Design and Analysis of Computer Experiments with Branching and Nested Factors



400 350

100

10

5

z₂

0 10

Department of Statistics Rufgers, the State University of New Jersey

Joint work with Roshan V. Joseph and S. N. Melkote

5

Ζ1

-5

0

Overview



Computer experiments with branching and nested factors.

- ✓ New class of designs.
- ✓ Optimality criteria.
- ✓ New metamodel.

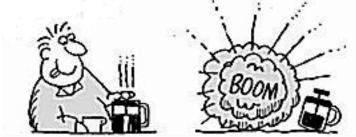
 First computer experiments were conducted at Los Alamos National Laboratory to study the behavior of nuclear weapons.

- First computer experiments were conducted at Los Alamos National Laboratory to study the behavior of nuclear weapons.
- Computer experiments are becoming popular because many physical experiments

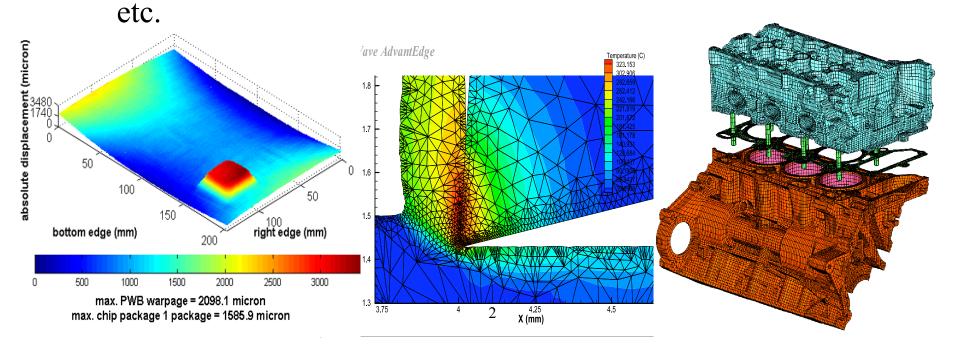
EBOOM EBOOM

are difficult or impossible to perform.

- First computer experiments were conducted at Los Alamos National Laboratory to study the behavior of nuclear weapons.
- Computer experiments are becoming popular because many physical experiments



are difficult or impossible to perform.Applications: biology, environmental science and engineering

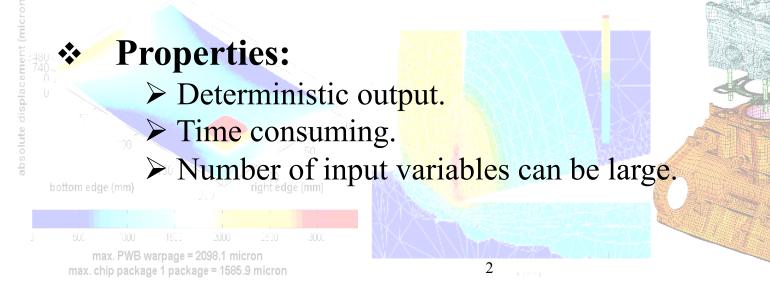


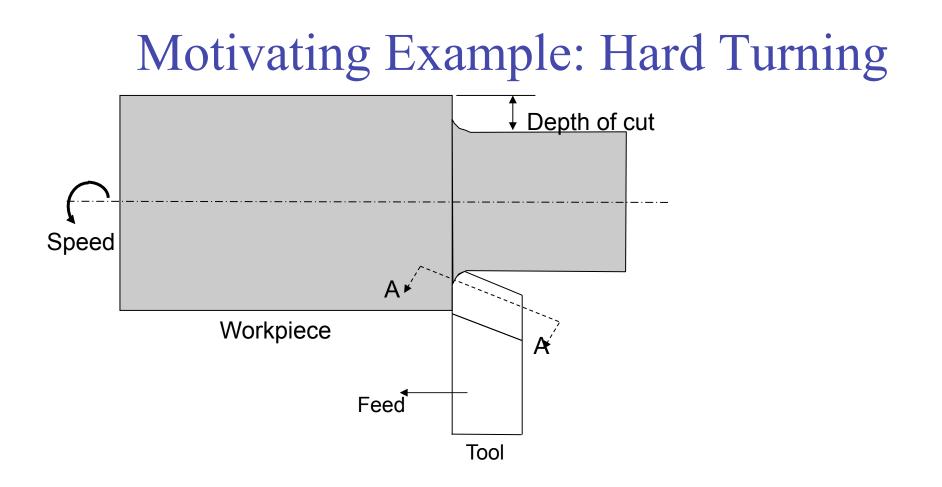
- First computer experiments were conducted at Los Alamos
 National Laboratory to study the behavior of nuclear weapons.
- Computer experiments are becoming popular because many physical experiments

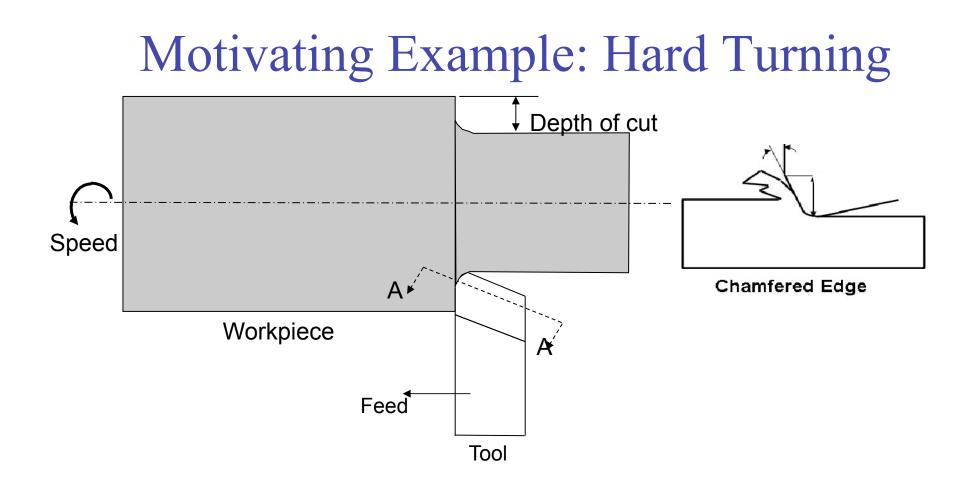


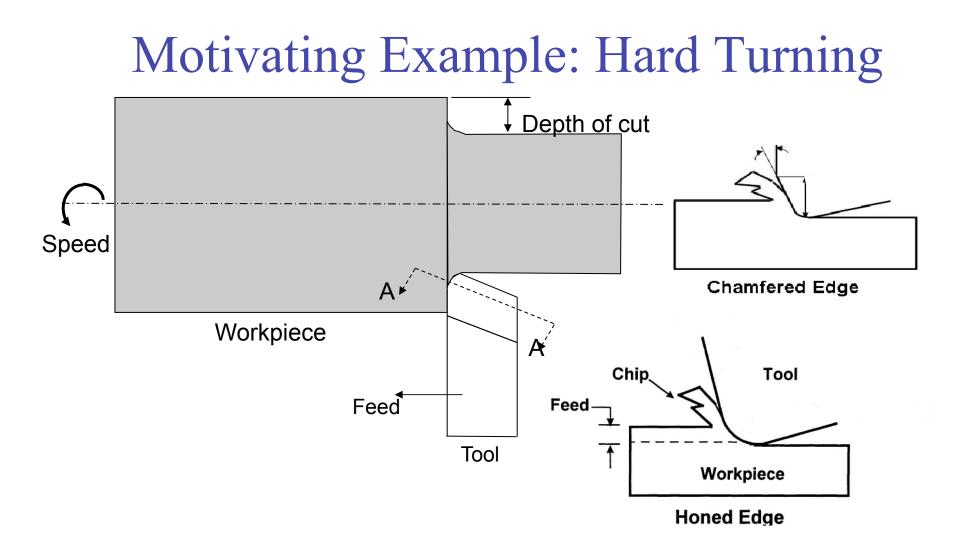
are difficult or impossible to perform. Applications: biology, environmental science and engineering

etc.

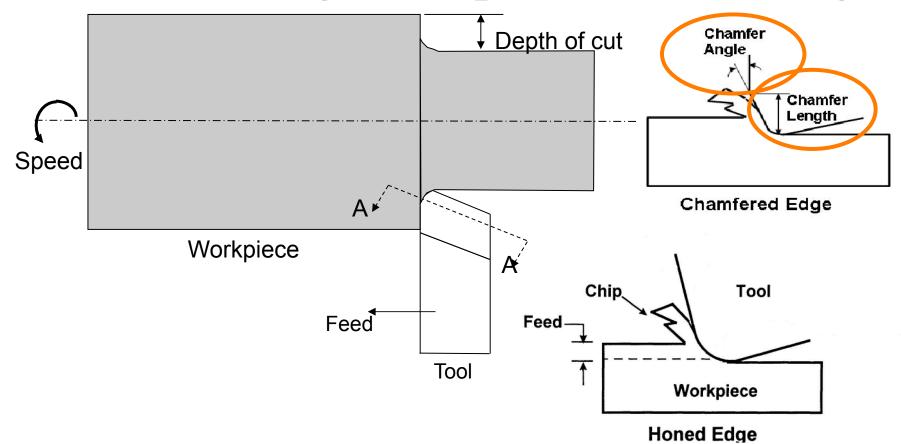








Motivating Example: Hard Turning



- Nested factor: A factor that can change with respect to the level of another factor.
- Branching factor: A factor within which other factors are nested.

- Nested factor: A factor that can change with respect to the level of another factor.
- Branching factor: A factor within which other factors are nested.

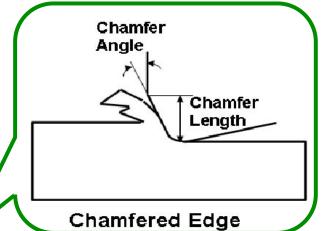
Branching factor	Z ₁	Tool

- Nested factor: A factor that can change with respect to the level of another factor.
- Branching factor: A factor within which other factors are nested.

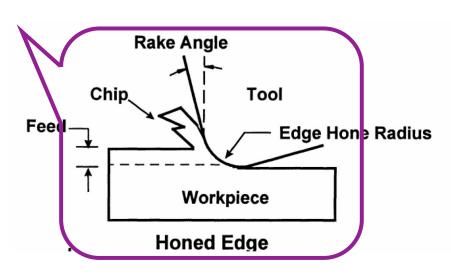
	Chamfer Angle Chamfer
$\langle [$	Length
	Chamfered Edge

Branching factor	Z ₁	Tool (chamfer)

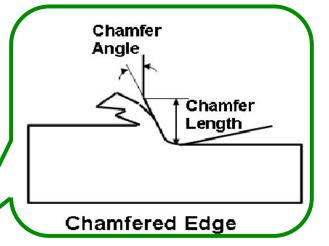
- Nested factor: A factor that can change with respect to the level of another factor.
- Branching factor: A factor within which other factors are nested.



Branching factor	Z ₁	Tool (chamfer & hone)



- Nested factor: A factor that can change with respect to the level of another factor.
- Branching factor: A factor within which other factors are nested.

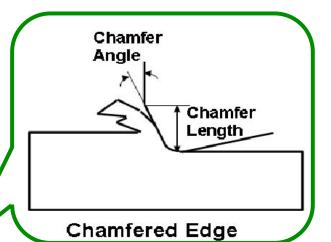


Hard turning experiment

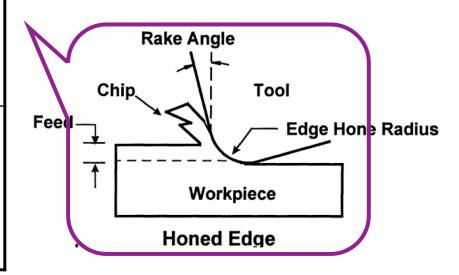
Branching factor	Z ₁	Tool (chamfer & hone)
Nested	v ₁ z ₁ =chamfer	Angle
factors	v ₂ z ₁ =chamfer	Length

4

- Nested factor: A factor that can change with respect to the level of another factor.
- Branching factor: A factor within which other factors are nested.



Паги	turning experim	ent	
Branching factor	Z ₁	Tool (chamfer & hone)	
Nested factors	$v_1 z_1=chamfer$ $v_2 z_1=chamfer$ $v_1 z_1=hone$ $v_2 z_1=hone$	Angle Length None None	



Computer Experiments in Machining

Type of Factor	Notation	Factor	Ranges
Branching factors	Z ₁	Tool (Cutting edge shape)	chamfer & hone
Nested factors	v ₁ z ₁ =chamfer	Angle	17~ 20
	v ₂ z ₁ =chamfer	Length	115~140
	v ₁ z ₁ =hone	None	None
	v ₂ z ₁ =hone	None	None
Shared factors	x ₁	Cutting edge radius	5~25
	x ₂	Rake angle	-15 ~ -5
	X ₃	Tool nose radius	0.4 ~ 1.6
	X ₄	Cutting speed	120 ~ 240
	x ₅	Feed	0.05 ~ 0.15
	x ₆	Depth of cut	0.1 ~ 0.25

- Properties
- Branching factors are qualitative, which cannot be divided into intervals.

• Properties

Branching factors are qualitative, which cannot be divided into intervals.

Branching-by-nested interaction can be significant.

• Properties

Branching factors are qualitative, which cannot be divided into intervals.

Branching-by-nested interaction can be significant. Printed circuit board (PCB) manufacturing

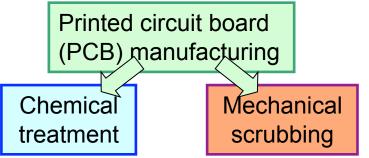
• Properties

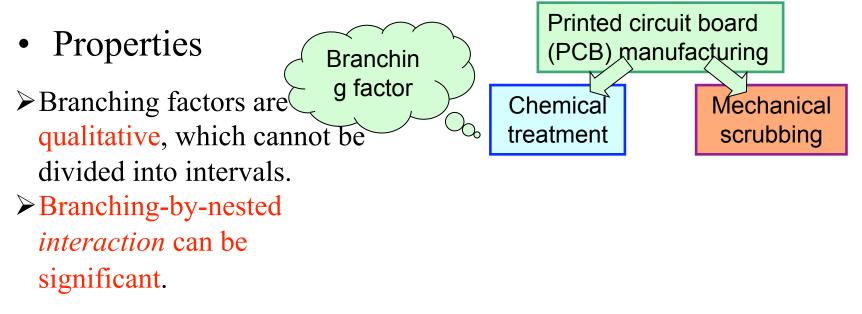
Branching factors are qualitative, which cannot be divided into intervals.

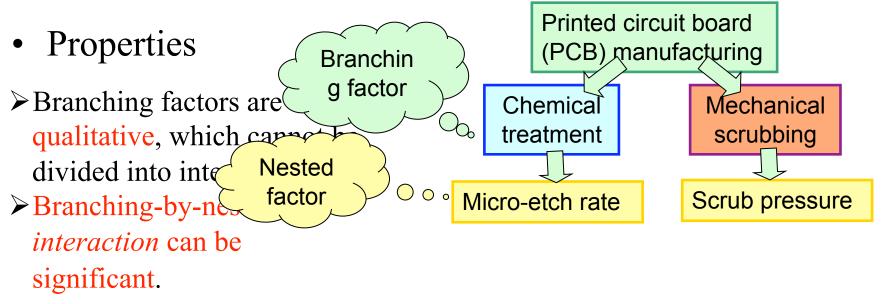
Branching-by-nested interaction can be significant. Printed circuit board (PCB) manufacturing

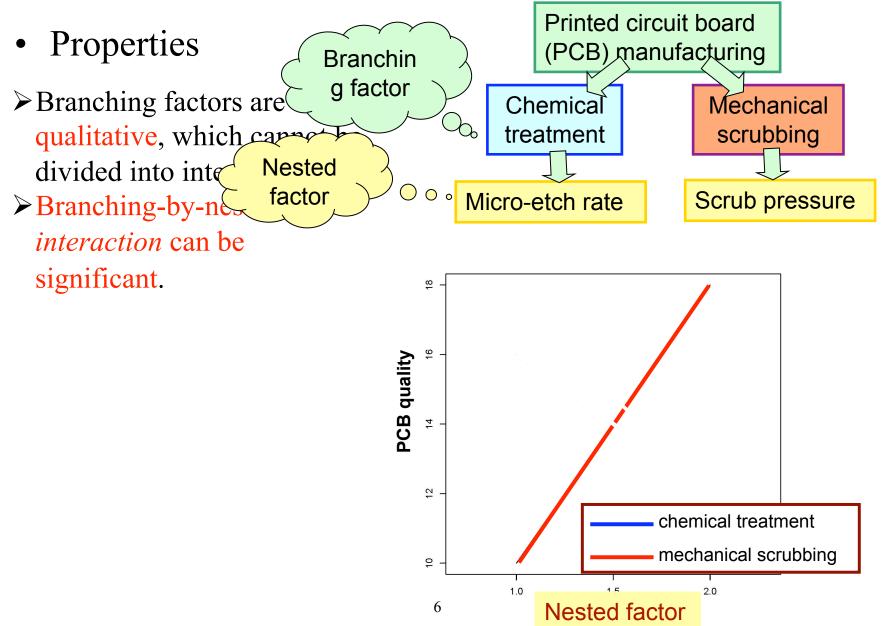
Two surface preparation methods

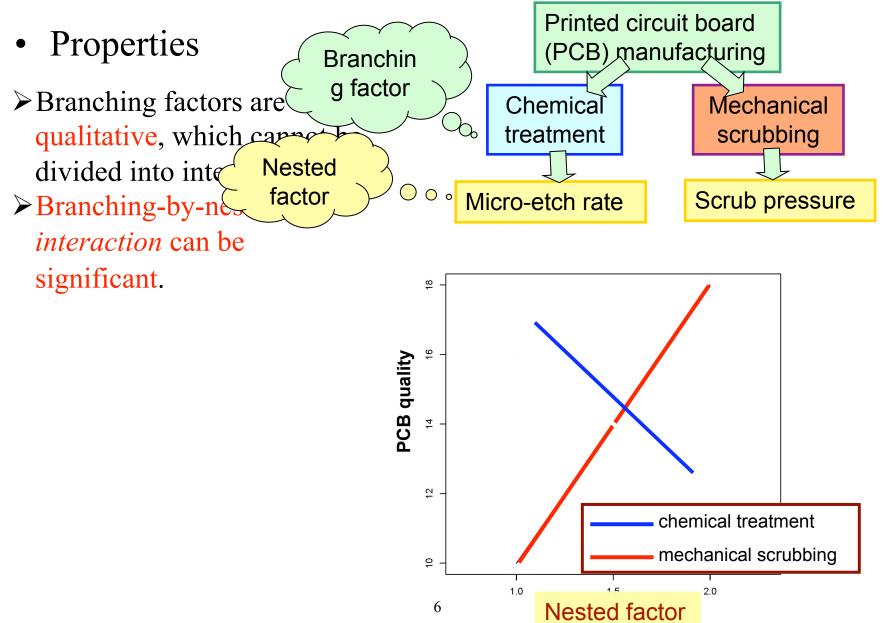
- Properties
- Branching factors are qualitative, which cannot be divided into intervals.
- Branching-by-nested interaction can be significant.

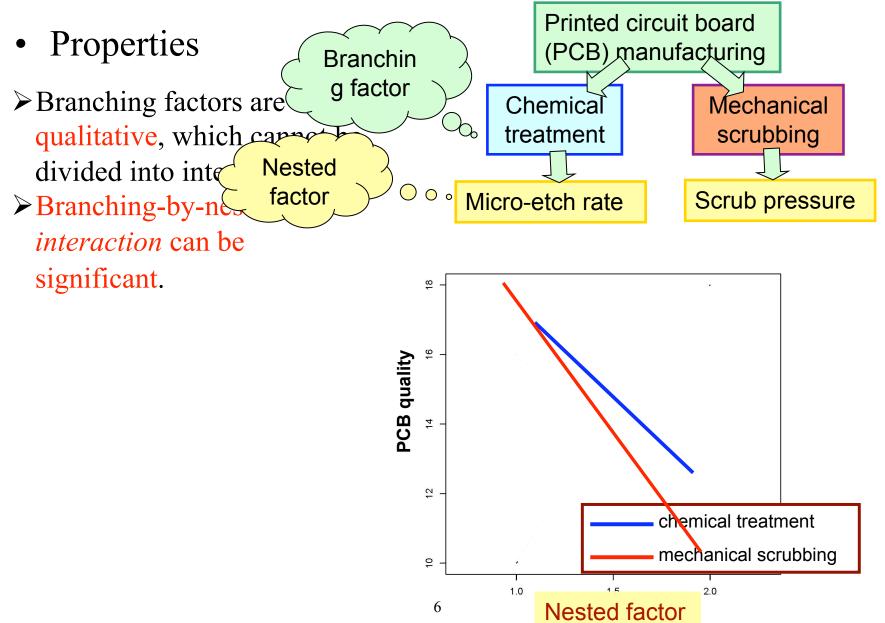


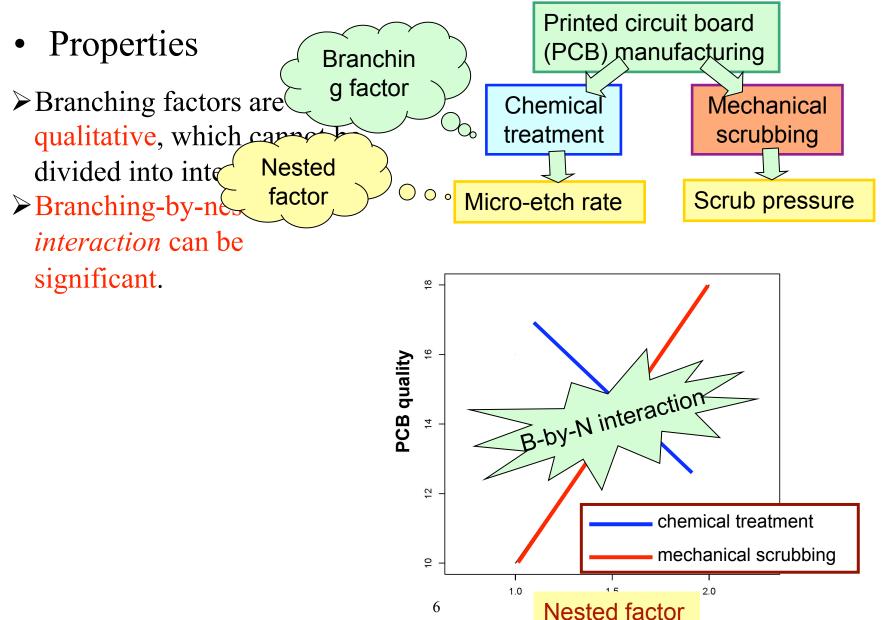


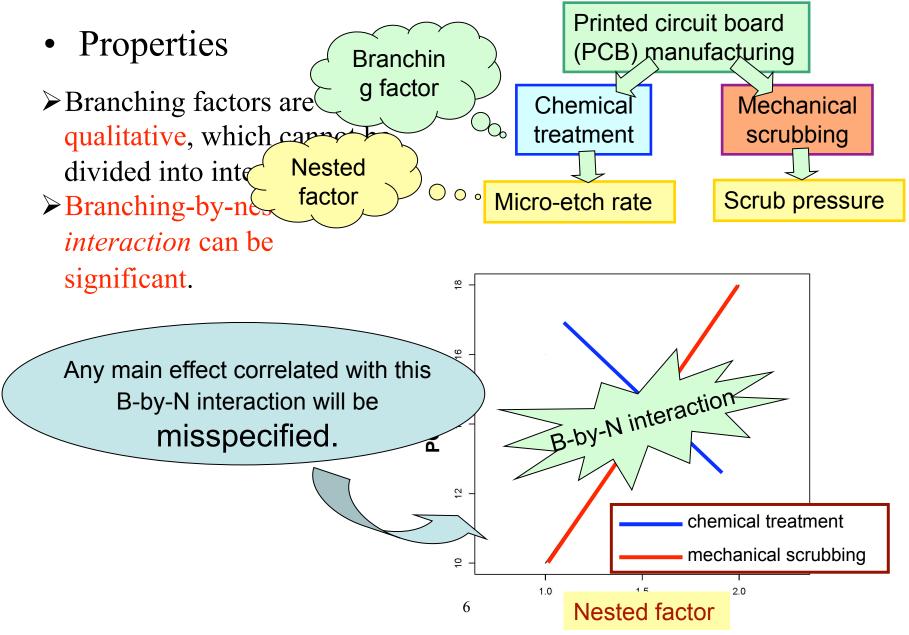












Computer Experiments with Branching and Nested Factors

- Branching and nested factors (Phadke, 1989; Taguchi, 1987).
- Challenges in design:

➢ Involve both quantitative and qualitative factors.

Some two-factor interactions are important.

• Challenges in modeling:

 \succ No correlation function defined for nested factors.

• First work on design and analysis of computer experiments with branching and nested factors.

Computer Experiments with Branching and Nested Factors

- Branching and nested factors (Phadke, 1989; Taguchi, 1987).
- Challenges in design:

No existing design can be applied directly.

• Challenges in modeling:

 \succ No correlation function defined for nested factors.

• First work on design and analysis of computer experiments with branching and nested factors.

Computer Experiments with Branching and Nested Factors

- Branching and nested factors (Phadke, 1989; Taguchi, 1987).
- Challenges in design:

No existing design can be applied directly.

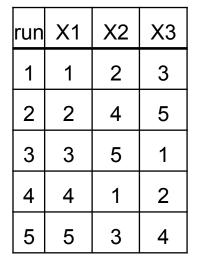
• Challenges in modeling:

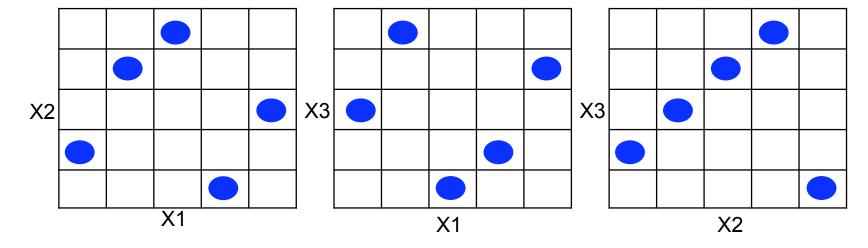
No suitable model available.

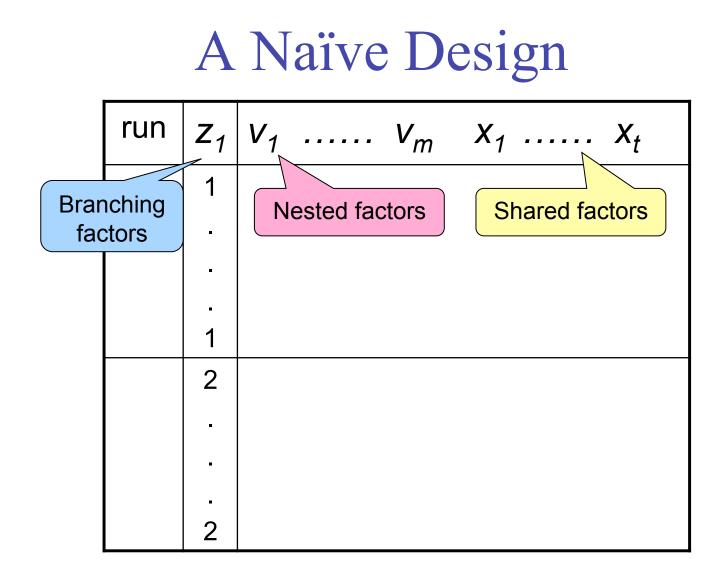
• First work on design and analysis of computer experiments with branching and nested factors.

Design of Computer Experiments with Branching and Nested Factors

- Latin hypercube design (LHD).
 - ≻ McKay, Beckman, Conover (1979).
 - \succ Easy to construct.
 - One-dimensional balance.
- LHD cannot be applied directly.

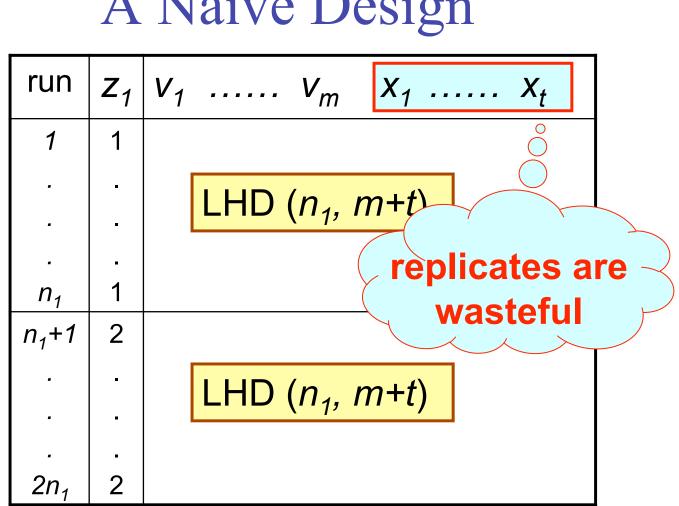




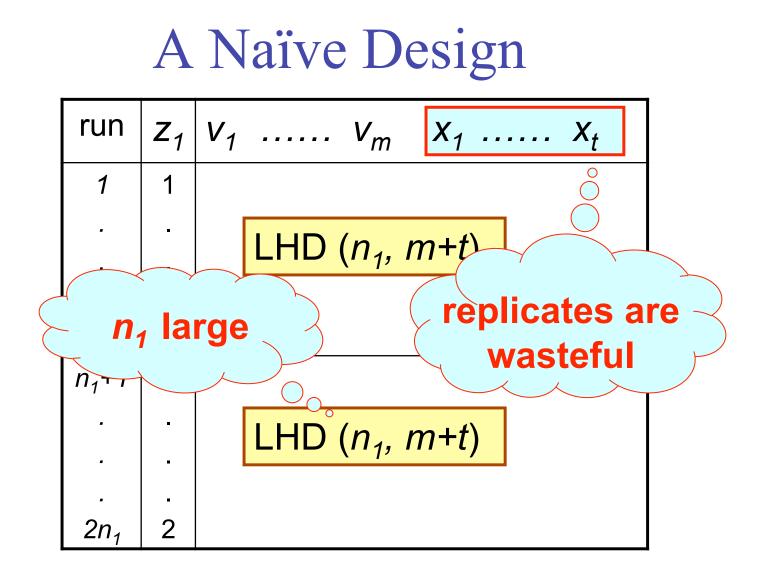


A Naïve Design

run	<i>Z</i> ₁	$V_1 \ldots V_m X_1 \ldots X_t$
1	1	
	-	LHD $(n_1, m+t)$
	•	
n ₁	1	
n ₁ +1	2	
	-	LHD $(n_1, m+t)$
	-	
2n ₁	2	



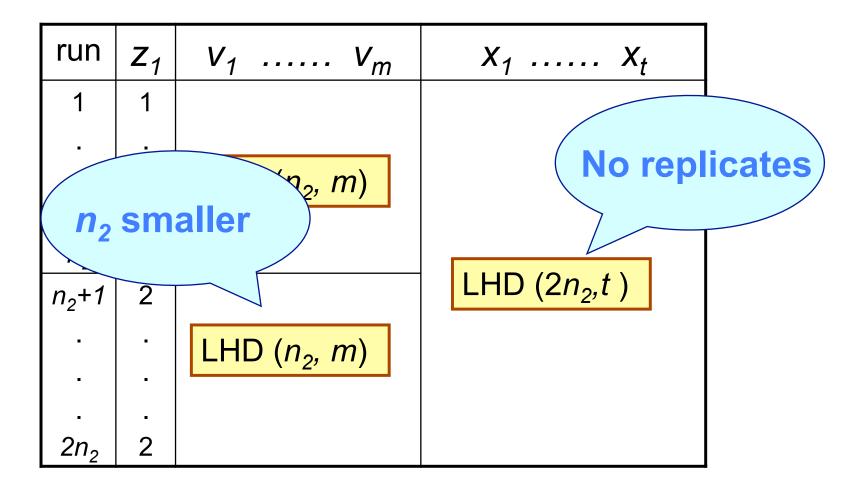
A Naïve Design

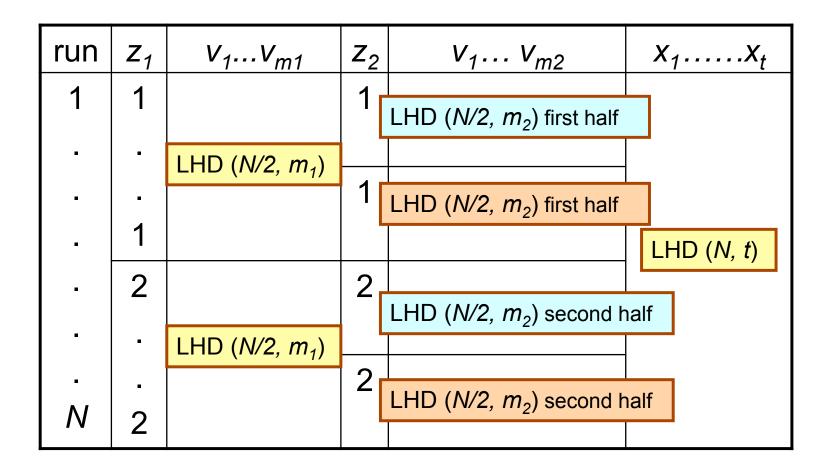


run	Z ₁	V ₁ V _m	X ₁ X _t
1	1		
	-		
	-		
	•		
n ₂	1		
n ₂ +1	2		
	-		
	-		
	-		
2n ₂	2		

run	Z ₁	V ₁ V _m	X ₁ X _t
1	1		
	-		
	-		
-	-		
<i>n</i> ₂	1		
n ₂ +1	2		LHD $(2n_2, t)$
•	-		
.	-		
	-		
2n ₂	2		

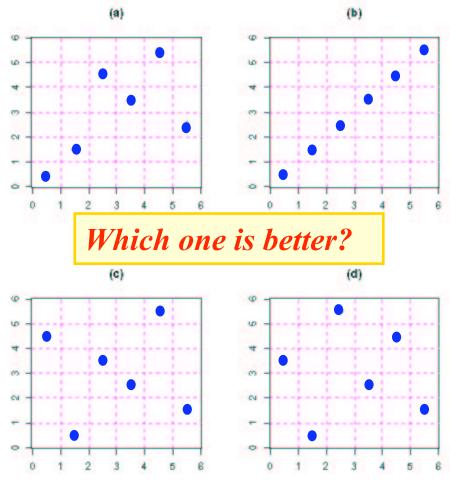
run	Z ₁	V ₁ V _m	X ₁ X _t
1	1		
	-		
	-	LHD (<i>n</i> ₂ , <i>m</i>)	
-	•		
n ₂	1		
n ₂ +1	2		LHD $(2n_2, t)$
	-	LHD (<i>n</i> ₂ , <i>m</i>)	
	-		
-	•		
2n ₂	2		





How to Find a "Good" Design?

• For *n* runs and *k* factors, we can obtain $(n!)^k$ LHDs.



Minimize Correlation

- Iman and Conover (1982), Owen (1994), and Tang (1998) proposed to find designs minimizing correlations among factors.
- Owen (1994)

$$\rho^{2} = \frac{\sum_{i=2}^{k} \sum_{j=1}^{i-1} \rho_{ij}^{2}}{k(k-1)/2},$$

where ρ_{ij} is the linear correlation between columns *i* and *j*.

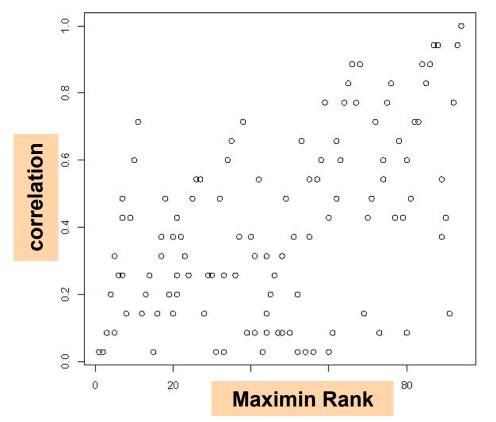
• Figure (c) shows the optimal LHD found by Tang (1998).

Maximize Minimum Distance

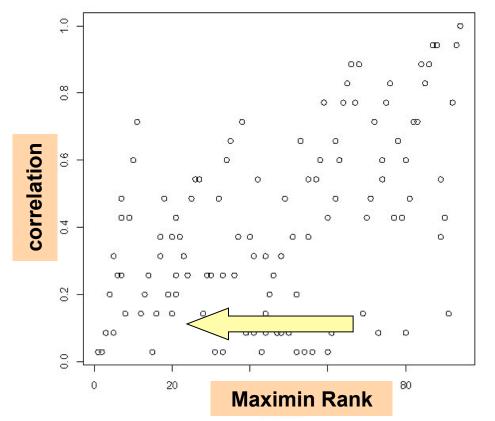
- Johnson, Moore, and Ylvisaker (1990) proposed the maximim distance criterion, which maximizes the minimum inter-site distance.
- Morris and Mitchell (1995) proposed to find the best LHD by maximizing the minimum distance between the points.
- Use a scalar-valued function to rank competing designs.

$$\phi_p = \left(\sum_{i=1}^{\binom{n}{2}} \frac{1}{d_i^p}\right)^{1/p}$$

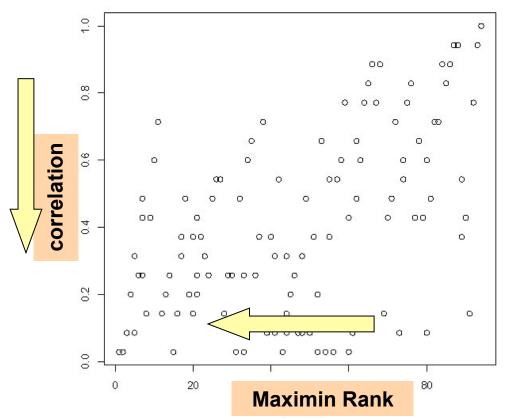
✓ d_i is the (rectangular or Euclidean) distance between two design points.



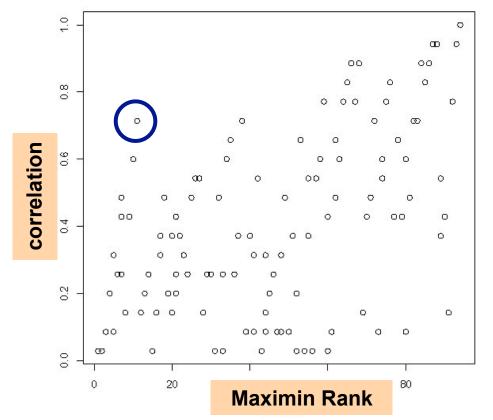
- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.



- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.

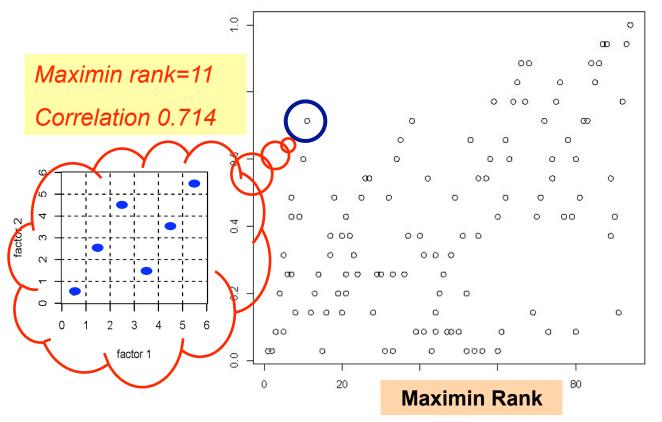


- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.



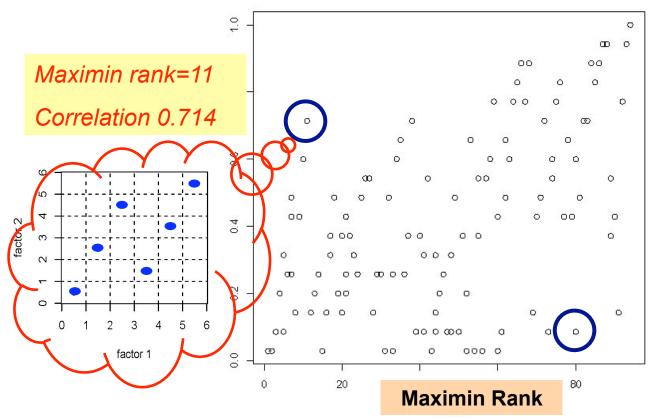
- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.

• Maximin rank vs. correlation in n=6, k=2 case.



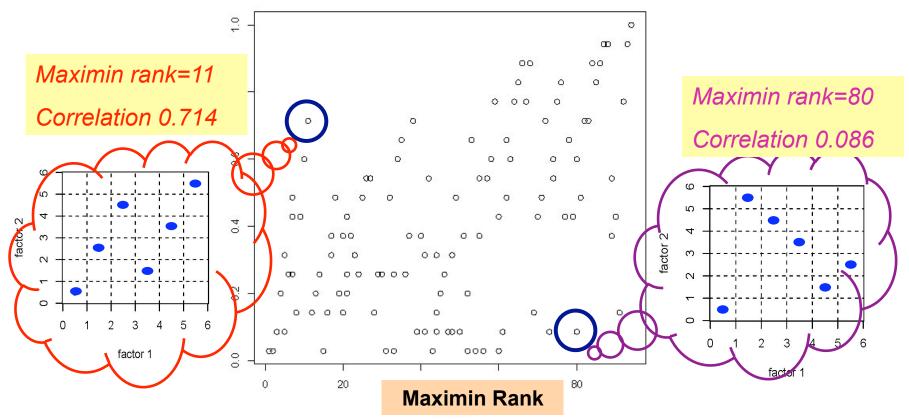
- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.

• Maximin rank vs. correlation in n=6, k=2 case.



- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.

• Maximin rank vs. correlation in n=6, k=2 case.



- These two criteria can give conflicting results.
- Minimization of one criterion may not lead to minimization of the other.

Orthogonal-Maximin Latin Hypercube Design

- Our objective is to find an LHD that minimizes both ρ^2 and ϕ_p .
- Proposition 1: For an $LHD(n,k), \phi_{p,L} \leq \phi_p \leq \phi_{p,U}$.

$$\phi_{p,L} = \left\{ \binom{n}{2} \left(\frac{\left\lceil \bar{d} \right\rceil - \bar{d}}{\left\lfloor \bar{d} \right\rfloor^p} + \frac{\bar{d} - \left\lfloor \bar{d} \right\rfloor}{\left\lceil \bar{d} \right\rceil^p} \right) \right\}^{1/p} \quad \text{and} \quad \phi_{p,U} = \left\{ \sum_{i=1}^{n-1} \frac{(n-i)}{(ik)^p} \right\}^{1/p}$$

• Define

$$\psi_p = w\rho^2 + (1-w)\frac{\phi_p - \phi_{p,L}}{\phi_{p,U} - \phi_{p,L}},$$

where $w \in (0, 1)$.

• A design that minimizes ψ_p is called an orthogonal-maximin Latin hypercube design (OMLHD).

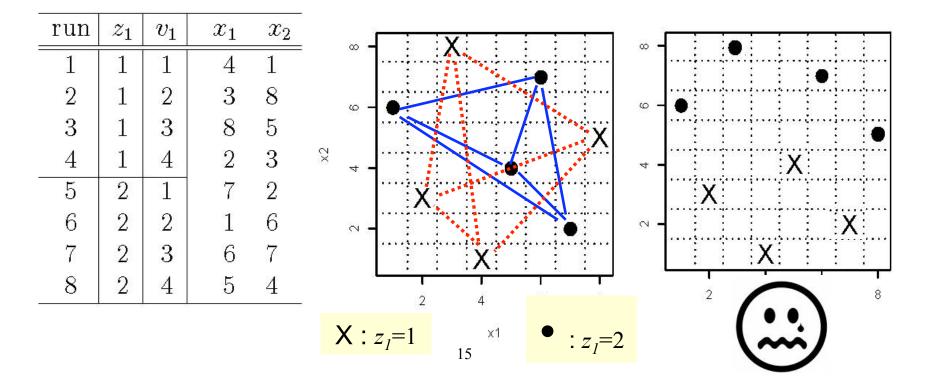
Reference: V. Roshan Joseph and Ying Hung (2008). Orthogonal-Maximin Latin Hypercube Designs, *Statistica Sinica*, 18, 171-186.

Orthogonal-Maximin BLHD

Orthogonal-Maximin BLHD

$$\rho^{2} = \frac{\sum_{i=2}^{p} \sum_{j=1}^{i-1} \rho_{ij}^{2} + \sum_{i=1}^{t} \sum_{j=1}^{\alpha} \tilde{\rho}_{ij}^{2}}{(p(p-1)/2) + \alpha t},$$

$$\square \longrightarrow \phi_{P} = \left(\sum_{g \neq h} \left[\frac{t}{d_{x}(g, h)}\right]^{P} + \sum_{i=1}^{k_{1}} \sum_{g_{t+1} = h_{t+1} = z_{1,i}} \left[\frac{1+t}{d_{v_{1}}(g, h) + d_{x}(g, h)}\right]^{P}\right)^{1/P},$$



Orthogonal-Maximin BLHD

• Orthogonal-Maximin BLHD

$$\rho^{2} = \frac{\sum_{i=2}^{p} \sum_{j=1}^{i-1} \rho_{ij}^{2} + \sum_{i=1}^{t} \sum_{j=1}^{\alpha} \tilde{\rho}_{ij}^{2}}{(p(p-1)/2) + \alpha t},$$

$$\phi_{P} = \left(\sum_{g \neq h} \left[\frac{t}{d_{x}(g, h)}\right]^{P} + \sum_{i=1}^{k_{1}} \sum_{g_{t+1} = h_{t+1} = z_{1,i}} \left[\frac{1+t}{d_{v_{1}}(g, h) + d_{x}(g, h)}\right]^{P}\right)^{1/P},$$

Proposition 2:

$$\begin{aligned} \phi_{P,L} &= 6\kappa \bigg[t(N^2 - 1) + \sum_{i=1}^{k_1} m_1(n_1^2 - 1) \bigg]^{-1}, \\ \phi_{P,U} &= N \Bigg[\sum_{i=1}^{k_1} \sum_{j=1}^{n_1 - 1} \frac{(n_1 - j)(t + m_1)^P}{j^P(t + k_1 m_1)^P} + \sum_{j=1}^{N-1} \frac{N - j}{j^P} \bigg]^{1/P}. \end{aligned}$$

Orthogonal-Maximin BLHD

• Orthogonal-Maximin BLHD

$$\rho^{2} = \frac{\sum_{i=2}^{p} \sum_{j=1}^{i-1} \rho_{ij}^{2} + \sum_{i=1}^{t} \sum_{j=1}^{\alpha} \tilde{\rho}_{ij}^{2}}{(p(p-1)/2) + \alpha t},$$

$$\phi_{P} = \left(\sum_{g \neq h} \left[\frac{t}{d_{x}(g, h)}\right]^{P} + \sum_{i=1}^{k_{1}} \sum_{g_{t+1} = h_{t+1} = z_{1,i}} \left[\frac{1+t}{d_{v_{1}}(g, h) + d_{x}(g, h)}\right]^{P}\right)^{1/P},$$

Proposition 2:

$$\phi_{P,L} = 6\kappa \left[t(N^2 - 1) + \sum_{i=1}^{k_1} m_1(n_1^2 - 1) \right]^{-1},$$

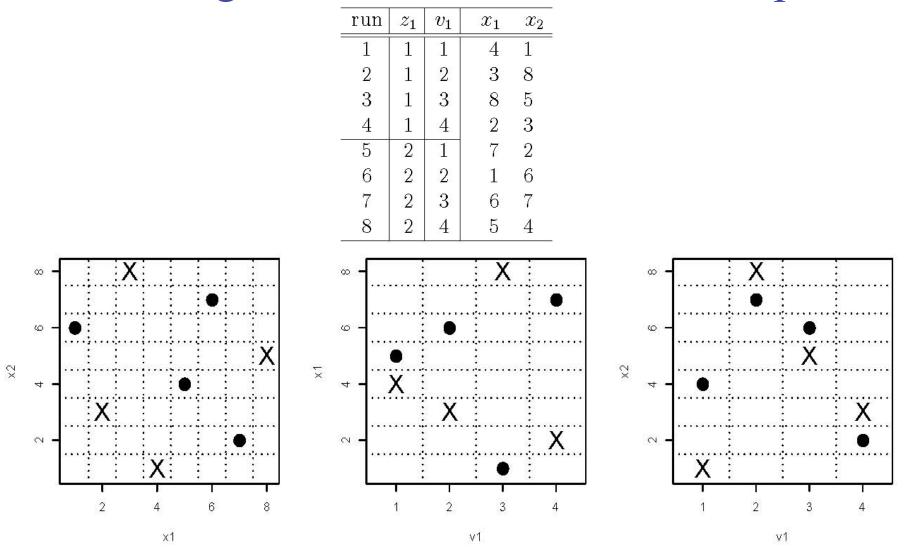
$$\phi_{P,U} = N \left[\sum_{i=1}^{k_1} \sum_{j=1}^{n_1 - 1} \frac{(n_1 - j)(t + m_1)^P}{j^P(t + k_1 m_1)^P} + \sum_{j=1}^{N-1} \frac{N - j}{j^P} \right]^{1/P}.$$

• A multi-objective criterion is to minimize

$$\psi_P = w\rho^2 + (1-w)\frac{\phi_P - \phi_{P,L}}{\phi_{P,U} - \phi_{P,L}}$$

• Heuristic algorithm: simulated annealing algorithm(SAA)

Orthogonal-Maximin BLHD Example



16

Model Fitting in Computer Experiments

- Kriging is first developed 40 years ago by G. Matheron and named in honor of D. Krige.
- Kriging is widely used in the analysis of computer experiments (Sacks et al. 1989; Santner et al. 2003) and in spatial statistics.
- Interpolating metamodel.
 - Interpolation property essential for *deterministic* computer experiments.
- Efficient in higher dimensions.

• Kriging:

$$Y(\mathbf{x}) = \mathbf{v}(\mathbf{x})' \mu_m + Z(\mathbf{x}),$$

 $Z(\mathbf{x})$ is assumed to be a weak stationary stochastic process with mean 0 and covariance function $\sigma_m^2 \psi$

• Product correlation function:

$$\gg w_1 = (x_{11}, x_{12}, z_{11}, v_{11}^{z_{11}})$$
, $w_2 = (x_{21}, x_{22}, z_{21}, v_{21}^{z_{21}})$.
 $\gg \operatorname{cor}(Y(w_1), Y(w_2))$:

• Kriging:

$$Y(\mathbf{x}) = \mathbf{v}(\mathbf{x})' \mu_m + Z(\mathbf{x}),$$

 $Z(\mathbf{x})$ is assumed to be a weak stationary stochastic process with mean 0 and covariance function $\sigma_m^2 \psi$

• Product correlation function:

$$> oldsymbol{w}_1 = (x_{11}, x_{12}, z_{11}, v_{11}^{z_{11}})$$
 , $oldsymbol{w}_2 = (x_{21}, x_{22}, z_{21}, v_{21}^{z_{21}})$.

 $\succ \operatorname{cor}(\boldsymbol{Y}(\boldsymbol{w}_1), \boldsymbol{Y}(\boldsymbol{w}_2)) = \varphi_1(x_{11}, x_{21})\varphi_2(x_{11}, x_{21}) \boldsymbol{\varpi}_1(v_{11}, v_{21}) \boldsymbol{\xi}_1(z_{11}, z_{21}).$

• Kriging:

$$Y(\mathbf{x}) = \mathbf{v}(\mathbf{x})' \mu_m + Z(\mathbf{x}),$$

 $Z(\mathbf{x})$ is assumed to be a weak stationary stochastic process with mean 0 and covariance function $\sigma_m^2 \psi$

• Product correlation function:

 $\succ \boldsymbol{w}_{1} = (x_{11}, x_{12}, z_{11}, v_{11}^{z_{11}}), \boldsymbol{w}_{2} = (x_{21}, x_{22}, z_{21}, v_{21}^{z_{21}}).$ $\succ \operatorname{cor}(\boldsymbol{Y}(\boldsymbol{w}_{1}), \boldsymbol{Y}(\boldsymbol{w}_{2})) = \varphi_{1}(x_{11}, x_{21})\varphi_{2}(x_{11}, x_{21}) \overline{\omega}_{1}(v_{11}, v_{21})\xi_{1}(z_{11}, z_{21}).$

• Gaussian correlation function for shared fag prs:

$$\psi(\boldsymbol{h}) = \exp(-\sum_{j=1}^{P} \theta_j h_j^2)$$

• Kriging:

$$Y(\mathbf{x}) = \mathbf{v}(\mathbf{x})' \mu_m + Z(\mathbf{x}),$$

 $Z(\mathbf{x})$ is assumed to be a weak stationary stochastic process with mean 0 and covariance function $\sigma_m^2 \psi$

• Product correlation function:

 $\succ \boldsymbol{w}_{1} = (x_{11}, x_{12}, z_{11}, v_{11}^{z_{11}}), \boldsymbol{w}_{2} = (x_{21}, x_{22}, z_{21}, v_{21}^{z_{21}}).$ $\succ \operatorname{cor}(\boldsymbol{Y}(\boldsymbol{w}_{1}), \boldsymbol{Y}(\boldsymbol{w}_{2})) = \varphi_{1}(x_{11}, x_{21})\varphi_{2}(x_{11}, x_{21})\overline{\omega}_{1}(v_{11}, v_{21})\xi_{1}(z_{11}, z_{21}).$

• Isotropic correlation for the branching factors:

$$\xi_1(z_{11}, z_{21}) = \exp\left\{-\theta_1 I_{[z_{11} \neq z_{21}]}\right\}.$$

New correlation function for the nested factors

n

• Define new correlation function for nested factors:

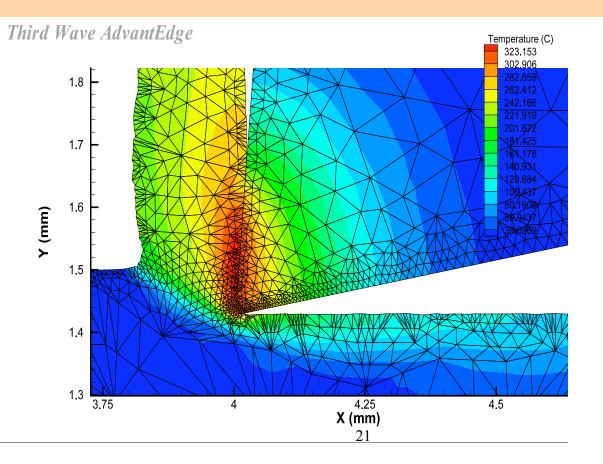
$$\varpi_1(v_{11}^{z_{11}}, v_{21}^{z_{21}}) = \exp\bigg\{-\sum_{j=1}^2 \gamma_j(v_{11}^{z_{11}} - v_{21}^{z_{21}})^2 I_{[z_{11}=z_{21}=j]}\bigg\}.$$

Computer Experiments in Machining

Type of Factor	Notation	Factor	Ranges	
Branching factors	Z ₁	Tool (Cutting edge shape)	chamfer & hone	
Nested factors	v ₁ z ₁ =chamfer	Angle	17~ 20	
	v ₂ z ₁ =chamfer	Length	115~140	
	v ₁ z ₁ =hone	None	None	
	v ₂ z ₁ =hone	None	None	
Shared factors	Shared factors x ₁		5~25	
	x ₂	Rake angle	-15 ~ -5	
	X ₃	Tool nose radius	0.4 ~ 1.6	
	X ₄	Cutting speed	120 ~ 240	
	x ₅	Feed	0.05 ~ 0.15	
	x ₆	Depth of cut	0.1 ~ 0.25	

Finite Element Based Simulation in Machining

- Hard turning experiments are simulated from *AdvantEdge*
- Time consuming: 12 hrs ~1 day per run



Finite Element Based Simulation in Machining

Run	z1	v1	v_2	×1	x_2	æβ	∞4	°Б	×6	force
1	1	1	6	15	23	7	9	18	10	162.1
2	1	2	11	25	3	25	14	25	19	284.9
3	1	3	3	4	20	18	18	5	26	160.3
4	1 1 1	4	14	9	6 8	6	27	ל 2	17	121.1
5	1	5	8	16	8	21	2	2	1	104.6
1 2 4 5 7 8 9	1 1 1	3 4 5 6	1	17	10	5	25	19	25	241.9
ל	1	ተ	12	29	26	15	5	14	12	195.4
8	1	7 8	5	26	16	30	22	15	6	159.5
9	1	9	15	ተ	13	26	ተ	11	27	241.6
10	1 1 1	10	10	1	29	20	23	6	5	88.33
11	1	11	2	20	21	27	10	20	29	320.4
12	1 1	12	7	8	11	14	4	29	21	218.8
13	1	13	13	22	9	1	24	27	9	193.5
14	1	14	4	10	2	24	28	13	13	198.6
15	1	15	9	28	25	13	17	3 8	28	155.1
16	2			19	5	9	1	8	20	164.4
17	2			14	28	17	6	21	24	323.6
18	2			6	17	4	16	12	4	109.1
19	2 2 2			11	1	12	15	4	8	115.4
20	2			27	22	8	30	24	16	254.8
21	2 2 2			21	14	23	19	10	22	217.0
22	2			23	18	22	12	28	3	243.7
28	2 2 2			8	27	3	3	26	14	131.5
24	2			13	15	19	29	16	30	258.7
25	2			24	12	2	11	1	18	109.3
26	2			18	24	28	8	17	2	174.8
27	2 2			12	30	11	26	9	11	157.0
28	2			2	4	16	13	30	15	133.1
29	2 2			30	7	10	20	23	7	210.1
30	2			5	19	29	21	22	23	278.8

Computer Experiments in Machining

• Fitted kriging model:

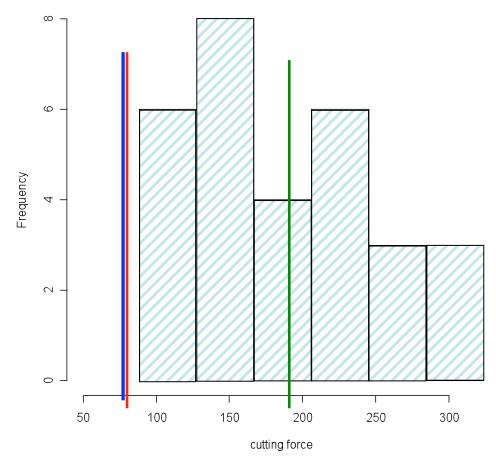
 $\hat{y}(\mathbf{w}) = \mathbf{5.1} + \mathbf{0.2x_{5l}} + \mathbf{0.2x_{6l}} - \mathbf{0.12x_{1l}x_{6l}} + \hat{\psi}(\mathbf{w})' \hat{\Psi}^{-1}(\mathbf{y} - \mathbf{V_3}\hat{\mu}_3).$

• Optimal setting:

Factors	setting			
Cutting edge shape	Chamfer angle: 18.74, length 128.13			
Cutting edge radius	5			
Rake angle	-13.8			
Tool nose radius	1.41			
Cutting speed	222			
Feed	0.067			
Depth of cut	0.123			

Computer Experiments in Machining

histogram of observed cutting force



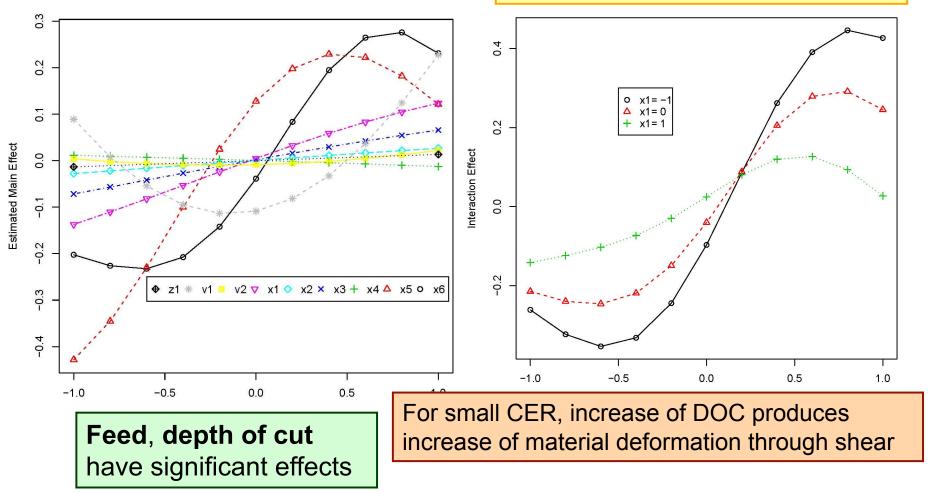
- Predicted minimum cutting force: 81 N (Lower than 191N, the observed average)
- Confirmation experiments: 79 N \square Confirms the validity of the optimal setting.

Sensitivity Analysis

Main effects plot

Interaction:

cutting edge radius & depth of cut



Concluding Remarks

- The first work in computer experiments with branching and nested factors.
- A new class of design, branching Latin hypercube design (BLHD), is proposed and optimal criteria are discussed.
- New metamodel: blind kriging and new correlation function.
- New method provides an efficient way to find optimal settings of branching factors, nested factors and shared factors simultaneously.

Thank you !