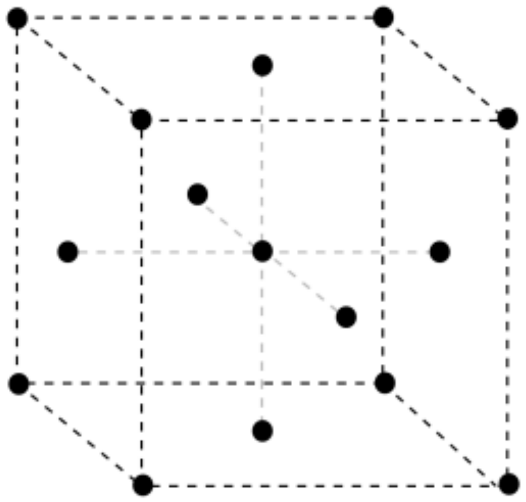
  Some interaction effects

  Some quadratic effects
- Best-known designs is:

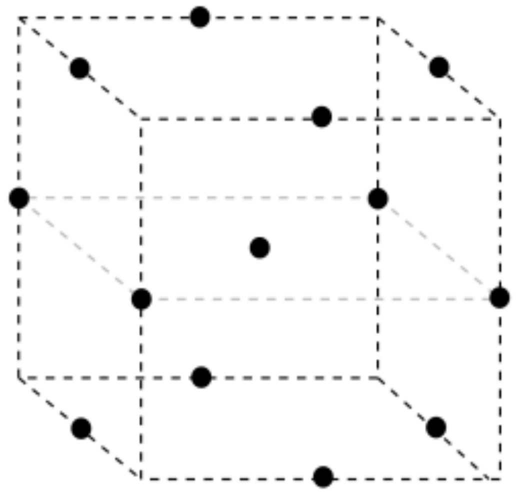
  Definitive Screening Design

# *The most popular 3-level designs*

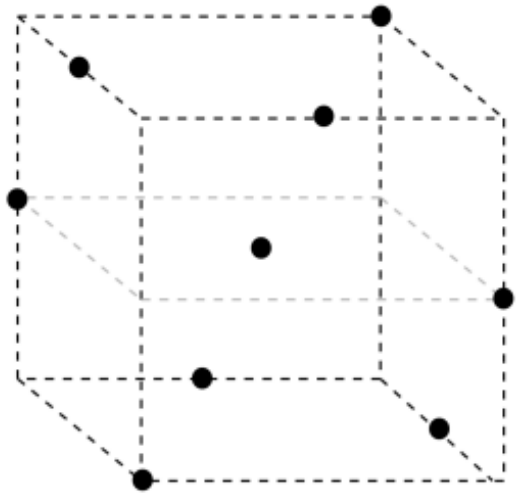
Response surface designs for three quantitative factors



Central composite design



Box-Behnken design



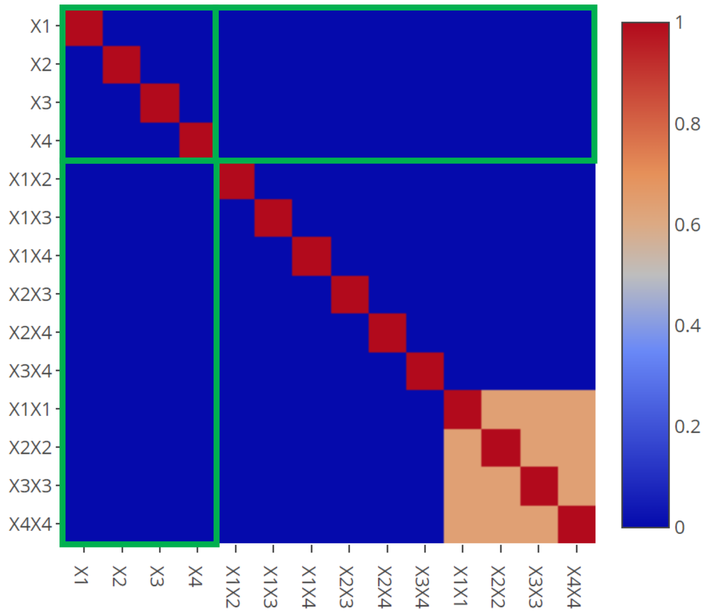
Definitive screening design

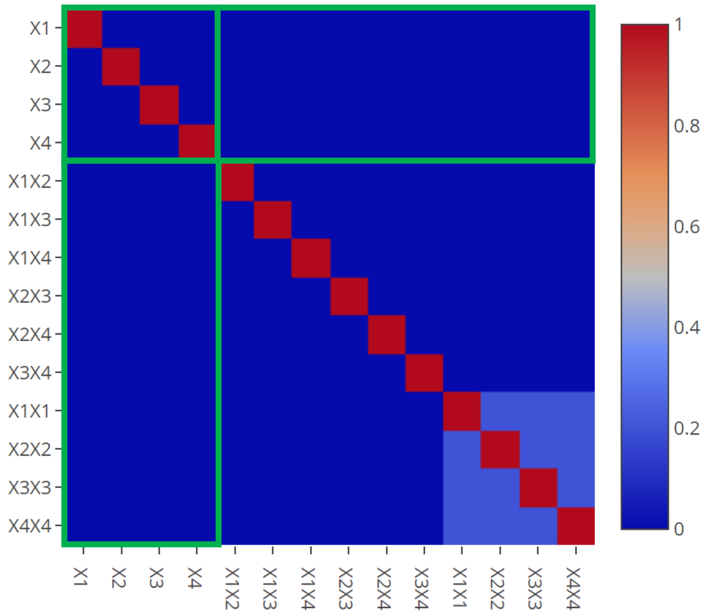
# The most popular 3-level designs

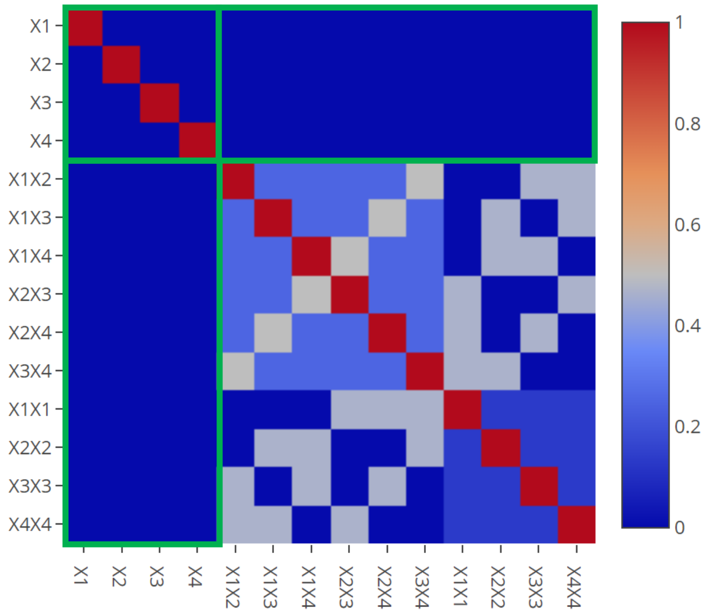
Central composite design

Box-Behnken design

Definitive screening design



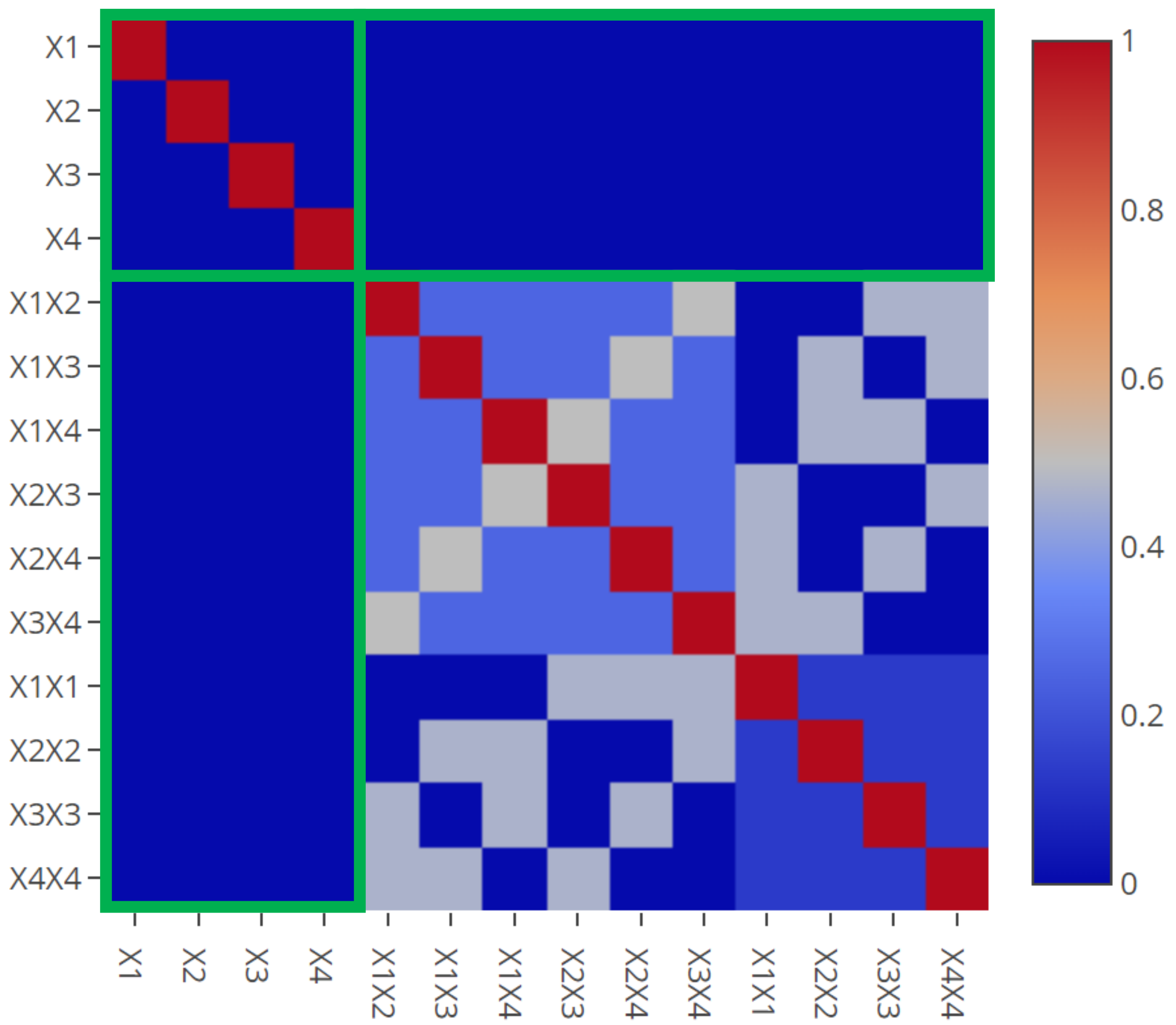




All these designs belong to the family of OMARS designs

[Núñez Ares, Goos 2020 & Núñez Ares, Schoen, Goos 2023]

# *OMARS designs*




***O****rthogonal*

main effects estimated independently

from each other

***M****inimally* ***A****liased*

main effects estimated independently

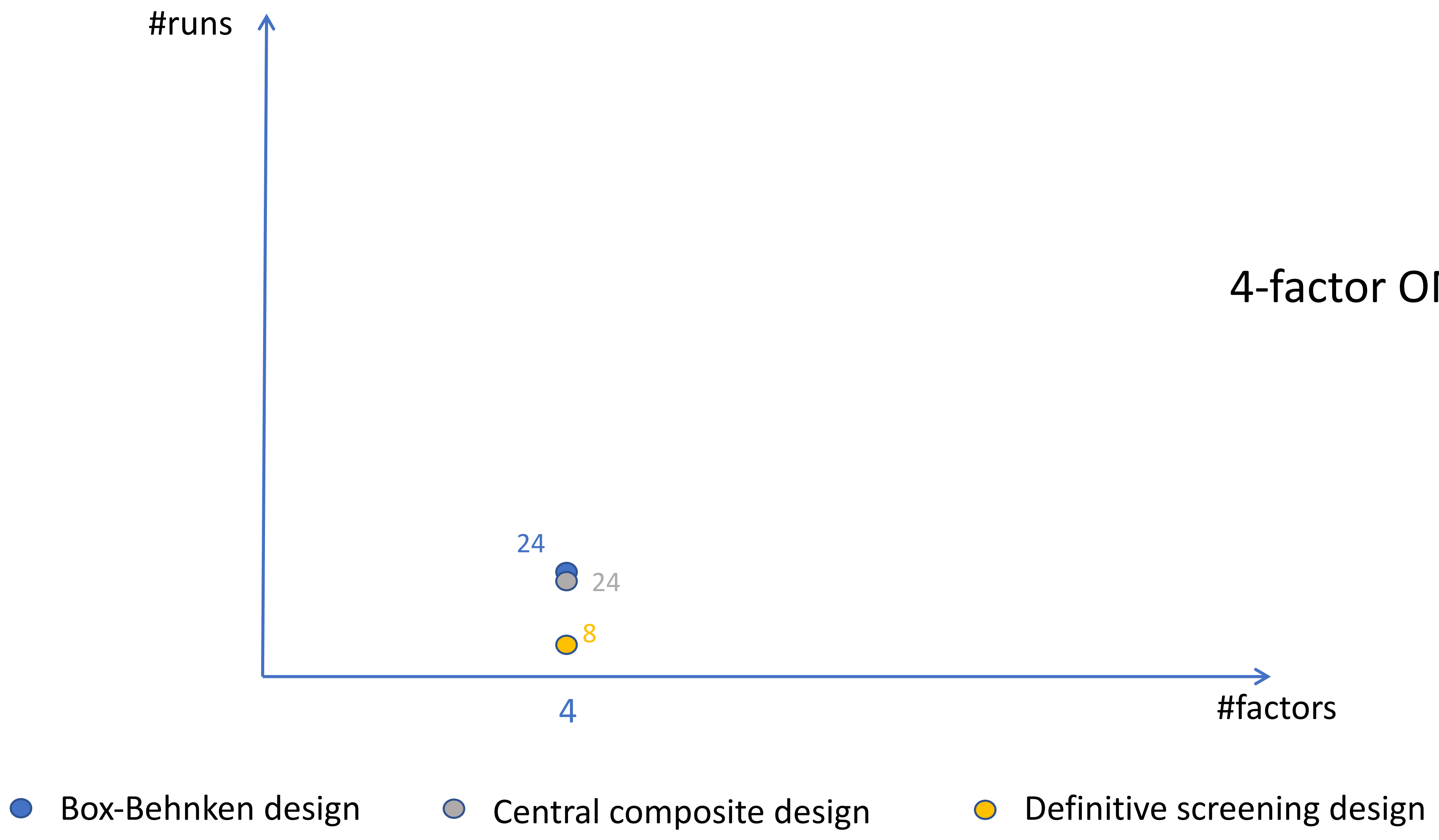
from all second-order effects

***R****esponse* ***S****urface* ***Designs***

allow the estimation of a partial or

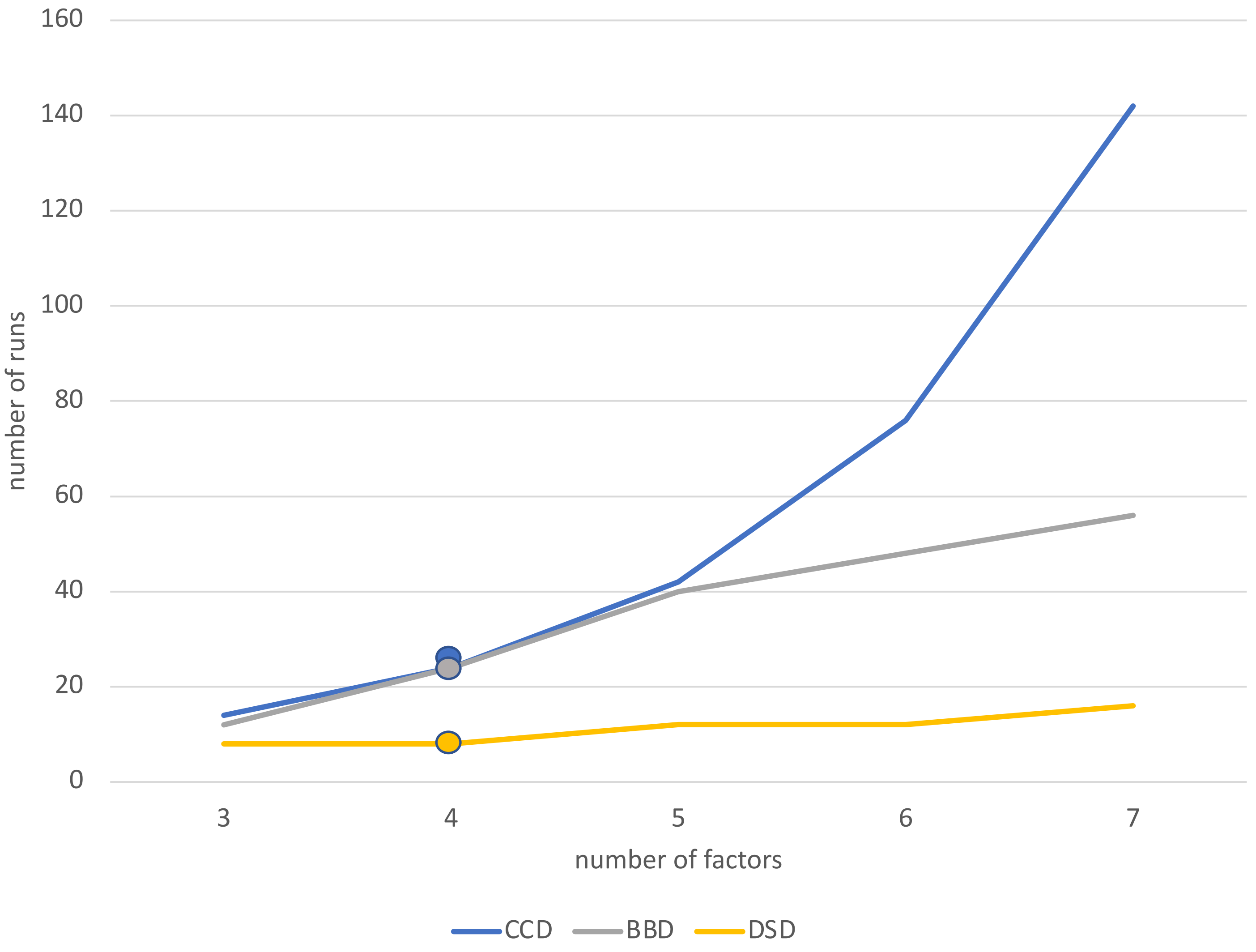
complete second-order effects model

# Choosing an OMARS design

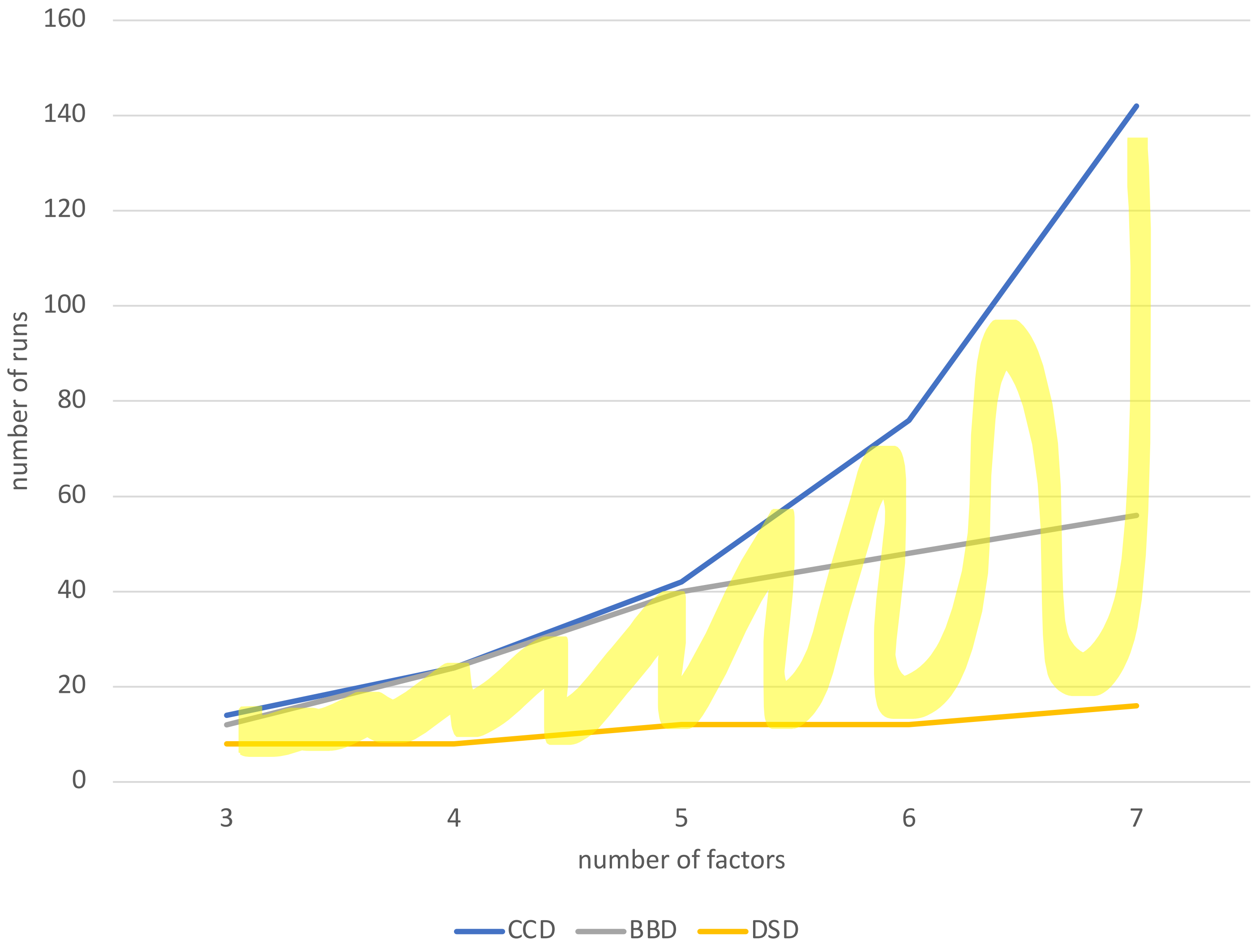



4-factor OMARS designs

# *Limited number of standard designs*



# Motivation of the present work




Are there more OMARS designs with a number of runs between the small DSDs and the large CCDs?

# *Positive answer: our OMARS catalog*

| #runs/#factors | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 14 | 46 | 128 | 11 | 4 | 2 | | | | | | | | | |
| 16 | 159 | 190 | 152 | 61 | 8 | 3 | | | | | | | | |
| 18 | 198 | 359 | 552 | 171 | 30 | 11 | | | | | | | | |
| 20 | 572 | 1,621 | 5,569 | 5,117 | 997 | 171 | 7 | 3 | | | | | | |
| 22 | 1,438 | 5,788 | 42,262 | 97,792 | 37,941 | 3,021 | 145 | 6 | | | | | | |
| 24 | 1,921 | 12,765 | 168,045 | 886,015 | 1,919,652 | 142,192 | 12,637 | 1,658 | 152 | 35 | | | | |
| 26 | 2,235 | 21,482 | 807,530 | 9,611,789 | 5,086,943 | 1,815,173 | 898,596 | 287,208 | 298,799 | 1,426 | 7 | | | |
| 28 | 492 | 3,285 | 91,111 | 1,022,895 | 1,255,206 | 265,213 | 37,228 | 7,676 | 1,505 | 487 | 93 | | | |
| 30 | 1,263 | 18,761 | 1,822,824 | 27,311,163 | 55,340,120 | 26,620,971 | 3,231,476 | 60,050 | 560 | 31 | 8 | 1 | | |
| 32 | 33 | 656 | 5,177 | 47,237 | 114,145 | 99,398 | 47,574 | 17,237 | 3,594 | 430 | | | | |
| 34 | 38 | 651 | 8,564 | 139,985 | 171,785 | 15,654 | 878 | 177 | 27 | 15 | 4 | 4 | 1 | 1 |
| 36 | 64 | 2,157 | 38,368 | 1,926,480 | 4,971,761 | 1,646,150 | 53,536 | 669 | 11 | 1 | 1 | 1 | | |
| 38 | 95 | 4,420 | 137,380 | 15,097,844 | 7,034,284 | 3,086,804 | 28,877 | 232 | 27 | 15 | 4 | 4 | 1 | 1 |
| 40 | 129 | 9,688 | 919,100 | 59,240,843 | 66,439,987 | 7,590,489 | 983,545 | 12,560 | 26 | 13 | 3 | 3 | 1 | 1 |
| TOTAL | 8,683 | 81,951 | 4,046,645 | 115,387,396 | 142,372,861 | 41,285,250 | 5,294,499 | 387,476 | 304,701 | 2,453 | 120 | 13 | 3 | 3 |
| | | | | | | | | | | | | | | |
| GRAND TOTAL | 309,172,054 | | | | | | | | | | | | | |

# *How did we find them?*

Our enumeration method:

**Properties**

- Enumerates non-isomorphic designs
- All enumerated designs are OMARS

**Approach**

- Integer programming
- High throughput computing

**Execution**

- HPC and HTC infrastructures
- High total computation time

# *4-factor 23-run OMARS design*

| | X1 | X2 | X3 | X4 |
|---|---|---|---|---|
| 1 | - | - | - | - |
| 2 | - | 0 | 0 | + |
| 3 | - | 0 | 0 | + |
| 4 | - | 0 | 0 | 0 |
| 5 | - | 0 | + | - |
| 6 | - | + | 0 | 0 |
| 7 | 0 | - | - | + |
| 8 | 0 | - | 0 | 0 |
| 9 | 0 | - | + | + |
| 10 | 0 | - | + | 0 |
| 11 | 0 | 0 | - | 0 |
| 12 | 0 | 0 | 0 | - |

| | X1 | X2 | X3 | X4 |
|---|---|---|---|---|
| 13 | 0 | 0 | + | - |
| 14 | 0 | + | - | - |
| 15 | 0 | + | - | + |
| 16 | 0 | + | 0 | - |
| 17 | 0 | + | + | + |
| 18 | + | - | 0 | - |
| 19 | + | 0 | - | 0 |
| 20 | + | 0 | 0 | - |
| 21 | + | 0 | 0 | + |
| 22 | + | 0 | 0 | + |
| 23 | + | + | + | 0 |

Non-foldover

Balanced for MEs

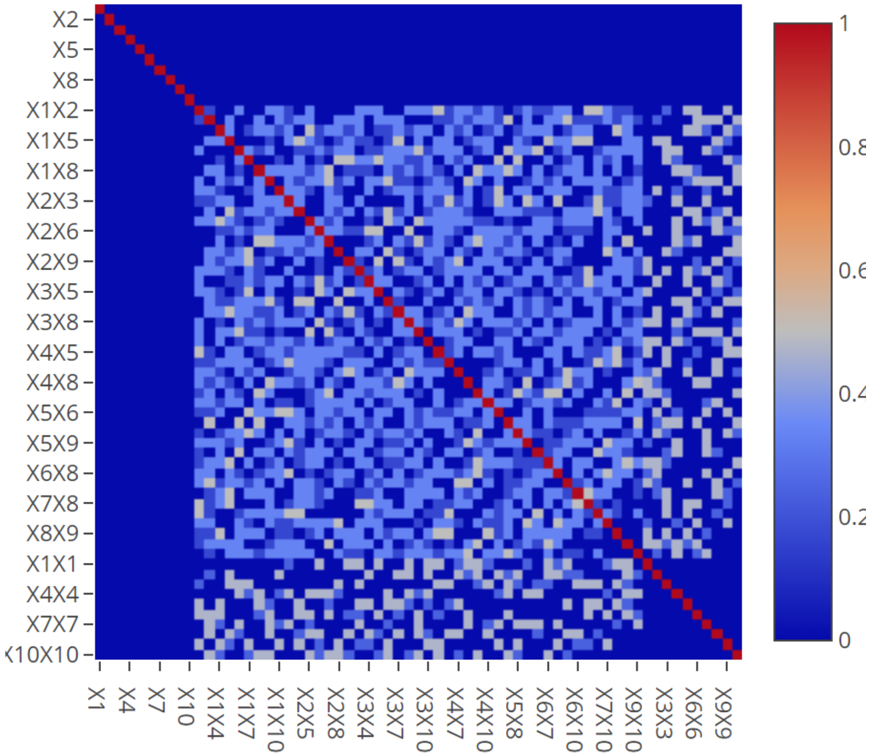
No center runs

# Application 1: complex problems

Experiment related to nuclear waste storage.

Extremely expensive and high estimation quality requirements

10-factor 27-run design with Projection estimation capacity = 4




Projection information capabilities

| #factor projections | D-eff | A-eff | G-eff | PV |
|---|---|---|---|---|
| 3 | 42.04 | 27.61 | 62.85 | 0.244 |
| 4 | 35.37 | 17.79 | 33.44 | 0.907 |

Max 4th order correlation = 0.5

Quadratic effects orthogonal to each other

Min power to detect an IE: 0.898

Min power to detect a QE: 0.629

# Application 2: mixed-level OMARS

iSEa

International Statistical Engineering Association

HOME | DISCOVER ISEA ▸ | EXPLORE STATISTICAL ENGINEERING ▸ | GET INVOLVED ▸ | EVENTS ▸ | STU HUNTER CONFERENCE ▸

Back

Add to my calendar

## A case study of a mixed-level OMARS design

When 20 Sep 2023
10:00 AM - 11:00 AM
Location Online

Maria Lanzerath, statistician at W. L. Gore & Associates, will give the next ISEA webinar on design of experiments and showcases how these are used in her work.

GORE

Very expensive experiment in production

6+2 experimental factors

24 runs

3 extra covariates

Analysis + optimization using specialized method

# *Application 3: bioessay*


GSK

6 quantitative factors

Bioreactor with capacity for 24 runs

# *Are they any good? How to choose?*

Using a multi-criteria selection methodology which takes into account

- The maximum budget for the experimental design,
- The goal of the experiment: optimization and/or screening,
- The existence of blocking factors and other experimental conditions

OMARS designs have been discovered while at KU Leuven university.

Since this last summer, a new spin-off company, EFFEX, markets the catalog of OMARS designs

# *Exhaustive design characterization*

D-, A-, G- and I-Efficiencies

Color map on correlations

Number of replicates and degrees of freedom for pure error estimation

Maximum 4$^{th}$ order correlation

Powers

Rank of several design matrices

Projection estimation properties

VIFs, Relative Error of Estimates


EFFEX
A collaborative platform
for your experimental plans
www.effex.app

# *Optimization experiments goals*

**Be able to fit a SOE model**

Only designs with a full rank second-order effects matrix are selected

A low cubic bias protects from unexpected active third order terms

**Quality of estimation & prediction**

D-, A-, G- and I-efficiencies are important

Minimum powers to detect main, interaction and quadratic effects

**Meaningful modeling**

A minimum correlation between the effects is an important quality criterion

Number of center points and number of replicates

Degrees of freedom for pure error estimation that allows lack-of-fit tests

**Run size matters**

The trade-offs between the run size and the mentioned criteria need to be assessed

For example, there is always a cheaper and better alternative to standard RSDs

# *Example 1: optimization experiment*

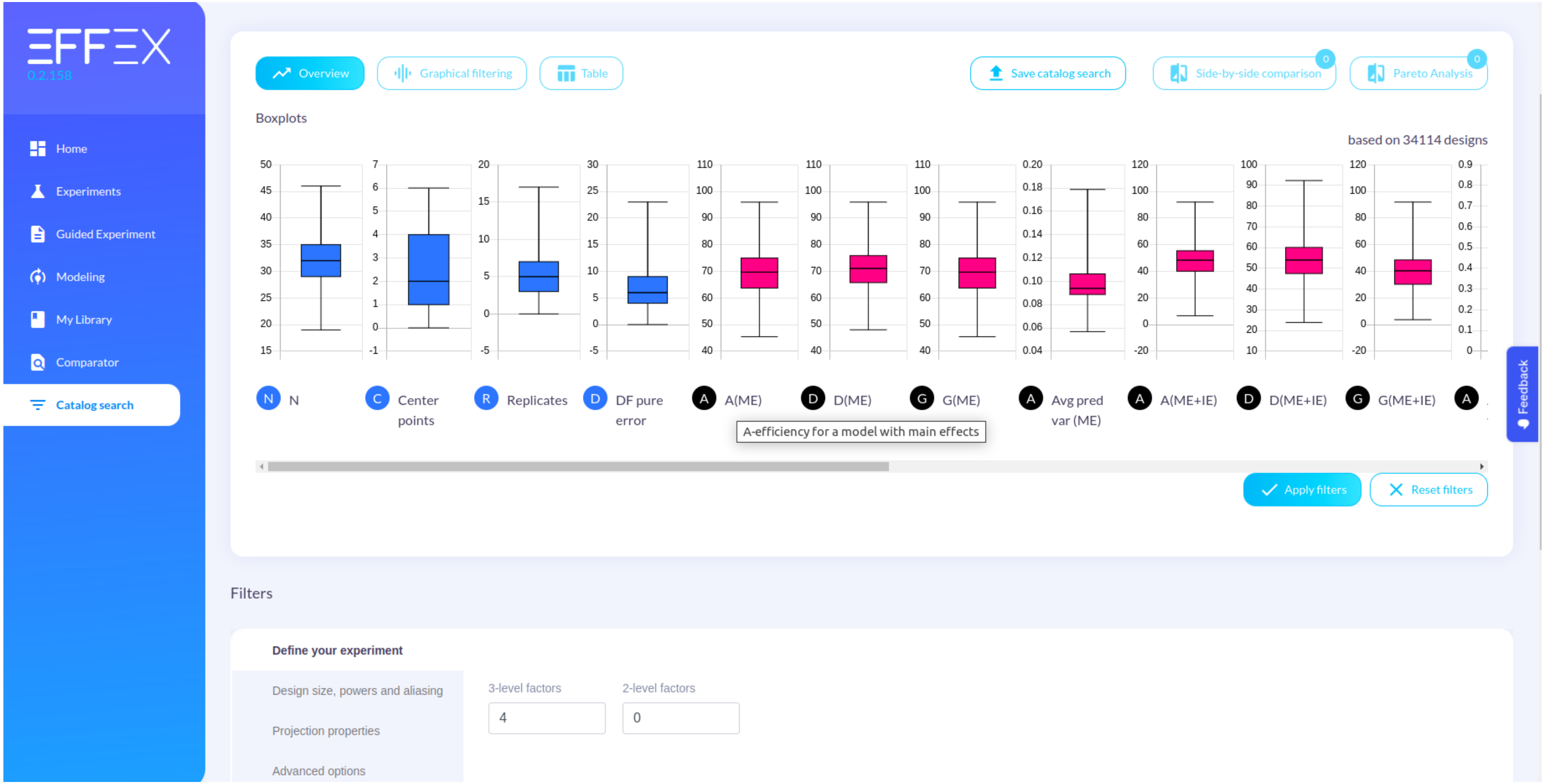
Optimization second-order design for 4 quantitative factors

Consider the standard designs:

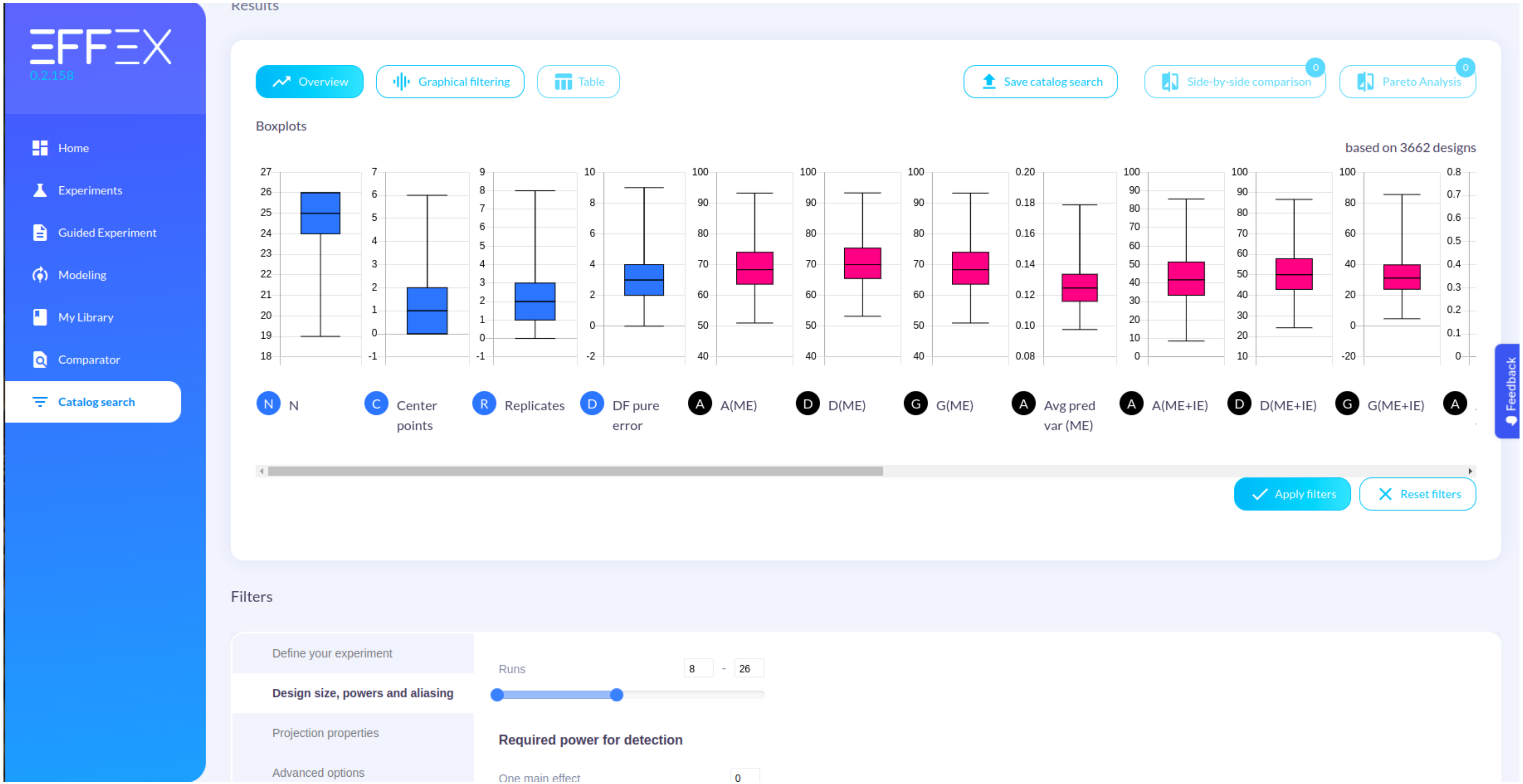
- Central composite design: 26 runs with 2 center points
- Box-Behnken design: 27 runs with 3 center points

| | CCD | BBD |
|---|---|---|
| Power interaction effect | 0.952 | 0.452 |
| Power quadratic effect | 0.309 | 0.564 |
| Maximum 4th order correlation | 0.639 | 0.2 |
| G-efficiency | 75.07 | 23.8 |
| Prediction variance | 0.325 | 0.4 |
| Pure error | YES | YES |

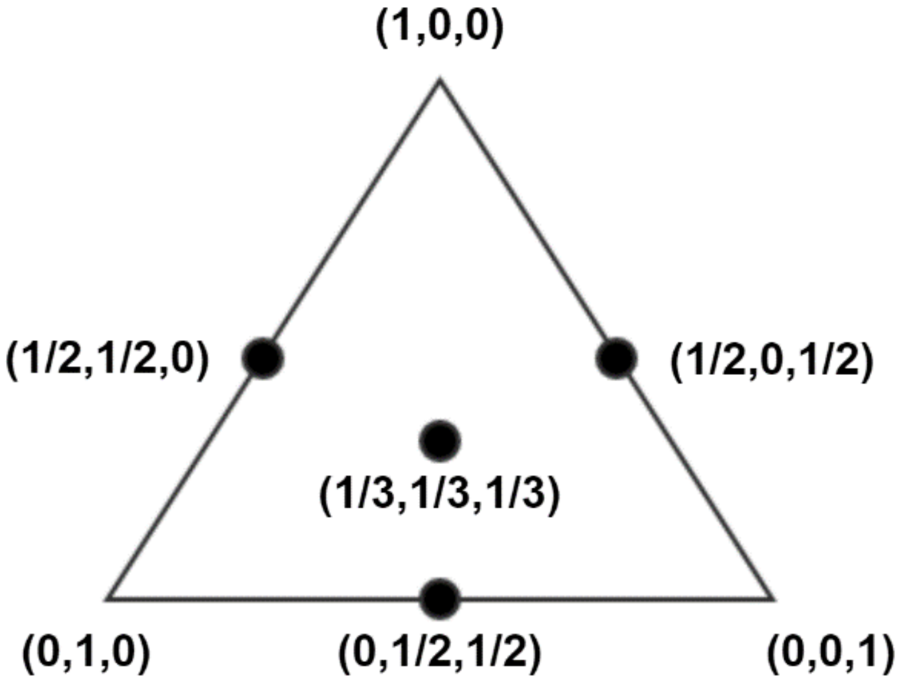
# *First filtering using EFFEX software*



# Second filtering



# *Ternary plot to compare designs*




Ternary plot = simplex plot = barycentric plot

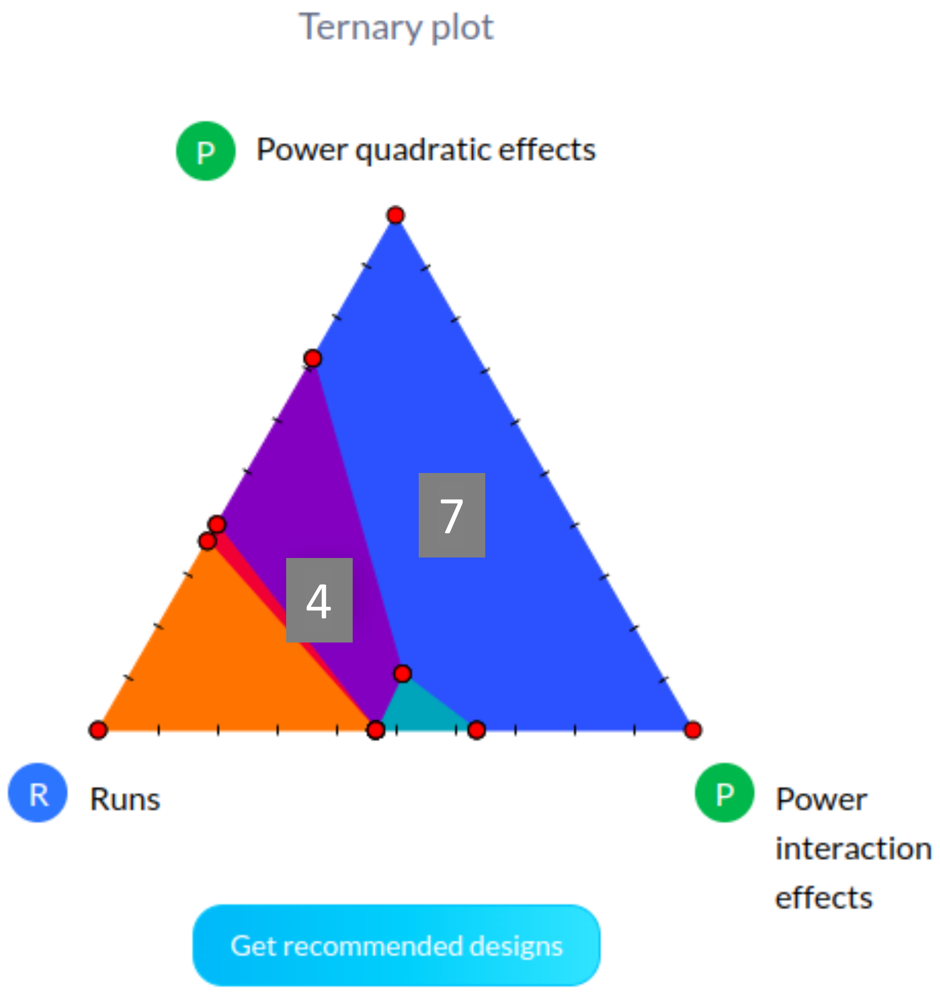
Three variables sum up to a constant

($w_a$, $w_b$, $w_c$) are three non-negative weights s.t

$$w_a + w_b + w_c = 1$$

In our case, the weights act on three design quality criteria

# *Ternary plot to compare designs*



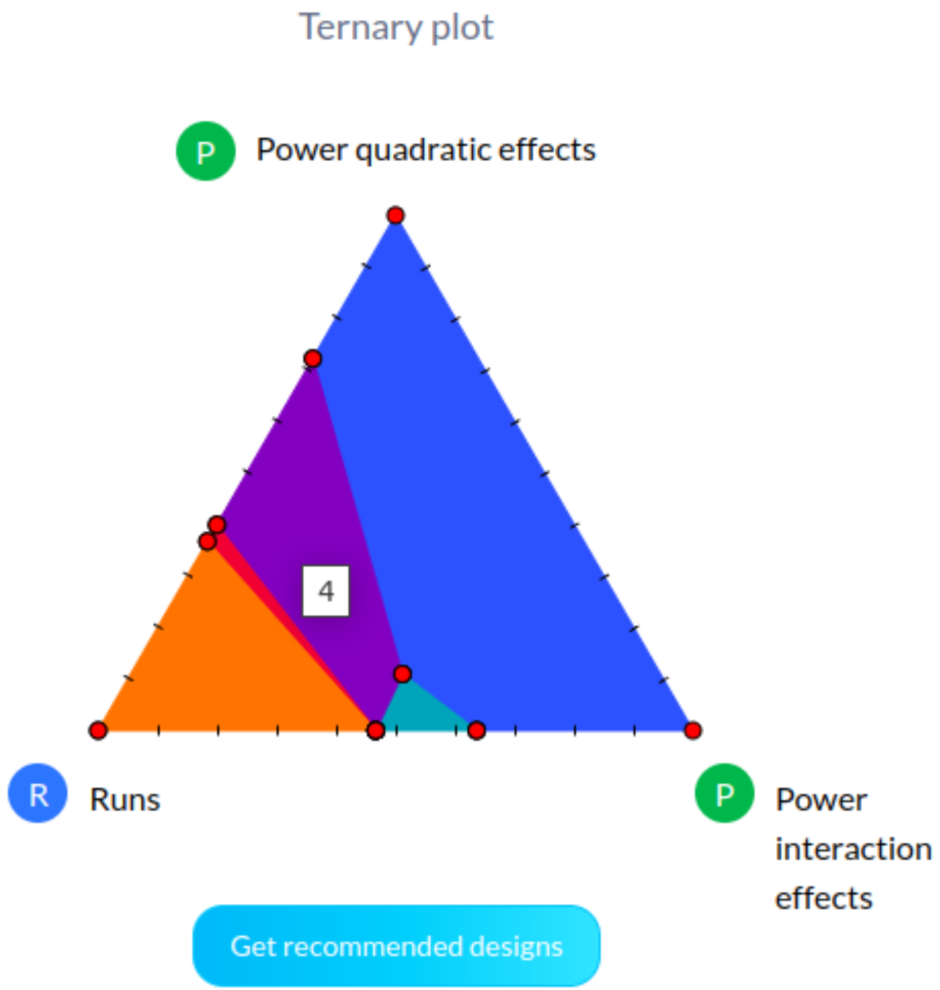

The colored areas represent the design(s) that best perform for the weights ($w_a$, $w_b$, $w_c$) contained in that area.

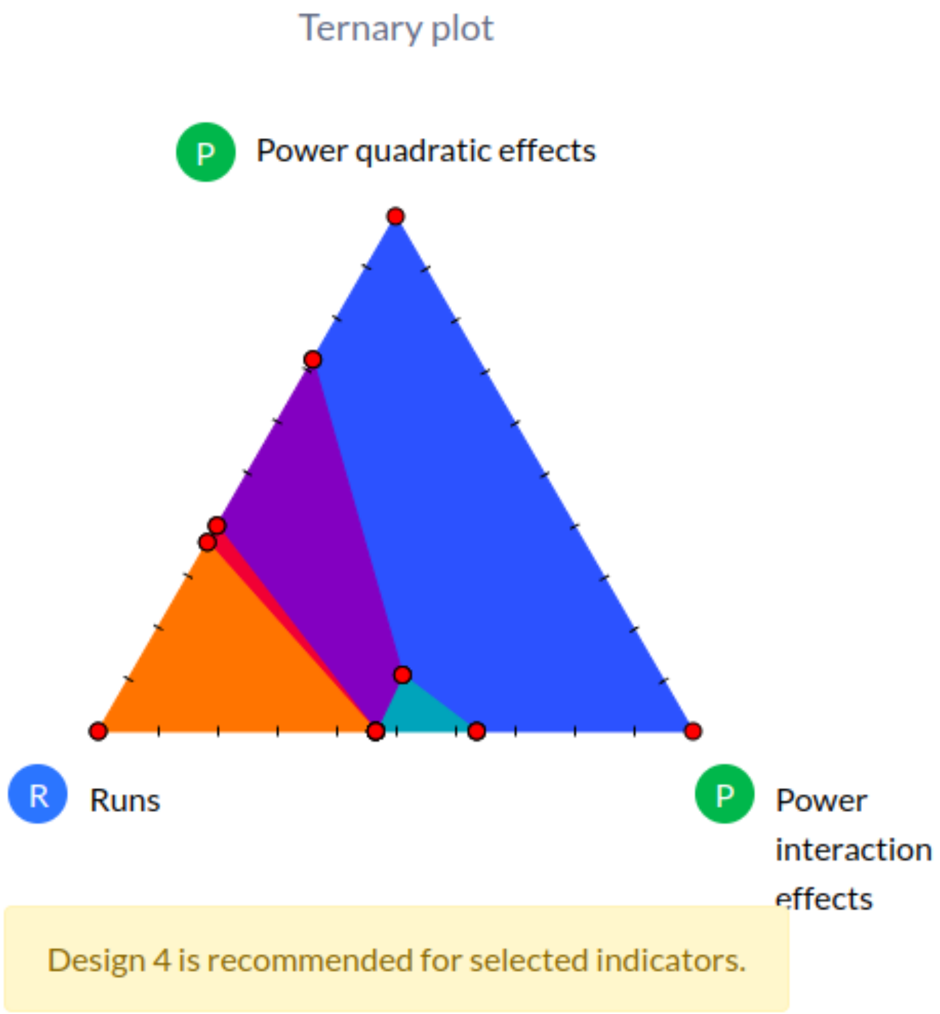
For example, the design 7 (blue area) performs best for the power to detect quadratic and interaction effects, but it is not the smallest.

Design 4 performs well for a scenario where all criteria have the same weights

Designs in the ternary plot are non dominated by any other design for the criteria considered

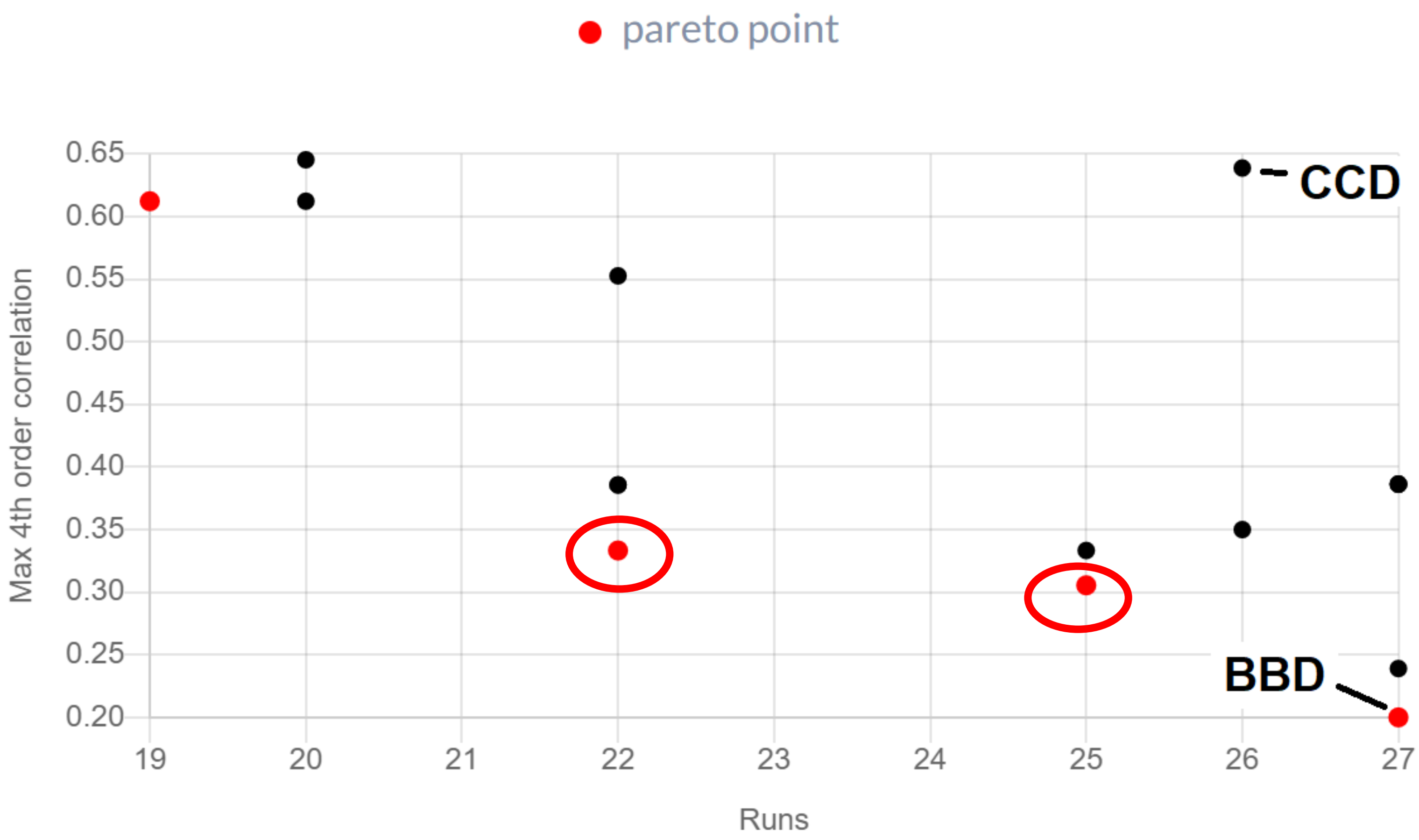
# Pareto analysis optimization experiment



# *Second-order designs for 4 factors*

We select 14 4-factor second-order OMARS designs, and we compare them to the CCD and the BBD

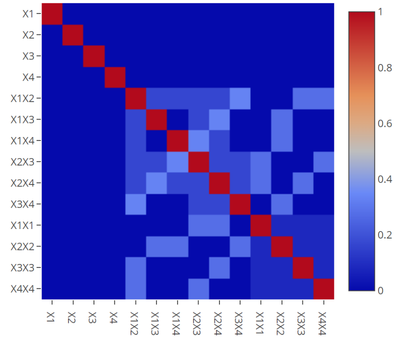


# *Two OMARS designs for optimization*

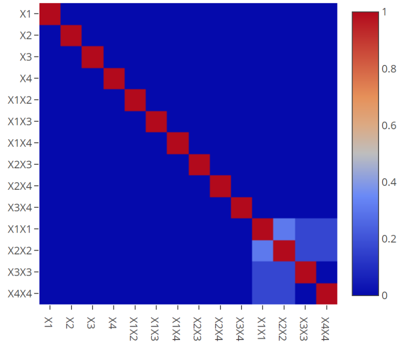
22- and 25-run OMARS designs

| | OMARS 1 | OMARS 2 | CCD | BBD |
|---|---|---|---|---|
| Number of runs | 22 | 25 | 26 | 27 |
| Power interaction effect | 0.699/0.641 | 0.948/0.876 | 0.952 | 0.452 |
| Power quadratic effect | 0.373/0.345 | 0.528/0.423 | 0.309 | 0.564 |
| Maximum 4th order correlation | 0.333 | 0.305 | 0.639 | 0.2 |
| G-efficiency | 40.21 | 73.13 | 75.07 | 23.8 |
| Prediction variance | 0.455 | 0.429 | 0.325 | 0.4 |
| Pure error (number of replicates) | NO | YES (2) | YES (1) | YES (1) |

22-run OMARS




25-run OMARS



# *Screening experiments goals*

Detect active main effects

Power to detect the main effects should be high.

Better when main effects are orthogonal with each other and with all second-order effects (OMARS)

Detect *some* SOE effects

How can we quantify ***some*** here?

- Projection estimation capacity
- Model matrix rank:

How do we estimate the quality in detecting the SOEs:

- Maximum correlation between SOEs
- Power

ONE-SHOT EXPERIMENT

Budget constraints

Of course, we want to do all this with a minimum run size…

Study the trade-off between the quality criteria and the run size

# *Example 2: a screening experiment*

Screening second-order design for 6 quantitative factors, no more than 22 runs.

Benchmark designs: Definitive screening designs with 16 to 22 runs.

Projection estimation capacity equals **3**:

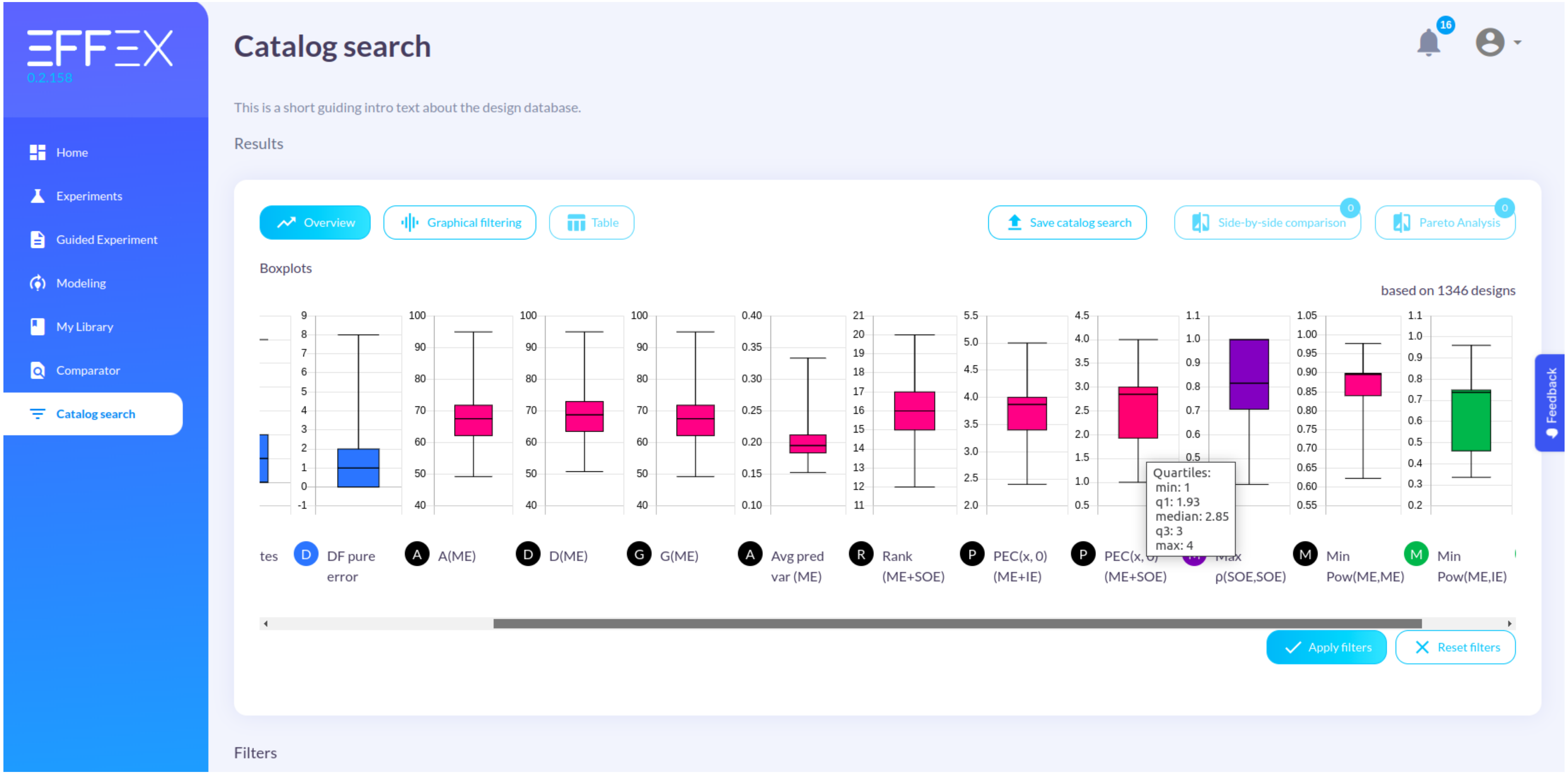
- 6 factors: X1, X2, X3, X4, X5, X6
- There are $\binom{6}{3} = 20$ subsets of 3 factors out of the six
- With these designs we can fit a full second-order effects model on any subset of **3** factors

# *Benchmark designs*

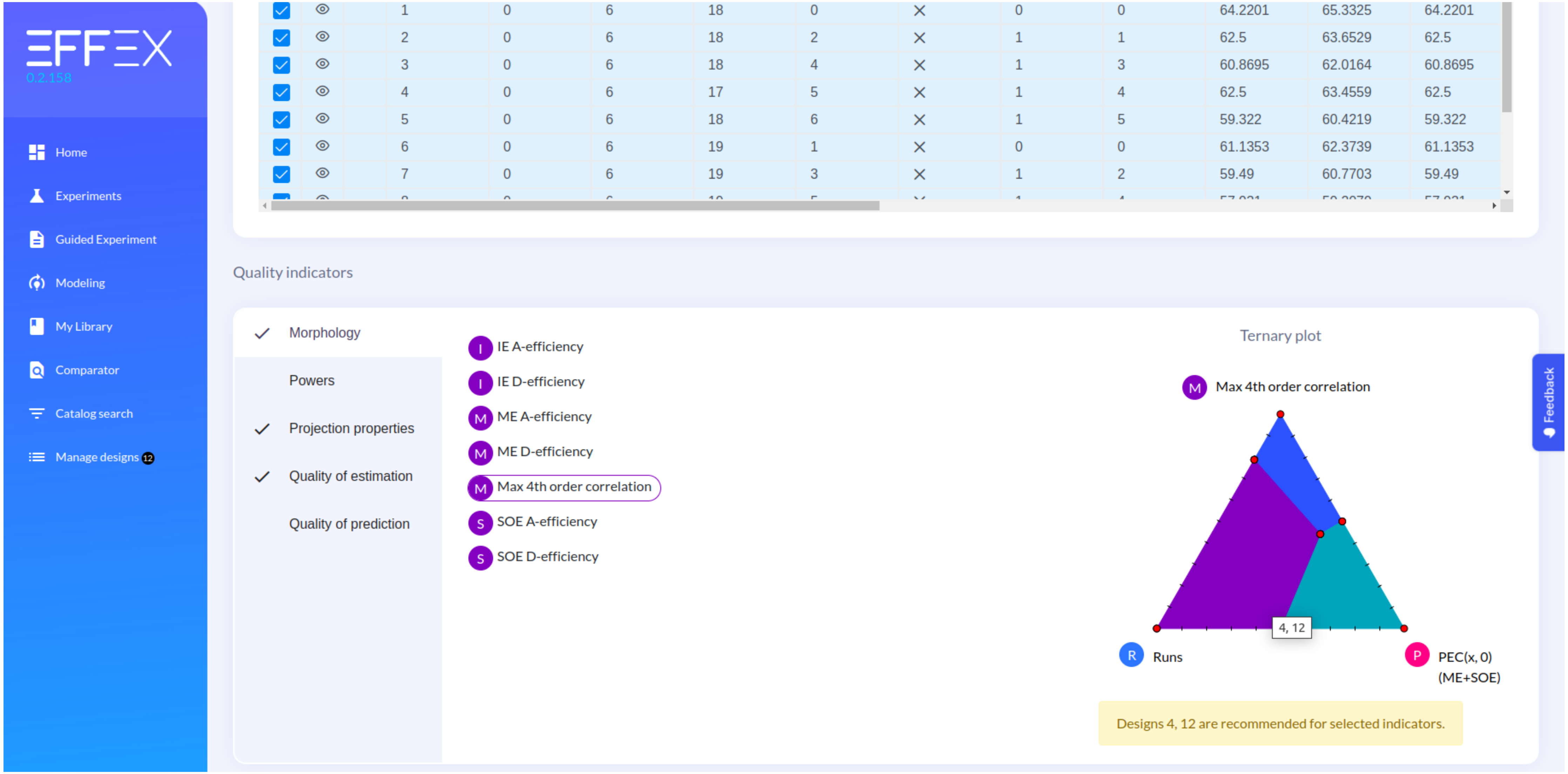
Characteristics of the benchmark definitive screening designs:

| | DSD#1 | DSD#2 | DSD#3 | DSD#4 | DSD#5 | DSD#6 | DSD#7 |
|---|---|---|---|---|---|---|---|
| Number of runs | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| Power interaction effect | 0.857 | 0.868 | 0.876 | 0.882 | 0.935 | 0.958 | 0.96 |
| Power quadratic effect | 0.215 | 0.29 | 0.357 | 0.417 | 0.29 | 0.318 | 0.391 |
| Maximum 4th order correlation | 0.666 | 0.666 | 0.666 | 0.666 | 0.75 | 0.75 | 0.75 |
| Projection estimation capacity | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| Projection information capacity D-eff (3) | 41.81 | 44.43 | 44.25 | 43.39 | 44.43 | 44.51 | 44.8 |
| Projection prediction variance (3) | 1.25 | 1.21 | 0.41 | 0.36 | 0.53 | 0.5 | 0.38 |

# *First filtering using EFFEX software*

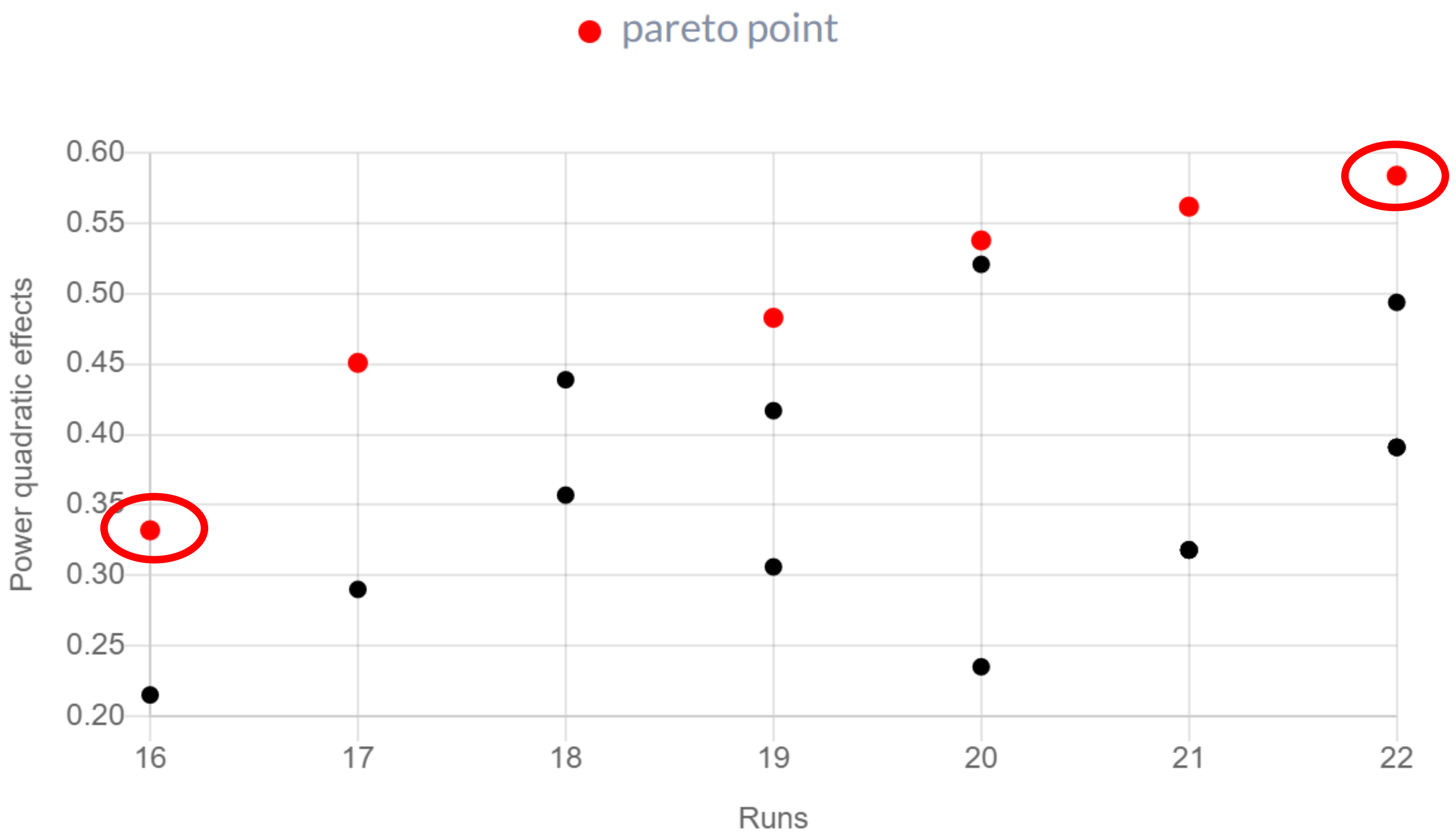


# *Pareto analysis example*



# *Pareto analysis*

We select 15 6-factor screening OMARS designs, and we compare them to the DSDs

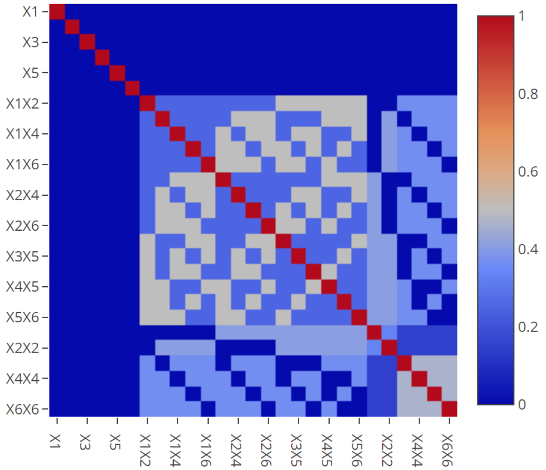



DSDs are not in the Pareto front

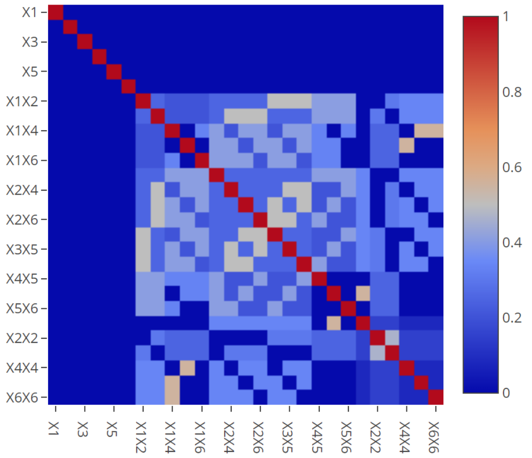
# Two OMARS for screening

| | OMARS 1 | DSD#1 | DSD#2 | DSD#3 | DSD#4 | DSD#5 | DSD#6 | DSD#7 | OMARS 2 |
|---|---|---|---|---|---|---|---|---|---|
| Number of runs | 16 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 22 |
| Power interaction effect | 0.698 | 0.857 | 0.868 | 0.876 | 0.882 | 0.935 | 0.958 | 0.96 | 0.748 |
| Power quadratic effect | 0.332 | 0.215 | 0.29 | 0.357 | 0.417 | 0.29 | 0.318 | 0.391 | 0.494 |
| Maximum 4th order correlation | 0.5 | 0.666 | 0.666 | 0.666 | 0.666 | 0.75 | 0.75 | 0.75 | 0.552 |
| Projection estimation capacity | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3.8 |
| Projection information capacity D-eff (3) | 40.34 | 41.81 | 44.43 | 44.25 | 43.39 | 44.43 | 44.51 | 44.8 | 39.99 |
| Projection prediction variance (3) | 0.46 | 1.25 | 1.21 | 0.41 | 0.36 | 0.53 | 0.5 | 0.38 | 0.33 |

16-run OMARS




22-run OMARS



# *Advantages of having a large design catalog*

Flexibility in terms of number of runs.
All standard OMARS designs are included in the catalog.
The best designs of the literature are available
Consider multiple criteria while choosing a design

**Optimization**

- There are cheaper alternatives to CCDs and BBDs.
- The weak points of CCDs and BBDs can be overcome
- More OMARS have 2FIs orthogonal to each other and to QEs.

**Screening**

- Higher projection estimation capacity than DSDs.
- Higher powers to detect curvature than DSDs
- Lower correlation between SOEs than DSDs

# *Example 3: a screening experiment in blocks*

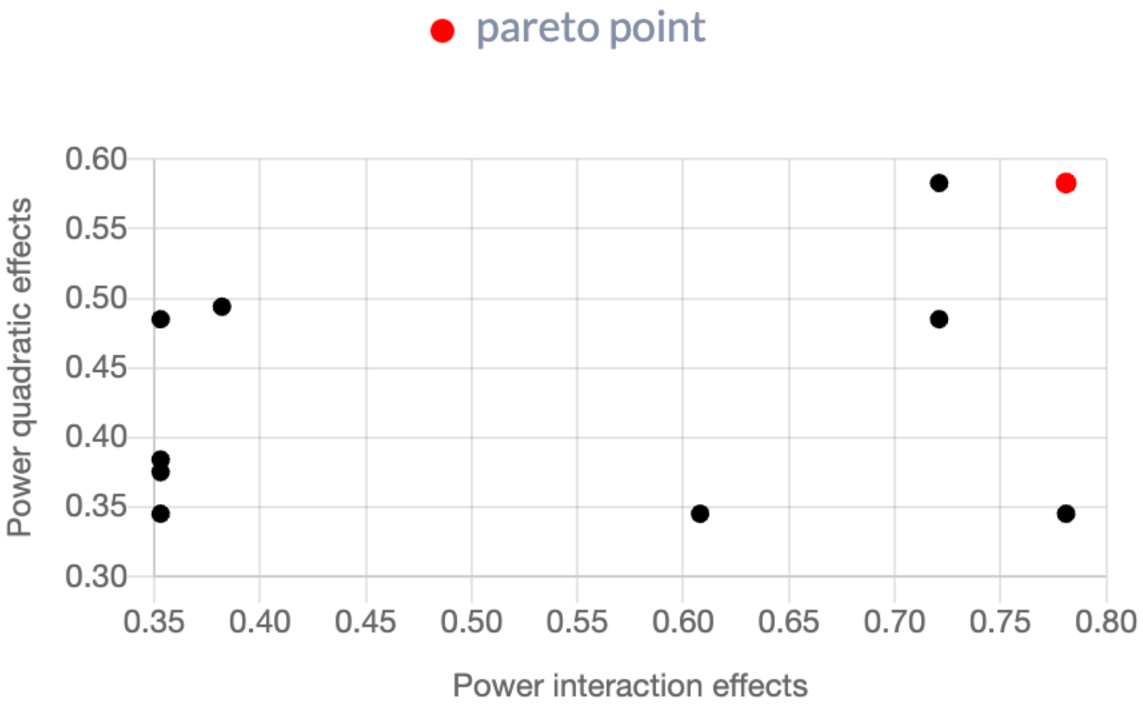
Screening second-order design for 6 quantitative factors and 1 two-level categorical factors, with 4 blocks of 6 runs each.

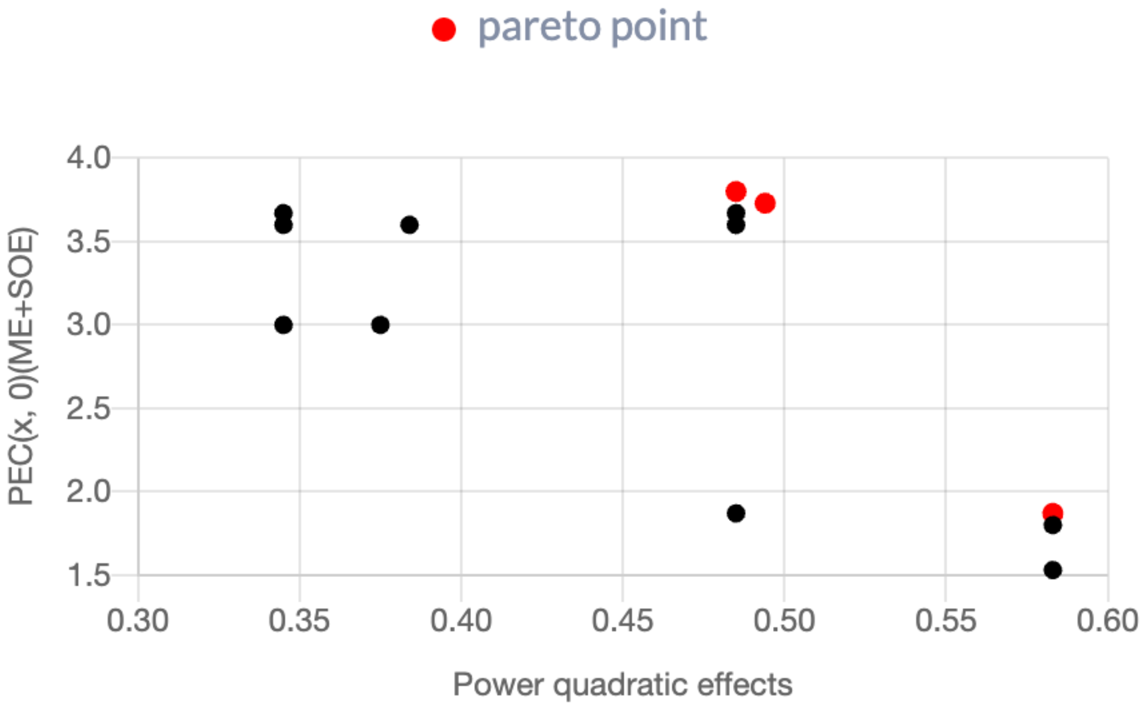
The must is having a high estimating projection capacity for second-order effects models that include quadratic effects.

Important criteria are: power to detect interaction and quadratic effects, projection properties, and aliasing between the different effects

# *Pareto analysis*

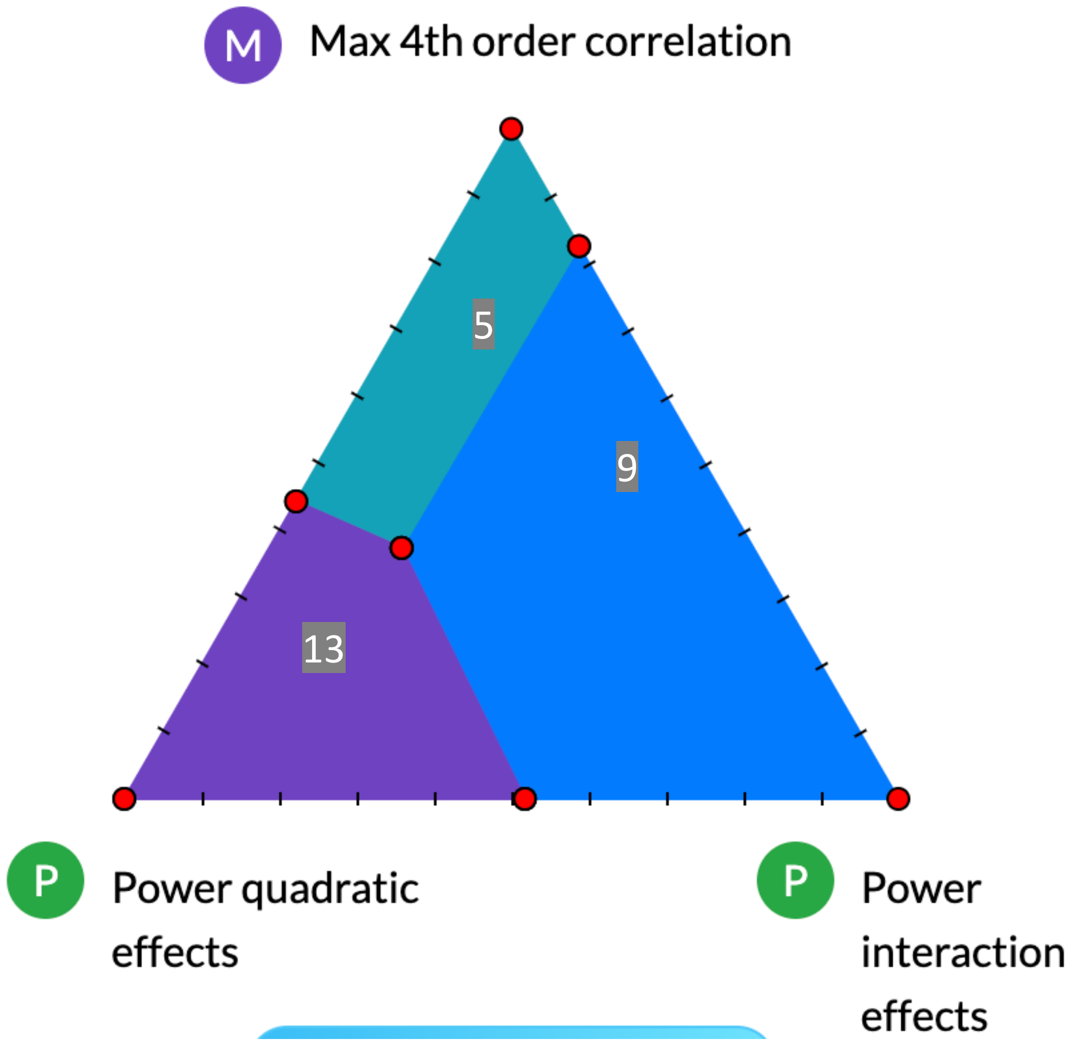
We select 24 7-factor screening blocked OMARS designs and perform a pareto analysis




There are competing designs regarding the three criteria considered

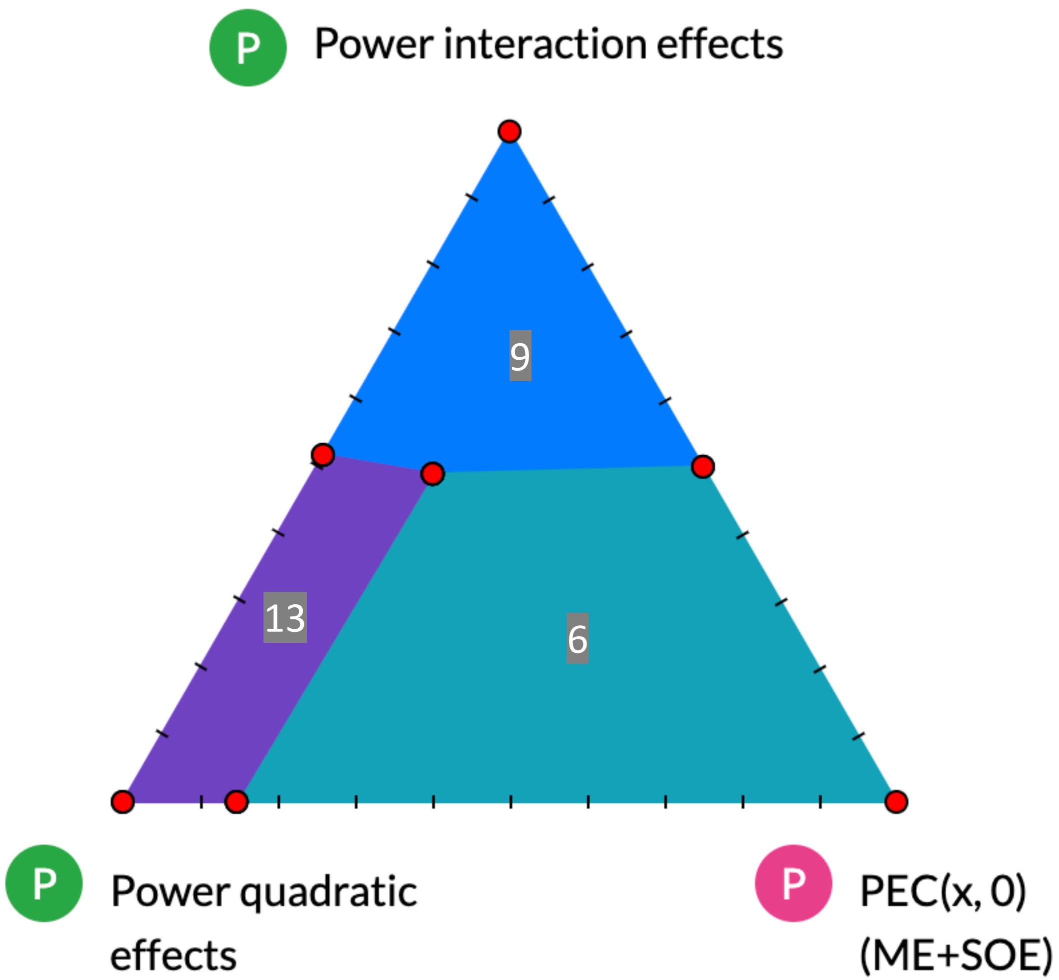
# *Pareto analysis*




Each area represents the design(s) that perform best for the three considered criteria at the corresponding weights.

Design 9 performs well for a scenario where all three criteria get a similar weight (importance)

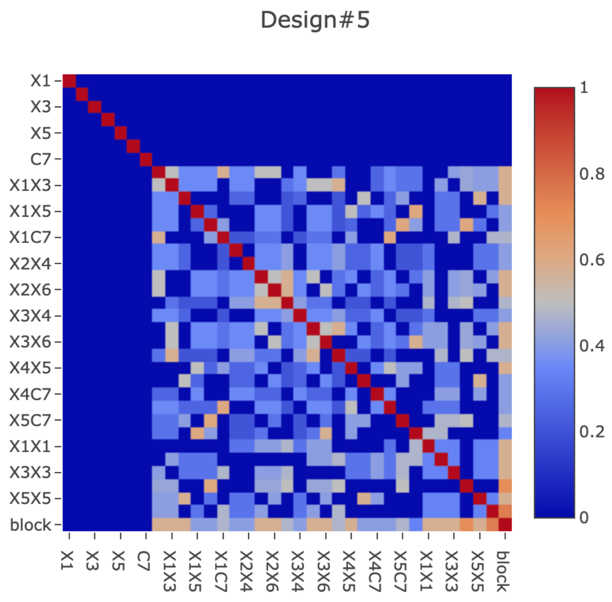
# *Pareto analysis*

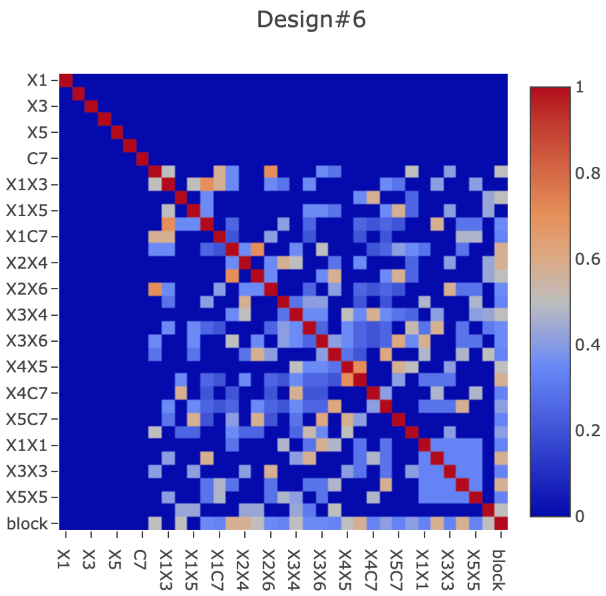


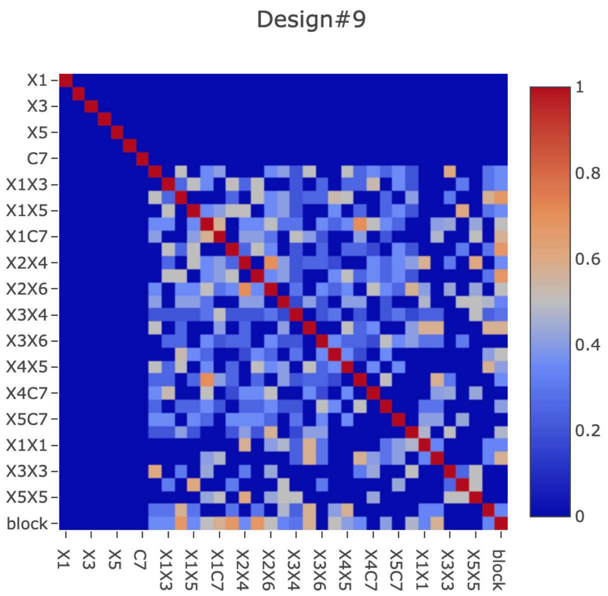

When considering the projection properties, design 6 appears as a good alternative

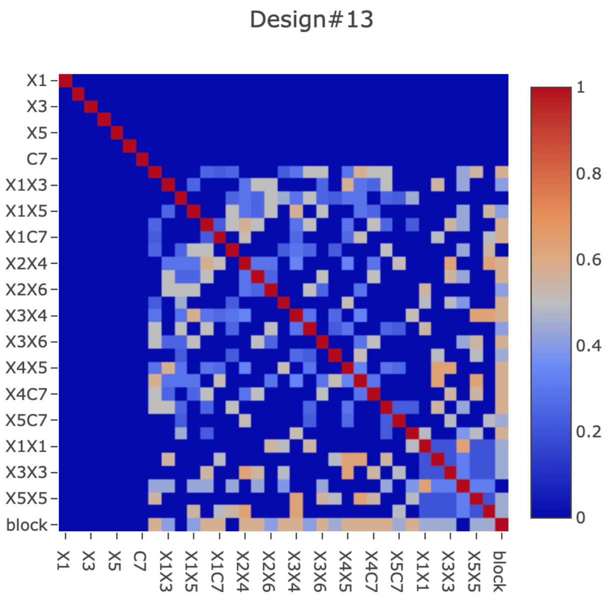
Let’s study them in more detail

# Detailed comparison



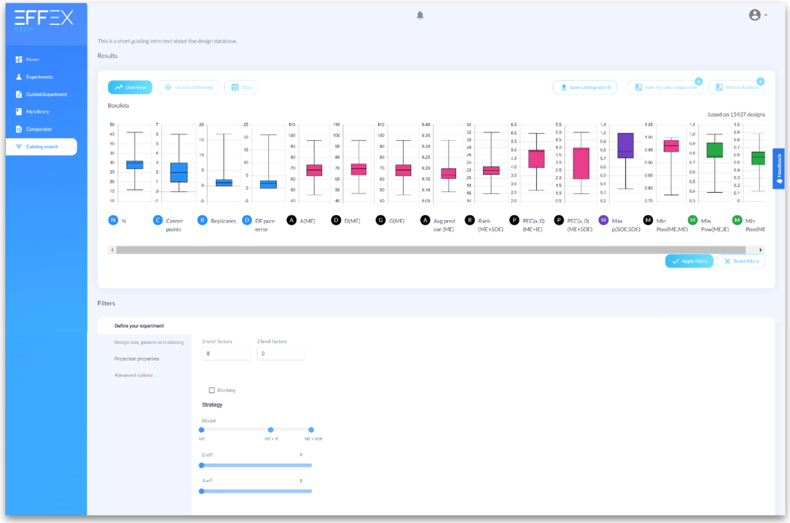




Comparing the color map on correlations can give us the decisive reasons to select one of the designs

# *A design for every challenge*



- OMARS®
- Blocking
- I-, A-optimal
- Split-plot
- Augmentation
- Bayesian
- Covariates

**From easy DoE to complex multi factor problems**
**EFFEX has the best solution at a minimal experimental time & cost**

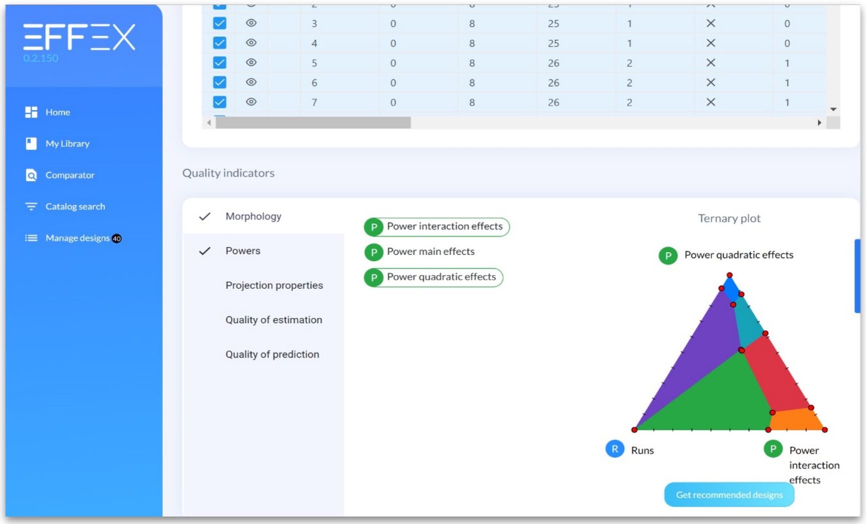
# *Getting the most out of your data*

Modeling capabilities

- All subsets selection
- (Bayesian) regression
- LASSO
- Dantzig
- Automatic model fitting

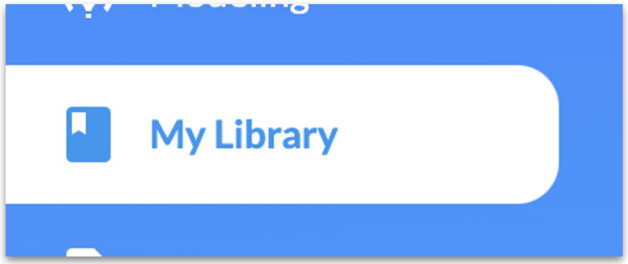
Optimization capabilities

- Probability of success plots
- Desirability
- Robust optimization
- Sensitivity analysis




**Thanks to our unique graphical modeling and optimization tools**

# *Your project information in one place*




We know how hard it is to keep all information on your DoE together. The Design, the conversations, the modelling and optimization. Why decisions have been made?

Therefore we have created a library where you can find an overview of of all your projects and DoE items.

Never search for your project information again!

# ***Software demo***

*"never give a software demo" (popular saying)*

*join.effex.app*

# *Summary*

The catalog offers much flexibility in choosing a design.

Often improves DSDs for screening and CCDs and BBDs for optimization.

Our catalog allows finding a design for novel problems, like, for example, a screening design in blocks with a high power to detect QEs.

The availability of a complete catalog allows us to develop a multi-criteria selection approach.

***Thank you!***

# *Extensions 1: mixed-level OMARS*

OMARS with quantitative and two-level categorical factors

The orthogonality structure is preserved

We improve the previous work on mixed-level RDSs.

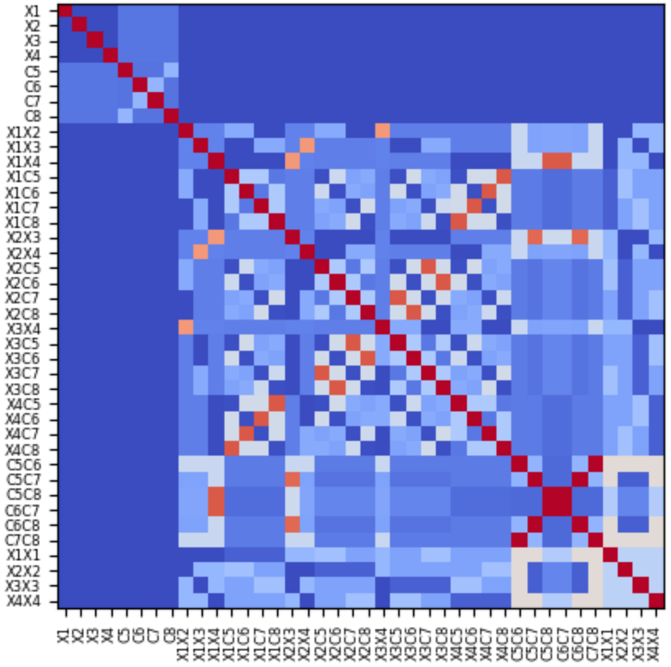
We built a large catalog of mixed-level OMARS for both screening and optimization.

A similar design selection approach can be followed for mixed-level designs too.

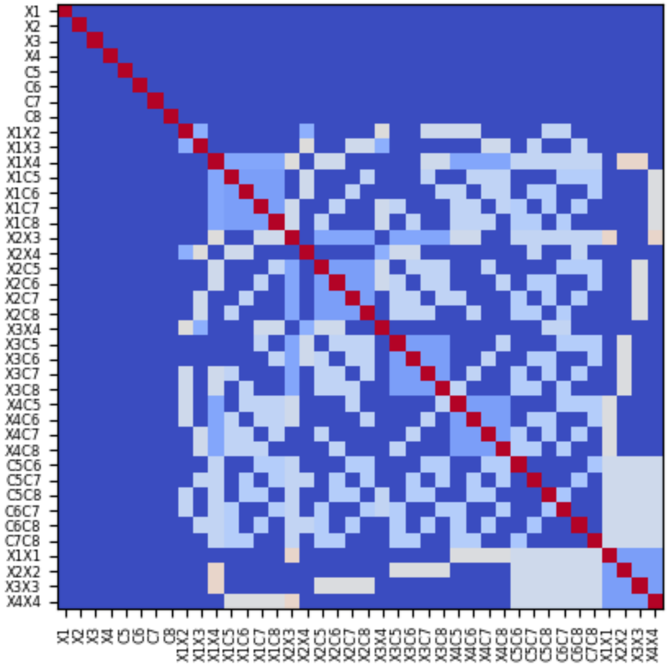
# *Mixed-level OMARS: example*

Screening design with 4 quantitative factors and 4 two-level categorical factors.
Two DSDs (22 and 26 runs) and one 24-run OMARS comparison.

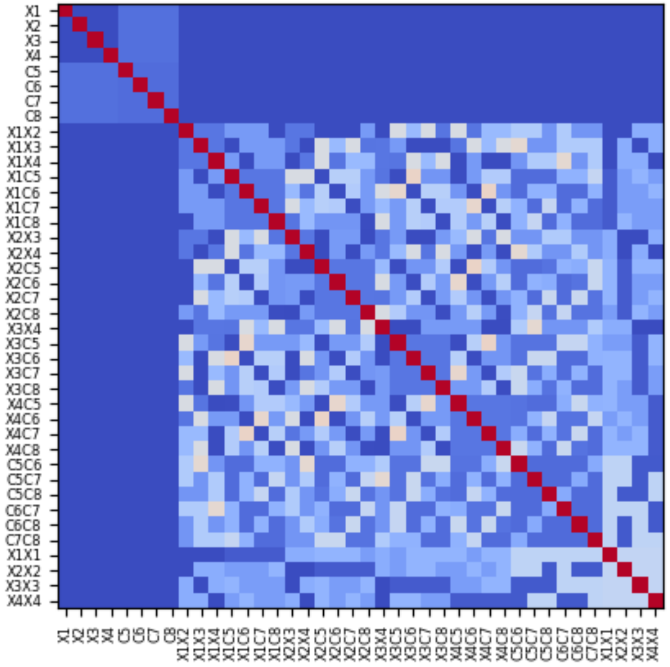
22-run DSD



24-run OMARS



26-run DSD



# Application 2: mixed-level design

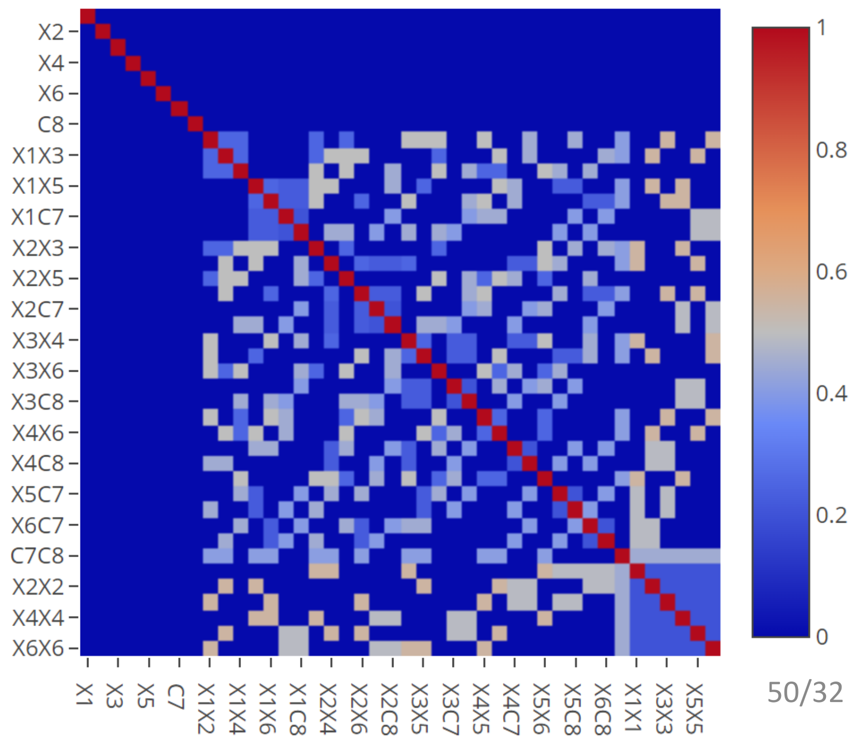
Chemical experiment in the health sector.

Screening + optimization experiment

6 quantitative factors and 2 two-level categorical factors

Budget of 24 runs

Max 4th order correlation = 0.54
Good projection properties
Twice the power to detect quadratic effects
than alternative DSDs from JMP



# *Extensions 2: blocked OMARS*

(Mixed-level) OMARS usually can be blocked in different ways.

Our blocked designs have the following properties:

- Block effect is orthogonal to main effects
- Minimal aliasing between the blocks and the second-order effects

Our approach is based on integer programming

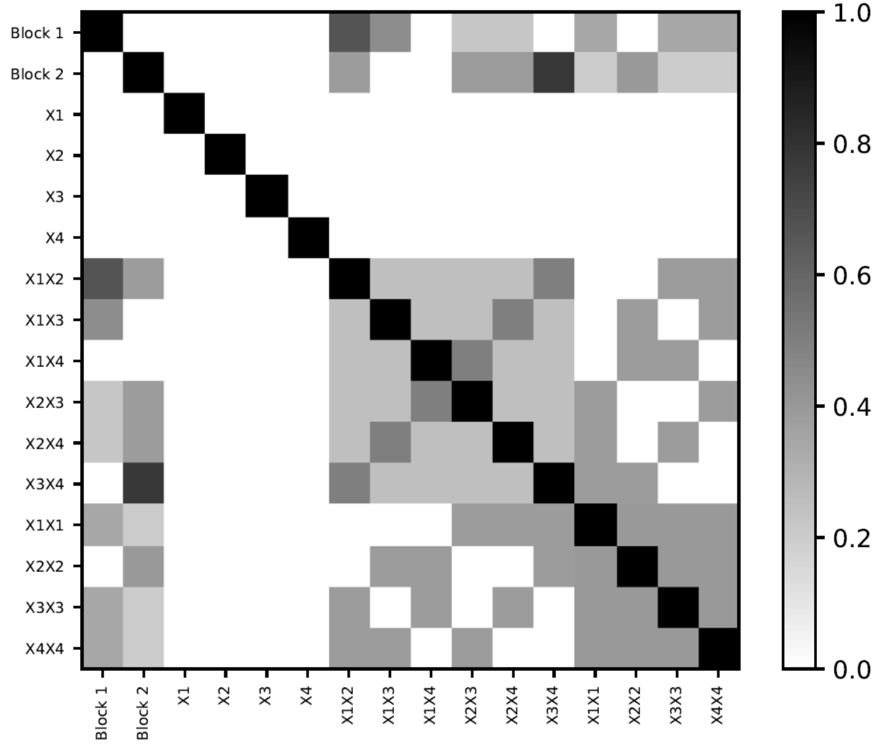
We built a large catalog of blocked (mixed-level) OMARS for both screening and optimization.

A similar design selection approach can be followed for mixed-level designs too.

# *Blocked OMARS: example*

4-factor 15-run definitive screening design

| block 1 | block 2 | block 3 |
|---|---|---|
| - - + 0 | + + + - | + - 0 - |
| 0 + + + | - 0 + - | + - + + |
| 0 - - - | + 0 - + | - + - - |
| + + - 0 | - - - + | - + 0 + |
| 0 0 0 0 | 0 0 0 0 | 0 0 0 0 |

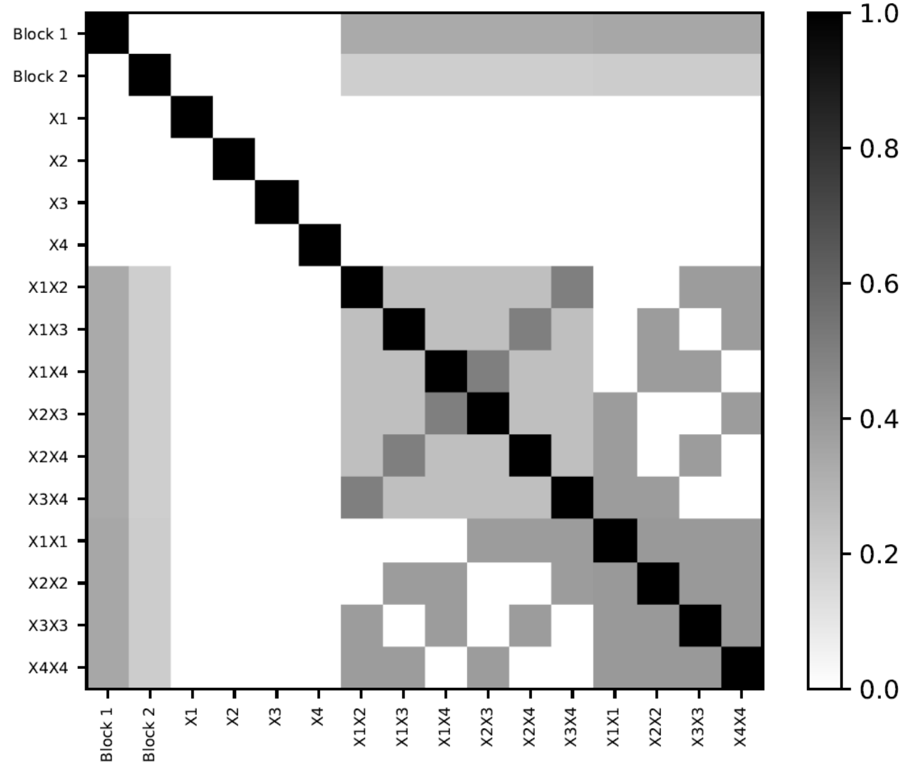



Blocking scheme using JMP16

# *Blocked OMARS: example*

4-factor 15-run definitive screening design

| block 1 | block 2 | block 3 |
|---|---|---|
| 0 + + + | 0 - - - | 0 0 0 0 |
| - - + 0 | + + - 0 | + + + - |
| + 0 - + | - 0 + - | - - - + |
| + - 0 - | + - + + | 0 0 0 0 |
| - + - - | - + 0 + | 0 0 0 0 |

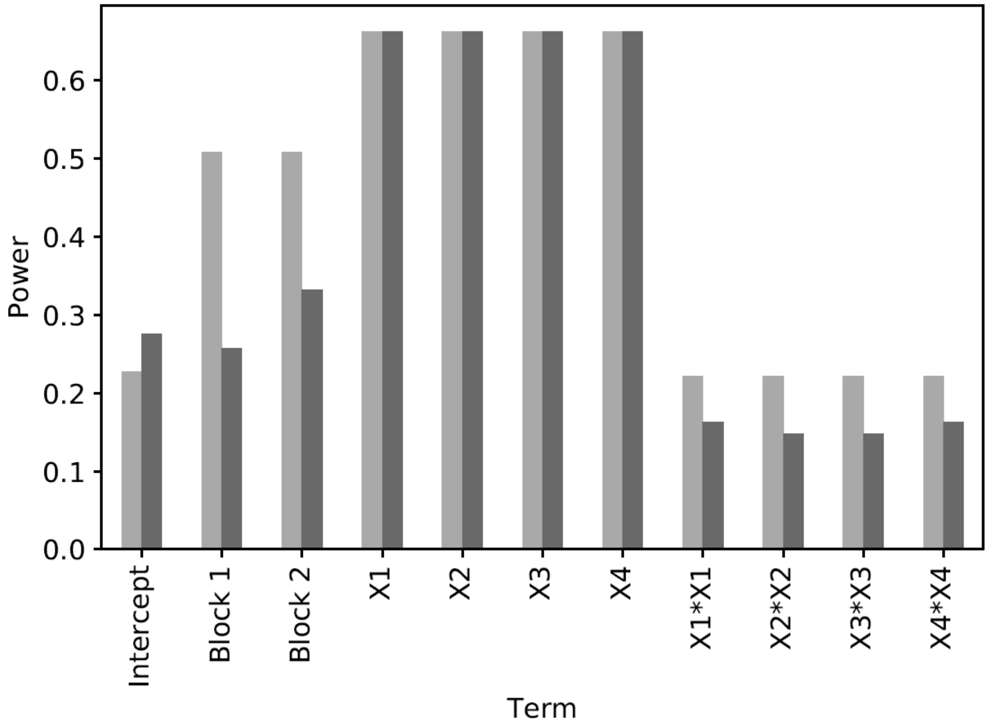



Blocking scheme using our approach

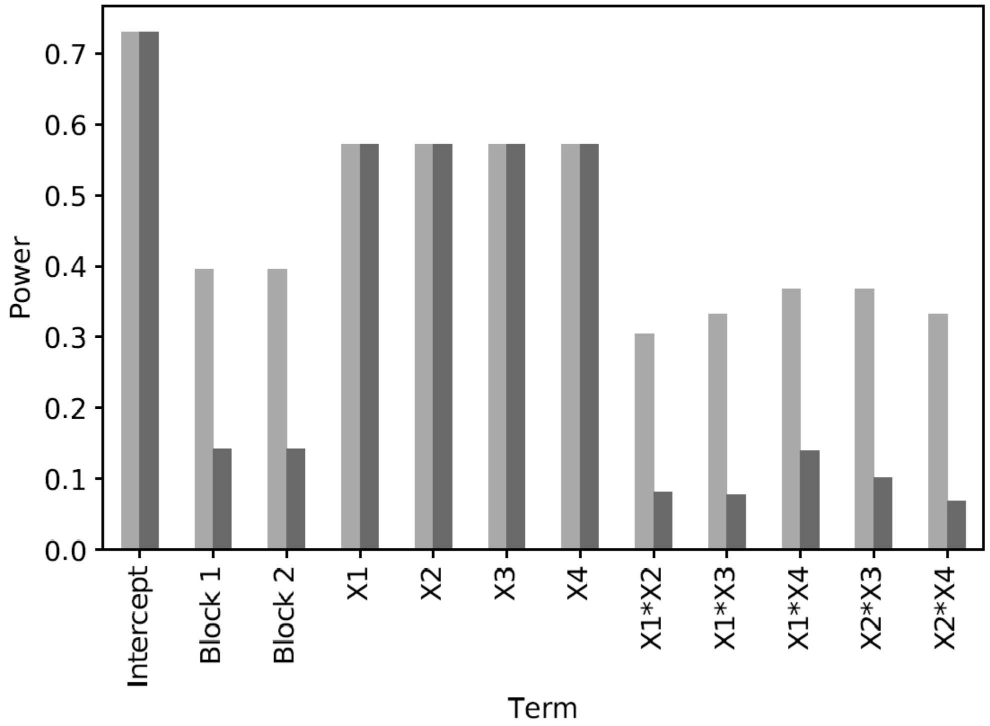
# Blocked OMARS: example

Powers to detect the effects in two models

JMP16 EFFEX




Model 1




Model 2

# *The best way to design an experiment*

Using Design of Experiments → Having an experimental design is a necessary condition.

Where to find an experimental design?

**Generate it on the fly**

- Flexible
- Single criterion optimization
- Uncertainty on:
  - Generation time
  - Quality

**Choose an existing design**

- In books, articles and online catalogs
- Nonflexible
- The best designs:
  - have a name, and
  - have been studied in detail.

# *Existing catalogs of experimental designs*

**Orthogonal arrays** in (Hedayat et al 1999), (Schoen et al 2010) (Eendebak et al. 2023)

**Definitive screening designs** in (Schoen, Eendebak, Vázquez & Goos 2022)

**OMARS designs** in (Núñez Ares and Goos 2020), (Núñez Ares, Schoen and Goos 2023), Núñez Ares and Goos 2023)

**D-optimal two-level screening designs with a number of runs not multiple of four** in (King et al 2020)

**What if we would have a catalog with all the best designs for the majority of the industrial cases?**

# References


Núñez Ares, J., & Goos, P. (2020). Enumeration and multicriteria selection of orthogonal minimally aliased response surface designs. *Technometrics*, *62*(1), 21-36.

Núñez Ares, J., Schoen, E. D., & Goos, P. (2023). Orthogonal minimally aliased response surface designs for three-level quantitative factors and two-level categorical factors. *Statistica Sinica.-Taiwan*, *33*(1), 107-126.

Núñez Ares, J., & Goos, P. (2023). Blocking OMARS designs and definitive screening designs. *Journal of Quality Technology*, 1-21.

Hedayat, A. S., Sloane, N. J. A., & Stufken, J. (1999). *Orthogonal arrays: theory and applications*. Springer Science & Business Media.

Schoen, E. D., Eendebak, P. T., & Nguyen, M. V. (2010). Complete enumeration of pure-level and mixed-level orthogonal arrays. *Journal of Combinatorial Designs*, *18*(2), 123-140.

King, C, Jones, B, Morgan, J, Lekivetz, R. Direct construction of globally d-optimal designs for factors at two levels and main effects models. *Qual Reliab Engng Int*. 2020; 36: 797–816

Eendebak, P. T., Schoen, E. D., Vazquez, A. R., & Goos, P. (2023). Systematic enumeration of two-level even-odd designs of strength 3. *Computational Statistics & Data Analysis*, *180*, 107678.