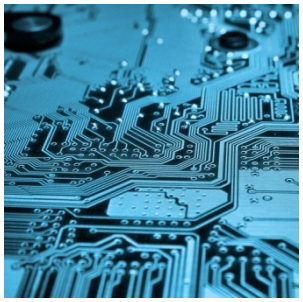
BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

Design of Experiments in R

Prof. Ulrike Grömping
Beuth University of Applied Sciences
Berlin



Outline of Talk

- Quick overview of history and status for Design of Experiments (DoE) in R
- Example-based review of DoE principles
- Further application areas and principles
- The landscape of R packages for DoE

History

- **conf.design** 2001, **IdDesign** 2003, **AlgDesign crossdes BsMD** all 2004
  - **IdDesign** has been archived in December 2013, since its maintainer Roderick D Ball from New Zealand was no longer contactable (happily retired? dead? …).
  - **AlgDesign maintained by Jerome Braun** after Bob Wheeler's death
- My own involvement
  - started small in 2007, aiming for some Minitab-like effects plots from R, which was the birth of **FrF2**
  - gained momentum in 2009 with a sabbatical semester (packages **DoE.base**, **DoE.wrapper**, **RcmdrPlugin.DoE**, later **FrF2.catlg128**)
- Apparently, DoE in R gained momentum around 2009.
- 109 packages in CRAN Task View on Experimental Design
  - boundary to other topics a bit wobbly (grey zone; when is a package too specialized or too unspecific?)

History

- **conf.design** 001, **IdDesign** 2003, **AlgDesign crossdes BsMD** all 2004
  - How to confound effects
  - chived in December 2013,
    since its maintainer Roderick D Ball from New Zealand was no longer contactable (happily retired? dead? …).
  - **AlgDesign maintained by Jerome Braun** after Bob Wheeler's death
- My own involvement
  - started small in 2007,
    aiming for some Minitab-like effects plots from R,
    which was the birth of **FrF2**
  - gained momentum in 2009 with a sabbatical semester (packages **DoE.base**, **DoE.wrapper**, **RcmdrPlugin.DoE**, later **FrF2.catlg128**)
- Apparently, DoE in R gained momentum around 2009.
- 109 packages in CRAN Task View on Experimental Design
  - boundary to other topics a bit wobbly
    (grey zone; when is a package too specialized or too unspecific?)

History

- **conf.design** 2001, **IdDesign** 2003, **AlgDesign crossdes BsMD** all 2004
  - **IdDesign** has been archived in December 2013, since its maintainer Roderick D Ball from New Zealand was no longer contactable (happily retired? dead? …).
  - **AlgDesign maintained by Jerome Braun** after Bob Wheeler's death
- My own involvement
  - started small in 2007, aiming for some Minitab-like effects plots from R, which was the birth of **FrF2**
  - gained momentum in 2009 with a sabbatical semester (packages **DoE.base**, **DoE.wrapper**, **RcmdrPlugin.DoE**, later **FrF2.catlg128**)
- Apparently, DoE in R gained momentum around 2009.
- 109 packages in CRAN Task View on Experimental Design
  - boundary to other topics a bit wobbly (grey zone; when is a package too specialized or too unspecific?)

History

- **conf.design** 2001, **ldDesign** 2003, **AlgDesign** **crossdes BsMD** all 2004
  - **ldDesign** has been archived in [illegible] since its maintainer Roderick D [illegible] contactable (happily retired? dead? …).
  - **AlgDesign** **maintained by Jerome Braun** after Bob Wheeler's death

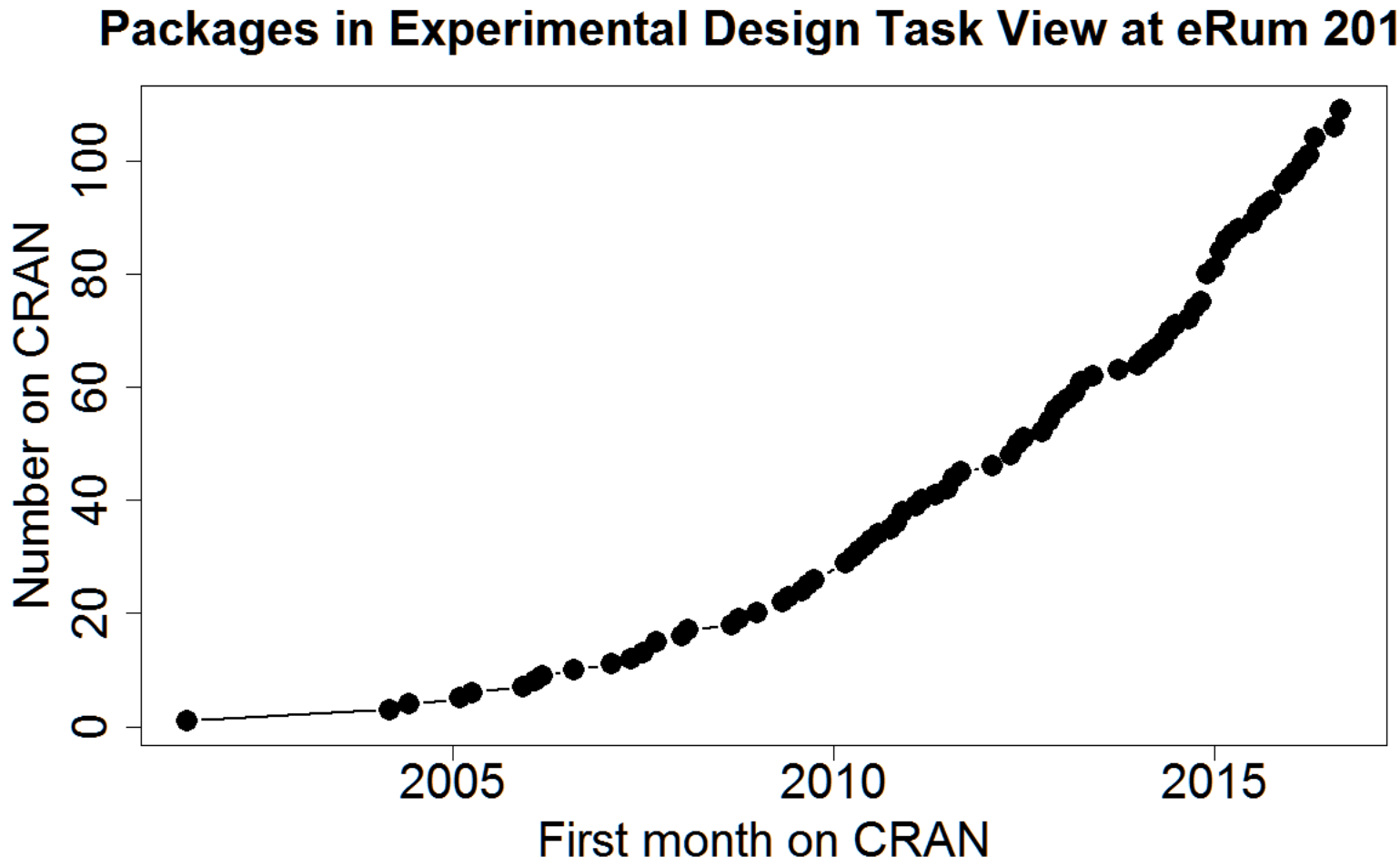
Optimal Design (with Fedorov alg.)

- My own involvement
  - started small in 2007,
    aiming for some Minitab-like effects plots from R,
    which was the birth of **FrF2**
  - gained momentum in 2009 with a sabbatical semester (packages **DoE.base**, **DoE.wrapper**, **RcmdrPlugin.DoE**, later **FrF2.catlg128**)
- Apparently, DoE in R gained momentum around 2009.
- 109 packages in CRAN Task View on Experimental Design
  - boundary to other topics a bit wobbly
    (grey zone; when is a package too specialized or too unspecific?)

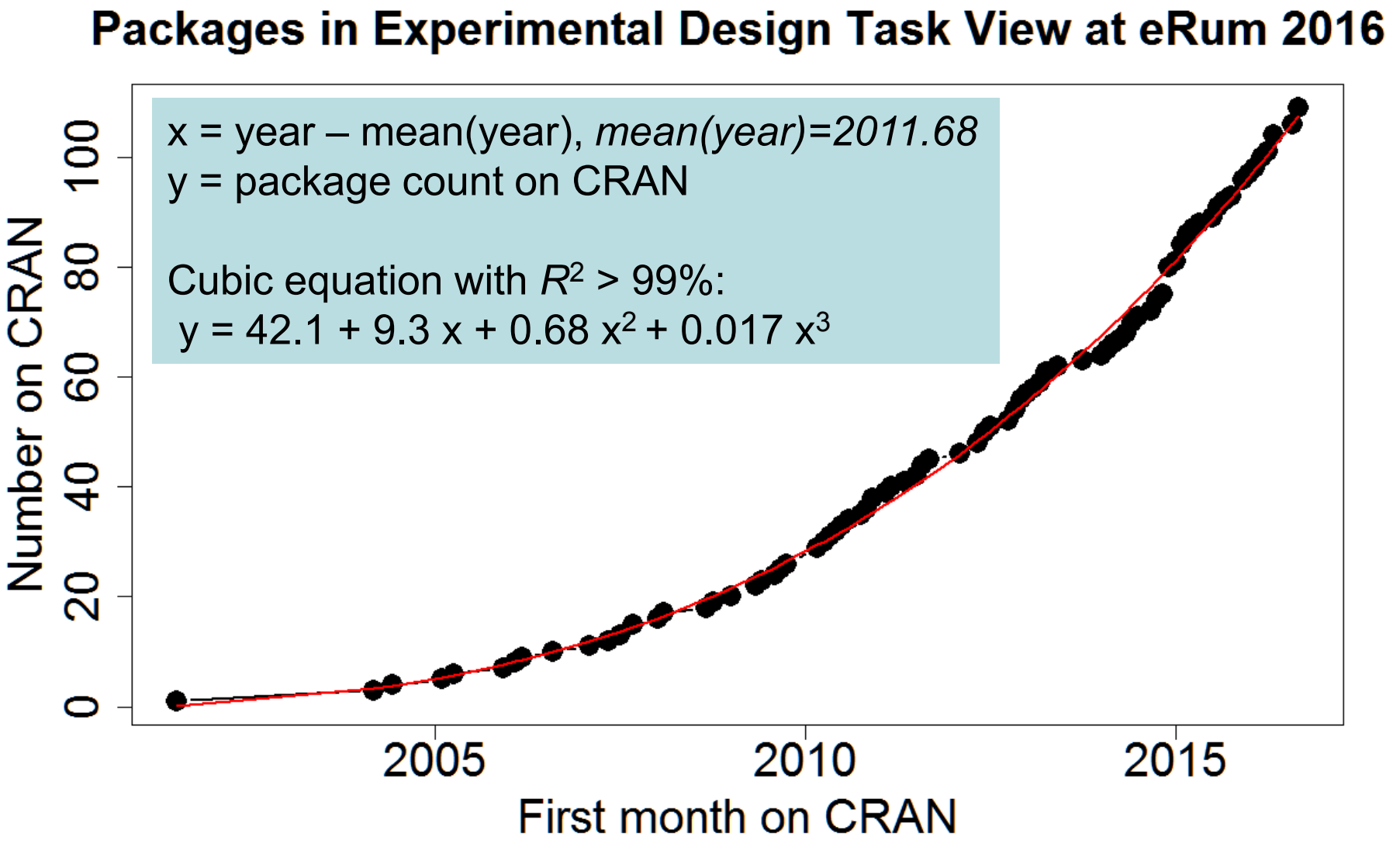
History

- **conf.design** 2001, **IdDesign** 2003, **AlgDesign crossdes BsMD** all 2004
  - **IdDesign** has been archived in December 2013, since its maintainer Roderick D Ball from New Zealand was no longer contactable (happily retired? dead? …).
  - **AlgDesign maintained by Jerome Braun** after Bob Wheeler's death
- My own involvement
  - started small in 2007, aiming for some Minitab-like effects plots from R, which was the birth of **FrF2**
  - gained momentum in 2009 with a sabbatical semester (packages **DoE.base**, **DoE.wrapper**, **RcmdrPlugin.DoE**, later **FrF2.catlg128**)
- Apparently, DoE in R gained momentum around 2009.
- 109 packages in CRAN Task View on Experimental Design
  - boundary to other topics a bit wobbly (grey zone; when is a package too specialized or too unspecific?)

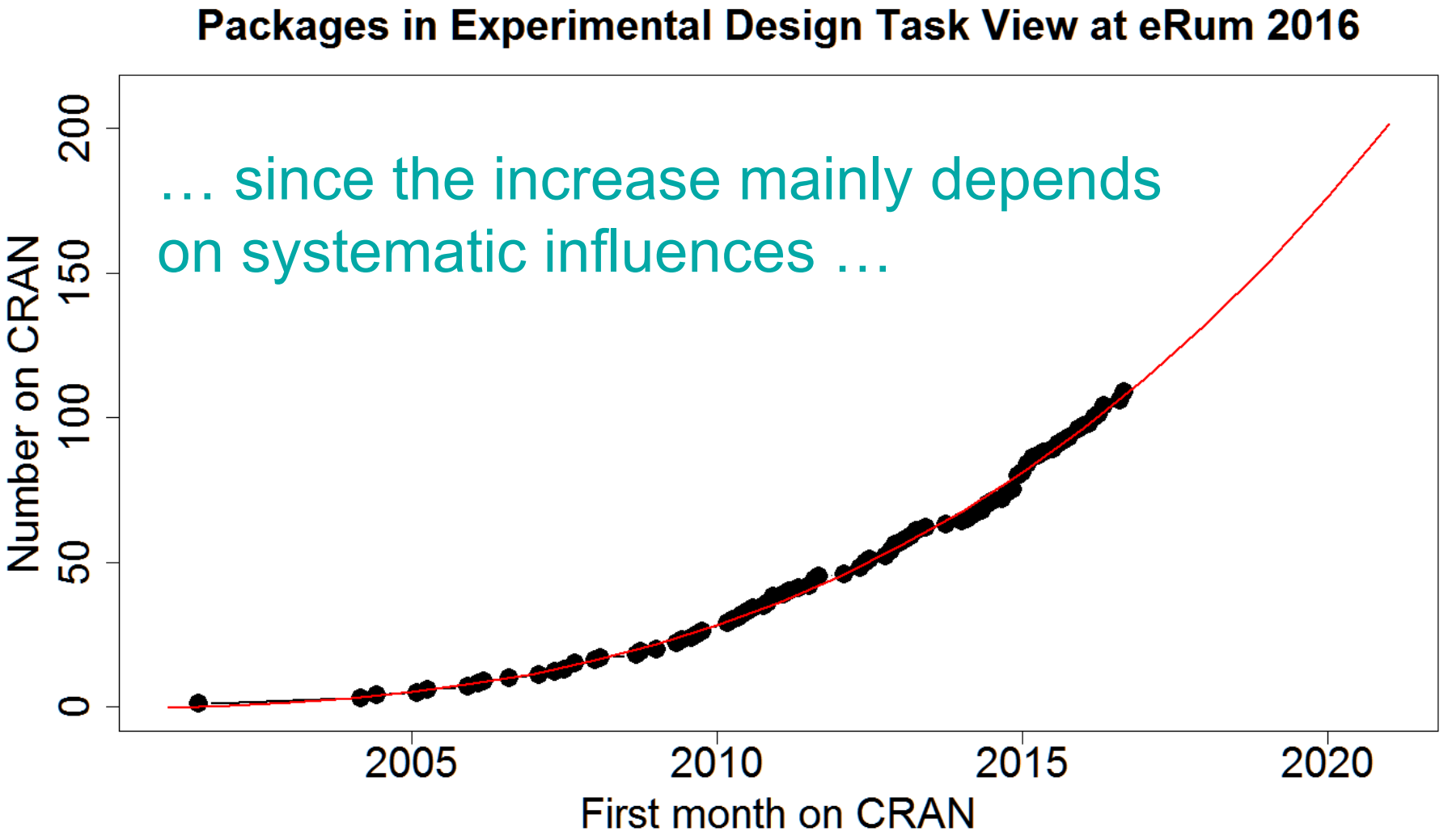
R continues embracing DoE



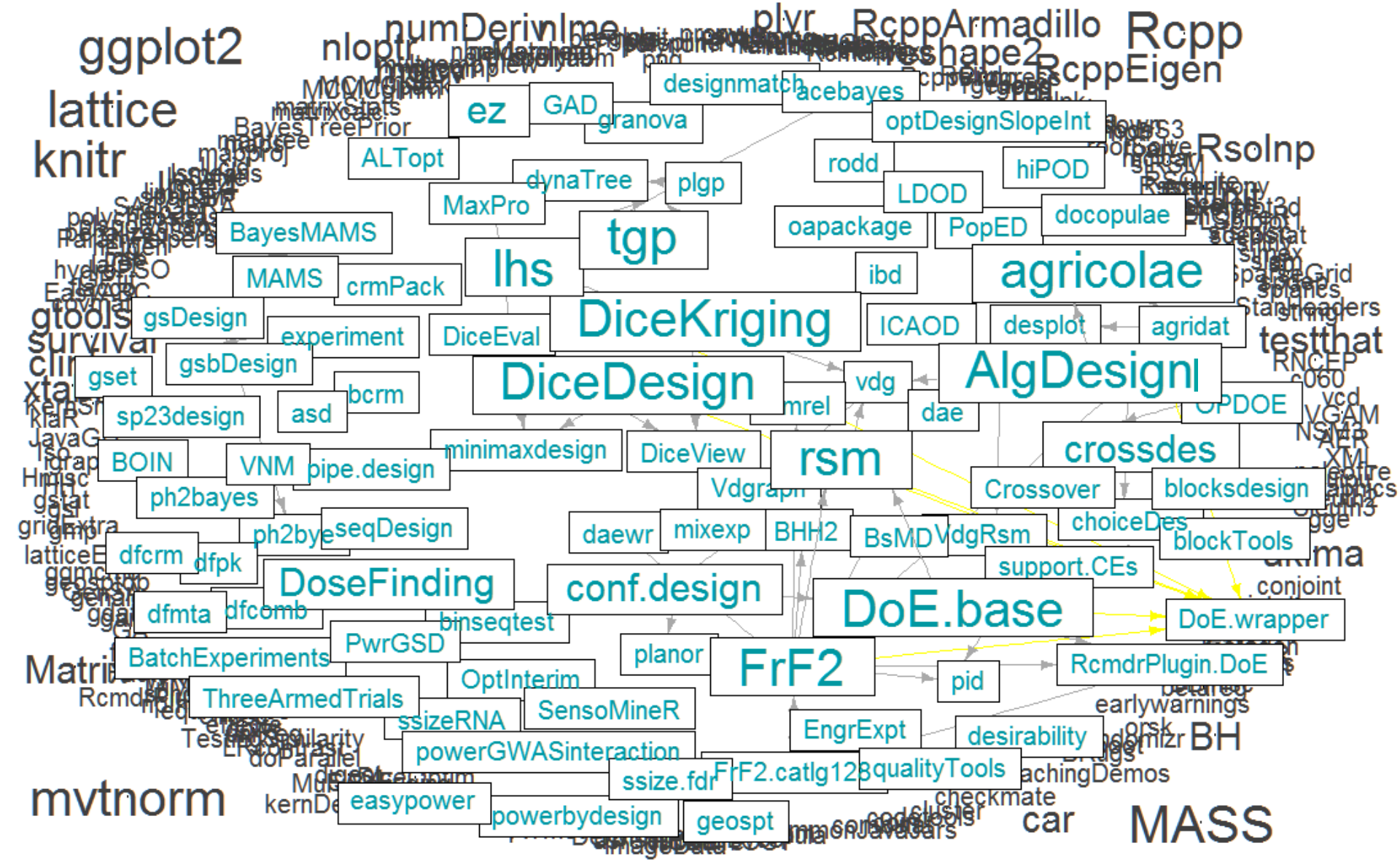
R continues embracing DoE



Prediction not very reliable …



Experimental Design Task View



Created with package **igraph**

DoE package popularity (**cranlogs**)

- Top 30 downloads from Rstudio servers Oct 07 2015 to Oct 06 2016

	count	year		count	year
agricolae	73232	2007	DoseFinding	6337	2010
AlgDesign	61519	2004	SensoMineR	6267	2005
ez	37405	2009	DiceKriging	6237	2010
lhs	22800	2006	agridat	5734	2011
DoE.base	19216	2009	crossdes	4932	2004
conf.design	15153	2001	qualityTools	4761	2010
FrF2	15144	2007	gsDesign	4724	2009
tgp	14540	2006	BatchExperiments	4597	2012
rsm	14284	2008	geospt	4620	2012
BsMD	12909	2004	desirability	4534	2007
BHH2	10268	2005	GAD	4102	2010
DiceDesign	8732	2010	asd	4080	2012
RcmdrPlugin.DoE	7256	2009	daewr	4024	2012
dae	7244	2010	bcrm	4018	2012
DoE.wrapper	7163	2009	blocksdesign	4006	2014

Examples: Obama 1

- Simple experiment from first Obama electoral campaign:
  What splash screen to choose on the campaign website
  for making people sign up for the Obama campaign ?

  - Factor **Button**
    with levels

    1 JOIN US NOW

    2 LEARN MORE

    3 SIGN UP NOW

    4 SIGN UP

  Factor **Medium**
  with levels

  1 Original (Obama pic)

  2 Family (pic w. Obama)

  3 Change (pic)

  4 Obama‘s video

  5 Sam‘s video

  6 Springfield video

- See more on results at https://blog.optimizely.com/2010/11/29/how-obama-raised-60-million-by-running-a-simple-experiment/
  (blog entry by Dan Siroker, Director of Analytics for the Obama 2008 campaign)

Examples: Obama 1

- **Full factorial design**:
  all 4 x 6 = 24 level combinations of the two factors can be run
- Large experiment: total of 310,382 people included,
  thus almost about 13,000 per combination
  - **replication**: the same combination is redone
  - **randomization**: the combination is randomly assigned to each visitor
    → avoiding unwanted side effects from any systematic assignment
- Reasons for having many replications:
  - individual replications were cheap
  - there is a lot of variability in the behavior of humans
    → replications useful
  - the response is a rate, the conversion rate,
    a sizable sample for each combination is helpful.

Examples: Obama 2

- From talk by Rayid Ghani, Chief Scientist for Obama 2012 Campaign:
- Many small experiments were used (small point in the talk), e.g. several experiments on fund raising by e-mail
  - send out mass e-mails
  - to those who signed up on the campaign website
  - asking for money
- Quick experiments, immediately exploit results!

- Interesting to me because of its result:
  - Factor: **How much** money should be asked for (**initially**)? E.g.: asking for 5$ in the subject line
  - Factor: **Donation options** offered in the link

Donate	Donate	Donate
O 5$	O 7$	O 7,50$
O 10$	O 10$	O 15$
O 15$	O 15$	O 20$

Examples: different situations require different experiments

- Obama examples: many simple experiments with many replications each
  - behavior of humans → a lot of variability,
  - results needed quickly,
  - cheap experimental runs,
  - sequential experimentation easy.

- Different situation: biotechnological experiment
  (dataset **`VSGFS`** in R package **DoE.base**,
  from Vasilev, Schmidt, Grömping, Fischer, Schillberg *PLoS ONE* 2014)
  - less (but still relevant) variability
  - less time pressure
  - more expensive experimental runs
  - logistics make sequential experimentation difficult:
    time period for lab facilities needs to be booked
    and used to as efficiently as possible.

Examples: VSGFS

- Transgenic Tobacco Cell Suspension Cultures (in flasks)
- Responses: Biomass [g],
  geraniol content [µg/g],
  geraniol yield [µg]
- Factors:

2	Light	Lght –	Lght +		
2	ShakFreq	SF –	SF +		
2	InocSize	IS –	IS +		
3	FilledVol	FV –	FV 0	FV +	
2	CM	CM –	CM +		
3	Carbo	Sucrose	Glucose	Mannitol	
4	Cyclodextrin	beta	methyl-beta	triacetyl-beta	none

- Unreplicated full factorial would need $2^4 \times 3^2 \times 4 = 576$ experimental runs
  → not affordable
- 72 experimental runs + some baseline control runs doable in one go
- Unreplicated
- Run order **randomized**

Examples: VSGFS OA

- Orthogonal array (OA) based on an OA(72, $2^{43}\,3^{8}\,4^{1}\,6^{1}$) from **DoE.base**:
  - all main effects can be estimated orthogonally to each other *per se*
  - column selection was optimized
    with function **`oa.design`** from package **DoE.base** such that

prio 1
  - confounding of main effects with 2-factor-interactions minimal

prio 2
  - confounding among 2-factor interactions or
    of main effects with 3-factor interactions subsequently minimized

  - no specific model assumed,
    only priority on main effect estimation

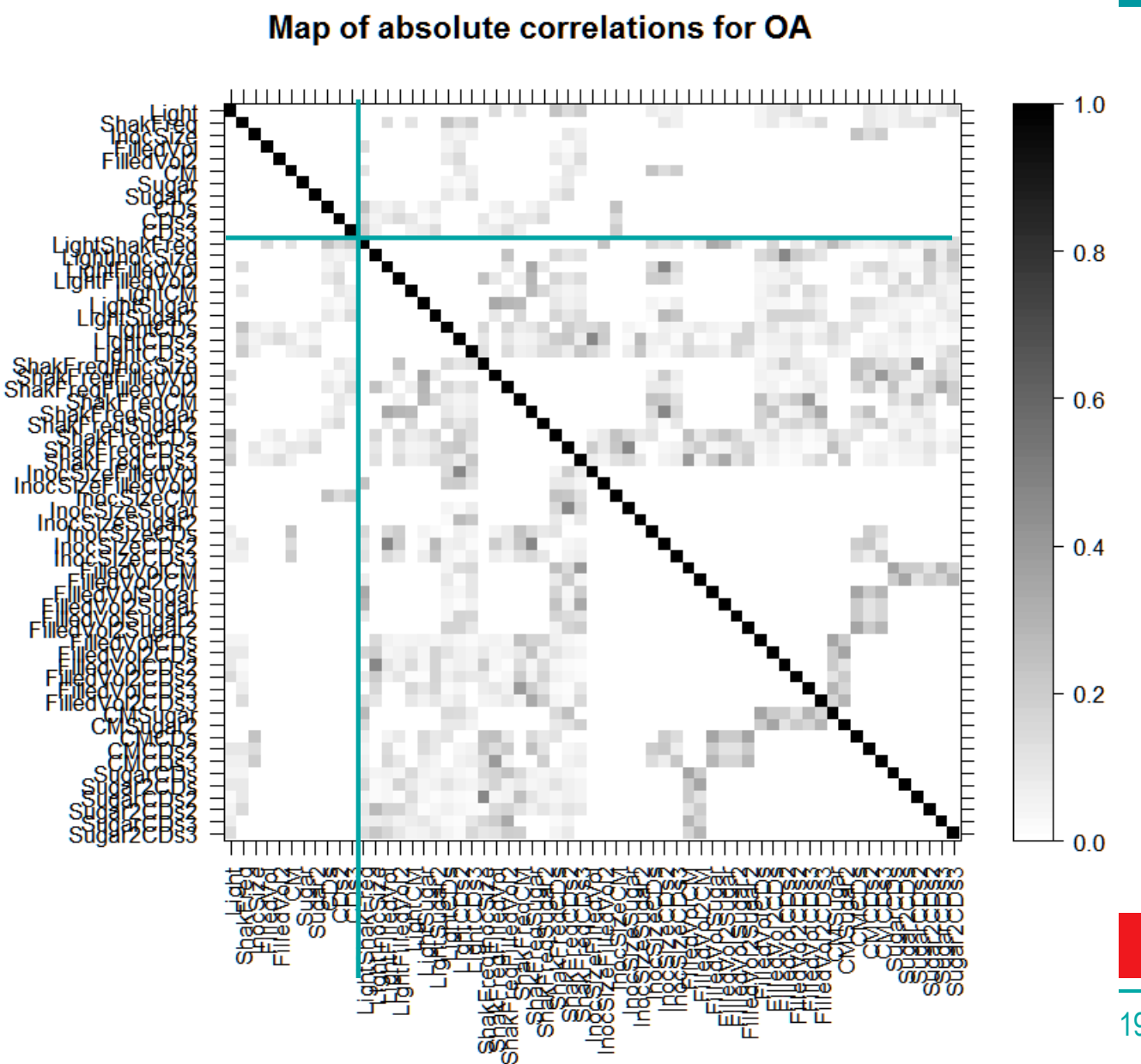
- **X** is model matrix for model with Intercept, main effects and 2fis

Examples: VSGFS OA correlations

OA:
little confounding
with main effects

more confounding
among 2fis

confounding
patterned

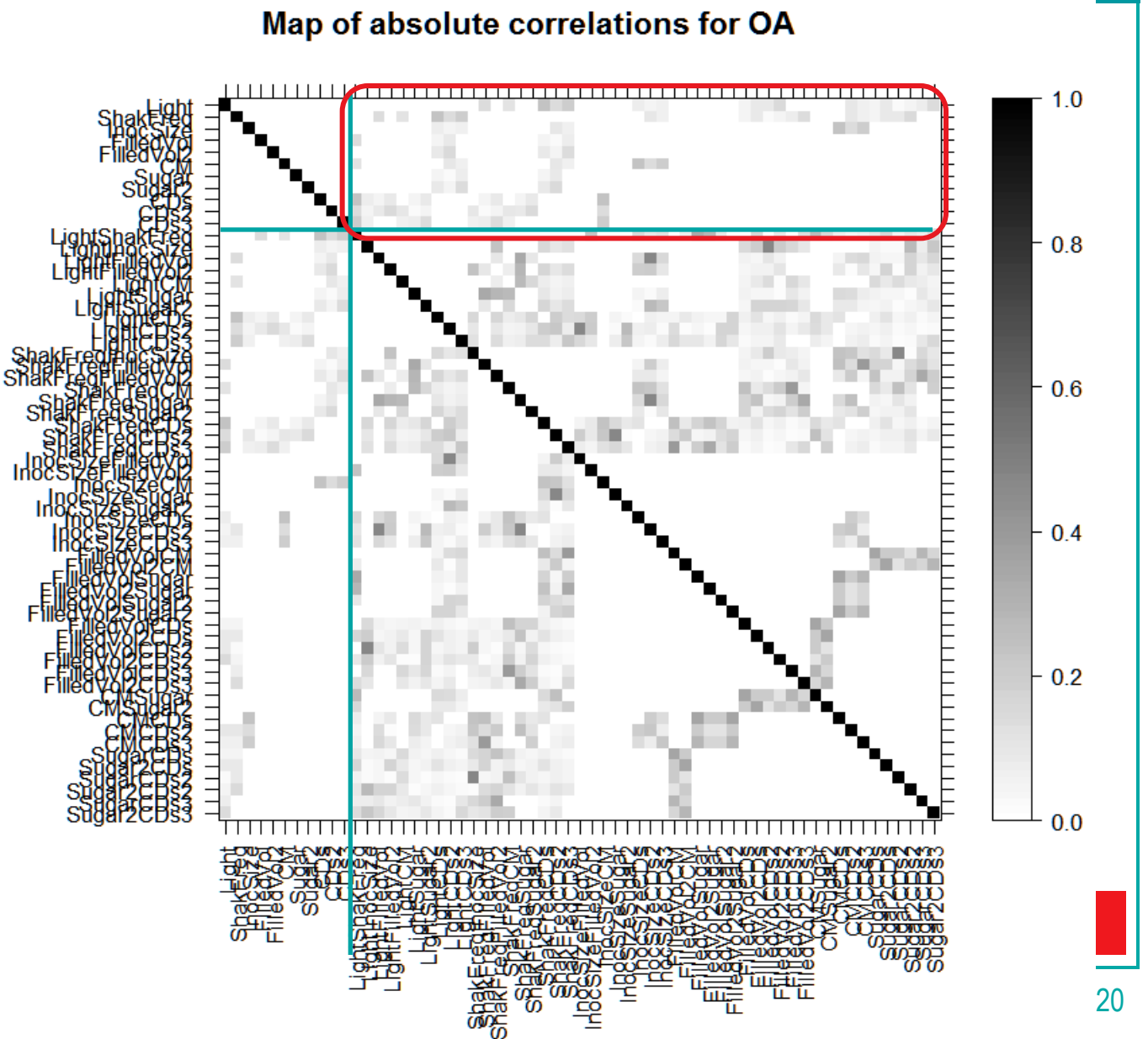


Examples: VSGFS OA correlations

OA:
little confounding
with main effects

more confounding
among 2fis

confounding
patterned



Examples: VSGFS OA-alternative

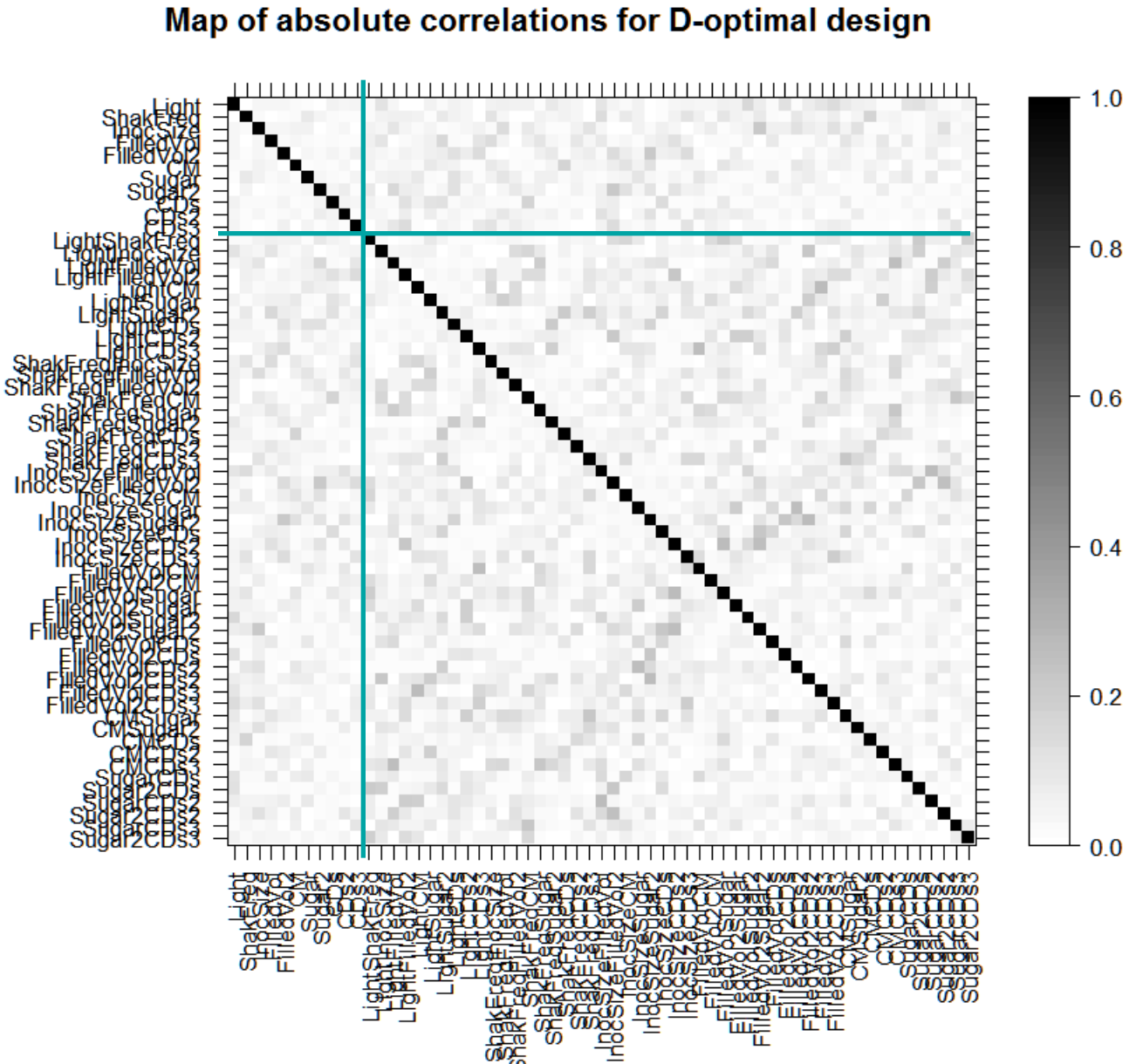
- The 72 run OA is (of course) not suitable for in-depth analysis of sophisticated models,
  it is rather a screening experiment (see also later).
  - doable: linear model with main effects only, 60 df for error
    perhaps linear model with main effects and some 2fis
    linear model with main effects and *all* 2fis, 10 df for error

- Alternative:
  D-optimal design for 72 runs
  assuming factorial model with main effects and all 2-factor interactions:
  - D-optimality: maximize the determinant of $\mathbf{X}^T\mathbf{X}$
    → minimize hypervolume of the confidence region of coefficient vector
  - Same priority on estimation of main effects and 2-factor interactions

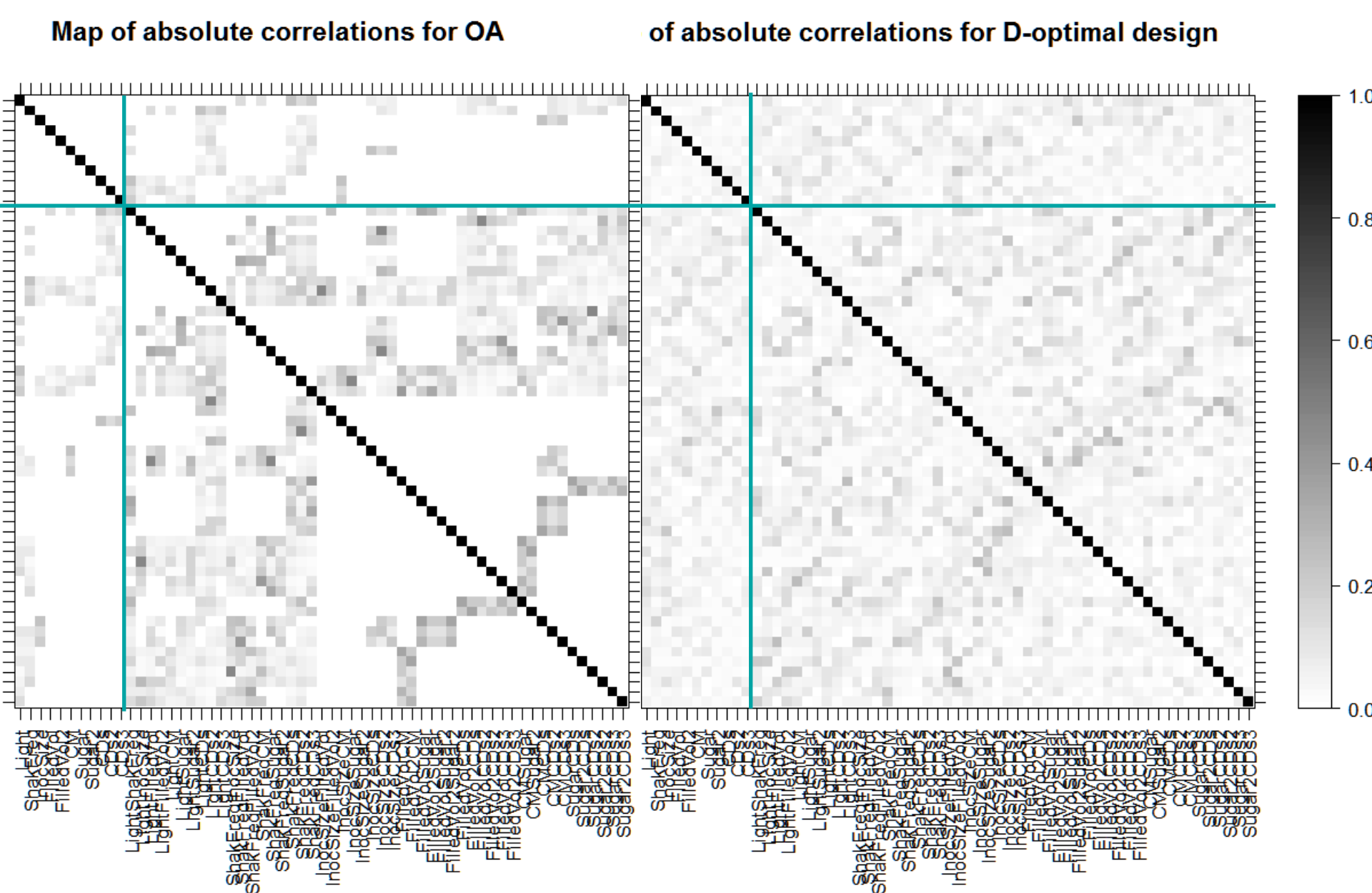
created with package **AlgDesign**,
using convenience function **`Dopt.design`**
from **DoE.wrapper**

Examples: VSGFS D-optimal

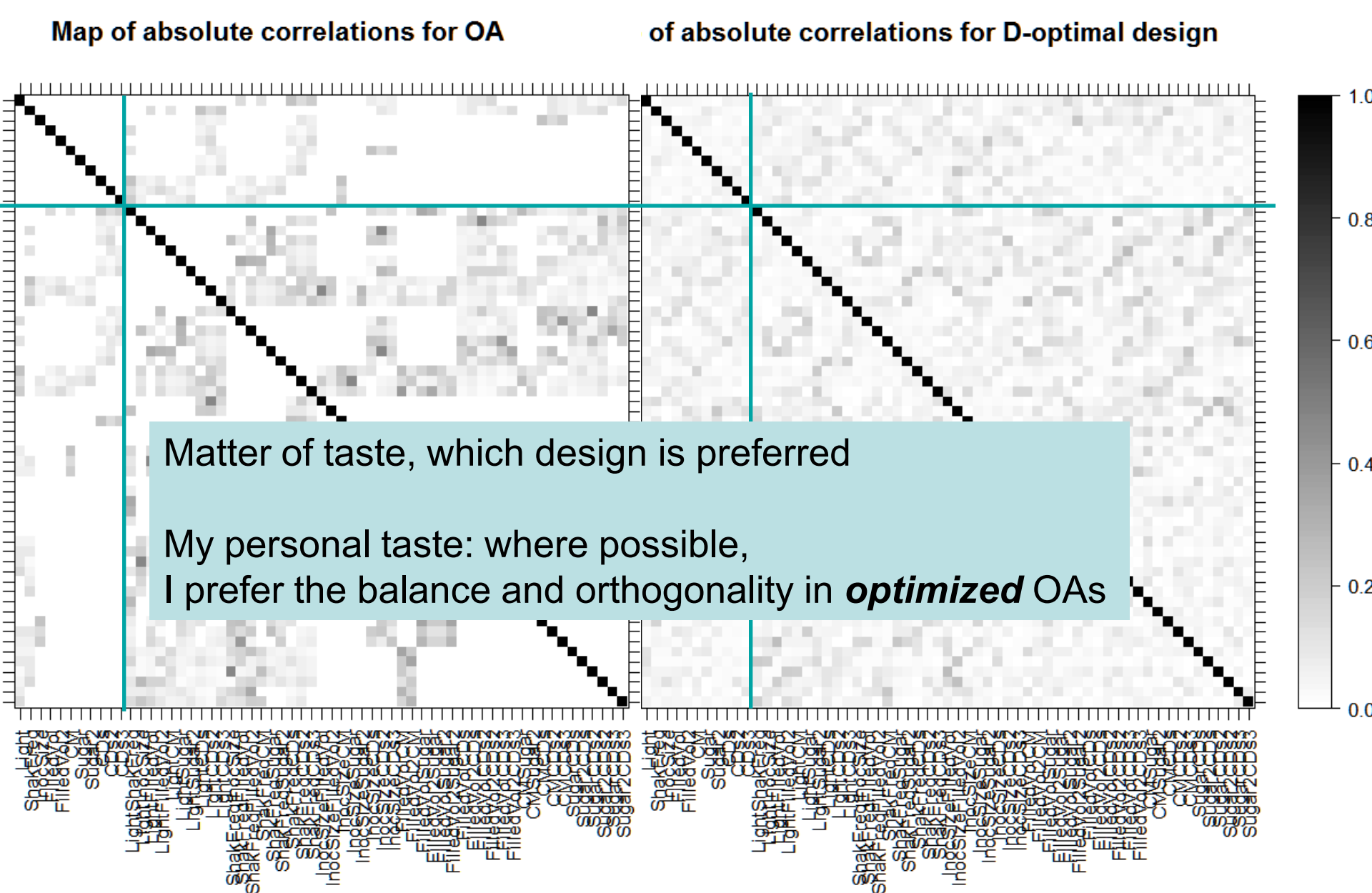
D-optimal:
confounding more
equally distributed



OA vs. D-optimal



OA vs. D-optimal



Matter of taste, which design is preferred

My personal taste: where possible,
I prefer the balance and orthogonality in ***optimized*** OAs

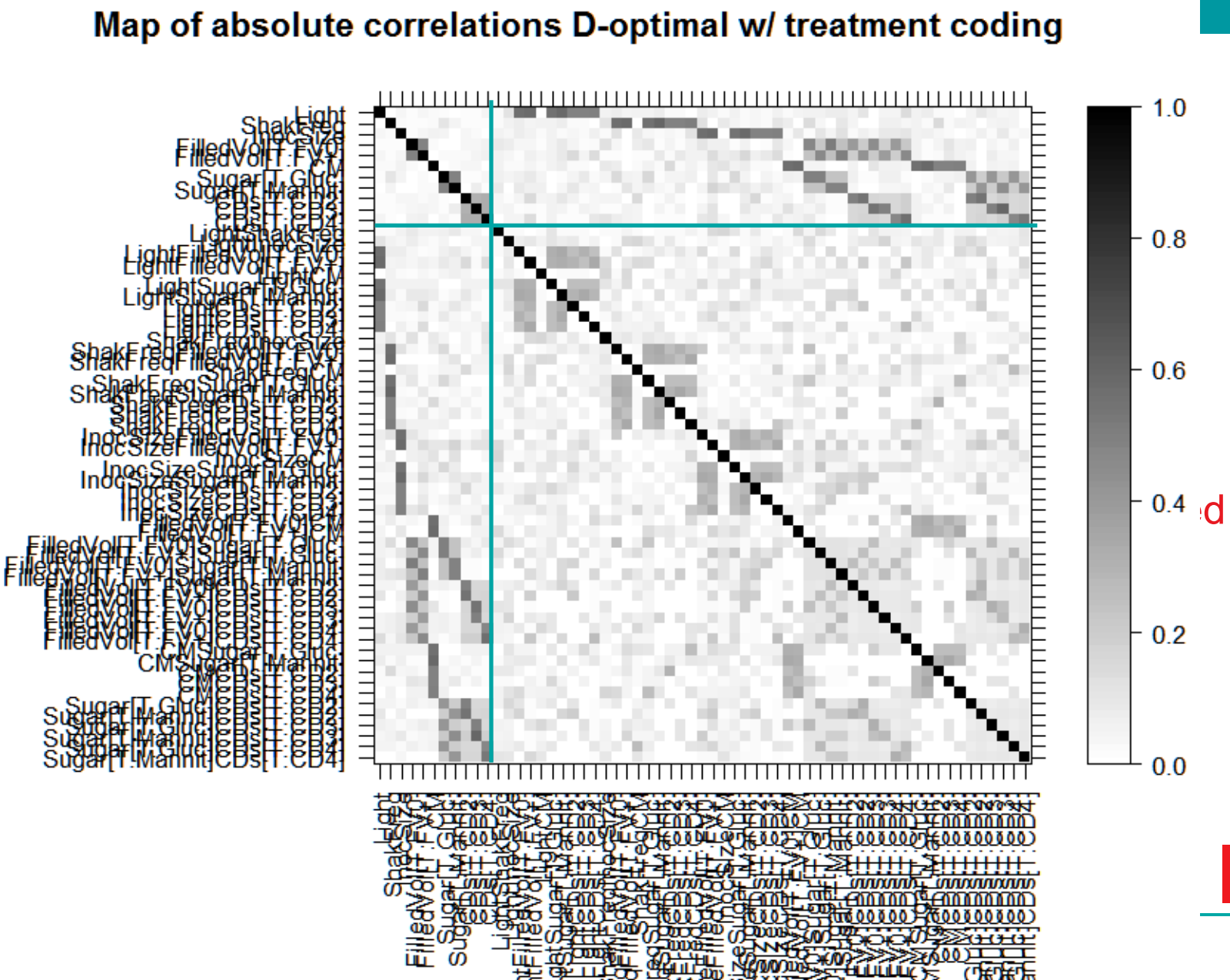
An aside: Coding of factors!

- Covariance matrix of coefficient estimates (of course) depends on the coding of factors
  !!! unwanted for qualitative factors !!!
  - Orthogonal coding with columns normalized such that, in a full factorial, all model coefficients
    - would have the same variance
    - and would be uncorrelated to each other

    `contr.XuWu` or **`contr.XuWuPoly`** in package **DoE.base**
  - If such recoding is not applied, the confounding pattern is dominated by artificial coding dependencies!

An aside: Coding of factors!

- Co
  de
  !!!



Application areas

- We‘ve seen examples from web design and mass communication, and from biotechnological research.

Application areas

- We‘ve seen examples from web design and mass communication, and from biotechnological research.
- **The classic application area**: agricultural experiments (Sir R. A. Fisher)
  Special aspects:
  - spatial structure with neighboring effects from plowing directions, soil similarities, …
  - blocks, plots, sub plots etc. – randomization restrictions are common

Application areas

- We‘ve seen examples from web design and mass communication, and from biotechnological research.
- **The classic application area**: agricultural experiments (Sir R. A. Fisher)
  Special aspects:
  - spatial structure with neighboring effects from plowing directions, soil similarities, …
  - blocks, plots, sub plots

Blocking is a principle of accounting for known relevant but uninteresting influences: Include the known influence as a block factor.

Application areas

- We‘ve seen examples from web design and mass communication, and from biotechnological research.
- **The classic application area**: agricultural experiments (Sir R. A. Fisher)
  Special aspects:
  - spatial structure with neighboring effects from plowing directions, soil similarities, …
  - blocks, plots, sub plots etc. – randomization restrictions are common

Randomization **hierarchy levels** need to be accounted for in the analysis.

Application areas

- We‘ve seen examples from web design and mass communication, and from biotechnological research.
- **The classic application area**: agricultural experiments (Sir R. A. Fisher) Special aspects:
  - spatial structure with neighboring effects from plowing directions, soil similarities, …
  - blocks, plots, sub plots etc. – randomization restrictions are common
- Further application areas (all with their special aspects) include e.g.
  - clinical trials (many R packages)
  - pharmacological experiments
  - population genetics
  - market research (choice designs …),
  - sensory studies
  - psychology
  - engineering

This is my main field of experience

Engineering approach to DoE

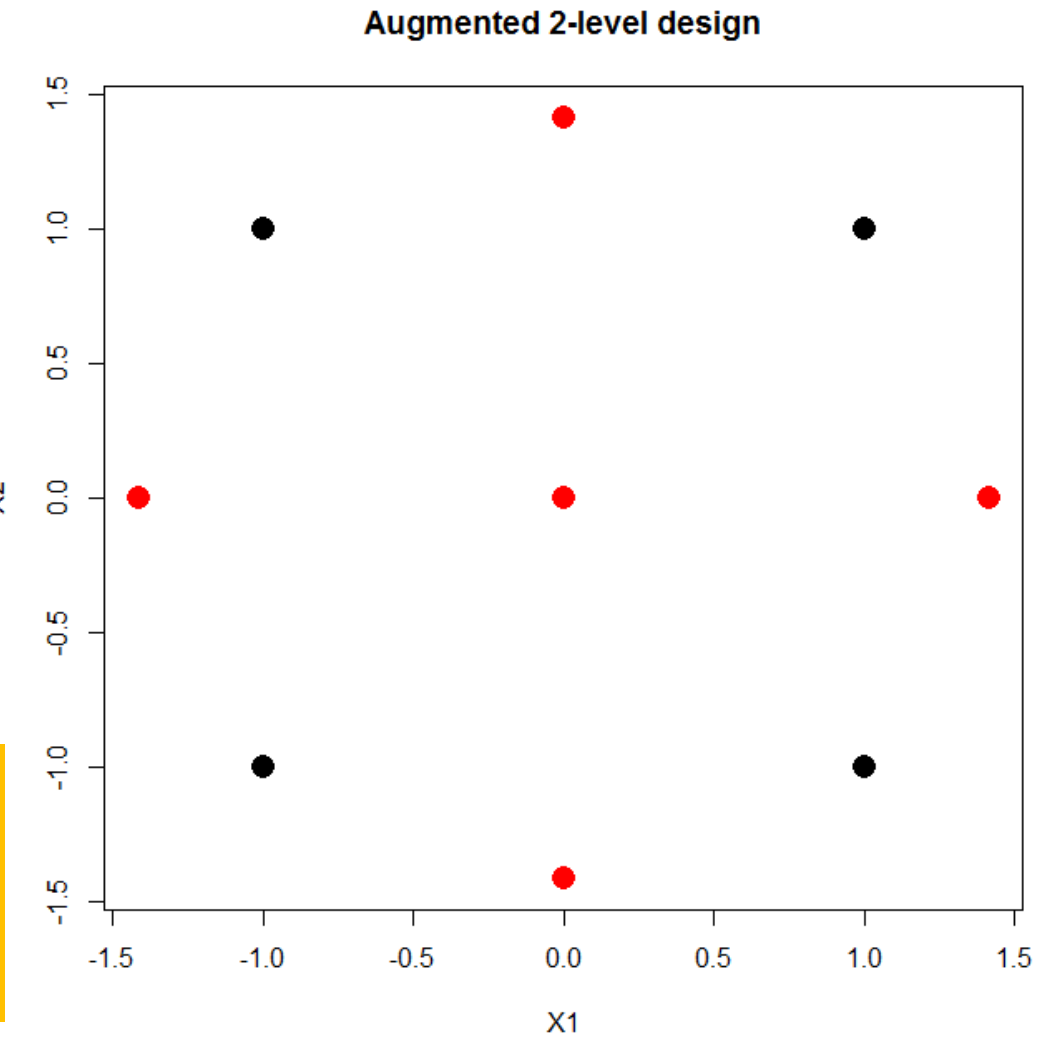
- Ideal procedure: sequential
  - screening designs: many factors, identify (most) relevant ones
    concentrate on main effects estimation
    (relevant packages **DoE.base**, **FrF2** (`pb`, `FrF2`), **BHH2**)
  - follow-up experiments after a screening phase
    - higher resolution factorial designs *(less confounding)*
      (**DoE.base**, **FrF2** `FrF2`, **planor**),
- Especially for quantitative factors

  2-level factorial design possibly with center points (**FrF2**)
  - possibly moving the design space towards more promising regions by steepest ascent / descent (package **rsm**)
    - response surface designs in the promising region for quantitative factors (**rsm**, `rsm`)
      (e.g., for full quadratic model; idea: Taylor approximation)

Engineering approach to DoE

- Ideal procedure: sequ
  - screening designs
    concentrate on ma
    (relevant packages
  - follow-up experime
    - higher resolutic
      (**DoE.base**, **Fr**
- Especially for quantita
  2-level factoria
    - possibly m
      regions by

Created with **`ccd`** of **rsm** through function **`ccd.design`** of **DoE.wrapper**



on)

Important in Engineering: Fractional factorial 2-level designs

Factors A	B	C	D ABC
–	–	–	–
+	–	–	+
–	+	–	+
+	+	–	–
–	–	+	+
+	–	+	–
–	+	+	–
+	+	+	+

Full factorial 2^3

full factorial in A, B, C

D allocated to the 3-factor interaction

Result is a half-fraction of a full factorial in the four factors.

D=ABC:
confounding by generating contrasts → regular designs

Important in Engineering: Fractional factorial 2-level designs

Factors A	B	C	D
			ABC
–	–	–	–
+	–	–	+
–	+	–	+
+	+	–	–
–	–	+	+
+	–	+	–
–	+	+	–
+	+	+	+

full factorial in A, B, C

D allocated to the 3-factor interaction

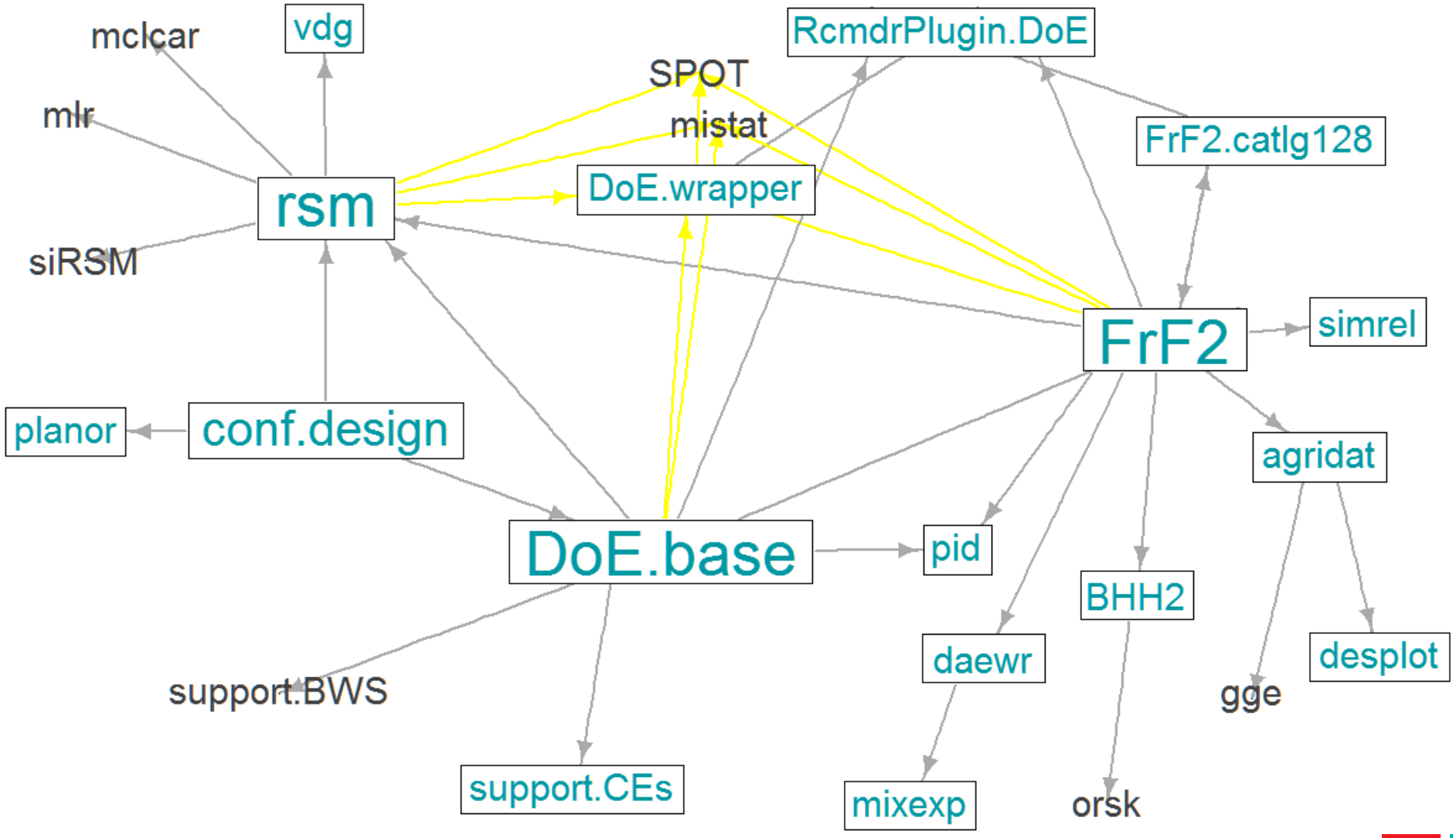
Result is a half-fraction of a full factorial in the four factors.

Full factorial 2^3

D=ABC:
confounding by generating contrasts → regular designs

This type of factor assignment
and its consequences is dealt with
by **conf.design** (more generally than only for 2-level designs)
and with more features by **FrF2** (2-level only) and **planor** (general).

Descendants of **conf.design**



Engineering: Implicit replication

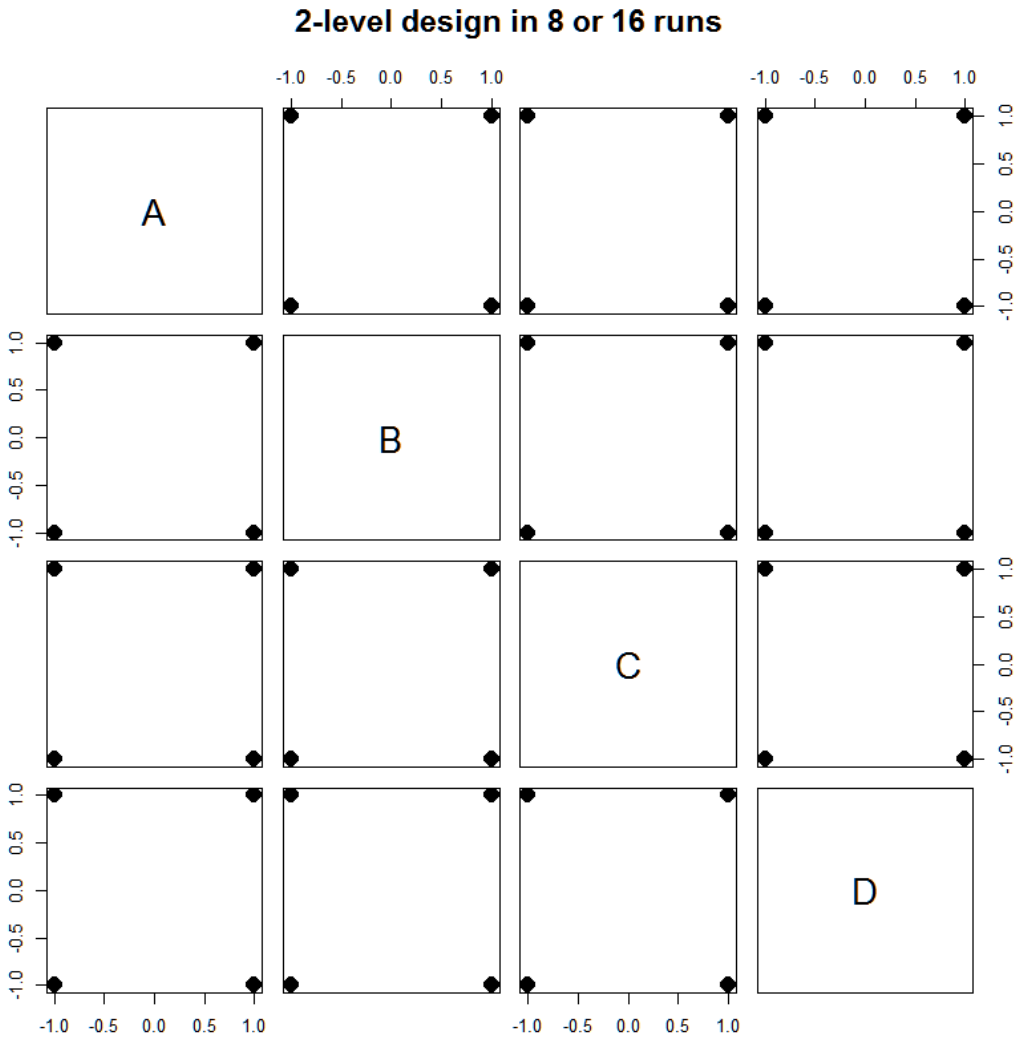
Factors A	B	C	D ABC
−	−	−	−
+	−	−	+
−	+	−	+
+	+	−	−
−	−	+	+
+	−	+	−
−	+	+	−
+	+	+	+

Now, assume that 16 runs are affordable:

- replicate the 8 run design (8 df for error)
- or run an unreplicated full factorial in the four factors?
  - advantage: no confounding between 2fis
  - likely at least 5 error df from the 4fi and the 3fis
  - If one of the factors proves irrelevant, the remaining 3-factor design has replicates: **implicit replication** in 3-factor projection.

Engineering: Implicit replication

Factors A	B
−	−
+	−
−	+
+	+
−	−
+	−
−	+
+	+



Now, assume that

- replicate the
- or run an ur
  - advanta
  - likely at

Design created with `FrF2` from **FrF2**

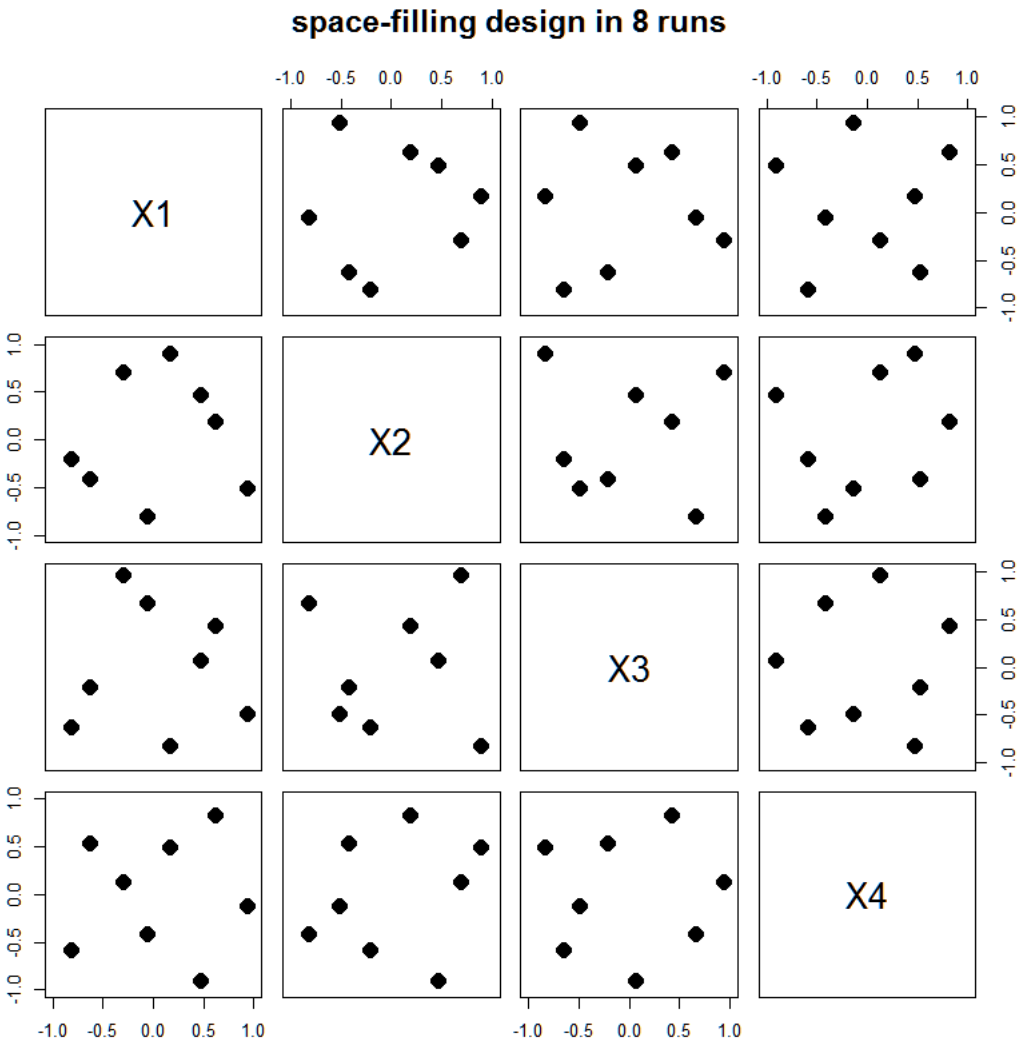
implicit replication in 3-factor projection.

Computer experiments: different perspective

- Examples up to now: (implicit) replication desirable,
  because of experimental variability:
  two runs with the same setup will yield different response values
- Computer models:
  - same input → same output → (implicit) replication not useful
- Computer models with quantitative factors:
  - often larger experimental space,
    low order polynomial approximation for response surface
    does not work very well
  - latin hypercube designs, space-filling or uniformity criteria

Computer experiments: different perspective

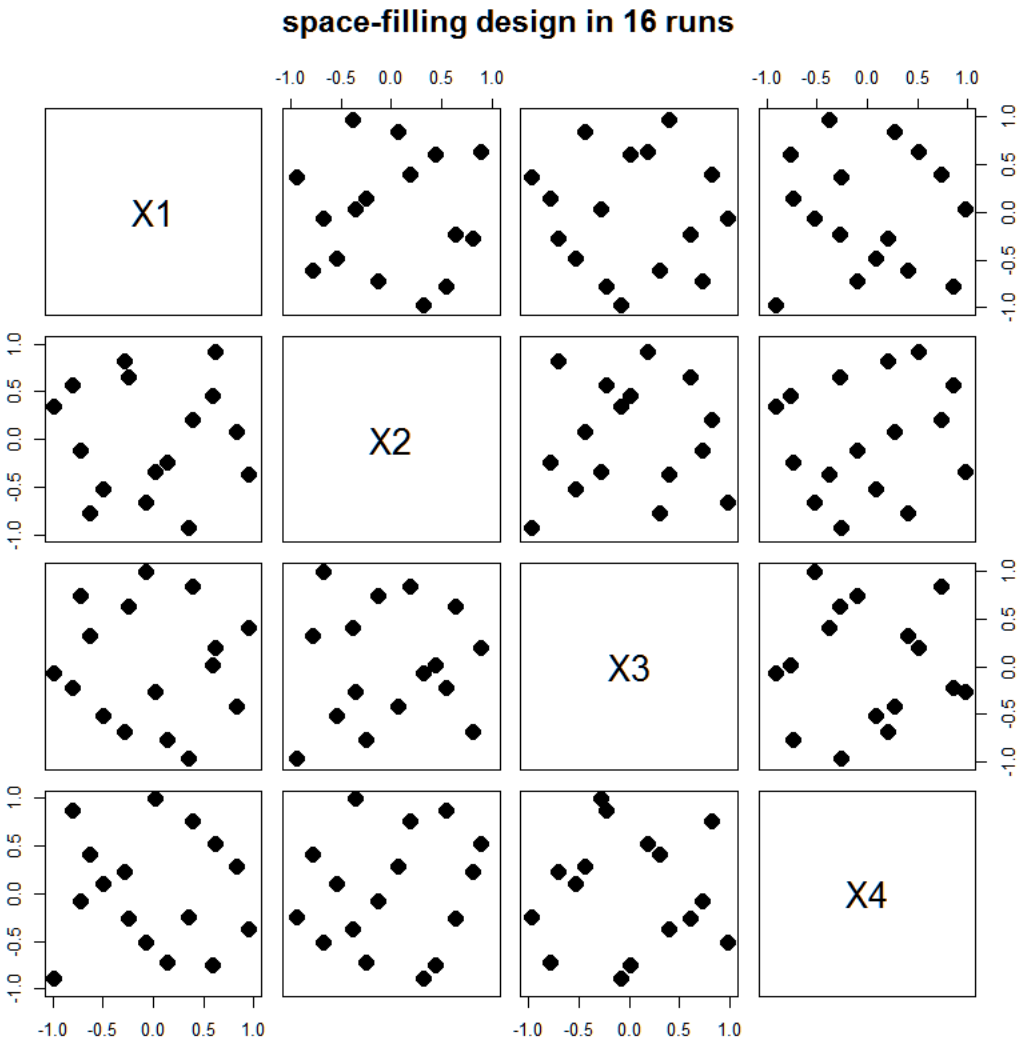
- Examples up to now: (i because of experiment two runs with the same
- Computer models:
  - same input → sam
- Computer models with
  - often larger experir low order polynomi
  - latin hypercube des



Design created with **`optimumLHS`** from **lhs** through function **`lhs.design`** from **DoE.wrapper**

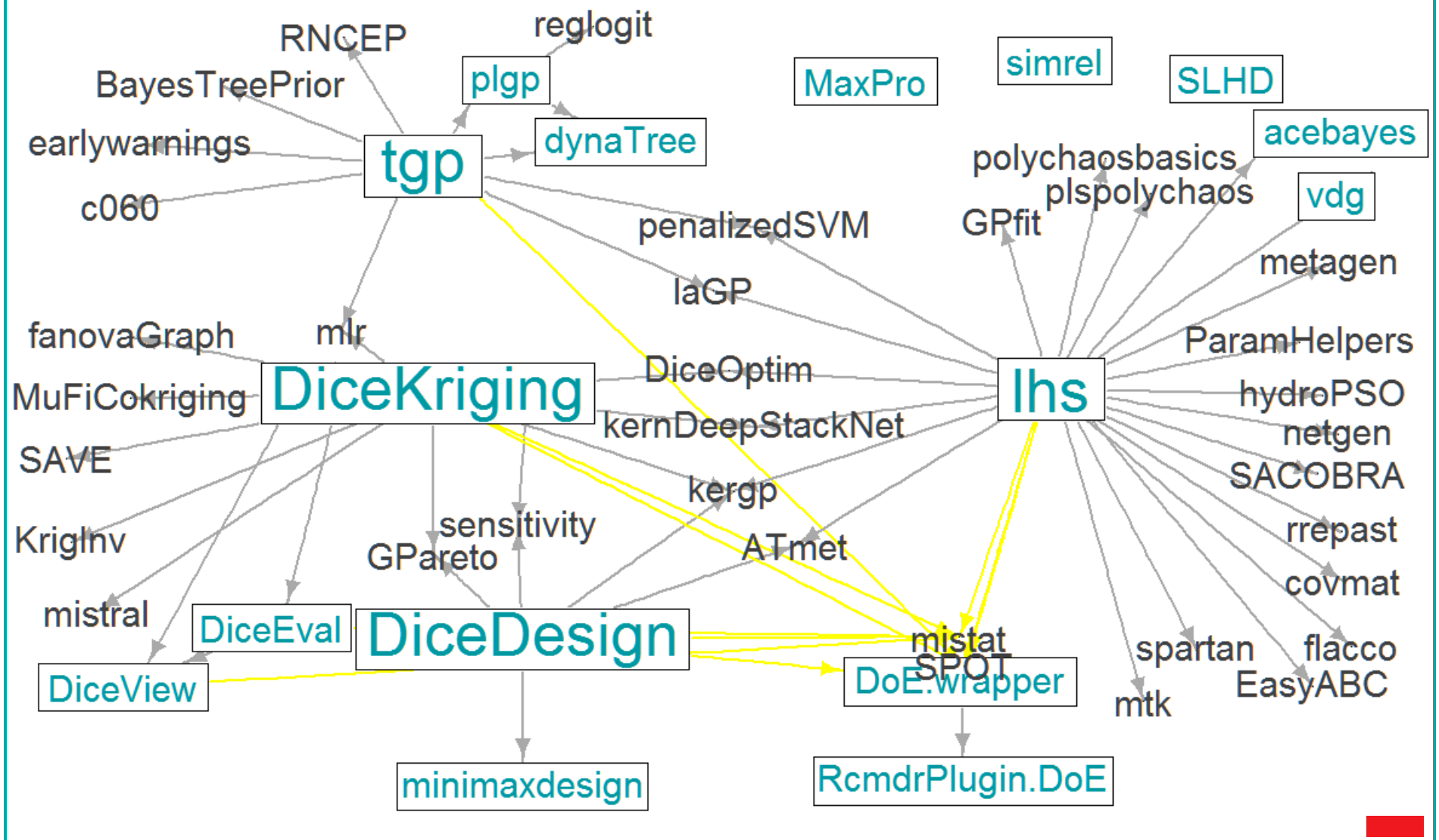
Computer experiments: different perspective

- Examples up to now: (i because of experiment two runs with the same
- Computer models:
  - same input → same
- Computer models with
  - often larger experir low order polynomi
  - latin hypercube des



Design created with **`optimumLHS`** from **lhs** through function **`lhs.design`** from **DoE.wrapper**

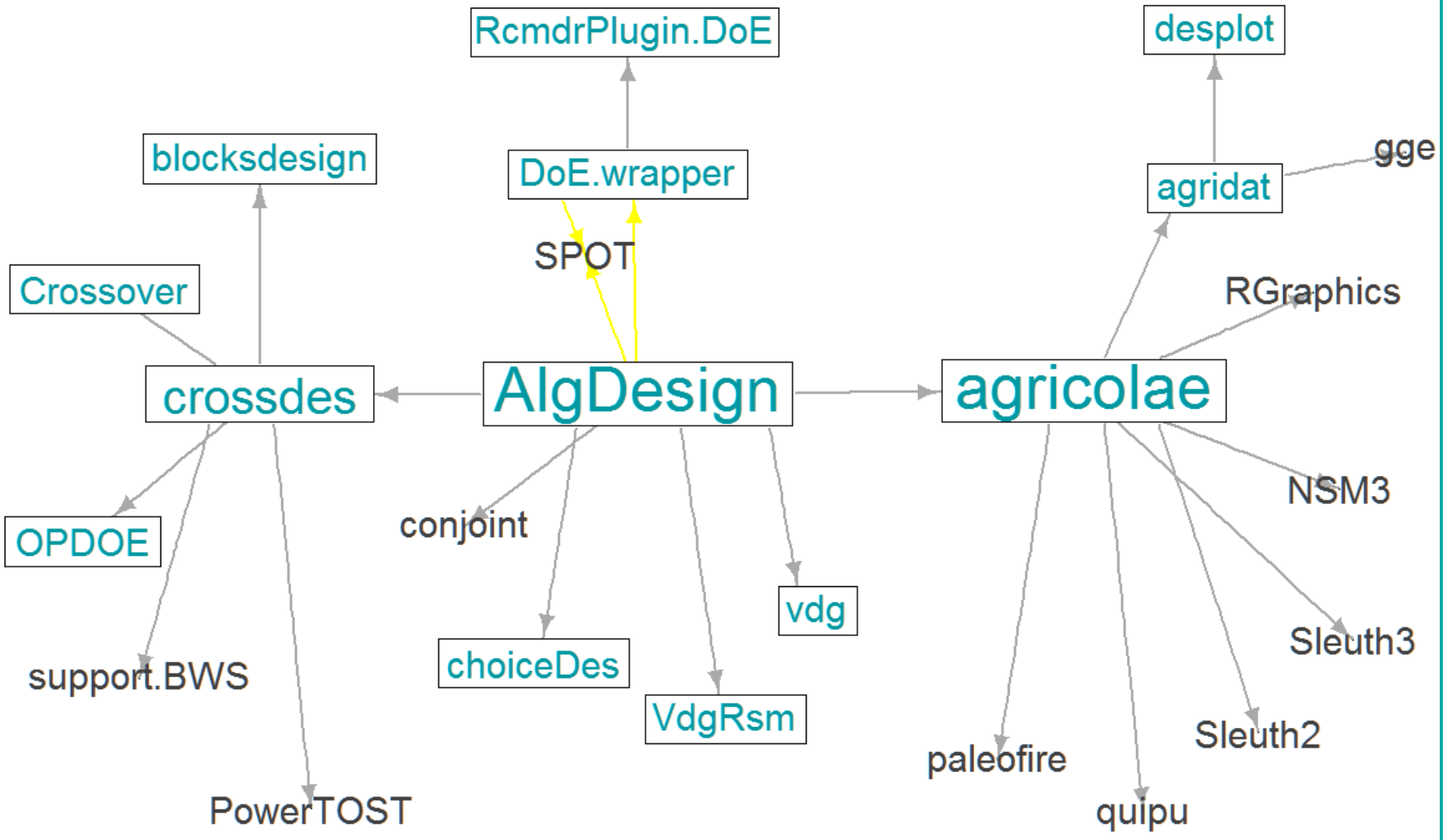
Packages for computer experiments and descendants



Some more info on optimal designs

- **AlgDesign** designs for parameter estimation within a known linear model (Optimal design, letter optimality criteria like D-, A- or I-optimality), constraints via candidate set
  - recent package **OptimalDesign** uses same optimality criteria but different algorithms,
    external software Gurobi (for quadratic integer programming, constrained optimization)
  - package **rodd** for T-optimal designs
  - …
- Packages for non-linear optimal design
  - **LDOD** for locally D-optimal designs for various specific model types, including popular generalized linear models
  - **popED** for certain situations in non-linear mixed models
  - **ICAOD** (imperialist competitive algorithm, very recent)
  - …

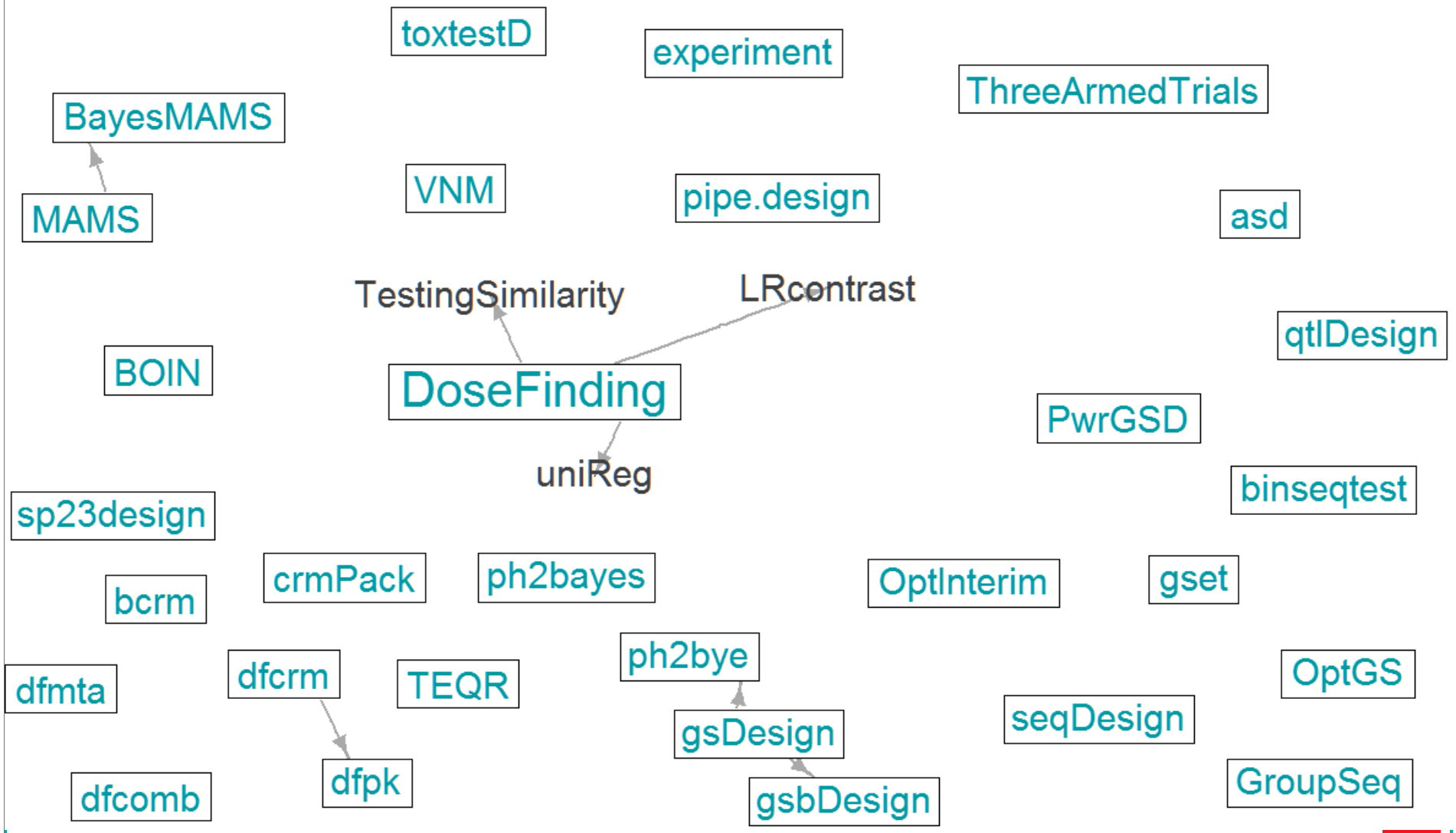
Descendants of **AlgDesign**



Bayesian designs

- quite a few Bayesian approaches,
  e.g. **tgp** or several clinical trial packages

Clinical trial related packages and descendants



Final remarks

- There is a lot of DoE in R.
- Some areas are better represented than others, feel free to contribute ...

- For the task view, please inform me about
  - new packages for the task view
  - trouble with task view packages (packages that should be removed)
  - possibly new DoE projects in R
  - co-author with complementary expertise ?

- I continue working on
  - criteria for the quality of (mixed level) non-regular orthogonal arrays
  - methods for the (semi-)automatic creation of good designs from them

- GUI programming (**RcmdrPlugin.DoE**) gets neglected (volunteers?)