ORTHOGONAL NEARLY LATIN HYPERCUBE DESIGNS

David M. Steinberg

Tel Aviv University

Workshop on Computer Experiments
The Isaac Newton Institute
Sept 2011, Cambridge

Joint work with Dennis Lin, Penn State University.

Ideas from Dizza Bursztyn, Ashkelon College.

OUTLINE

1. Designing Computer Experiments
2. Latin Hypercube Designs
3. Rotated Factorial Designs
4. Orthogonal Nearly LHD's from Rotated Factorials
5. Conclusions

Designing a Computer Experiment

Computer experiments require good designs.

The goals might be:

- To identify the important factors.
- To develop a “prediction equation”.
- To locate “good operating conditions”.

Latin Hypercube Designs

Latin Hypercubes are the most popular class of experimental plan.

LHD's place the input levels for each factor on a uniform grid.

Then "mate" the levels across factors.

McKay, Beckman and Conover, *Technometrics*, 1979.

Latin Hypercube Designs

Other mating schemes have been suggested to obtain columns with low correlation.

Ye showed how to get $2m-2$ fully orthogonal columns with 2^m runs.

Butler showed how to get orthogonality with respect to a trigonometric regression model and 2^m runs.

Devon Lin presented clever methods for using OA's to get some orthogonal LHD's.

Latin Hypercube Designs

Steinberg and Lin showed how to get close to *n-1* orthogonal columns in an *n*-run LHD.

BUT – the method works only for very special values of *n*.
For example, 16; 256; 65,536.

Latin Hypercube Designs

Some known results on orthogonal LHD's.

n	K	Source
16	12	SL
32	16	SLL
48	12	LBST
64	32	LMT
80	12	LBST
96	24	LBST
112	12	LBST
128	48	LBST

SL – Steinberg & Lin

SLL – Sun, Liu & Lin

LBST – Lin, Bingham, Sitter & Tang

LMT – Lin, Mukerjee & Tang

Orthogonal *Nearly* LHD's

Difficult to meet the twin goals of

- LHD (perfect univariate spacing)
- Orthogonality (bivariate property)

We compromise the univariate spacing to maintain the bivariate orthogonality.

Rotated Factorial Designs

Bursztyn and Steinberg developed experimental plans with many levels in which linear effects are orthogonal.

Start with a "standard" first-order orthogonal design, like a 2^{k-p} fractional factorial: *D*.

"Rotate" the design using a rotation matrix *R*: $D \rightarrow DR$.

Then $(DR)'(DR) = R'D'DR = nR'R = nI$.

LHD's as Rotated Factorial Designs

Steinberg and Lin showed how to rotate two-level factorials into Latin Hypercube designs with a large number of first-order orthogonal columns.

Steinberg and Lin combined a rotation idea in Bursztyn and Steinberg with another rotation idea developed by Lin and Beattie.

LHD's as Rotated Factorial Designs

Lin and Beattie: rotate 2^k factorials to Latin Hypercube designs. The intuition:

- Columns in a LHD are an arithmetic sequence.
- Columns in *DR* are linear combinations of the rows of *D* (the 2^k design).
- The rows of *D* are a binary expansion of the odd integers.
- Using appropriate powers of 2 as the elements in *R*, each column in *DR* is an integer sequence.

LHD's as Rotated Factorial Designs

-1	-1	-1
1	-1	-1
-1	1	-1
1	1	-1
-1	-1	1
1	-1	1
-1	1	1
1	1	1

LHD's as Rotated Factorial Designs

Weights	2	-4	1
	-1	-1	-1
	1	-1	-1
	-1	1	-1
	1	1	-1
	-1	-1	1
	1	-1	1
	-1	1	1
	1	1	1

LHD's as Rotated Factorial Designs

Weights	2	-4	1	
	-1	-1	-1	1
	1	-1	-1	5
	-1	1	-1	-7
	1	1	-1	-3
	-1	-1	1	3
	1	-1	1	7
	-1	1	1	-5
	1	1	1	-1

Weighted Sums

LHD's as Rotated Factorial Designs

Lin and Beattie: rotate 2^k factorials to Latin Hypercube designs.

- Can we organize weights for multiple columns in a rotation matrix R?
- Yes – provided R is 2^m by 2^m.
- A simple recursive scheme gives the rotation matrices.

LHD's as Rotated Factorial Designs

Lin and Beattie: rotate 2^k factorials to Latin Hypercube designs.

$$R_0 = [1]$$

$$R_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

$$R_{j+1} \propto \begin{bmatrix} R_j & -2^{2^j} R_j \\ 2^{2^j} R_j & R_j \end{bmatrix}$$

LHD's as Rotated Factorial Designs

Steinberg and Lin:

$$DR = \left[D_1 \mid \cdots \mid D_t\right] \begin{bmatrix} R & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & R \end{bmatrix}$$
$$= \left[D_1 R \mid \cdots \mid D_t R\right]$$

Bursztyn & Steinberg: create 2^m sets of columns, each a full factorial

Lin & Beattie

The resulting design is an orthogonal Latin hypercube.

Orthogonal *Nearly* LHD's

The construction requires that each set of columns be a full factorial design.

What if we rotate a set of columns that is ***not*** a full factorial?

We get a rotated column that may have good univariate properties, but *not* perfect spacing.

Orthogonal *Nearly* LHD's

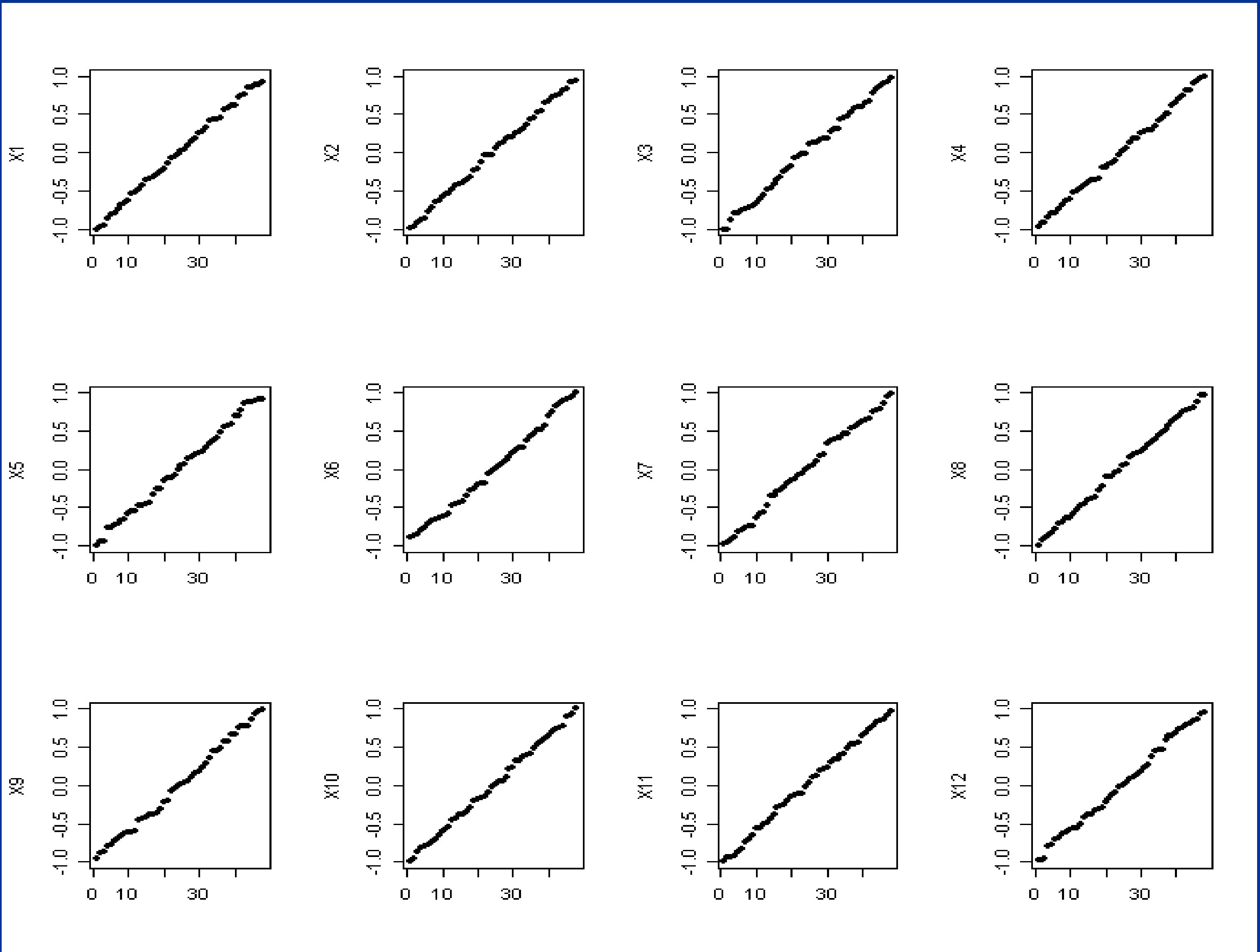
The idea for the orthogonal *nearly* LHD's is to rotate sets of columns from a Plackett-Burman design.

For example, with a 48-run PB design, we can rotate 5 sets of 8 columns each.

A set of 48 runs will not include each of the 2^8 factorial points; so the rotated column won't be perfectly uniform.

Orthogonal *Nearly* LHD's

Some uniform qq-plots with 12 factors and 48 runs.



Orthogonal *Nearly* LHD's

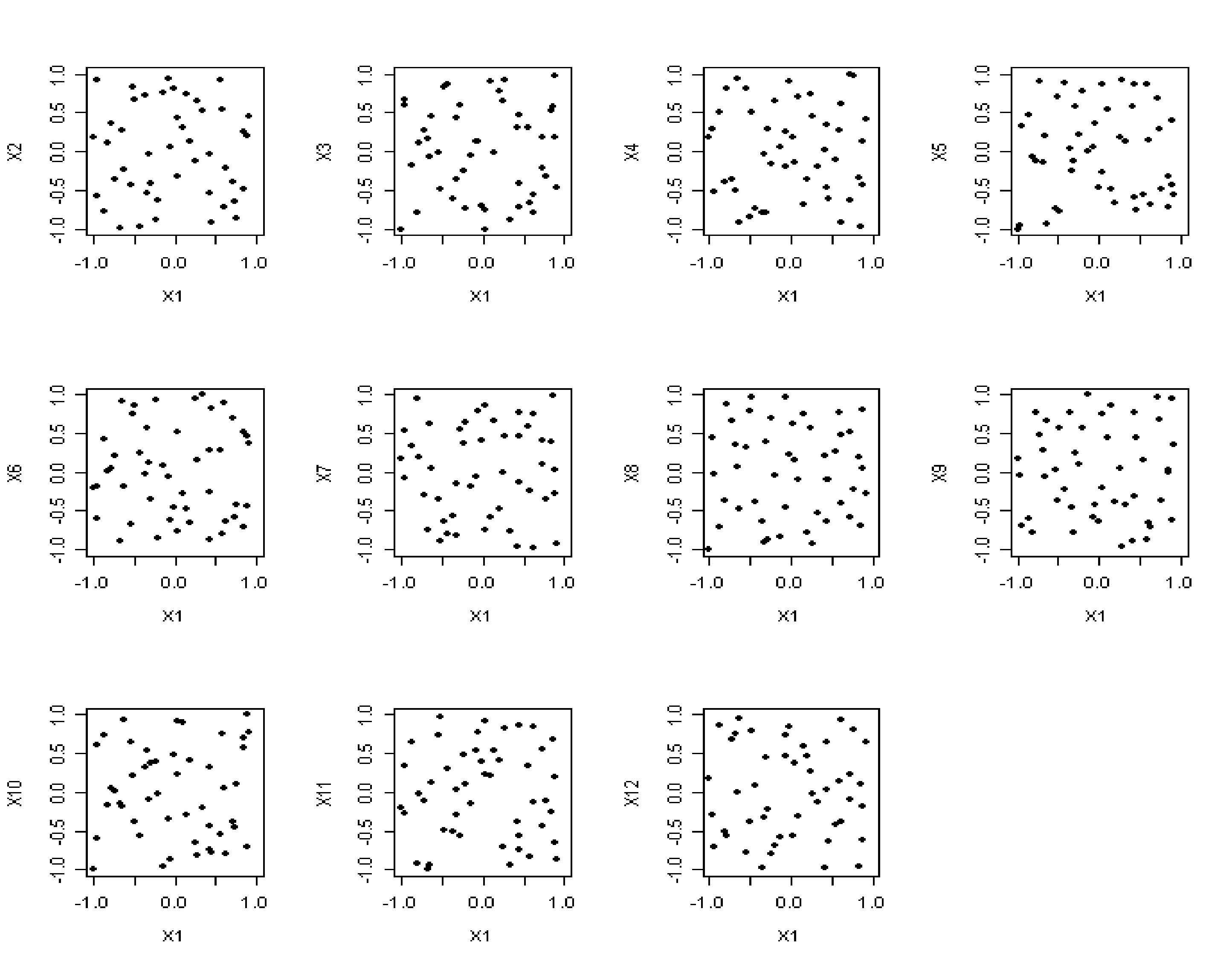
For a true LHD, all univariate inter-point distances are 0.043.

For the orthogonal design, some distances are 0.

The longest distances range from 0.094 (X11) to 0.172 (X12).

Orthogonal *Nearly* LHD's

Some two-factor projections of the design.



Orthogonal *Nearly* LHD's

Both univariate and bivariate properties are related to projection properties of the original PB design.

Denote the columns of the PB design by $c_1,,,,c_{47}$.

Orthogonal *Nearly* LHD's

Suppose we jointly rotate the first 8 columns.

The first rotated column (unscaled) has the form:

$$r_1 = \pm 2^7 c_1 \pm 2^6 c_2 \pm 2^5 c_3 \pm \ldots \pm c_8$$

The "pattern" of points for this factor is largely determined by the projection of the PB design on the leading columns in the sum.

Orthogonal *Nearly* LHD's

$$r_1 = \pm 2^7 c_1 \pm 2^6 c_2 \pm 2^5 c_3 \pm \ldots \pm c_8$$

The balance on c_1 and c_2 guarantees that ¼ of the values of r_1 are in each quartile.

If there is also perfect balance for the c_1, c_2, c_3 projection, 1/8 of the points will be in each octile.

There is a repeat value if the projection of the PB design on all 8 factors has a repeat run.

Orthogonal *Nearly* LHD's

Now look at pairs of columns.

$$r_1 = \pm 2^7 c_1 \pm 2^6 c_2 \pm 2^5 c_3 \pm \ldots \pm c_8$$

$$r_2 = \pm 2^7 d_1 \pm 2^6 d_2 \pm 2^5 d_3 \pm \ldots \pm d_8$$

For two rotated columns in the same rotation set, the columns in the sum are identical, but permuted.

Otherwise, the two sets of columns are fully distinct.

Orthogonal *Nearly* LHD's

$$r_1 = \pm 2^7 c_1 \pm 2^6 c_2 \pm 2^5 c_3 \pm \ldots \pm c_8$$

$$r_2 = \pm 2^7 d_1 \pm 2^6 d_2 \pm 2^5 d_3 \pm \ldots \pm d_8$$

The bivariate projection is dominated by the two lead PB columns for each rotated column.

Generally, good bivariate projections will correspond to good projections of these PB factors.

The balance on c_1 and d_1 guarantees that ¼ of the values are in each quadrant – a U-design based on a 2-level OA.

Orthogonal *Nearly* LHD's

$r_1 = \pm 2^7 c_1 \pm 2^6 c_2 \pm 2^5 c_3 \pm \ldots \pm c_8$

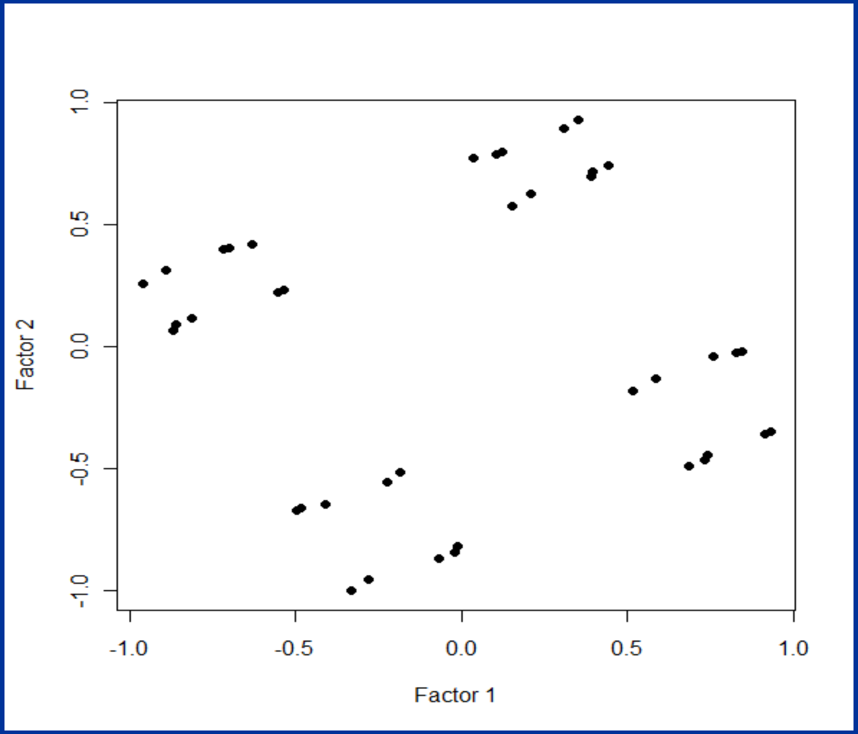
$r_2 = \pm 2^7 d_1 \pm 2^6 d_2 \pm 2^5 d_3 \pm \ldots \pm d_8$

If there is also perfect balance for the c_1,c_2,d_1,d_2 projection, 1/16 of the points will be in each part of a 4*4 grid – a U-design from a 4-level OA.

If the projection *includes* all 16 design points, there will be points in each 4*4 grid square in the r_1,r_2 plane.

Orthogonal *Nearly* LHD's

Adjacent columns share the same first two columns. The bivariate projection has points in only 4 of the 16 squares.



Orthogonal *Nearly* LHD's

The 48-run PB design (Paley) has 7 different 4-factor projection patterns.

0	1	2	3	4	5	6	%
0	0	4	8	4	0	0	39.4
0	1	4	6	4	1	0	30.3
0	2	4	4	4	2	0	18.2
1	0	6	4	1	4	0	3.0
1	2	3	4	3	2	1	3.0
0	0	8	0	8	0	0	3.0
0	4	1	4	6	0	1	3.0

Orthogonal *Nearly* LHD’s

- About 94% include all 16 points.
- NO projections have 3 repeats of each run.

Orthogonal *Nearly* LHD's

- Choosing rotation sets in which all pairs of 2 lead columns have the first pattern should give good designs.
- We do not have an algorithm for generating good pairs of columns.

Orthogonal *Nearly* LHD's

- For the 48 run PB, a trial-and-error search found only 5 pairs of columns, all of which have the best projection pattern.
- Including pairs with the two best patterns leads to 12 pairs of columns; i.e. to a rotated design with 12 columns.
- The design has points in all cells of a 4*4 grid for each two-factor projection, but never 3 points in each cell.
- Rotation sets can be completed so that each factor has points in each 1/16 bin.

Orthogonal *Nearly* LHD’s

The 52-run PB design (Williams) has 425 different 4-factor projection patterns!

Some of these occur for as few as 10 quartets.

Most include all 16 points; some as few as 12.

Some points are repeated as many as 9 times in a single projection.

Twelve projections have 4 singletons, 4 points repeated four times, and 4 repeated eight times.

Orthogonal *Nearly* LHD's

What criteria should be used to decide which are the best projections?

One choice is to minimize the sum of the squares of the repeat numbers.
Another is to eliminate projections with high or low repeat numbers.

Some Design Comparisons

Suppose you use the design to fit a simple first-order regression model, to “screen” the most influential factors:

$$Y = X\beta + \varepsilon.$$

But the true dependence involves additional regression terms:

$$Y = X\beta + Z\gamma.$$

Then $\hat{\beta}$-hat $= \beta + (X'X)^{-1}X'Z\gamma = \beta + A\gamma$.

The matrix A is known as the *alias matrix*.

Some Design Comparisons

The alias matrix depends on the design, the model used for screening, and the extra terms in Z.

A good screening design should have small values in A for simple screening models and somewhat more complex extra terms.

Some Design Comparisons

Compare our 48-run, 10-factor design to a maximin distance design given by a team at U. of Tilburg.

Criteria: minimum distance and sum of squared alias matrix entries.

Case 1: fit first-order terms, need all second order and pure cubics.

Case 2: fit main effects of first and second order, need all second and third order and pure quartics.

Some Design Comparisons

	Maximin	ONLHD
Min Dist	2.22	1.36
Case 1	3.42	11.63
Case 2	357.73	161.22

Conclusions

- Rotations of standard designs can lead to effective screening designs with the ability to also fit higher-order effects.
- Initial work suggest that these designs can be effective for identifying the most important factors.
- Although we have no algorithm for generating rotation sets, most 4-factor projections appear to be relatively well-balanced in small PB designs.