Akaike Memorial Lecture Award 2016

Date&Time:September 5, 2016, 15:30-17:30

Place:Room 101, Human and Social Science Hall 1, Kakuma Campus,
Kanazawa University

Program

Chairman :

Manabu Iwasaki (President, Japan Statistical Society)

1. **Explanation of Akaike Memorial Lecture Award and awards ceremony**
   Tomoyuki Higuchi (Director-General, The Institute of Statistical Mathematics)
2. **Akaike Memorial Lecture**
   "A fresh look at effect aliasing and interactions : some new wine in old bottle"
   C.F. Jeff Wu (Professor and Coca-Cola Chair in Engineering Statistics, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA)
3. **Discussion by two discussants**
   Ryo Yoshida (The Institute of Statistical Mathematics)
   Chien-Yu Peng (Academia Sinica, Taiwan)
4. **Prof. Wu's rejoinder and floor discussion**

Handouts

[1] Introduction of the Akaike Memorial Lecture Award

[2] Press release of the Akaike Memorial Lecture Award on July 14, 2016

[3] Handouts of the Akaike Memorial Lecture by Professor C.F. Jeff Wu.

[4] Introduction of the two Discussants

Introduction of the Akaike Memorial Lecture Award

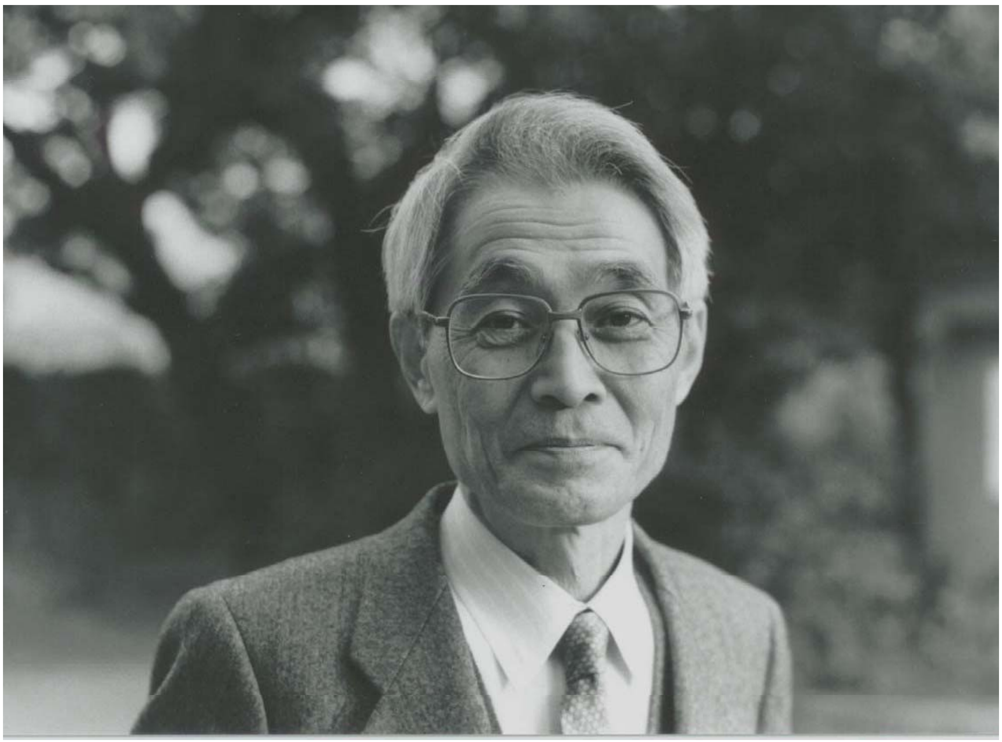
Prof. Tomoyuki Higuchi

Director-General,
The Institute of Statistical Mathematics (ISM)

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

Hirotugu Akaike （1927－2009）

赤池弘次



*This slide was made by Dr. Genshiro Kitagawa, President of the Research Organization of Information and Systems (ROIS)

Brief History

1952 Graduated from Math. Dept., Tokyo University Researcher of the **Institute of Statistical Mathematics**

1962 Head of 2nd Section, 1st Division

1973 Director of **5th Division**

1985 Director of **Dept. of Prediction and Control**

1986 Director-General of ISM (-1994)

1988 Member of **Science Council of Japan** (- 1991)
Chair of Dept. of Statistical Science, Graduate University of Advanced Study

1994 Prof. Emeritus, ISM
Prof. Emeritus, Graduate Univ. for Advance Study

Prizes

1972 Ishikawa Prize
(Establishment of statistical analysis and control method for dynamic systems)

1980 Okochi Prize
(Research and realization of optimal steam temperature control of thermal electric plant)

1989 Asahi Prize
(Research on statistics, in particular theory and applications of AIC)

The Purple Ribbon Medal
(Statistics, in particular time series analysis and its applications)

1996 The 1st Japan Statistical Society Prize
(Contributions to statistical theory and its applications)

2000 The Order of the Sacred Treasure

2006 Kyoto Prize
(Major contribution to statistical science and modeling with the development of AIC)

Fellow of ASA, RSS, IMS, IEEE, JSS

For more information of Dr. Hirotugu Akaike, please visit our website.

http://www.ism.ac.jp/akaikememorial/index-e.html

Laureate of 22nd **Kyoto Prize**


稲盛財団
INAMORI FOUNDATION
2006.6.9

プレスリリース

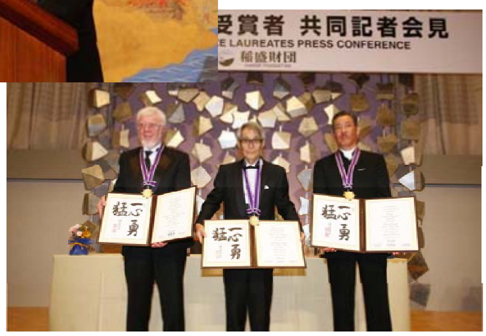
第22回（2006）京都賞受賞者の決定

財団法人稲盛財団（理事長・稲盛和夫）は第22回（2006）京都賞の受賞者を決定しました。本年の受賞者は、以下の3名です。

■先端技術部門
本年授賞対象分野：バイオテクノロジー及びメディカルテクノロジー
レナード・アーサー・ハーツェンバーグ博士（アメリカ・74歳・スタンフォード大学教授）

■基礎科学部門
本年授賞対象分野：数理科学
赤池弘次博士（日本・78歳・統計数理研究所名誉教授）

"Major contribution to statistical science and modeling with the development of the Akaike Information Criterion (AIC)"



*This slide was made by Dr. Genshiro Kitagawa, President of the Research Organization of Information and Systems (ROIS)

The Institute of Statistical Mathematics   Japan Statistical Society

July 14, 2016

The Institute of Statistical Mathematics and the Japan Statistical Society Joint Inaugural Akaike Memorial Lecture Award, Selection of the First Awardee

◆ Introduction

The Institute of Statistical Mathematics (ISM) and the Japan Statistical Society (JSS) have inaugurated the Akaike Memorial Lecture Award under their joint sponsorship. The lecture will be presented during the Japanese Joint Statistical Meeting, a combined meeting of the organizations involved in the statistical sciences, and will feature a biennial lecturer recognized for his or her research accomplishments in this field. A memorial to the legacy of Dr. Hirotugu Akaike, we hope that this lecture will be a valuable stimulus to the minds of younger colleagues and contribute to the development of the statistical sciences.

The first lecture will be held as the planning session of the JSS for the 2016 Japanese Joint Statistical Meeting, which will be held at Kanazawa University from Sunday, September 4 to Wednesday, September 7.

As the awardee, ISM and JSS are proud to announce to have retained Prof. C.F. Jeff Wu of Georgia Institute of Technology, School of Industrial and Systems Engineering.

◆ Overview of the Akaike Memorial Lecture Award

The Akaike Memorial Lecture Award has been planned since 2014 under the joint sponsorship of ISM and JSS. We have named this lecture award after Dr. Hirotugu Akaike, who left a wide-reaching and influential legacy of research in the statistical sciences, and intend for these events to be both opportunities for exchange among statistical researchers from inside and outside Japan and to provide inspiration to young and talented researchers, contributing to further advances in this field.

Every two years, one lecturer is selected under the standards of the Akaike Memorial Lecture Award Nominating Committee from among those individuals who have, like Dr. Akaike, stood out as being ahead of their time, exercising an international influence over a wide range of fields in the statistical sciences (including mathematical sciences such as control and optimization, and mathematical engineering) and applied fields. The awardee receives a ¥100,000 honorarium, an award plaque, and travel expenses.

To promote the education of students and young researchers, the Akaike Memorial Lecture features a selected board of representatives who will engage in discussions after the lecture. The lecture and follow-up discussion will be published as an invited paper in the *Annals of the Institute of Statistical Mathematics* (AISM) or the *Journal of the Japan Statistical Society* (JJSS).

◆ First Awardee: Prof. Chien-Fu Jeff Wu

【Experience】

1949	Born in Taiwan
1971	B.Sc. (Mathematics) National Taiwan University
1976	Ph.D. (Statistics) University of California, Berkeley, USA
1976-1977	Lecturer, Department of Statistics University of California, Berkeley, USA
1977-1980	Assistant Professor, Department of Statistics, University of Wisconsin, Madison, USA
1980-1983	Associate Professor, Department of Statistics, University of Wisconsin, Madison
1983-1988	Professor, Department of Statistics, University of Wisconsin, Madison
1988-1993	Professor and GM/NSERC Chair in Quality and Productivity, Department of Statistics and Actuarial Science, University of Waterloo, Canada
1993-2003	Professor and Chair (1995-1998), Department of Statistics and Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, USA
1997-2003	H. C. Carver Professor of Statistics, University of Michigan, Ann Arbor
2003-	Professor and Coca-Cola Chair in Engineering Statistics, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA

【Research Accomplishments】

Prof. C.F. Jeff Wu has been a vigorous pioneer in the theory of experimental design, EM algorithms and resampling, especially bootstrapping. His research has addressed a broad spectrum of topics in statistics; let us describe some of his particularly notable accomplishments below.

He proposed a general optimal design algorithm using the fact that approximate optimal design problems become constrained convex problems and proved that they converge asymptotically to the optimal design. He examined the convergence of EM algorithms and obtained results under conditions that are applicable to most practical problems. He also made key contributions to the justification of the bootstrapping and jackknife methods from the viewpoint of mathematical statistics.

As experimental design methods were quickening the development of new products and technologies in

Japanese industry in the early 1990s, Prof. Wu used advances in statistical methods to propose fundamental revisions to experimental design, based on the agricultural methods of R. A. Fisher and the robust parameter design method of Genichi Taguchi. He developed a new method called “conditional main effect analysis” for distinguishing among the effects of configuration, sparsity, the principle of transmission, and aliasing in factorial experiments, thereby contributing greatly to the development of the technometrics field.

Recently, Prof. Wu has devoted attention to virtual experiments on computers, in search of principles beyond those identified by Fisher for problems examined with numerical experiments.

【Reasons for Award】

Prof. Wu has conducted vigorous and pioneering work on the theory of experimental design, EM algorithms and resampling. His support of industry has also been highly valued and he has received many awards in statistical quality control. He has long recognized the importance of data science; on entering his post as H. C. Carver Professor at Michigan University in 1997, he gave a speech titled “Statistics = Data Science?” in which he emphasized the role of analysis of large volumes of data and cooperation with people in fields outside of statistics. More recently, he has proposed new methods for experimental design, adapted to the procedures of experiments performed on computers (simulations). Prof. Wu has maintained an exemplary balance among theory, procedure and applications in his research. Since he first came to Japan in 1987 together with Prof. G. E. P. Box to observe quality control in industries, he has visited this country many times and continued exchanges with Japanese statisticians and the industrial sector.

Prof. Wu has also visited ISM on several occasions to lecture and engage in discussions and debates with our young researchers. On the strength of Prof. Wu’s record of research achievements as a statistician and his strong links with ISM and JSS, the nominating committee was proud to recommend Prof. Wu as an entirely appropriate awardee to deliver the first Akaike Memorial Lecture.

◆ **Akaike Memorial Lecture 2016**

Speaker: Prof. C.F. Jeff Wu
(Georgia Institute of Technology, School of Industrial and Systems Engineering)
Title: A fresh look at effect aliasing and interactions: some new wine in old bottle
Date and Time: September 5, 2016, 15:30-17:30
Place: Kakuma Campus, Kanazawa University
http://www.kanazawa-u.ac.jp/e/campuses/

◆ Biography of Dr. Hirotugu Akaike

Born on Nov. 5, 1927 in Shizuoka Prefecture, Japan, Hirotugu Akaike graduated from the First Higher School, the Imperial Naval Academy, and the Tokyo University Science Department Mathematics Faculty. He was accepted into the Institute of Statistical Mathematics in 1952.

He led the way in the field of time series analysis, with R&D resulting in software packages such as TIMSAC for spectral analysis, multivariate time series models, and statistical control methods. In the 1970s, he advocated for what was named the Akaike Information Criterion, a standard for data volume, establishing a new, prediction-centered paradigm for statistical modeling differing from conventional statistical theory. His research influenced a sweeping variety of research fields. In the 1980s, he participated in the development of practical implementations of Bayesian modeling, and played a leading role in finding new data processing methods suitable for the high-information age in which we now live. His research results were held in the highest esteem by his colleagues and earned him many prizes, including the Medal of Honor (Purple Ribbon), the Second Class Order of the Sacred Treasure, and the Kyoto Prize. Citations of his works continue to grow.

Dr. Akaike took the position of Director-General of ISM in 1986. While overseeing the operation of the Institute, he also took part in establishing and teaching on the Statistical Studies program at the Graduate University for Advanced Studies. His term as Director-General ended in 1994. He was appointed Professor Emeritus at the Graduate University for Advanced Studies but never lost his passion for research; rather than resting on his well-deserved laurels, he continued his work, publishing studies on Bayesian models and of the golf swing. He also served as the 19th president of the Japan Statistical Society from January, 1989 to December, 1990. He passed away in Ibaraki Prefecture on August 4, 2009 (age 81).

Hirotugu Akaike Memorial Website:

http://www.ism.ac.jp/akaikememorial/index_e.html

A fresh look at effect aliasing and interactions: some new wine in old bottles

C. F. Jeff Wu
Industrial and Systems Engineering
Georgia Institute of Technology

- Traditional view of effect aliasing and interactions.
- De-aliasing of "aliased effects": using reparametrization and exploiting nonorthogonality in parametrization .
- De-aliasing strategies for:
  - two-level (regular) fractional factorial designs;
  - nonregular FFDs (e.g., Plackett-Burman designs);
  - three-level FFDs.
- Applications in machine learning: bi-level variable selection.
- A historical perspective.

A 2^{4-1} design example

- Consider a 2^{4-1} design with I = ABCD

A	B	C	D	AB	=	CD
-	-	-	-	+		+
-	-	+	+	+		+
-	+	-	+	-		-
-	+	+	-	-		-
+	-	-	+	-		-
+	-	+	-	-		-
+	+	-	-	+		+
+	+	+	+	+		+

Aliasing of effects

- The two-factor interactions (2fi's) AB and CD are said to be aliased (Finney, 1945) because they represent the same contrast (same column in matrix); mathematically similar to *confounding* between treatment and block effects (Yates, 1937).
- Traditional wisdom: The pair of effects cannot be disentangled, and are thus *not estimable*. They are said to be fully aliased.
- A provocative question: can they be de-aliased *without* adding runs?
- Hint: view AB as part of the 3d space of A, B, AB; similarly for C, D, CD; joint space has 5 dimensions, not 6; then reparametrize each 3d space.

Two-factor Interaction via conditional main effects

- Define the conditional main effect of A given B at level +: $ME(A|B+) = \bar{y}(A+|B+) - \bar{y}(A-|B+)$

  similarly, $ME(A|B-) = \bar{y}(A+|B-) - \bar{y}(A-|B-)$

  Then $AB = [ME(A|B+) - ME(A|B-)]/2.$
- can view the conditional main effect $ME(A|B+)$, $ME(A|B-)$ as interaction components*.*
- Original ideas in my 2011 Fisher Lecture, later in JASA, 2015; fully developed ideas and methodology in Su and Wu, 2017, *J. Quality Tech.* to appear.)

Defining relations of cme's

- In effect estimation, we have
  - $ME(A) + INT(A,B) = cme(A|B+)$.
  - $ME(A) - INT(A,B) = cme(A|B-)$.
- In short hand notation, we have
  - $(A|B+) = \frac{1}{2}(A + AB)$
  - $(A|B-) = \frac{1}{2}(A - AB)$

<table>
<tr><th>A</th><th></th><th>AB</th><th></th><th>A|B-</th></tr>
<tr><td>+</td><td rowspan="4">−</td><td>+</td><td rowspan="4">=</td><td>0</td></tr>
<tr><td>+</td><td>-</td><td>+</td></tr>
<tr><td>-</td><td>-</td><td>0</td></tr>
<tr><td>-</td><td>+</td><td>-</td></tr>
<tr><td></td><td></td><td></td><td>/2</td><td></td></tr>
</table>

- Terminology:

  A: parent effect; AB: interaction effect

Orthogonal modeling

- For a 2^{k-p} design with *k* factors, the set of candidate effects consists of $4 \times \binom{k}{2}$ cme's, k main effects and $\binom{k}{2}$ 2fi's.
- Without any restriction, it is hard to find good models from such a large candidate set, i.e., can lead to many incompatible models.
- In this work, we restrict the model search to *orthogonal models*, i.e., effects in a candidate model are orthogonal to each other.

Orthogonality relations I

- cme's are orthogonal to all the traditional effects, except for their parent effects and interaction effects.
- cme's having the same parent effect and interaction effect are *twins.*
- Twin cme's are orthogonal.

Georgia Institute of Technology

ISyE INDUSTRIAL & SYSTEMS ENGINEERING

Rule 1

- Substitute a pair of 2fi and its parental main effect with similar magnitude by one of the corresponding twin cme's.
  - If the pair have the same sign
    - $cme(A|B+) = ME(A) + INT(A,B)$ will have larger magnitude than both A and AB
    - Replace A and AB with (A|B+)
  - If the pair have the opposite signs
    - $cme(A|B-) = ME(A) - INT(A,B)$ will have larger magnitude than both A and AB
    - Replace A and AB with (A|B-)

Georgia Institute of Technology

ISyE INDUSTRIAL & SYSTEMS ENGINEERING

Orthogonal relations II

- cme's having the same parent effect but different interaction effects are *siblings*.
- Siblings are **NOT** orthogonal
- cme's having the same or fully aliased interaction effects are said to belong to the same *family.*
- Non-twin cme's in the family are non-orthogonal, which is the key to the success of the CME analysis strategy.

A|B+	A|B-	B|A+	B|A-	C|D+	C|D-	D|C+	D|C-
+	0	+	0	+	0	+	0
+	0	+	0	0	-	0	-
0	+	-	0	0	+	-	0
0	+	-	0	-	0	0	+
-	0	0	+	0	+	-	0
-	0	0	+	-	0	0	+
0	-	0	-	+	0	+	0
0	-	0	-	0	-	0	-

Rules 2 and 3

- Rule 2: Only one cme among its siblings can be included in the model. Only one cme from a family can be included in the model.
- cme's having different parent effect and interaction effect are orthogonal to each other.
- Rule 3: cme's with different parent effects and different interaction effects can be included in the same model.

CME Analysis

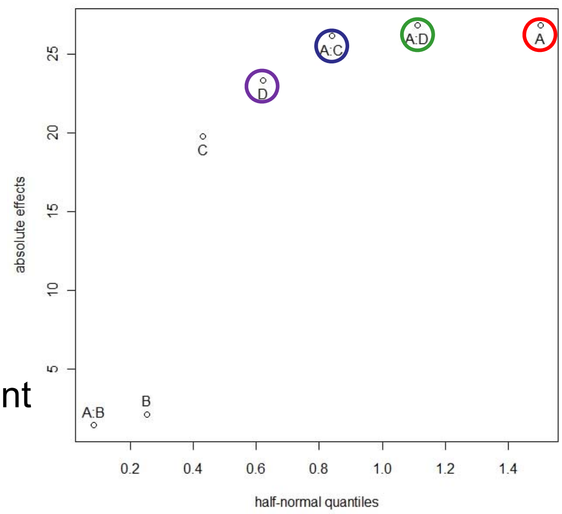
- Based on the three rules, we propose the CME analysis:
  – (i). Use the traditional analysis methods such as ANOVA or half-normal plot, to select significant effects, including aliased pairs of effects. Go to (ii).
  – (ii). Among all the significant effects, use Rule 1 to find a pair of fully aliased 2fi and its parental main effect, and substitute them with an appropriate cme. Use Rules 2 and 3 to guide the search and substitution of other such pairs until they are exhausted.

Example (Filtration)

- 2_{IV}^{4-1} design with $I = ABCD$

- The CME analysis

$y \sim$ (A|D+) + (D|B-) $+ C$



- Step (ii)
  – A and AD are both significant
  – Consider either (A|D+) or (A|D-)
  – D and DB(=AC) are both significant
  – Consider either (D|B+) or (D|B-)

Summary of Example

- In the traditional analysis, we have:
  - $y \sim A(0.45\%) + AD(0.45\%) + AC(0.47\%) + D(0.59\%) + C(0.82\%). (R^2 = 99.79\%)$
- In the CME analysis, we have:
  - $y \sim (A|D+)\ (0.013\%) + AC\ (0.039\%) + D\ (0.055\%) + C\ (0.089\%). (R^2 = 99.79\%)$
  - $y \sim (A|D+)\ (1.96 \times 10^{-5}) + (D|B-)\ (2.72 \times 10^{-5}) + C\ (0.026\%). (R^2 = 99.66\%)$
- The third model is the best in terms of p values for significant effects. All three models have nearly the same R^2 values.
- The cme's (A|D+) and (D|B-) have good engineering interpretations, while AD and AC in first model are fully aliased, thus no good interpretation.

Regular Fractional Factorial Designs

- Regular (2^{n-k}, 3^{n-k} designs):

  algebraic definition: columns of the design matrix form a group over a finite field; $\Rightarrow$ the interaction between any two columns is among the columns.

  statistical definition: any two factorial effects are either *orthogonal* or *fully aliased* (WH book).
- Until the mid-80s, regular FFDs dominated the theory and practice of FFD.

Nonregular Fractional Factorial Designs

- Nonregular designs:
  some pairs of factorial effects can be *partially aliased* (i.e., non-orthogonal nor fully aliased*);*
  $\Rightarrow$ more complex aliasing pattern.
- Its practice in the west was popularized by G. Taguchi when he introduced his favored orthogonal arrays like L_{18} and L_{36} in the mid-80's to the US. His motivation was *practical*.
- I got interested in this class of designs for their flexibility in sample size but later discovered their capability in estimating interactions.
- An inspirational moment in the summer heat of Nagoya (Central Japan Quality Association) in 1986, during our delegation visit (led by G. Box) to Japan to learn its quality practice.

Design Matrix OA(12, 2^7) and Lifetime Data

$OA(12,2^{11})$ (Hadamard matrix of order 12)

Lifetime data (Hunter et al., 1982, Metallurgical Trans.)

Run	Factor *A*	*B*	*C*	*D*	*E*	*F*	*G*	8	9	10	11	Logged Lifetime
1	+	+	−	+	+	+	−	−	−	+	−	6.058
2	+	−	+	+	+	−	−	−	+	−	+	4.733
3	−	+	+	+	−	−	−	+	−	+	+	4.625
4	+	+	+	−	−	−	+	−	+	+	−	5.899
5	+	+	−	−	−	+	−	+	+	−	+	7.000
6	+	−	−	−	+	−	+	+	−	+	+	5.752
7	−	−	−	+	−	+	+	−	+	+	+	5.682
8	−	−	+	−	+	+	−	+	+	+	−	6.607
9	−	+	−	+	+	−	+	+	+	−	−	5.818
10	+	−	+	+	−	+	+	+	−	−	−	5.917
11	−	+	+	−	+	+	+	−	−	−	+	5.863
12	−	−	−	−	−	−	−	−	−	−	−	4.809

Blood Glucose Experiment

$OA(18, 2^1 3^7)$ (Masuyama 增山元三郎, 1957; Taguchi, 田口玄一，1987)

Run	A	G	B	C	D	E	F	H	Mean Reading
				Factor					
1	0	0	0	0	0	0	0	0	97.94
2	0	0	1	1	1	1	1	1	83.40
3	0	0	2	2	2	2	2	2	95.88
4	0	1	0	0	1	1	2	2	88.86
5	0	1	1	1	2	2	0	0	106.58
6	0	1	2	2	0	0	1	1	89.57
7	0	2	0	1	0	2	1	2	91.98
8	0	2	1	2	1	0	2	0	98.41
9	0	2	2	0	2	1	0	1	87.56
10	1	0	0	2	2	1	1	0	88.11
11	1	0	1	0	0	2	2	1	83.81
12	1	0	2	1	1	0	0	2	98.27
13	1	1	0	1	2	0	2	1	115.52
14	1	1	1	2	0	1	0	2	94.89
15	1	1	2	0	1	2	1	0	94.70
16	1	2	0	2	1	2	0	1	121.62
17	1	2	1	0	2	0	1	2	93.86
18	1	2	2	1	0	1	2	0	96.10

Partial and Complex Aliasing

- For the 12-run Plackett-Burman design OA(12, 2^{11})

$$\mathrm{E}\hat{\beta}_i = \beta_i + \frac{1}{3}\sum_{j,k\neq i} \pm \beta_{jk}$$

partial aliasing: coefficient $\pm\frac{1}{3}$

complex aliasing: $45\ (=\binom{10}{2})$ partial aliases.

- In partial aliasing, interactions and main effects are *not* orthogonal to each other; non-orthogonality is the key to success of our analysis strategy.
- Traditionally, complex aliasing was considered to be a disadvantage (called “hazards” ☹ by C. Daniel).Standard texts (until WH) pay little attention to this type of designs.

A paradigm shift

- Traditionally experiments with complex aliasing were used for **screening** purpose, i.e., estimating main effects only.
- A paradigm shift: using *effect sparsity* and *effect heredity*, Hamada-Wu (1992) recognized that complex aliasing can be turned into an advantage for studying interactions.
- Allows interactions to be studied *without* making additional runs.

Guiding Principles for Factorial Effects

- **Effect Hierarchy Principle**:
  - Lower order effects more important than higher order effects;
  - Effects of same order equally important.
- **Effect Sparsity Principle**: Number of relatively important effects is small.
- **Effect Heredity Principle**: for an interaction to be significant, at least one of its parent factors should be significant.

(Wu-Hamada "Experiments", 2000, 2009; Wu, 2015)

HW analysis strategy

- Use effect sparsity to realize that the size of true model(s) is much smaller than the nominal size.
- Use effect heredity to rule out many incompatible models in model search.
- Use the Bayesian variable selection method to perform efficient search over a large space; Chipman's (1996) Bayesian formulation incorporating such design principles.
- Effective if the number of significant interactions is small.

Analysis Results: Cast Fatigue Experiment

- Main effect analysis: F (R^2=0.45)
  F, D (R^2=0.59)
  - Original experimenters dissatisfied with result: wrong sign of D effect, and suggested a DE interaction, claim design did not have enough information.

- HW analysis: F, FG (R^2=0.89)
  F, FG, D (R^2=0.92)
  - 95% CI for D contains positive effect (true by engineering), also the identified FG is partially aliased with the suspected DE. Better fit (R^2 doubled) and correct engineering interpretation.

Analysis results: Blood Glucose Experiment

- Frequentist analysis (HW strategy):
  Main effect analysis: E_q, F_q (R^2=0.36)
  HW analysis: B_l, $(BH)_{lq}$, $(BH)_{qq}$ (R^2=0.89)
- Bayesian analysis also identifies B_l, $(BH)_{ll}$, $(BH)_{lq}$, $(BH)_{qq}$ as having the highest posterior model probability.
- Main effect analysis gave very poor fit, completely missed the important factors, and incapable of finding interactions.

Useful Orthogonal Arrays

- Collection in Wu-Hamada book:
  $OA(12,2^{11})$*, $OA(12,3^1 2^4)$, $OA(18,2^1 3^7)$*, $OA(18,6^1 2^6)$, $OA(20,2^{19})$, $OA(24,3^1 2^{16})$, $OA(24,6^1 2^{14})$, $OA(36,2^{11} 3^{12})$*, $OA(36,3^7 6^3)$, $OA(36,2^8 6^3)$, $OA(48,2^{11} 4^{12})$, $OA(50,2^1 5^{11})$, $OA(54,2^1 3^{25})$.
- Run Size Economy:
  $OA(12,2^{11})$ vs. 16-run 2^{k-p} designs, $8 \le k \le 11$,
  $OA(18,2^7)$ vs. 27-run 3^{k-p} designs, $5 \le k \le 7$,
  $OA(36, 2^{11} 3^{12})$: saturated (i.e., use up all degrees of freedom).
- Taguchi called * $L_{12}(2^{11})$, $L_{18}(2^1 3^7)$, $L_{36}(2^{11} 3^{12})$.

OA(36, 3^{12}) (Seiden, 1954)
OA(36, 2^{11} 3^{12}) (Taguchi, 1987)

Run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0	0	0	1	1	0	0	1	0	2	2	0	1	0	1	0	0	0	1	1	1	0	1
2	0	0	0	0	2	0	2	0	2	0	0	1	1	1	0	1	0	0	0	1	1	1	0
3	0	0	1	0	0	2	1	2	0	0	1	0	0	1	1	0	1	0	0	0	1	1	1
4	0	0	2	2	0	1	0	0	1	1	0	0	1	0	1	1	0	1	0	0	0	1	1
5	0	1	2	2	0	0	1	1	2	0	2	2	1	1	0	1	1	0	1	0	0	0	1
6	0	1	2	1	2	1	2	2	2	2	1	0	1	1	1	0	1	1	0	1	0	0	0
7	0	1	0	0	2	2	0	2	1	1	2	2	0	1	1	1	0	1	1	0	1	0	0
8	0	1	1	2	1	2	2	0	0	2	0	2	0	0	1	1	1	0	1	1	0	1	0
9	0	2	1	2	1	0	0	2	2	1	1	1	0	0	0	1	1	1	0	1	1	0	1
10	0	2	1	0	0	1	2	1	1	2	2	1	1	0	0	0	1	1	1	0	1	1	0
11	0	2	2	1	2	2	1	1	0	1	0	1	0	1	0	0	0	1	1	1	0	1	1
12	0	2	0	1	1	1	1	0	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	2	2	1	1	2	1	0	0	1	1	0	1	0	0	0	1	1	1	0	1
14	1	1	1	1	0	1	0	1	0	1	1	2	1	1	0	1	0	0	0	1	1	1	0
15	1	1	2	1	1	0	2	0	1	1	2	1	0	1	1	0	1	0	0	0	1	1	1
16	1	1	0	0	1	2	1	1	2	2	1	1	1	0	1	1	0	1	0	0	0	1	1
17	1	2	0	0	1	1	2	2	0	1	0	0	1	1	0	1	1	0	1	0	0	0	1
18	1	2	0	2	0	2	0	0	0	0	2	1	1	1	1	0	1	1	0	1	0	0	0
19	1	2	1	1	0	0	1	0	2	2	0	0	0	1	1	1	0	1	1	0	1	0	0
20	1	2	2	0	2	0	0	1	1	0	1	0	0	0	1	1	1	0	1	1	0	1	0
21	1	0	2	0	2	1	1	0	0	2	2	2	0	0	0	1	1	1	0	1	1	0	1
22	1	0	2	1	1	2	0	2	2	0	0	2	1	0	0	0	1	1	1	0	1	1	0
23	1	0	0	2	0	0	2	2	1	2	1	2	0	1	0	0	0	1	1	1	0	1	1
24	1	0	1	2	2	2	2	1	2	1	2	0	0	0	0	0	0	0	0	0	0	0	0
25	2	2	2	0	0	2	2	0	2	1	1	2	1	0	1	0	0	0	1	1	1	0	1
26	2	2	2	2	1	2	1	2	1	2	2	0	1	1	0	1	0	0	0	1	1	1	0
27	2	2	0	2	2	1	0	1	2	2	0	2	0	1	1	0	1	0	0	0	1	1	1
28	2	2	1	1	2	0	2	2	0	0	2	2	1	0	1	1	0	1	0	0	0	1	1
29	2	0	1	1	2	2	0	0	1	2	1	1	1	1	0	1	1	0	1	0	0	0	1
30	2	0	1	0	1	0	1	1	1	1	0	2	1	1	1	0	1	1	0	1	0	0	0
31	2	0	2	2	1	1	2	1	0	0	1	1	0	1	1	1	0	1	1	0	1	0	0
32	2	0	0	1	0	1	1	2	2	1	2	1	0	0	1	1	1	0	1	1	0	1	0
33	2	1	0	1	0	2	2	1	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1
34	2	1	0	2	2	0	1	0	0	1	1	0	1	0	0	0	1	1	1	0	1	1	0
35	2	1	1	0	1	1	0	0	2	0	2	0	0	1	0	0	0	1	1	1	0	1	1
36	2	1	2	0	0	0	0	2	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0

Implications and follow-up work

- Success in the HW analysis strategy led to research on the *hidden projection* properties of nonregular designs. Commonly used arrays like OA(12, 2^{11}), OA(18, 3^7), have desirable projection properties (i.e., a number of interactions can be estimated with good efficiency); Lin-Draper (1993), Wang-Wu (1995).
- It has rejuvenated research and opened a new field on optimal nonregular designs, including extensions of the minimum aberration design theory to nonregular designs; generalized minimum aberration (Tang-Deng, Deng-Tang, 1999, Xu-Wu, 2001), minimum moment aberration (Xu, 2003), etc.

Effect Heredity Principle

- Coined by Hamada-Wu (1992), used to rule out incompatible models in model search. Original motivation: application to analysis of experiments with complex aliasing.
- Strong (both parents) and weak (single parent) versions defined by Chipman (1996) in Bayesian framework; strong heredity is the same as the marginality principle by McCullagh-Nelder (1989) but their motivation was to keep model invariance.

A computational challenge in variable selection

- Select the important subset of variables:

$$Y = \beta_1 X_1 + \cdots + \beta_q X_q + \beta_{11} X_1^2 + \beta_{12} X_1 X_2 + \cdots + \beta_{qq} X_q^2 + \epsilon$$

- A very difficult optimization problem when q is large.
  - $2^{2q+q(q-1)/2}$ possible models.
  - Even for q = 5, there are a million models.
- Stepwise regression techniques: unstable.

Use of heredity principle in variable selection

- Aliasing leads to infinite number of optima for least squares minimization. Heredity rule helps to break the aliases and reduce the number of local minima. This helps the search for best models through optimization techniques.
- A digression: nonnegative garrote:
  $\frac{1}{2} \parallel Y - X\hat{\beta}^{LS}\theta \parallel^2$, subject to $\sum_{j=1}^{p} \theta_j \leq M$ and $\theta_j \geq 0\ \forall j$,
  where $\hat{\beta}_j^{LS}$ is the least squares estimate of β_j in the regression model.
- Since both constraints and objective are *convex*, this allows much faster computations using quadratic programming techniques.

Use of heredity principle in variable selection (continued)

- Yuan, Joseph, Zou (2009) used nonnegative garrote to reformulate the strong heredity principle by using the convexity constraints:
  $$\theta_i \leq \theta_j\ \forall\ j \in D_i,$$
  where D_i = set of parent effects of θ_i. Example: $\theta_i = \{AB\}, D_i = \{A, B\}$. Why does it imply strong heredity?
  Ans: If $\theta_i > 0$ (i.e. active), then $\theta_j > 0$ (active) $\forall\ j \in D_i$.
- Similarly, for weak heredity, they used the convexity constraints:
  $$\theta_i \leq \sum_{j\in D_i} \theta_j\ .$$
  Following the same example, if $\theta_i > 0$ (active), then at least one of j's in D_i has $\theta_j > 0$ (active).

Three-level fractional factorial designs: Seat belt experiment

- An experiment to study the effect of four factors on the pull strength of truck seat belts
- 27 runs were conducted; each one was replicated three times

	Factor	Level 0	Level 1	Level 2
A.	pressure (psi)	1100	1400	1700
B.	die flat (mm)	10.0	10.2	10.4
C.	crimp length (mm)	18	23	27
D.	anchor lot (#)	P74	P75	P76

Design matrix and response data, seat-belt experiment (first 14 runs)

- a 3^{4-1} design with $I=ABCD$, $D=ABC$, etc.

Run	Factor *A*	*B*	*C*	*D*	Strength			Flash		
1	0	0	0	0	5164	6615	5959	12.89	12.70	12.74
2	0	0	1	1	5356	6117	5224	12.83	12.73	13.07
3	0	0	2	2	3070	3773	4257	12.37	12.47	12.44
4	0	1	0	1	5547	6566	6320	13.29	12.86	12.70
5	0	1	1	2	4754	4401	5436	12.64	12.50	12.61
6	0	1	2	0	5524	4050	4526	12.76	12.72	12.94
7	0	2	0	2	5684	6251	6214	13.17	13.33	13.98
8	0	2	1	0	5735	6271	5843	13.02	13.11	12.67
9	0	2	2	1	5744	4797	5416	12.37	12.67	12.54
10	1	0	0	1	6843	6895	6957	13.28	13.65	13.58
11	1	0	1	2	6538	6328	4784	12.62	14.07	13.38
12	1	0	2	0	6152	5819	5963	13.19	12.94	13.15
13	1	1	0	2	6854	6804	6907	14.65	14.98	14.40
14	1	1	1	0	6799	6703	6792	13.00	13.35	12.87

Design matrix and response data, seat-belt experiment (next 13 runs)

Run	Factor A	B	C	D	Strength			Flash		
15	1	1	2	1	6513	6503	6568	13.13	13.40	13.80
16	1	2	0	0	6473	6974	6712	13.55	14.10	14.41
17	1	2	1	1	6832	7034	5057	14.86	13.27	13.64
18	1	2	2	2	4968	5684	5761	13.00	13.58	13.45
19	2	0	0	2	7148	6920	6220	16.70	15.85	14.90
20	2	0	1	0	6905	7068	7156	14.70	13.97	13.66
21	2	0	2	1	6933	7194	6667	13.51	13.64	13.92
22	2	1	0	0	7227	7170	7015	15.54	16.16	16.14
23	2	1	1	1	7014	7040	7200	13.97	14.09	14.52
24	2	1	2	2	6215	6260	6488	14.35	13.56	13.00
25	2	2	0	1	7145	6868	6964	15.70	16.45	15.85
26	2	2	1	2	7161	7263	6937	15.21	13.77	14.34
27	2	2	2	0	7060	7050	6950	13.51	13.42	13.07

ANOVA analysis result

- Based on the p values, *A*, *C* and *D* are significant.
- Also two aliased sets of effects are significant, $AB=CD^2$ and $AC=BD^2$, but
  - aliased interaction components cannot be de-aliased,
  - meaning of *AB*, *AC*, etc.?

Orthogonal components (OC) system: decomposition of $A \times B$ interaction

- $A \times B$ has 4 degrees of freedom; it has two component denoted by AB and AB^2, each having 2 df's;
- Let the levels of A and B be denoted by x_1 and x_2 respectively;
- AB represents the contrasts whose x_1 and x_2 satisfy $x_1+x_2=0, 1, 2 \pmod 3$; the other interaction component AB^2 is similarly defined. All components are *orthogonal* to each other, thus the name OC system.
- Note: this is the classical and prevailing approach but it is deficient.

Representation of AB and AB^2 in a Latin Square

- Factor A and B combinations (x_1 level of A, x_2 levels of B)

x_1	x_2: 0	1	2
0	$\alpha i\ (y_{00})$	$\beta k\ (y_{01})$	$\gamma j\ (y_{02})$
1	$\beta j\ (y_{10})$	$\gamma i\ (y_{11})$	$\alpha k\ (y_{12})$
2	$\gamma k\ (y_{20})$	$\alpha j\ (y_{21})$	$\beta i\ (y_{22})$

- α, β, γ correspond to (x_1, x_2) with $x_1+x_2=0, 1, 2 \pmod 3$ resp. Their SS is AB.
- i, j, k correspond to (x_1, x_2) with $x_1+2x_2=0, 1, 2 \pmod 3$ resp. Their SS is AB^2.
- Difficult to interpret the meaning of significance of AB or AB^2.

A reparametrization: linear-quadratic (LQ) system

- The 2 df's in a quantitative factor, say A, can be decomposed into the linear and quadratic components. Letting y_0, y_1 and y_2 represent the observations at level 0, 1 and 2, the *linear effect* is defined as $y_2 - y_0$

  and the *quadratic effect* as $(y_2 + y_0) - 2y_1$

  which is the difference between two consecutive linear effects $(y_2 - y_1) - (y_1 - y_0)$
- The linear and quadratic effects are represented by two mutually orthogonal vectors:

$$A_l = \frac{1}{\sqrt{2}}(-1,0,1) \qquad A_q = \frac{1}{\sqrt{6}}(1,-2,1)$$

Interactions in linear-quadratic system

- The 4 df's in $A \times B$ can be decomposed into four mutually orthogonal terms:

  $(AB)_{ll}$, $(AB)_{lq}$, $(AB)_{ql}$, $(AB)_{qq}$, which are defined as follows: for i, j = 0, 1, 2,

$$\begin{aligned}
(AB)_{ll}(i,j) &= A_l(i)B_l(j), \\
(AB)_{lq}(i,j) &= A_l(i)B_q(j), \\
(AB)_{ql}(i,j) &= A_q(i)B_l(j), \\
(AB)_{qq}(i,j) &= A_q(i)B_q(j).
\end{aligned}$$

- They are called the *linear-by-linear, linear-by-quadratic, quadratic-by-linear* and *quadratic-by-quadratic* interaction effects, and denoted as $l \times l$, $l \times q$, $q \times l$ and $q \times q$

Designs with resolution *III* and *IV*

- In traditional wisdom, interactions in *III* or *IV* designs are not estimable. A more elaborate analysis method is required to extract the maximum amount of information from data.
- Consider the 3^{3-1} design with $C=AB$, whose design matrix is given below

Run	A	B	C
1	0	0	0
2	0	1	1
3	0	2	2
4	1	0	1
5	1	1	2
6	1	2	0
7	2	0	2
8	2	1	0
9	2	2	1

- Its main effects and two-factor interactions have the aliasing relations:

$$A=BC^2,\ B=AC^2,\ C=AB,\ AB^2=BC=AC$$

Analysis of designs with resolution *III*

- In addition to estimating the 6 df's in A, B and C, there are 2 df's left for estimating the three aliased effects AB^2, BC and AC.
- Instead, consider using the remaining 2 df's to estimate any pair of the $l\times l$, $l\times q$, $q\times l$ or $q\times q$ effects between A, B and C.
- Suppose that the two interaction effects taken are $(AB)_{ll}$ and $(AB)_{lq}$. Then the 8 df's can be represented by the following model matrix:

Run	A_l	A_q	B_l	B_q	C_l	C_q	$(AB)_{ll}$	$(AB)_{lq}$
1	-1	1	-1	1	-1	1	1	-1
2	-1	1	0	-2	0	-2	0	2
3	-1	1	1	1	1	1	-1	-1
4	0	-2	-1	1	0	-2	0	0
5	0	-2	0	-2	1	1	0	0
6	0	-2	1	1	-1	1	0	0
7	1	1	-1	1	1	1	-1	1
8	1	1	0	-2	-1	1	0	-2
9	1	1	1	1	0	-2	1	1

Analysis of designs with resolution *III*(contd)

- Because any component of $A \times B$ is orthogonal to A and to B, there are only four non-orthogonal pairs of columns whose correlations are: $\pm \frac{1}{\sqrt{8}}$ or $\pm \sqrt{\frac{3}{8}}$
- Because the last four columns in the matrix are non-orthogonal, they can't be estimated with full efficiency. However, non-orthogonality is the saving grace ☺ because it leads to estimability.
- The estimability of $(AB)_{ll}$ and $(AB)_{lq}$ demonstrates an *advantage* of LQ system over OC system. The *AB* interaction component cannot be estimated because it is aliased with *C.* (Further theory using indicator functions in Sabbaghi-Dasgupta-Wu, 2014).

Variable selection strategy

1. For a quantitative factor, say *A*, use A_l and A_q for the *A* main effect.
2. For a qualitative factor, say *D,* select two contrasts from D_{01}, D_{02} and D_{12} for the *D* main effect.
3. For *X* and *Y,* use the products of the two contrasts of *X* and the two contrasts of *Y* to represent the 4 df's in $X \times Y$*.*
4. Using the contrasts in 1-3 as candidate variables, perform a stepwise regression or subset selection procedure to identify a suitable model. Use effect heredity principle to *rule out* incompatible models.

Analysis of seat-belt experiment

- Using these 39 contrasts as the candidate variables, variable selection led to the following model with R^2=0.811:

$$\begin{aligned}\hat{y}_1 &= 6223.0741 + 1116.2859A_l - 190.2437A_q + 178.6885B_l \\ &\quad -589.5437C_l + 294.2883(AB)_{ql} + 627.9444(AC)_{ll} \\ &\quad -191.2855D_{01} - 468.4190D_{12} - 486.4444(CD)_{l,12}\end{aligned}$$

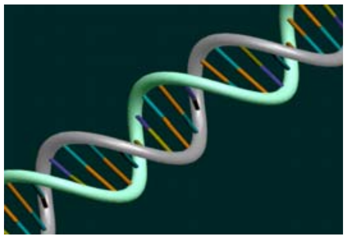
- This model obeys effect heredity. *A*, *B*, *C* and *D* and *A*×*B*, *A*×*C* and *C*×*D* are significant. And each of the three interaction components is interpretable. In contrast, the ANOVA analysis identified *A*, *C* and *D* and the *AC*(=BD^2) and *AB*(=CD^2) interaction components as significant.

What have we learned for 3-level designs?

- ANOVA analysis is inadequate; the proposed strategy can extract information on interactions even for resolution *III* and *IV* designs; this casts in doubt the use of "resolution" (Box-Hunter, 1961) in choosing designs ☹.
- Prevailing advice on using resolution V designs for 3-level experiments is too conservative and misguided.
- The linear-quadratic parametrization creates non-orthogonality, the key to its success.
- Materials first available in chapter 6 of Wu-Hamada book (2000, 2009), not in any papers.

Use of conditional main effects (cme's) for variable selection

- Interpretability of cme's also makes it a useful tool for variable selection in observational studies.
- cme's provide intuitive basis functions for many applications:
  - **Genome markers:**
    - A|B+ indicates gene A is active only when gene B is active;
  - **Clinical trials:**
    - A|B+ indicates drug A is effective only when drug B is used.



Two distinctions from designed experiments

1. Orthogonal framework never occurs in observational data:
   - Initial groupings of twin, sibling and family effects motivated from an *orthogonal* model.
   - New groupings needed to capture effect correlations in the non-orthogonal setting.
2. Goal not to disentangle aliased effects, but to separate active effects from correlated groups of inert effects:
   - Bi-level selection is needed which performs between-group and within-group effect selection.

New effect groupings

- **Main effect (me) pairs**:
  - e.g., A and B
- **Inverse pairs:**
  - A cme pair with parent and conditioned effects swapped
  - e.g., A|B+ and B|A+
- **Parent-child pairs:**
  - A cme and its parent effect
  - e.g., A and A|B+
- **Uncle-nephew pairs:**
  - A cme and its conditioned effect
  - e.g., B and A|B+

Bi-level CME selection

$$\min_{\beta} Q(\beta) \equiv \frac{1}{2}||\boldsymbol{y} - \boldsymbol{X}\beta||^2 + \sum_{\mathcal{G}\in\mathbb{G}} \sum_{g\in\mathcal{G}} f_o \left\{ \sum_{k\in g} f_i(\beta_k) \right\}$$

- Observations $\boldsymbol{y}$, model matrix $\boldsymbol{X}$, coefficients β.
- Effect group g, collection of effect groups $\mathcal{G}$, set of all collections $\mathbb{G}$:
  - e.g., one group of siblings, collection of sibling groups, etc.
- Outer penalty f_o:
  - Controls between-group selection (e.g., selecting sibling groups); allows for effect coupling.
- Inner penalty f_i:
  - Controls within-group selection (e.g., selecting within a sibling group).

Effect coupling

- Effect coupling: selecting an effect in group g allows other effects in g to enter the model more easily.
- This is intuitive for cme's:
  - If A|B+ is active, then its siblings A|C+, A|D+, … are more likely to be active.
  - When many sibling pairs are in the model, the criterion encourages the selection of their parent effect instead.

− Full paper in Mak and Wu (2016).

A historical perspective

- I had the basic ideas of conditional main effect (cme) analysis in 1988 but did not fully realize its implications. The 2011 Fisher lecture gave me the courage and opportunity to develop and publish it. This last piece also benefited from the new perspectives I got from the 1993 and 2000 work.
- Hamada-Wu (1993) showed some interactions in 2-level nonregular designs can be estimated.
- Inference about interactions for 3-level fractional factorial designs using the linear-quadratic system followed naturally from the 1993 work; first appeared in the 2000 WH book.
- Common thread: non-orthogonality in the parametrization.

Common theme: reparametrization and non-orthogonality

- 2-level regular FFD: use conditional main effect as the new parametrization, which induces non-orthogonality among some effects.
- 2-level nonregular FFD: nonregularity is the inherent property of these designs, which leads to non-orthogonality.
- 3-level regular FFD: regular designs (i.e., OC system) are orthogonal, no hope! The linear-quadratic system gives the new parametrization and non-orthogonality.

Further remarks

- CME's provide a class of new basis functions in bi-level variable selection, ongoing work. Potential impact outside physical experiments, e.g., in medical and social studies.
- Need *design-theoretic* work to give more fundamental understanding on how and why the new CME analysis method works (Sabbaghi, 2016, using theory of indicator functions).
- The work collectively serve as a *transition* from orthogonal experiments to non-orthogonal experiments/studies like optimal designs or observational studies. Potential impact in big data. Need further exploration.

Introduction of the Discussant 1

Ryo Yoshida
(Associate Professor of The Institute of Statistical Mathematics)

[Profile]
Education
2004, Ph.D., Graduate University for Advanced Studies
Professional Experience
2011 - present, Associate Professor, The Institute of Statistical Mathematics
2007 - 2011, Assistant Professor, The Institute of Statistical Mathematics
2005 - 2007, Project Assistant Professor, The Institute of Medical Science, The University of Tokyo
2013 - present, Visiting Researcher, Advanced Telecommunications Research Institute International (ATR)
2015 - present, Invited Researcher, National Institute of for Materials Science (NIMS)

[Research Interest]
Bioinformatics
Materials Informatics

[Selected Publication]
Toyoshima *et al.*, Accurate automatic detection of densely distributed cell nuclei in 3D space, *PLoS Compt Biol*, 2016.
Ikebata *et al*, Repulsive parallel MCMC algorithm for discovering diverse motifs from large sequence sets, *Bioinform*, 2015.
Yamashita *et al.*, Atom environment kernels on molecules, *J Chem Inf Model*, 2014.
Tokunaga *et al*. Automated detection and tracking of many cells by using 4D live-cell imaging data, *Bioinform*, 2014.
Yoshida *et al.*, Bayesian experts in exploring reaction kinetics of transcription circuits, *Bioinform*, 2010
Yoshida *et al.*, Bayesian learning in sparse graphical factor models via variational mean-field annealing, *J Mach Learn Res*, 2010

Introduction of the Discussant 2

Chien-Yu Peng
(Associate Research Fellow of Institute of Statistical Science, Academia Sinica)

[Profile]

Education

2004 – 2008, Ph.D., Statistics, National Tsing-Hua University
2000 – 2002, M.A., Statistics, National Tsing-Hua University
1996 – 2000, B.S., Applied Mathematics, National Cheng-Kung University

Professional Experience

2016 – present, Associate Research Fellow with Tenure, Academia Sinica
2009 – 2016, Assistant Research Fellow, Academia Sinica
2008 – 2009, Postdoctoral Research Associate, National Tsing-Hua University

Professional Service

2016 – present, Associate Editor for *Technometrics*

[Research Interest]

Reliability (degradation modeling and analysis)
Design of experiments (indicator functions and computer experiments)
Mathematical statistics (theory of reliability)

[Selected Publication]

Peng, C. Y.* (2015/11), “Optimal classification policy and comparisons for highly reliable products.” *Sankhya B*, 77(2), 321–358.
Peng, C. Y.* (2015/03), “Inverse Gaussian processes with random effects and explanatory variables for degradation data.” *Technometrics*, 57(1), 100–111.
Peng, C. Y. and Wu, C. F. Jeff* (2014/02), “On the choice of nugget in Kriging modeling for deterministic computer experiments.” *Journal of Computational and Graphical Statistics*, 23(1), 151–168.
Peng, C. Y. and Tseng, S. T.* (2013/06), “Statistical lifetime inference with skew-Wiener linear degradation models.” *IEEE Transactions on Reliability*, 62(2), 338–350. Best Paper Award at ISQM (2009)
Cheng, Y. S. and Peng, C. Y.* (2012/06), “Integrated degradation models in R using iDEMO.” *Journal of Statistical Software*, 49(2), 1–22.
Peng, C. Y.* (2012/04), “A note on optimal allocations for the second elementary symmetric function with applications for optimal reliability design.” *Naval Research Logistics*, 59(3–4), 278–284.