Sliced Full Factorial-based LHDs as a Framework for a Batch Sequential Design Algorithm

McCormick
Northwestern Engineering

NORTHWESTERN UNIVERSITY

Weitao Duan and Bruce Ankenman

Department of Industrial Engineering and Management Sciences, Northwestern University

Abstract

When building a metamodel for a complex computer simulation, the experiment design should exhibit good properties of both projectivity and orthogonality. We present a batch sequential experiment design method that produces uniformity in higher dimensions especially at certain stages of the design, but approximates those properties after every stage.

Problems with a Design with Fixed Design Points

- ❑ More data may be needed if desire precision is not met
- ❑ A waste of effort if fewer experiments suffice
- ❑ Information collected cannot be used to improve design

Advantages of Sequential Designs

- ❑ Can still be implemented if budget is unknown.
- ❑ Can terminate at any time

Overview of SFFLHD

sFFLHD uses the idea of sliced space filling designs by (Qian, 2009) and extends Loeppky, Moore and Williams's batch sequential design (Loeppky et al, 2010)

- ❑ Space filling properties at certain stages of the design (golden stages).
- ❑ Good orthogonality and projectivity at intermediate stages.
- ❑ More batches can easily be sampled if needed.

An orthogonal array OA(L^2,m,L,2) can be partitioned into L levels
Rows 1-3,4-6,7-9 are the three layers.
Columns 2-4 of each layer form a Latin hypercube
Each layer is used as a batch

1	1	1	1
1	2	2	3
1	3	3	2
2	1	2	3
2	2	3	2
2	3	1	1
3	1	3	2
3	2	1	1
3	3	2	3

Construction

$W_{.1}$	$W_{.2}$	$W_{.3}$	A^i	A^i_p	$V_{.1}$	$V_{.2}$	$V_{.3}$	$X_{.1}$	$X_{.2}$	$X_{.3}$	Run	Batch
1	1	1			1	2	2	0.074	0.213	0.131	1	
2	2	3		A^1_1	5	4	8	0.555	0.392	0.821	2	1
3	3	2			9	9	5	0.971	0.951	0.511	3	
1	2	3			2	6	7	0.163	0.565	0.669	4	
2	3	2	A^1	A^1_2	4	8	6	0.430	0.815	0.573	5	2
3	1	1			7	1	1	0.710	0.068	0.015	6	
1	3	2			3	7	4	0.274	0.695	0.443	7	
2	1	1		A^1_3	6	3	3	0.629	0.333	0.303	8	3
3	2	3			8	5	9	0.851	0.516	0.924	9	

$W_{.1}$	$W_{.2}$	$W_{.3}$	A^i	A^i_p	$V_{.1}$	$V_{.2}$	$V_{.3}$	$X_{.1}$	$X_{.2}$	$X_{.3}$	Run	Batch
1	1	1			2	6	4	0.074	0.213	0.131	1	
2	2	3		A^1_1	15	11	23	0.555	0.392	0.821	2	1
3	3	2			27	26	14	0.971	0.951	0.511	3	
1	2	3			5	16	19	0.163	0.565	0.669	4	
2	3	2	A^1	A^1_2	12	22	16	0.430	0.815	0.573	5	2
3	1	1			20	2	1	0.710	0.068	0.015	6	
1	3	2			8	19	12	0.274	0.695	0.443	7	
2	1	1		A^1_3	17	9	9	0.629	0.333	0.303	8	3
3	2	3			23	14	25	0.851	0.516	0.924	9	
1	1	2			9	4	17	0.309	0.117	0.606	10	
2	2	1		A^2_1	13	15	2	0.455	0.537	0.053	11	4
3	3	3			19	21	27	0.674	0.769	0.967	12	
1	2	1			1	13	7	0.028	0.471	0.244	13	
2	3	3	A^2	A^2_2	14	24	20	0.490	0.873	0.724	14	5
3	1	2			21	5	10	0.747	0.178	0.336	15	
1	3	3			6	20	26	0.203	0.707	0.947	16	
2	1	2		A^2_3	11	3	15	0.402	0.094	0.530	17	6
3	2	1			22	12	6	0.812	0.432	0.193	18	
1	1	1			7	1	8	0.225	0.008	0.280	19	
2	2	3		A^3_1	18	17	24	0.648	0.616	0.871	20	7
3	3	2			25	27	13	0.895	0.985	0.470	21	
1	2	3			4	10	22	0.113	0.340	0.791	22	
2	3	2	A^3	A^3_2	16	25	18	0.588	0.905	0.644	23	8
3	1	1			24	8	5	0.880	0.276	0.171	24	
1	3	2			3	23	11	0.100	0.826	0.399	25	
2	1	1		A^3_3	10	7	3	0.359	0.251	0.088	26	9
3	2	3			26	18	21	0.961	0.640	0.772	27	

The big grid design **W**, non-overlapping OAs, preserves orthogonality.

The small grid design **V**, a (near) LH, preserves LHD projectivity.

Levels of **V** are determined according to **W** and expand if necessary to accommodate more design points.

Sampling **X** is drawn according to the near LH **V** similar to LHD method

The sequential design can continue indefinitely.

Variance Reduction of Mean Estimators

Each batch achieves the same variance reduction as an ordinary LHD comparing to random sampling

$$Var(\overline{Y^r_p}) = \sum_{|u|\geq 2} M^r_p(u,|u|)L^{-2}Var(\alpha_u(\mathbf{x})) + o(L^{-1})$$

More variance reduction is achieved when the sequential design is an OA-based LHD

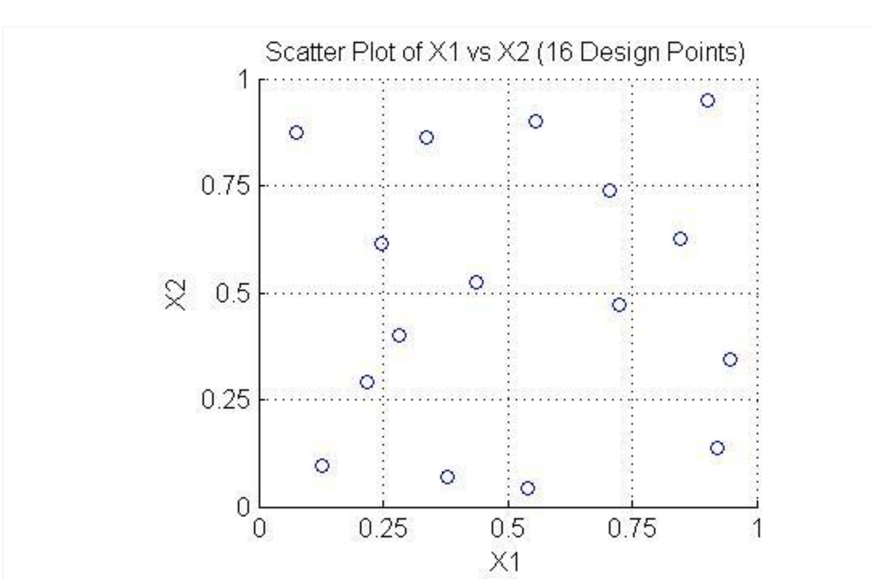
$$Var(\overline{Y}) = \sum_{|u|\geq 3} M(u,|u|)l^{-2}Var(\alpha_u(\mathbf{x})) + o(l^{-1})$$

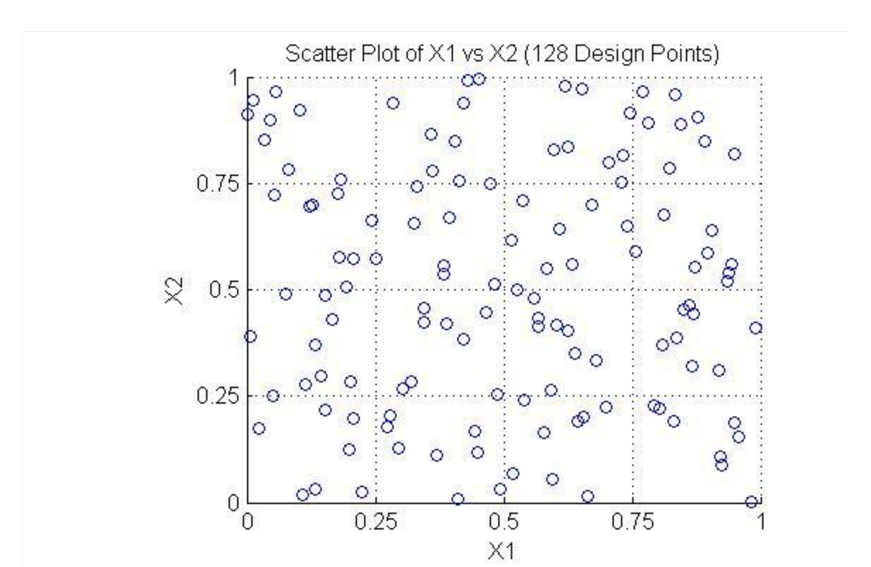
Certain degree of variance reduction at other stages

Batch	1	4	**8**	12	**16**
Design Points	8	32	**64**	96	**128**
sFFLHD	9.386	1.253	**0.051**	0.178	**0.020**
bMmLHD	245.810	46.817	**18.310**	5.051	**0.436**
MmDist	13.834	32.051	**14.866**	6.425	**4.510**
rsFFLHD	155.487	43.463	**0.047**	1.347	**0.019**

Table : Comparison of RMSE of $\hat{\mu}$ for each design scheme for Borehole function

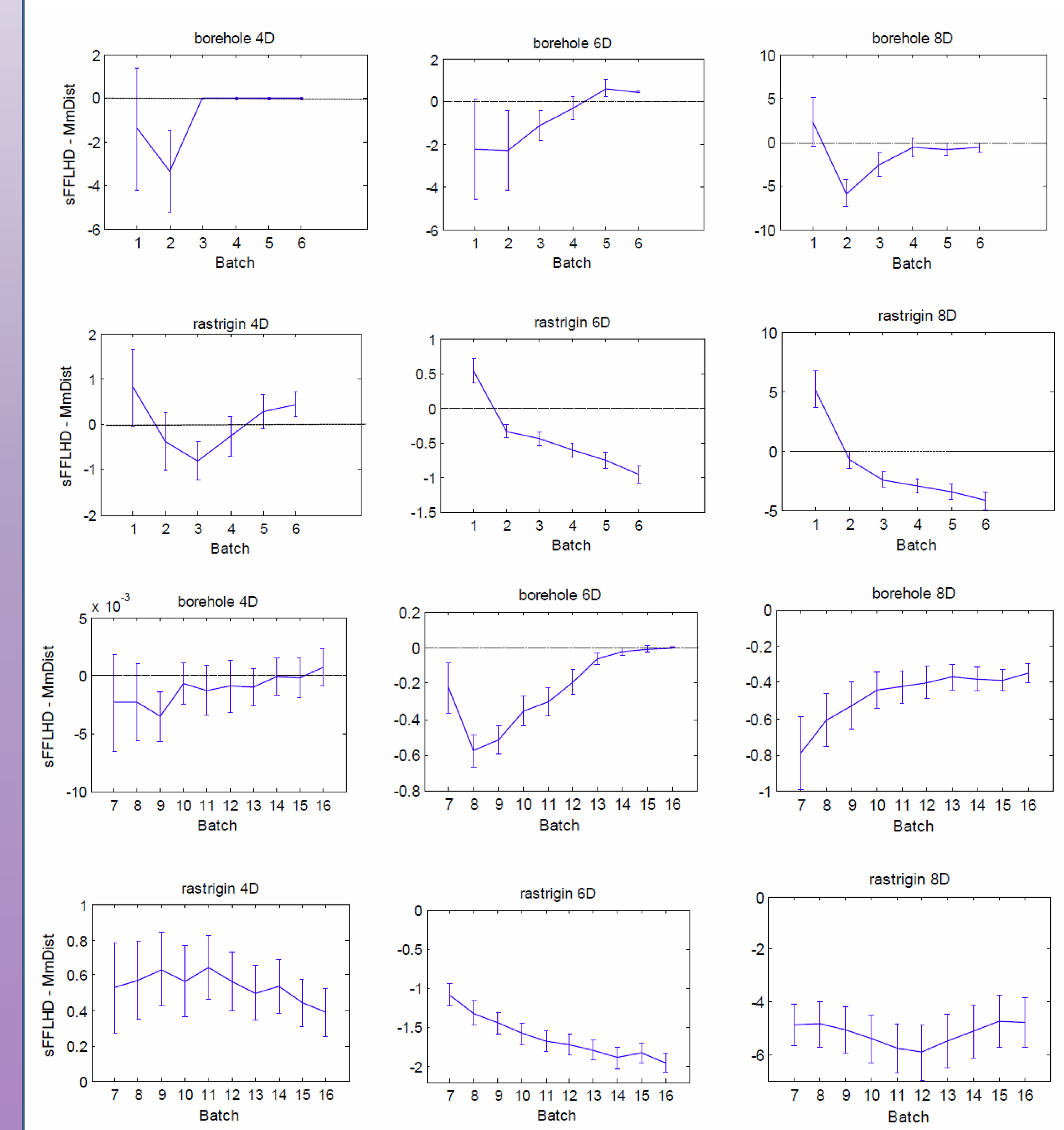
Good Space Filling Properties





MSE in Gaussian Model Fitting

A negative confidence interval for the difference in MSE indicates sFFLHD fits Gaussian model better



Conclusions

- In low dimensional examples, sFFLHD performs as well or nearly as well in terms of RMSE of the GP model fit.
- In high dimensional examples, sFFLHD produces lower RMSE of the GP model fit
- sFFLHD dominates the other tested designs in estimating the mean.
- Each batching being an LHD contributes substantially to sFFLHD's good performance if the design does not reach the orthogonal stages.

Reference

Loeppky, J.L., Moore, J.L.,Williams, B.J., 2010. Batch sequential designs for computer experiments. *Journal of Statistical Planning and Inference*, Volume 140, Issue 6, 1452-1464.

Qian, P.Z.G., Wu, C.F.J., 2009, Sliced Space-Filling Designs. {\it Biometrika}, 96, 945-956.