QUALITY ENGINEERING
2019, VOL. 31, NO. 1, 69–72
https://doi.org/10.1080/08982112.2018.1549423

Rejoinder

Dennis K. J. Lin and Jiayu Peng

Department of Statistics, Pennsylvania State University, State College, Pennsylvania

Many thanks to Joseph Voelkel (JV) and Christine Anderson-Cook & Lu Lu (AL) for their valuable discussions. In this rejoinder, we will respond to JV's and AL's comments in the first two sections, respectively. In the final section of this rejoinder, we present some new thoughts motivated by their discussions.

Comments in response to Joseph Voelkel (JV)'s discussion

JV shared that his work was motivated by some conversations with industrial statisticians (e.g., from Kodak and Lipton). These are interesting stories, which show that systematic order-of-addition (OofA) designs are desirable in practice and that the existing techniques were inadequate. We are pleased to see that the design of OofA experiments have been receiving rapidly increasing attention from researchers. Many important results have been obtained in the last few years, since the pioneer work of Voelkel (2017).

As commented by JV, the optimality result of Peng, Mukerjee, and Lin (2017) fills up the gap between the availability of intuitively-favorable designs and theoretical support to justify their optimality. We would like to especially thank JV because our result was motivated by JV's exploration of pairwise-order (PWO) designs—Voelkel (2017) first derived the moment matrix of the full PWO design. In contrast to the moment matrices of full designs in traditional design problems, the moment matrix of the full PWO design is not compound symmetric. We were fascinated by the algebraic structure of PWO designs. This has motivated us to establish the first optimality result in the OofA literature.

We would emphasize that the knowledge of optimal designs, as a benchmark for assessing any other design, is useful in practice. For example, suppose one has already obtained a design with $|M| = 0.56$. If we know theoretically that a D-optimal design has $|M| = 0.6$, then the obtained design is quite favorable. If, on the other hand, a D-optimal design has $|M| = 0.8$, then there will be a need to improve the obtained design. In addition, the theoretical conditions of a design to be optimal provide guidance for closed-form constructions of optimal designs.

JV pointed out that the minimal-point PWO designs given in our article are not that efficient, compared with the ones obtained by Fedorov's exchange algorithm. We fully agree that the construction of such minimal-point designs needs further improvement. This construction is a first attempt in finding minimal-point PWO designs, and it, in particular, justifies the existence of nonsingular minimal-point PWO designs for any m. The structure of H herein, on the other hand, may form a basis of constructing of more efficient minimal-point designs.

JV has tabulated the number of runs of the optimal fractional PWO designs we presented in the paper. As he commented, these designs are expensive even with moderately large m (say, $m = 8$). Initial results have been obtained to reduce the number of runs from these designs, which works for small m. The idea is that the optimal designs we constructed are naturally blocked, with a blocking variable indicating the choice of the $m/2$ components in each subblock B_u (see Peng, Mukerjee, and Lin (2017) for the definition of B_u). The optimal designs we presented correspond to a complete block design, in the sense that the blocking variable takes all its $\binom{m}{m/2}$ levels, that is, all the possible choices of the $m/2$ components in B_u. Strategies from balanced incomplete block designs (BIBDs) are thus helpful in reducing the number of runs. For example, when $m = 8$, an optimal design with n being only 168 has been obtained. We hope to report some systematic results in the near future.

JV mentioned an interesting result, proved by Voelkel (2017), that an n-run PWO design is OofA-OA

CONTACT Jiayu Peng pengjy08@gmail.com Department of Statistics, Pennsylvania State University, Thomas Building, University Park, State College, PA 16802, USA.
This article was presented at the Sixth Stu Hunter Research Conference in Roanoke, Virginia, March 2018.

m	$p = \binom{m}{2} + 1$	$2\binom{m}{2}$	In-between size(s) divisible by 12
4	7	12	12
5	11	20	12
6	16	30	24
7	22	42	24, 36
8	29	56	36, 48
9	37	72	48, 60, 72
10	46	90	48, 60, 72, 84

of strength 2 (3) only if n is divided by 12 (24). JV noted that the run size $n = 2\binom{m}{2}$ that we choose in Algorithmic construction of efficient PWO designs with small number of runs section is not necessarily divisible by 12. In view of this, we have listed below multiples of 12 that are between the minimal size and $n = 2\binom{m}{2}$, with m up to 10. Construction for designs of such sizes should be worth particular attention.

The combined design of both process and OofA variables, as first suggested by JV, is indeed important in practice. Here is a comment regarding the optimal design under a model where the mean response is additive of the process factors as well as the PWO factors. We note that the full design will be optimal under this model, following the proof of the optimality result in Peng, Mukerjee, and Lin (2017). Explicitly, the full design here is the Kronecker product of a full factorial design and a full OofA design. In principle, exchange algorithms, such as Federov exchange or a modification of the proposed bubblesort exchange algorithm, can be exploited to find an efficient combined design of process and OofA variables, with a small number of runs. The closed-form construction of such designs seems infeasible, however.

Last but not least, JV emphasized the importance of studying the design robustness to multiple models. Voelkel (2017) made a first attempt along this line: in his search of optimal PWO designs, he considered a number of secondary criteria regarding the strength-3 orthogonality. In fact, AL's discussion on our article made emphasis on the design robustness as well, as will be discussed below.

Comments in response to Anderson-Cook and Lu (AL)'s discussion

We fully agree with AL's statement that "in the absence of a generally interpretable model that connects to the underlying mechanism driving the relationship between ingredient order and response, it seems prudent to explore alternative models." In fact, the mechanisms behind some OofA experiments are so obscure that it is impossible to determine a "most appropriate" model beforehand; for instance, in the experiment to find optimal taxa order in a phylogenetic tree. Thus at the stage of planning the experiment, a design robust to multiple models would be desirable. For the design construction, one approach adopted by Voelkel (2017) and Yang, Sun, and Xu (2017) is to first obtain a number of (nearly-)optimal designs under a main single criterion, and then select from the obtained designs via a secondary criterion. There are other design strategies via a multi-objective optimization. Among them, the Pareto front methodology developed by AL and their collaborators is a powerful tool when faced with a tradeoff between different objective functions.

AL has brought up an important practical concern on whether finding an optimal order via an OofA experiments is likely to yield sufficient rewards. Before any attempt to determine an optimal order, an economic pilot design is desirable for quantifying the impact of orders on the responses.

AL recommended a Graeco-Latin square as a pilot design. We would like to comment that the idea of constructing Latin-squares with nearly-balanced PWO factors has been introduced in Voelkel (2017) as well, in his review of Williams (1949)'s design. AL further recommended to randomly permuting the rows or columns of Latin square to improve the balancedness of PWO factors. Such a row/column permutation is indeed an effective strategy for improving the model robustness. Using a similar strategy, Yang, Sun, and Xu (2017) has obtained a number of designs that are optimal under their component-position model and meanwhile highly efficient under the PWO model.

In a pilot study, the exploratory data analysis is quite important. AL has developed two types of graphs for this purpose, as shown respectively in their Figures 1(b) and 2. Overall, Figure 2 exhibits the contrasts in a PWO model, while Figure 1(b) corresponds to Yang, Sun, and Xu (2017)'s component-position model. It is noted that if we re-scale the circles in either Figure 1(b) or Figure 2, the plot can convey quite different messages and sometimes be misleading. It is important to standardize the radius of each circle in a consistent manner.

AL proposed alternative models for a pilot study. For any order $\boldsymbol{a} = a_1a_2...a_m$, a first-order model based on "simple orders" treats each a_j $(1 \leq j \leq m)$ as a predictor. AL suggested to fit m univariate linear regression models, with the predictor being $a_1, ..., a_m$, respectively. They also suggested to centralize the predictors by replacing a_j with $l_j = a_j - (m+1)/2$, so that the jth of the m regression models reads as

$$\tau(\boldsymbol{a}) = \beta_0 + \beta_1 l_j, \quad [1]$$

with $\tau(\boldsymbol{a})$ being the mean response arising from $\boldsymbol{a}$. The fitted $\hat{\beta}_1$'s for these m simple linear regressions can be visualized by the plots in AL's Figure 3. Likewise, for $1 \le j \le m$, the jth second-order model based on simple orders can be formulated as

$$\tau(\boldsymbol{a}) = \beta_0 + \beta_1 l_j + \beta_2 q_j. \quad [2]$$

The fitted $\hat{\beta}_1$'s for these m regressions can be summarized in one plot, like Figure 3(a), while the fitted $\hat{\beta}_2$'s for these m regressions shall be summarized in another plot. Here q_j is defined as $2, 1, 0, -1, -2, -2, -1, 0, 1, 2$ for $a_j = 1, \ldots, 10$, respectively, when $m = 10$. It is not clear to us what is the general formula used by AL to produce such q_j. We believe that q_j's can, instead, be simply chosen as the centered values of $a_j^2, j = 1, \ldots, 10$, or explicitly, $q_j = a_j^2 - (1/m)\sum_{j=1}^m j^2 = a_j^2 - (m+1)(2m+1)/6$.

AL suggested to fit m separated regressions rather than a single regression model with all or multiple l_j's included (consider the first order model for simplicity). This is perhaps because the model including multiple l_j's is not estimable or yields high standard errors, under a pilot design with a very limited number of runs. However, when the number of runs is not that small, say, $n > 2m$, a single regression model on multiple l_j's is more natural for both design and analysis. Details will be given in the next section.

Some new thoughts

Following the discussion in the last section, we consider the definition of a model including all the simple orders a_j's as predictors. For simplicity, the centralized contrasts l_j's and q_j's are no longer considered in this section.

The naïve model is $\tau(\boldsymbol{a}) = \beta_0 + \sum_{j=1}^m \beta_j a_j$. However, this model is not estimable due to the constraint

$$\sum_{j=1}^m a_j = \sum_{j=1}^m j = m(m+1)/2.$$

An estimable model can be defined by removing a_1, that is,

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{j=2}^m \beta_j a_j. \quad [3]$$

Likewise, a model including all linear and quadratic terms on simple orders can be defined as

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{j=2}^m \beta_j a_j + \sum_{j=2}^m \beta_{jj} a_j^2. \quad [4]$$

By modifying the proof of Theorem 1 in Peng, Mukerjee, and Lin (2017), it can be shown that the full OofA design is D-optimal under either model [3] or [4] above. Efficient designs can be constructed under both models (by the exchange algorithm, for example). Note that model [3] requires a minimum of m runs while model [4] requires a minimum of $2m-1$ runs. As long as more than $2m$ runs are affordable for a pilot study, we think that it is more appropriate to conduct design and analysis under models [3] and [4], instead of AL's models [1] and [2].

A conventional second order model includes all the interaction terms as well. In view of [3] and [4], a second order model including two-way interactions can be defined as

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{j=2}^m \beta_j a_j + \sum_{j=2}^m \beta_{jj} a_j^2 + \sum_{2 \le j < k \le m} \beta_{jk} a_j a_k.$$

However, the above equation turns out be an unidentifiable model. It actually contains 1 extra degree of freedom. Thus, by removing the last interaction, the resulting model

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{j=2}^m \beta_j a_j + \sum_{j=2}^m \beta_{jj} a_j^2 + \sum_{2 \le j < k \le m, j \ne m-1} \beta_{jk} a_j a_k. \quad [5]$$

becomes estimable (the proof is skipped here). In fact, modeling of OofA experiments along this line is worth further investigation, and we hope to report some systematic conclusions in the near future.

We close this section by noting that the a_j's in models [3], [4] or [5] can be replaced by r_j's, where r_j indicates the rank of the jth component. A model based on r_j's may be more favorable for the interpretation purpose. As a simple example, in a first-order model, a positively-significant coefficient of r_j indicates that we would place the jth component at the end, instead of the beginning of the addition sequence, if the purpose is to increase the response.

Finally, we would like to thank Joe, Christine, and Lu for their discussions which have substantially enhanced the value of our article. We would like to thank Joe again for bringing our attention into the broad field of order-of-addition problems.

About the authors

Dennis K.J. Lin is a university distinguished professor of supply chain and statistics at Penn State University. His research interests are quality assurance, industrial statistics, data mining, and response surface. He has published near 200 SCI/SSCI papers in a wide variety of journals. He currently serves or has served as associate editor for more than 10 professional journals and was co-editor for Applied Stochastic Models for Business and Industry. Dr. Lin is an elected fellow of ASA, IMS and ASQ, an elected member of ISI, a lifetime member of ICSA, and a fellow of RSS. He is an honorary chair professor for various universities, including

Renmin University of China (as a Chang-Jiang Scholar), Fudan University, The University of Hong Kong, and National Chengchi University (Taiwan). His recent awards including, the recipient of the 2004 Faculty Scholar Medal Award (Penn State), the Youden Address (ASQ, 2010), the Shewell Award (ASQ, 2010), the Don Owen Award (ASA, 2011), the Loutit Address (SSC, 2011), the Hunter Award (ASQ, 2014), and the Shewhart Medal (2015).

Jiayu Peng is a PhD student at the Department of Statistics, Penn State University. His thesis advisor is Dr. Dennis K.J. Lin. Mr. Peng received a bachelor's degree in Mathematics from Tsinghua University, China. Mr. Peng's research interest includes design of experiments, non-parametric statistics, and quantile regression. He has received the Honorable Mention in JSM 2017 for the SPES Outstanding Presentation Awards, and the Student Scholarships for JSM 2015 SPES sponsored roundtable, and the 2011 First-class scholarship at Tsinghua University.

References

Peng, J., R. Mukerjee, and D. K. Lin. 2017. Design of order-of-addition experiments. *Biometrika*, Forthcoming.

Voelkel, J. G. 2017. The design of order-of-addition experiments. arXiv preprint arXiv:1701.02786.

Williams, E. J. 1949. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry* 2 (2):149–68. doi:10.1071/CH9490149.

Yang, J., F. Sun, and H. Xu. 2017. Component orthogonal arrays for order-of-addition experiments. Submitted for publication.