# Optimal Compound Orthogonal Arrays and Single Arrays for Robust Parameter Design Experiments

Peng Zeng

Department of Mathematics and Statistics
Auburn University

(joint work with Yu Zhu and Kristofer Jennings)

July 2006

# Outline

- Introduction
- Optimal compound orthogonal arrays
- Optimal single arrays
- Compound orthogonal arrays versus single arrays
- Conclusion

# Robust Parameter Design

- Quality improvement via variation reduction (Taguchi 1986).
- Two types of factors.
  - *Control factors* include the design parameters in a system.
  - *Noise factors* are hard to control. For example, temperature, etc.
- Key idea:
  - Explore the effects of control factors, noise factors, and *their interactions*, and choose control factor settings to simultaneously bring the system's mean response on target and reduce the performance variation.
- Both control and noise factors are systematically varied during experimentation.

# Experimental Plans and Modeling Strategies I

- Cross array or inner-outer product array (Taguchi 1986) is a cross product of an array for control factors and an array for noise factors.

control array / noise array / cross array / COA

$$\begin{pmatrix} + & + \\ + & - \\ - & + \\ - & - \end{pmatrix} \qquad M = \begin{pmatrix} + & + & + \\ + & - & - \\ - & + & - \\ - & - & + \end{pmatrix} \qquad \begin{pmatrix} + & + & M \\ + & - & M \\ - & + & M \\ - & - & M \end{pmatrix} \qquad \begin{pmatrix} + & + & M_1 \\ + & - & M_2 \\ - & + & M_3 \\ - & - & M_4 \end{pmatrix}$$

- Compound orthogonal array (COA) is a generalization of cross array (Rosenbaum 1994, 1996, 1998). It does not require crossing structure.
- At each fixed control setting, response mean and dispersion can be calculated. Use location-dispersion modeling to identify best settings of control factors.

# Experimental Plans and Modeling Strategies II

- Single array or combined array (Welch, Yu, Kang, and Sacks 1990; Shoemaker, Tsui, and Wu 1991) means one single design (array) to accommodate both type of factors.

$$\begin{pmatrix} A & B & a & b & c \\ + & + & + & + & + \\ + & + & - & + & - \\ + & - & + & - & - \\ + & - & - & - & + \\ & \cdots & \cdots & \cdots & \\ - & - & - & + & + \end{pmatrix}$$

- Use response modeling to model response as a function of control effects, noise effects and their interactions. Derive mean response model and variance transmitted model for optimization.

# Some Existing Optimality Criteria

- Optimal compound orthogonal arrays.
  - based on strength vector (Hedayat and Stufken 1999)
- Optimal single arrays.
  - based on modified wordlength pattern (Bingham and Sitter 2003)
  - based on new effects ordering principle and weighted combination of wordlength patterns (Wu and Zhu 2003)

  not sensitive or too complex
- Simple and direct approaches are needed in order to construct optimal experimental plans for robust parameter designs.

# $2^{l-p}$ Designs

- Regular two-level fractional factorial designs are generated by $p$ independent defining words.
- Defining contrast subgroup $\mathcal{G}$.
- Wordlength pattern, $W = (W_1, W_2, \ldots, W_l)$, where $W_i$ is the number of defining words of length $i$ in $\mathcal{G}$.
- Resolution is the smallest $i$ such that $W_i > 0$, and

$$\text{strength} = \text{resolution} - 1$$

- Minimum aberration (MA) design sequentially minimizes $W_i$ and considered to be optimal.
- An effect is clear if it is not aliased with main effects and two-factor interactions.

# $2^{(l_1+l_2)-p}$ Designs with Two Groups of Factors

- $l_1$ factors belong to Group I and $l_2$ factors belong to Group II.
- $p$ independent defining words and defining contrast subgroup $\mathcal{G}$.
- Wordtype pattern matrix $(A_{i.j})_{0\le i\le l_1;0\le j\le l_2}$, where $A_{i.j}$ is the number of words in $\mathcal{G}$ involving $i$ Group I factors and $j$ Group II factors.
- $A_{i.j}$ can be arranged into a sequence

$$W_t = (A_{3.0}, A_{2.1}, A_{1.2}, A_{0.3}, A_{4.0}, A_{3.1}, A_{2.2}, A_{1.3}, A_{0.4}, \ldots)$$

- For robust parameter design, $l_c$ control factors form Group I and $l_n$ noise factors form Group II.
- Both compound orthogonal arrays and single arrays can be viewed as $2^{(l_c+l_n)-p}$ designs.

# Roadmap for Constructing Optimal Plans for RPD

- For experiment size $N = 2^k$, define

$$S(2^k) = \{(i, j) : \lceil \log_2(i+1) \rceil + \lceil \log_2(j+1) \rceil \leq k\}.$$

- A cross array with $l_c$ control factors and $l_n$ noise factors exists if and only if $(l_c, l_n) \in S(2^k)$.
- When $k = 4$,

| $(i, j)$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 2 | ✓ | ✓ | ✓ | × | × | × | × |
| 3 | ✓ | ✓ | ✓ | × | × | × | × |
| 4 | ✓ | × | × | × | × | × | × |
| 5 | ✓ | × | × | × | × | × | × |
| 6 | ✓ | × | × | × | × | × | × |
| 7 | ✓ | × | × | × | × | × | × |

- We will proposed optimality criteria for experimental plans under different conditions.

| | $(l_c, l_n) \in S(2^k)$ | $(l_c, l_n) \notin S(2^k)$ |
|---|---|---|
| location-dispersion modeling | COA | — |
| response modeling | prodigal SA | economical SA |

- For compound orthogonal arrays, assume $A_{i.1} = A_{i.2} = 0$ for $1 \leq i \leq l_c$. Otherwise, they are not able to identify dispersion effects (Rosenbaum 1996).

# Compound Orthogonal Arrays

- Strength vector $T = (t_c, t_n, t_a)$, where

$$t_c = \min(l_c, \min\{i - 1 : A_{i.0} \neq 0 \text{ and } i \geq 1\})$$
$$t_n = \min(l_n, \min\{j - 1 : A_{i.j} \neq 0 \text{ and } j \geq 1\})$$
$$t_a = \min\{i + j - 1 : A_{i.j} \neq 0 \text{ and } i + j \geq 1\}$$

  Designs with same strength vector can be quite different.

- Example: Consider four 64-run designs with four control factors $(A, B, C, D)$ and six noise factors $(a, b, c, d, e, f)$. The first three have the same strength vector $T = (3, 2, 3)$, and the last one $T = (3, 2, 2)$.

| Design | Clear effects (C, n, CC, Cn, nn) |
|---|---|
| $I = ABCD = Dabd = Dace = Dbcf$ | (4, 6, 0, 18, 0) |
| $I = ABCD = abde = ABacd = ACabf$ | (4, 6, 0, 24, 9) |
| $I = ABCD = abce = abdf = ACacd$ | (4, 6, 0, 24, 0) |
| $I = ABCD = abd = ace = bcf$ | (4, 0, 0, 24, 0) |

# Order of Aliasing Severity

- For defining words of length 4,

$$Cnnn \lhd CCCC \lhd nnnn \quad \text{or} \quad A_{1.3} \lhd A_{4.0} \lhd A_{0.4}$$

- For defining words of length 5,

$$CCnnn \lhd Cnnnn \lhd CCCCC \lhd nnnnn \quad \text{or} \quad A_{2.3} \lhd A_{1.4} \lhd A_{5.0} \lhd A_{0.5}$$

- $A_{i_1.j_1}$ is more severe than $A_{i_2.j_2}$, denoted by $A_{i_1.j_1} \lhd A_{i_2.j_2}$, if

$$(i)\ i_1 + j_1 < i_2 + j_2; \quad \text{or}$$

$$(ii)\ |i_1 - j_1| < |i_2 - j_2| \text{ and } i_1 + j_1 = i_2 + j_2; \quad \text{or}$$

$$(iii)\ i_1 > i_2 \text{ and } |i_1 - j_1| = |i_2 - j_2| \text{ and } i_1 + j_1 = i_2 + j_2.$$

- $W_c$ sequence

$$W_c = (A_{3.0}, A_{0.3}, A_{1.3}, A_{4.0}, A_{0.4}, A_{2.3}, A_{1.4}, A_{5.0}, A_{0.5}, A_{3.3}, \ldots)$$

# $W_c$-Aberration

- $D_1$ is said to have less $W_c$-aberration than $D_2$ if $W_c(D_1)$ and $W_c(D_2)$ first differ at $A_{i_0.j_0}$ and $A_{i_0.j_0}(D_1) < A_{i_0.j_0}(D_2)$.
- Example (continued): For the four designs,

$$W_c(D_1) = (0, 0, 4, 1, 3, 0, 0, 0, 0, 4, 0, \dots)$$
$$W_c(D_2) = (0, 0, 0, 1, 1, 8, 0, 0, 0, 0, 4, \dots)$$
$$W_c(D_3) = (0, 0, 0, 1, 3, 8, 0, 0, 0, 0, 0, \dots)$$
$$W_c(D_4) = (0, 4, 0, 1, 3, 0, 0, 0, 0, 0, 0, \dots)$$

$D_2$ is the minimum $W_c$-aberration design.

# Table of Optimal COAs with $l_n \geq 3$

| Design | Generator | Strength $(t_c, t_n, t_a)$ | Clear Effects $(C, n, CC, Cn, nn)$ |
|---|---|---|---|
| 16-run: | | | |
| (3, 3) | *ABC Aabc* | (2, 2, 2) | (0, 3, 0, 6, 0) |
| ... | ... | ... | ... |
| 32-run: | | | |
| (2, 5) | *abcd ABabe* | (2, 2, 3) | (2, 5, 1, 10, 4) |
| (3, 4) | *ABC Aabcd* | (2, 3, 2) | (0, 4, 0, 12, 6) |
| (4, 3) | *ABCD ABabc* | (3, 2, 3) | (4, 3, 0, 12, 3) |
| ... | ... | ... | ... |
| 64-run: | | | |
| (3, 6) | *abce ACabd Bacdf* | (3, 2, 3) | (3, 6, 3, 18, 9) |
| (4, 5) | *ABCD ABabd ACace* | (3, 2, 3) | (4, 5, 0, 20, 10) |
| (5, 4) | *ABD ACE BCabcd* | (2, 3, 2) | (0, 4, 0, 20, 6) |
| ... | ... | ... | ... |

# Economical Single Arrays

- A single array $2^{(l_c+l_n)-p}$ with $(l_c, l_n) \notin S(N)$ is called an Economical Single Array.
  - Experimental plan is not big enough to allow crossing structure.
  - Only response modeling can be used to analyze data.
  - Must prioritize crucial effects such as control main effects, control-by-noise interactions.
  - Have to discriminate against unimportant effects such as noise effects.
- Different trade-off schemes lead to different criteria.

# $W_s$ sequence and $W_s$-Aberration

- Recall that $W_c$ sequence does not contain $A_{i.1}$ and $A_{i.2}$.
- For example, consider words of length 3 and 4,

$$CCn \lhd Cnn \lhd CCC \lhd nnn$$

$$CCnn \lhd CCCn \lhd Cnnn \lhd CCCC \lhd nnnn$$

- Properly including $A_{i.1}$ and $A_{i.2}$ in $W_c$ leads to

$$W_s = (A_{2.1}, A_{1.2}, A_{3.0}, A_{0,3}, A_{2.2}, A_{3.1}, A_{1.3}, A_{4.0}, A_{0.4}, A_{3.2}, A_{2.3}, \dots)$$

- $W_s$ partially preserves the hierarchical ordering principle.
- $W_s$-aberration and minimum $W_s$-aberration criterion can be proposed.
- Due to limited capacity in ESA, $W_s$ is too restrictive.

# Split Wordtype Pattern and $(W_{sm}, W_{sn})$-Aberration

- Split $W_s$ into two separate sequences

$$W_{sm} = (A_{2.1}, A_{1.2}, A_{3.0}, A_{2.2}, A_{3.1}, A_{1.3}, A_{4.0}, A_{3.2}, A_{2.3}, \ldots)$$
$$W_{sn} = (A_{0.3}, A_{0.4}, A_{0.5}, \ldots)$$

- Note that $W_{sn}$ contains patterns involving noise factors only.
- Select ESAs first according to $W_{sm}$, then further use $W_{sn}$, or in short, according to $(W_{sm}, W_{sn})$-aberration.
- Advantage: the selected ESAs possess clear combinatorial structure.
- Disadvantage: the trade-off between important and unimportant effects may be too extreme.

# Shifted Wordtype Pattern and $W_{ss}$-Aberration

- $nnn$ induces that $Cn$ is aliased with $Cnn$, which is similar to $CCnnn$. Thus $A_{0.3}$ should be immediately before $A_{2.3}$.
- Start with $W_s$, shift the patterns involving only noise factors rightward to proper positions

$$W_{ss} = (A_{2.1}, A_{1.2}, \ldots, A_{4.0}, A_{3.2}, A_{0.3}, A_{2.3}, A_{4.1}, \ldots, A_{4.2}, A_{0.4}, A_{2.4}, \ldots)$$

- $W_{ss}$-aberration and minimum $W_{ss}$-aberration design can be defined.
- Other possible milder trade-off schemes could also be used.
- Often lead to the same optimal ESA as the split wordtype patterns.

The difference between $W_{ss}$ and the list of rank-ordered effects given in Bingham and Sitter (2003) is that $W_{ss}$ further distinguishes wordtype patterns with the same modified wordlength.

$$W_{DR} = (A_{2.1} + A_{1.2}, A_{3.0} + A_{2.2}, A_{3.1} + A_{1.3}, A_{4.0} + A_{3.2} + A_{0.3} + A_{2.3}, \ldots)$$

Because $W_{DR}$ is relatively a coarse sequence, it may not be able to discriminate designs that may have different aliasing and structural properties.

# One Example

Consider 32-run designs with two control factors $(A, B)$ and five noise factors $(a, b, c, d, e)$. Three minimum $W_{DR}$-aberration arrays in Bingham and Sitter (2003) are

$$D_1: \ I = abcd = ABabe = ABcde$$

$$D_2: \ I = abc = ABade = ABbcde$$

$$D_3: \ I = abc = ade = bcde$$

But they have different shifted wordtype patterns. $D_1$ is the minimum $W_{ss}$-aberration single array.

All the arrays guarantee the clear estimation of control main effects and $Cn$ 2fi's. In addition, $D_1$ guarantees the clear estimation of noise main effects.

# Table of Optimal ESAs with Minimum $(W_{sm}, W_{sn})$-Aberration

| Design | Generator | Clear effects |
|---|---|---|
| 16-run: | | |
| (2, 4) | $abc\ ABad$ | (2, 1, 0, 4, 2) |
| ... | ... | ... |
| 32-run: | | |
| (4, 6) | $abd\ ace\ bcf\ ACabc\ ABDa$ | (4, 0, 0, 8, 0) |
| (5, 5) | $abd\ ace\ ACbc\ BDbc\ ABEa$ | (5, 0, 0, 4, 0) |
| (6, 4) | $abc\ ABDa\ ACEa\ BCFa\ ABCbd$ | (6, 1, 0, 6, 2) |
| ... | ... | ... |
| 64-run: | | |
| (4, 8) | $abe\ acf\ bcg\ Aabch\ ACad\ BDabcd$ | (4, 2, 5, 20, 4) |
| (8, 4) | $Cabc\ Dabd\ AEacd\ BFacd\ ABGab\ ABHbcd$ | (8, 4, 12, 24, 0) |
| (4, 9) | $abe\ acf\ bcg\ adh\ Aabci\ ACbd\ ABDacd$ | (4, 1, 5, 20, 2) |
| ... | ... | ... |

# Prodigal Single Arrays

- When cross arrays exist, single arrays can also be considered. We call them prodigal single arrays in order to distinguish the single arrays when cross arrays do not exist.
- It is not conclusive which criterion we should use for selecting optimal prodigal single array.
- Use all $W_s$-, $(W_{sm}, W_{sn})$-, and $W_{ss}$-aberrations and compare the optimal designs.
- Other considerations may be used to determine which optimal design to use in practice.

# Table of Optimal Prodigal Single Arrays

| Design | Generator | $W_s$ | $(W_{sm}, W_{sn})$ | $W_{ss}$ | $COA$ | $MA$ |
|---|---|---|---|---|---|---|
| 16-Run: | | | | | | |
| (3, 3) | $ABCa\ ABbc$ | √ | | | | √ |
| (3, 3) | $ABCa\ abc$ | | √ | √ | | |
| . . . | . . . | . . . | . . . | . . . | . . . | . . . |
| 32-Run: | | | | | | |
| (3, 4) | $abcd\ ABCab$ | √ | | | | √ |
| (4, 3) | $ABCD\ ABabc$ | √ | | | √ | √ |
| (4, 3) | $abc\ ABCDa$ | | √ | √ | | |
| (5, 2) | $ABCD\ ABEab$ | √ | √ | √ | | √ |
| . . . | . . . | . . . | . . . | . . . | . . . | . . . |
| 64-run: | | | | | | |
| (3, 5) | $ABabd\ Cabce$ | √ | | | √ | √ |
| (4, 4) | $Aabcd\ ABCDa$ | √ | | | | √ |
| . . . | . . . | . . . | . . . | . . . | . . . | . . . |

# One Example

Four control factors $(A, B, C, D)$ and five noise factors $(a, b, c, d, e)$ in a RPD experiment.

- If we can afford to conduct 64 runs, the optimal COA is generated by

$$I = ABCD = ABabd = ACace$$

- If we can only afford to conduct 32 runs, then 32-run COA does not exist, and optimal economical single array is generated by

$$I = ABac = ABbd = Aabd = BCDab$$

# Conclusion

- Propose criteria for selecting optimal compound orthogonal arrays and single arrays.
- Construct tables for optimal arrays under different conditions.
- Need to further study the theoretical properties of the optimal arrays.

**Thank you for your patience!**