Approximate Model Spaces for Model-Robust Experiment Design

Byran Smucker and Nathan Drew

Miami University, Oxford, OH

October 19, 2012
Design and Analysis of Experiments 2012, Athens, GA

Outline

Optimal Design

Facilitates:

- Precise estimation (e.g. $\mathcal{D}$-optimality); or
- Precise prediction (e.g. $\mathcal{IV}$-optimality).

Also facilitates tailor-made designs, e.g.:

- Constrained design space
- Mixture of continuous and categorical factors
- Sample size constraints

A Drawback

The form of the model between response and factors must be specified before design is constructed.

A Model-Robust Approach

Instead of focusing on a single model (optimal design), specify a set of models and find design that is "good" for all models of interest, if possible.

A Model-Robust Approach

Instead of focusing on a single model (optimal design), specify a set of models and find design that is "good" for all models of interest, if possible.

$\mathcal{D}$-optimal design, starting with arbitrary n-run design ξ_n:

$$\xi_n \xrightarrow{f} \mathbf{X}(\xi_n, f) \xrightarrow{\mathbf{X}'\mathbf{X}} \mathbf{M}(\xi_n, f) \rightarrow \xi_n^* = \arg\max_{\xi_n} |\mathbf{M}(\xi_n, f)|$$

A Model-Robust Approach

Instead of focusing on a single model (optimal design), specify a set of models and find design that is "good" for all models of interest, if possible.

$\mathcal{D}$-optimal design, starting with arbitrary n-run design ξ_n:

$$\xi_n \xrightarrow{f} \mathbf{X}(\xi_n, f) \xrightarrow{\mathbf{X}'\mathbf{X}} \mathbf{M}(\xi_n, f) \rightarrow \xi_n^* = \arg\max_{\xi_n} |\mathbf{M}(\xi_n, f)|$$

—

Model-robust design with respect to a set of models $\mathcal{F} = (f_1, f_2, \ldots, f_r)$:

$$\xi_n \xrightarrow{\mathcal{F}} \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_r\} \rightarrow \{\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_r\} \rightarrow \xi_n^* = \arg\max_{\xi_n} g[|\mathbf{M}_1|, |\mathbf{M}_2|, \ldots, |\mathbf{M}_r|]$$

Many Different Model Spaces

Main effects plus g two-factor interactions (MEPI):

- All models consisting of all k main effects and g out of $k(k-1)/2$ two-factor interactions
- Literature
  - Sun (1993)
  - Li and Nachtsheim, *Technometrics* (2000)
  - Smucker, del Castillo, and Rosenberger, forthcoming in *Technometrics*

Many Different Model Spaces

Supersaturated (SS)

- All models consisting of g out of k main effects, where $k > n - 1$ and $g \leq n - 1$.
- Literature
  - Jones, Li, Nachtsheim, and Ye, *JSPI* (2009)

Other Model Spaces

Projective

- Loeppky, Sitter, and Tang, *Technometrics* (2007)
- Smucker, del Castillo, and Rosenberger, forthcoming in *Technometrics*

All possible submodels of a maximal model (effect heredity can be enforced if desired)

- Tsai and Gilmour, *Technometrics* (2010)
- Smucker, del Castillo, and Rosenberger, *JQT* (2011)

Exploiting Effect Sparsity: MEPI Model Space

Consider a five-factor experiment in 12 runs.

- Full two-factor interaction model has $1 + 5 + 10 = 16$ parameters and can't be fit.
- Instead, assume that no more than $g = 3$ two-factor interactions will be active.
- There are 120 models which include 3 two-factor interactions.
- Design strategy: Find a design that can efficiently estimate all 120 models.

Advantage: More efficient designs in fewer runs, compared to resolution III or IV fractions.

A Drawback to Set-of-Models Approach

The model spaces are too large for many experiments of interest.

The MEPI Model Space Explodes

k	g	r
6	2	105
6	4	1,365
8	4	20,475
8	6	376,740
10	6	8,145,060
10	8	215,553,195
12	10	210,980,549,208

So does the Supersaturated Model Space

n	k	g	r
6	10	3	120
6	10	4	210
10	15	5	3,003
10	15	7	6,435
14	23	5	33,649
14	23	12	1,352,078
16	30	6	593,775
16	30	14	145,422,675
25	45	10	3,190,187,286

Current Methods Can't Handle Such Large Model Spaces

MEPI

- The largest model space Li and Nachtsheim (2000) consider includes less than 400,000 models.
- They do not give computation time for their designs for large model spaces.

Supersaturated

- The largest model space Jones et al. (2009) indicate includes fewer than 40,000 models.
- They state: "... the required computing time can become prohibitively large when n and [k] are large ... [L]arger designs can definitely be constructed ... with more computing power."

Outline

Notation

- $\mathcal{F}$: Full set of models.
- $\mathcal{S}_1$: Small sample of s_1 models chosen from $\mathcal{F}$, called the approximate model space.
- $\mathcal{S}_2$: Larger sample of s_2 models chosen from $\mathcal{F}$, used to evaluate a design with respect to $\mathcal{F}$.

Overview of Proposed Methodology

1. Select the approximate model space $\mathcal{S}_1$ at random from the full model space $\mathcal{F}$.
2. Construct n_t designs that are robust for the models in $\mathcal{S}_1$, via coordinate exchange.
3. Evaluate the n_t designs with respect to $\mathcal{F}$. If $\mathcal{F}$ is too large, select a larger sample $\mathcal{S}_2$ from $\mathcal{F}$ and evaluate the design with respect to $\mathcal{S}_2$.

The design that performs the best with respect to $\mathcal{F}$ (or $\mathcal{S}_2$) is chosen.

Ramifications of the Proposed Methodology

- Dramatically reduces computation time.
- Estimation capacity and efficiency of designs may be (slightly) inferior.

Step 1: Selecting the Approximate Model Space $\mathcal{S}_1$

Take a simple random sample from $\mathcal{F}$.

- An empirical study suggests $s_1 = 64$ is adequate, regardless of the size of $\mathcal{F}$.
- An alternative would be to choose the models in $\mathcal{S}_1$ systematically.
  - We tried this, using a maximin criterion.
  - It didn't show clear improvement, and increased the complexity of the procedure.

Step 2: Constructing Designs

Optimize the design with respect to $\mathcal{S}_1$ via a two-step process:

1. Maximize the number of models in $\mathcal{S}_1$ the design can estimate (i.e. maximize estimation capacity (EC)).
2. If $EC = 1$, maximize the average $\mathcal{D}$-efficiency of the design with respect to the models in $\mathcal{S}_1$.

This is accomplished via an algorithm that uses coordinate exchange [Meyer and Nachtsheim (1995)].

Step 3: Evaluating the Designs with respect to $\mathcal{F}$

- If $\mathcal{F}$ is small—say a few thousand—evaluate each design with respect to $\mathcal{F}$.
- If $\mathcal{F}$ is large, take a sample $\mathcal{S}_2$ from $\mathcal{F}$ and evaluate each design with respect to $\mathcal{S}_2$.
  - Can perform inference on EC and average $\mathcal{D}$-efficiency with respect to $\mathcal{F}$.

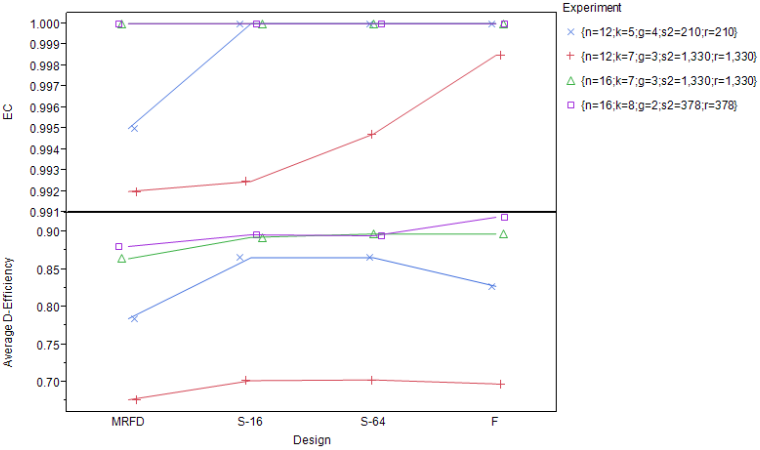
Outline

Comparing Designs

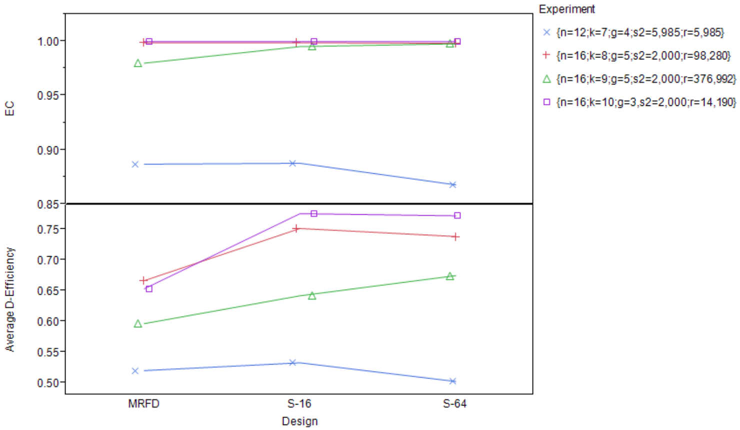
The following designs are compared:

- $\mathcal{S}$ designs. Designs constructed via the three step procedure described above.
  - $\mathcal{S}$-16 designs are based on a random sample of $s_1 = 16$.
  - $\mathcal{S}$-64 designs are based on a random sample of $s_1 = 64$.
- $\mathcal{F}$ designs. Designs constructed via the three step procedure with $\mathcal{S}_1 = \mathcal{F}$.
- MRFD. Designs from Li and Nachtsheim (2000) for the MEPI model space. They utilize $\mathcal{F}$ and require column-balance.

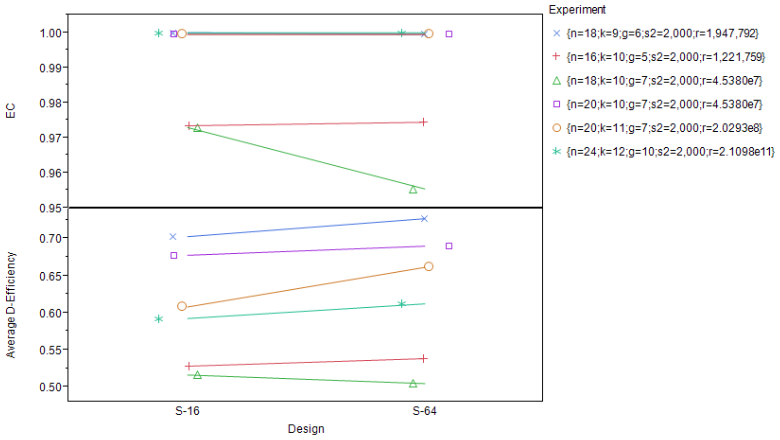
MEPI: Small Experiments



MEPI: Medium-sized Experiments



MEPI: Large Experiments



MEPI: Design Construction Times

n	k	g	s_2	r	Design	Time
16	7	3	1,330	1,330	$\mathcal{S}$-16	0.145
16	7	3	1,330	1,330	$\mathcal{S}$-64	0.449
16	7	3	1,330	1,330	$\mathcal{F}$	7.94
16	8	2	378	378	$\mathcal{S}$-16	0.161
16	8	2	378	378	$\mathcal{S}$-64	0.552
16	8	2	378	378	$\mathcal{F}$	2.96
16	9	5	2,000	376,992	$\mathcal{S}$-16	0.281
16	9	5	2,000	376,992	$\mathcal{S}$-64	0.813
20	10	7	2,000	4.5380e7	$\mathcal{S}$-16	0.566
20	10	7	2,000	4.5380e7	$\mathcal{S}$-64	1.54
24	12	10	2,000	2.1098e11	$\mathcal{S}$-16	1.08
24	12	10	2,000	2.1098e11	$\mathcal{S}$-64	3.67

Times are in minutes per algorithm try.

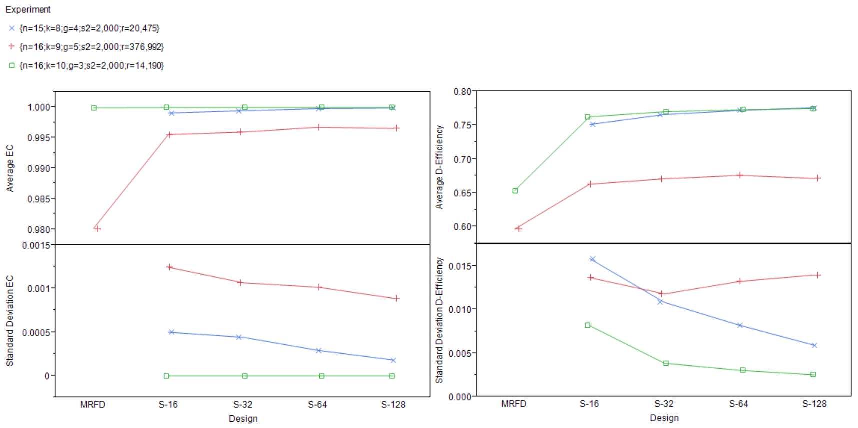
Outline

Setup

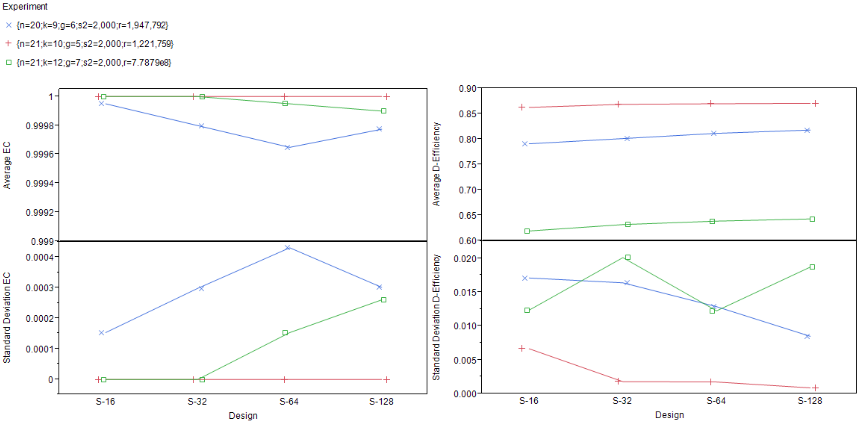
Since $\mathcal{S}_1$ is chosen randomly, we wish to study the effectiveness of the procedure over multiple randomly chosen $\mathcal{S}_1$'s. In what follows:

- 20 model sets were chosen.
- For each, $n_t = 50$ designs were constructed and the best chosen.
- The average EC and average $\overline{E}_{\mathcal{D}}$ were calculated, along with standard deviations of these quantities.

MEPI: Medium-sized Experiments



MEPI: Large Experiments



Outline

Overall Comments

Based on the presentation:

- Proposed designs are competitive in terms of design efficiency, and constructed in fraction of the time.
- Proposed designs are more efficient than Li and Nachtsheim designs and constructed in fraction of the time.
- If at least a few degrees of freedom above saturation, larger approximate model spaces are more efficient and less variable.
- Approximate model space size of 64 seems adequate regardless of the size of $\mathcal{F}$.

Overall Comments

Based on other work we have done:

- Procedure is effective for other model spaces (supersaturated; all possible submodels).
- These designs do not appear to give up a significant amount in model discriminating capabilities.

Acknowledgments

- Co-author and former graduate student Nathan Drew;
- John Bailer and Steve Wright, for their feedback;
- Miami University's Summer Research Grant.

Outline

Experiments with Large Model Spaces May Be Needed

- Li, Sudarsanam, and Frey (*Complexity*, 2006) empirically examine effect sparsity.
- Conclusions
  - Likely between 37% and 46% (mean: 41%) of main effects will be active.
  - Likely between 9% and 14% (mean: 11%) of two-factor interactions will be active.
- Caveat: Only full factorial designs with 7 or fewer factors.

Model-Robust is not the same as Model-Discriminating

It is possible for two models to be estimable but indistinguishable from each other.

- One way to measure model discrimination is the subspace angle (Jones et al. 2007).
- Given a design and a pair of models, the subspace angle is the angle between the subspaces spanned by the columns of the expanded design matrices $\mathbf{X}_1$ and $\mathbf{X}_2$.
- If the angle is close to 90 degrees, the models are close to orthogonal.
- If the angle is close to 0, the models are close to indistinguishable (nearly aliased).

With such large model spaces, model discrimination is of concern.

Each pair of models should be compared, and when the model space is large this is computationally prohibitive.

- Thus, we again sample, this time pairs of models.
- We can sample a large enough number that inference is very sharp.

Tabular Comparison

Model Space	n	k	g	Pairs Compared	Total Pairs	Design	Average	Minimum
MEPI	16	9	5	124,999*	7.106e10	$\mathcal{S}_1$-64	68.07	0
	16	9	5	125,000	7.106e10	MRFD	65.71	0
MEPI	24	12	10	125,000	2.226e22	$\mathcal{S}_1$-64	75.58	0
	28	12	10	125,000	2.226e22	$\mathcal{S}_1$-64	83.56	0
SS	8	12	3	24,090	24,090	$\mathcal{S}_1$-64	83.75	45
	8	12	3	24,090	24,090	MRSS	83	44
SS	8	12	5	313,236	313,236	$\mathcal{S}_1$-64	77.64	0
	8	12	5	313,236	313,236	MRSS	79.15	0
SS	16	30	6	125,000	1.763e11	$\mathcal{S}_1$-64	86.20	38.80
	16	30	10	125,000	4.514e14	$\mathcal{S}_1$-64	85.19	0

* If the two models to be compared were the same, this comparison was disregarded.

Graphical Comparison

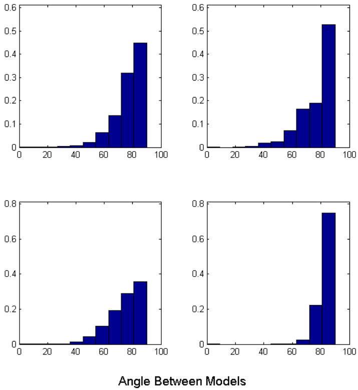


Figure: Top row: $\{n = 8; k = 12; g = 5\}$ supersaturated experiment with $\mathcal{S}_1$-64 design (top left) and MRSS design (top right). Bottom row: $\{n = 24; k = 12; g = 10\}$ MEPI experiment with $\mathcal{S}_1$-64 (bottom left) and $\{n = 28; k = 12; g = 10\}$ MEPI experiment with $\mathcal{S}_1$-64 design (bottom right).

Conclusion

- The $\mathcal{S}$-64 designs are competitive in terms of model discrimination (subspace angle) to designs in the literature.
- If the average subspace angle is too small, it can be increased by increasing the sample size.

Supersaturated: Small Experiments

n	k	g	s_2	r	Design	$EC_{\mathcal{F}}$	$\bar{E}_{\mathcal{F}}$
6	10	3	120	120	MRSS	1	0.9
6	10	3	120	120	$\mathcal{S}$-16	0.933	0.823
6	10	3	120	120	$\mathcal{S}$-64	1	0.903
6	10	3	120	120	$\mathcal{F}$	1	0.903
6	10	4	210	210	$\mathcal{S}$-16	0.986	0.747
6	10	4	210	210	$\mathcal{S}$-64	1	0.727
6	10	4	210	210	$\mathcal{F}$	1	0.756
10	15	6	5,005	5,005	MRSS	1	0.82
10	15	6	5,005	5,005	$\mathcal{S}$-16	1	0.812
10	15	6	5,005	5,005	$\mathcal{S}$-64	1	0.835
10	15	6	5,005	5,005	$\mathcal{F}$	1	0.837
10	15	7	6,435	6,435	$\mathcal{S}$-16	1.000	0.716
10	15	7	6,435	6,435	$\mathcal{S}$-64	1	0.777
10	15	7	6,435	6,435	$\mathcal{F}$	1	0.779

Supersaturated: Large Experiments

n	k	g	s_2	r	Design	$\bar{E}_{\mathcal{S}_2}$	$\bar{E}_{\mathcal{F}}$ LCL	$\bar{E}_{\mathcal{F}}$ UCL
14	23	5	2,000	33,649	$\mathcal{S}$-16	0.894	0.892	0.897
14	23	5	2,000	33,649	$\mathcal{S}$-64	0.913	0.911	0.915
14	23	12	2,000	1,352,078	$\mathcal{S}$-16	0.626	0.622	0.629
14	23	12	2,000	1,352,078	$\mathcal{S}$-64	0.635	0.632	0.638
16	30	6	2,000	593,775	$\mathcal{S}$-16	0.879	0.877	0.881
16	30	6	2,000	593,775	$\mathcal{S}$-64	0.892	0.890	0.894
16	30	14	2,000	1.454e8	$\mathcal{S}$-16	0.583	0.580	0.586
16	30	14	2,000	1.454e8	$\mathcal{S}$-64	0.599	0.596	0.602
25	45	10	2,000	3.190e9	$\mathcal{S}$-16	0.865	0.863	0.866
25	45	10	2,000	3.190e9	$\mathcal{S}$-64	0.880	0.878	0.882

For all designs, $EC = 1$.

Supersaturated: Multiple Model Sets