Received: 26 April 2019 | Revised: 24 October 2019 | Accepted: 29 December 2019

DOI: 10.1002/qre.2638

RESEARCH ARTICLE

WILEY

Comparison of conditional main effects analysis to the analysis of follow-up experiments for separating confounded two-factor interaction effects in 2^{k-p}_{IV} fractional factorial experiments

John Lawson

Department of Statistics, Brigham Young University, Provo, Utah

Correspondence
John Lawson, Department of Statistics, Brigham Young University, Provo, UT.
Email: lawson@byu.edu

Abstract

Two-factor interactions, where the effect of one factor depends on the level of another factor, are common, and understanding them is often the key to solving quality problems or making process improvements using designed experiments. Resolution IV 2^{k-p} fractional factorial designs are efficient and require fewer experiments or runs than resolution V or full factorial experiments. However, two-factor interactions are confounded with other two-factor interactions in resolution IV designs and their effects cannot be separated. Follow-up experiments have been recommended in the literature to separate the effects of significant but confounded strings of two-factor interactions in resolution IV designs.

Recently, an analysis based on conditional main effects (or CMEs) has been shown to be useful in determining which interaction in a confounded string of two-factor interactions is actually causing the significance without the need for follow-up experiments. In this article, I investigate the value of this method of analysis by comparing its use with the analysis of follow-up experiments using the data from three published experiments where follow-up experiments were used to "de-alias" confounded interactions.

KEYWORDS

aliased effects, CME analysis, confounding, effect heredity, effect sparsity, hierarchical ordering

1 | INTRODUCTION

When process troubleshooting to identify root causes, or when conducting process improvement studies, including many potential factors will increase the chances of reaching a solution in a timely manner. Factor effects and interactions of all orders can be estimated when studying all combinations of factor levels in a full factorial design. However, even if each of the k factors are studied with only two levels, the number of runs or test combinations in a full factorial plan may be infeasible if k is large. Normally, only a fraction of the total 2^k runs are used in practice.

One of the consequences of fractionating a 2^k factorial design is that factor effects and interactions become partially or completely confounded. In a regular fractional factorial, all estimable effects are orthogonal but completely confounded with other effects. For example, each estimable effect in a $\frac{1}{2}$ fraction is completely confounded with one other effect, and each estimable effect in a $\frac{1}{4}$ fraction is completely confounded with three other effects, etc.[1] In Plackett-Burman or a model robust design, each effect is partially confounded with many other effects.[2,3] When analyzing the data from a

fractional factorial design, the principles of *effect sparsity*, *hierarchical ordering*, and *effect heredity*[4] provide guidance in interpreting confounded effects.

The effect sparsity principle suggests that only a small subset of the main effects and interactions will be important or statistically significant. This would be especially true if the factors included in the experiment were only hypothesized to have potentially important effects. The hierarchical ordering principle implies that the lower-order effects (like main effects and two-factor interactions) are generally more important than higher-order effects (such as three-way and higher-order interactions). Finally, the effect heredity principle would lead one to believe that interactions involving significant main effects are more likely than interactions that do not involve at least one significant main effect. A two-factor interaction that does not involve a significant main effect would indicate that the effect of each factor involved in the interaction was equal but opposite in sign depending on the level of the other factor involved in the interaction. The study of Li et al[5] of 113 published factorial experiments showed this is a rare occurrence. The term *strong heredity* means that both main effects involved in an interaction should be significant if the interaction is significant. *Weak heredity* means that if an interaction is significant, at least one of the two main effects involved in that interaction should be significant as well.

When analyzing data from fractional designs (such as Plackett-Burman designs,[6] model robust designs,[7] or alternative screening designs),[8] there are many more main effects, two-factor interactions, and higher-order interactions than there are runs in the design. However, the complex aliasing in these designs[2,3] means that each two-factor interaction is only partially confounded with many main effects rather than being completely confounded with one main effect. The effect sparsity principle indicates only a subset of these effects and interactions are likely to be important, and models involving a subset of these partially confounded effects can be fit with a regression subset procedure such as stepwise regression or forward selection.[9] The list of candidate terms for the regression subset procedure is usually restricted to main effects and all possible two-factor interactions, as would be suggested by the hierarchical ordering principle. Jones and Nachtsheim[10] recommend a forward stepwise regression analysis that incorporates the *strong effect heredity* principle by forcing both main effects into the model at any step where an interaction involving these main effects enters the model. This procedure is incorporated in the *Combine* option in the JMP[11] forward stepwise regression and in the `ihstep` and `fhstep` functions in the R package `daewr`.[12]

When analyzing the data from a regular $\frac{1}{p}$th fraction of a 2^k design (denoted 2^{k-p}), a saturated model including $2^{k-p}-1$ orthogonal effects can be fit but each effect estimated is completely confounded with $p-1$ other effects and there is no residual mean square to test the effects. In order to judge which of the $2^{k-p}-1$ confounded sets of effects are significant, graphical techniques like the normal plot, half-normal plot, or Pareto diagram are employed. These graphical techniques are effective due to the effect sparsity principle. Once a subset of the confounded effects are identified as significant using graphical methods, the hierarchical ordering principle and effect heredity principle are useful in determining which effect in each confounded set is active.

In a resolution IV design, each main effect is unconfounded with other main effects or two-factor interactions. However, each main effect is completely confounded with one or more three-factor interactions and possibly additional higher-order interactions. In addition, each two-factor interaction is confounded with one or more two-factor interactions and additional higher-order interactions. When analyzing a resolution IV design, any confounded set of effects that is judged significant using graphical methods and contains a main effect would be interpreted to be a main effect. That interpretation is due to the hierarchical ordering principle. On the other hand, the interpretation may not be so clear if the lowest-order effect in a significant confounded set is a two-factor interaction.

For example, if the three confounded sets of effects in a resolution IV 2^{7-2} design that appeared to be significant were C, E, and $CE + FG$ (omitting three-factor and higher-order confounded effects in each set); the effect heredity principle would imply that the confounded set of two-factor interactions $CE + FG$ represents CE, since both main effects involved in this interaction were also significant. However, if the three significant sets of effects were F, G, and $CE + FG$, effect heredity would imply that the interaction represented FG. Conversely, if three significant sets of effects were E, F, and $CE + FG$, the effect heredity principle would not help in determining which of the two interactions was the active one.

In this situation, the two interactions CE and FG are completely confounded and, according to traditional wisdom, they cannot be separated without additional experiments.[13] An additional foldover design consisting of 16 more runs with the signs changed on just one of the factors C, E, F, or G would make the interactions CE and FG orthogonal[14] but would double the number of runs in the combined set of experiments. Box-Hunter and Hunter show that as few as three additional runs could be used to obtain separate estimates of CE and FG.[15] Using their method, the interactions would not

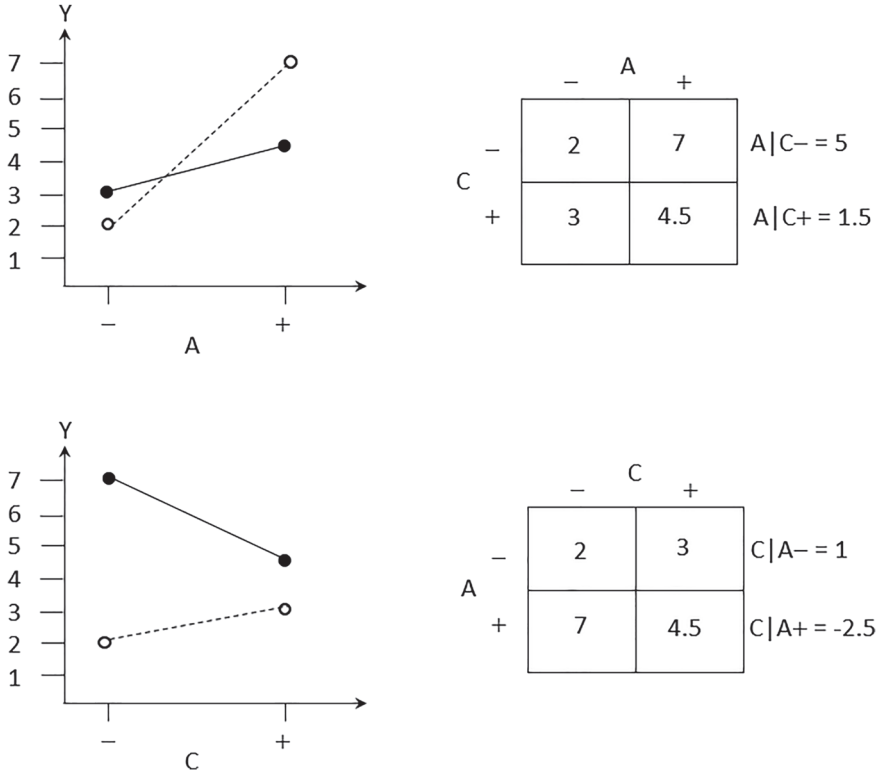

FIGURE 1 Illustration of conditional main effects (CMEs)

be orthogonal and must be estimated by the method of least-squares. Another method, using the D-optimality criterion, can find a set of follow-up experiments that will result in less correlated estimates of CE and FG[1,4]. However, in 2011, Wu[13] showed that confounded two-factor interactions could be separated without additional experiments using what he defined as analysis of conditional main effects.

The remainder of this paper is organized as follows. In Section 2, conditional main effects and the analysis of conditional main effects will be described. In Sections 3 to 5, using three sets of published data from experiments where follow-up experiments were included, the analysis of conditional main effects will be compared with the the analysis of follow-up experiments for the purpose of separating the confounded interactions. Section 6 summarizes the comparisons made in Sections 3 to 5. Section 7 describes a simulation study that helps clarify the results in Section 6. Finally, Section 8 gives conclusions and recommendations.

2 | CONDITIONAL MAIN EFFECTS

Wu[13] showed that a two-factor interaction effect is the difference of two conditional main effects, where the conditional main effect is the effect of one factor conditional on the other factor being at the + or − level. For example, Figure 1 illustrates graphically and numerically four conditional main effects (abbreviated CMEs). The CME of A when $C = -$ is labeled $A|C-$; the CME of A when $C = +$ is labeled $A|C+$; the CME of C when $A = -$ is labeled $C|A-$; finally, the CME of C when $A = +$ is labeled $C|A+$. On the left side of the figure, the CMEs at the + level of the other factor are represented as the black circles and lines while the CMEs at the − level of the other factor are represented by the open circles and dashed lines. The interaction $AC = CA$ is then equal to $(A|C+) - (A|C-) = 1.5 - 5 = -3.5$ or $(C|A+) - (C|A-) = -2.5 - 1 = -3.5$.

Table 1 shows a 2^{4-1}_{IV} design with defining relation $I = ABCD$. In this table, it can be seen that the interactions AC and BD are completely confounded and their calculation columns are identical. The column of signs for the CME of A given $C = +$ (labeled $A|C+$) is defined as factor A times an indicator function of the set where $\{C = +\}$, ie, $A \times I_{\{C=+\}}$. This column of signs for $A|C+$ can be used to calculate the effect of factor A conditional that $C = +$. The columns for the other CMEs $A|C-$, $C|A+$, etc, that are shown in the table, are defined similarly. The interaction $AC = (A|C+) - (A|C-)$ or $AC = (C|A+) - (C|A-)$. Likewise, the interaction $BD = (B|D+) - (B|D-)$ or $BD = (D|B+) - (B|D-)$.

Although the two interactions AC and BD are completely confounded in this this design, the CMEs that can be used to calculate the AC interaction are not completely confounded with the CMEs that can be used to create the

TABLE 1 2_{IV}^{4-1} Design with conditional main effects (CMEs)

Standard Order	*A*	*B*	*C*	*D*	*AC*	*BD*	*A|C+*	*A|C-*	*C|A+*	*C|A-*	*B|D+*	*B|D-*	*D|B+*	*D|B-*
1	-	-	-	-	+	+	0	-	0	-	0	-	0	-
2	+	-	-	+	-	-	0	+	-	0	-	0	0	+
3	-	+	-	+	+	+	0	-	0	-	+	0	+	0
4	+	+	-	-	-	-	0	+	-	0	0	+	-	0
5	-	-	+	+	-	-	-	0	0	+	-	0	0	+
6	+	-	+	-	+	+	+	0	+	0	0	-	0	-
7	-	+	+	-	-	-	-	0	0	+	0	+	-	0
8	+	+	+	+	+	+	+	0	+	0	+	0	+	0

BD interaction. For example, the correlation matrix for all the CME columns shown in Table 1 is shown as follows:

$$\begin{array}{c} \\ \boldsymbol{A|C+} \\ \boldsymbol{A|C-} \\ \boldsymbol{C|A+} \\ \boldsymbol{C|A-} \\ \boldsymbol{B|D+} \\ \boldsymbol{B|D-} \\ \boldsymbol{D|B+} \\ \boldsymbol{D|B-} \end{array} \begin{array}{c} \begin{array}{cccccccc} \boldsymbol{A|C+} & \boldsymbol{A|C-} & \boldsymbol{C|A+} & \boldsymbol{C|A-} & \boldsymbol{B|D+} & \boldsymbol{B|D-} & \boldsymbol{D|B+} & \boldsymbol{D|B-} \end{array} \\ \begin{bmatrix} 1.0 & 0.0 & 0.5 & -0.5 & 0.5 & -0.5 & 0.5 & -0.5 \\ 0.0 & 1.0 & -0.5 & 0.5 & -0.5 & 0.5 & -0.5 & 0.5 \\ 0.5 & -0.5 & 1.0 & 0.0 & 0.5 & -0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.0 & 1.0 & -0.5 & 0.5 & -0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 & -0.5 & 1.0 & 0.0 & 0.5 & -0.5 \\ -0.5 & 0.5 & -0.5 & 0.5 & 0.0 & 1.0 & -0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 & -0.5 & 0.5 & -0.5 & 1.0 & 0.0 \\ -0.5 & 0.5 & -0.5 & 0.5 & -0.5 & 0.5 & 0.0 & 1.0 \end{bmatrix} \end{array}.$$

In this matrix, it can be seen that the CMEs ($A|C+$, $A|C-$, $C|A+$ and $C|A-$), which can be used to create *AC*, have a correlation coefficient of ± 0.5 with each of the CMEs ($B|D+$, $B|D-$, $D|B+$, and $D|B-$), which can be used to create the interaction *BD*.

2.1 | Analysis with conditional main effects

Since the CMEs are only partially confounded (like the two-factor interaction columns in a Plackett-Burman design, model robust design, or alternative screening design), Wu[13] has suggested a stepwise or forward selection procedure be used to identify which subset of the main effects and CMEs (that can be used to calculate the confounded interactions) are significant. He calls this *CME analysis*.

In his paper, Wu[13] comments that "the trick lies in the choice of a candidate set of effects ...." He illustrated the choice of the candidate set with an example. In this example, he identified four effects in an unreplicated 2_{IV}^{6-2} design that appeared to be significant on the half-normal plot of effects. One of the effects identified was a string of two confounded two-factor interactions. He formed a candidate set for forward selection that included all six main effects and all eight possible conditional main effects associated with the two confounded two-factor interactions. Next, he used a forward selection procedure to identify the first three terms (one less than the four terms identified on the half normal plot) to enter the model. The three-term model explained an equivalent amount of variation in the data, and the three terms were more significant than the terms in a four-term model including all effects identified on the half normal plot. The conditional main effect that entered the model indicated which of the confounded two-factor interactions was causing the significance, and it had a meaningful engineering interpretation.

This conditional main effect analysis complements the effect heredity principle because of the following situation: if confounded interactions involve at least one significant main effect, one of the CMEs (that can be used to create the confounded interaction) is likely to be larger than the others (like $A|C-$ that is shown in Figure 1).

Su and Wu[16] defined some additional properties of CMEs and extended the CME analysis to fitting models in orthogonal effects not requiring subset selection procedures. They noted that each CME (ie, $A|C+$) is related to a main effect (A) and an interaction (AC). They defined the main effect as the *parent effect* of the CME and the interaction as the *interaction* of the CME. They further defined the main effect being conditioned (ie, C) as the *conditioning effect*. They designated two CMEs that differ only by their conditioning level (ie, $A|C+$ and $A|C-$) as *twins* and CMEs that have the same parent effect but different interaction effects (ie, $A|C+$ and $A|B+$) as *siblings*. Finally, if two interactions are confounded

TABLE 2 Factors and levels for analytical laboratory experiment

Label	Factors	Levels -	Levels +
A	Sample Preparation	Method 1	Method 2
B	Moisture Measurement	Volume	Weight
C	Mixing Speed (rpm)	800	1600
D	Mixing Time (hrs)	0.5	3.0
E	Healing Time (hrs)	1	2
F	Spindle Number	1	2
G	Protective Lid	Absent	Present

(like *AC* and *BD* in Table 1), then all the CMEs that can be used to calculate either of the confounded interactions are members of the same *family*.

Based on the definitions, Su and Wu showed that every CMEs is orthogonal to traditional effects except for its parent effect and its interaction effect. In addition, twin CMEs are orthogonal, as are CMEs with different parents and interaction effects. However, sibling CMEs are not orthogonal, nor are non-twin CMEs in the same family. Based on these definitions and properties, Su and Wu proposed a new CME analysis defined by the following two steps.

1. Use the normal graphical techniques (such as the normal plot, half normal plot, or Pareto diagram) to identify the significant sets of confounded effects in a 2^{k-p}_{IV} fractional factorial design, and determine all the aliases for each set of confounded two-factor interactions.
2. Search for a model that does not have confounded effects by substituting one of the twin CMEs for its parent main effect and its interaction effect.

As an example of step 2, consider the design shown in Table 1 in the previous section. If the seven calculated effects from response values for this design were plotted on a half normal plot and main effects, *A*, *D*, and the two confounded interactions $AC = BD$ appeared to be significant, then an orthogonal model could be defined by substituting the CME $A|C+$ for its parent main effect *A* and its two-factor interaction $AC = BD$. This would result in a model containing just two terms, *D* and $A|C+$.

Su and Wu[16] advised to substitute the CME $A|C+$ for *A* and $AC = BD$ when the calculated effects for *A* and $AC = BD$ have the same sign (ie, both positive or both negative). On the other hand, if *A* and $AC = BD$ have opposite signs, they advise to substitute $A|C-$ for *A* and $AC = BD$.

Another two-term orthogonal model can be formed by substituting either $D|B+$ or $D|B-$ for the parent effect *D* and its two factor interaction $DB = AC$. Again, the choice of $D|B+$ or $D|B-$ would depend on whether the signs of the calculated effects *D* and *DB* had the same or opposite signs.

Looking at the R^2 statistic and the significance of the individual terms in the models will determine which of the two models just described best represents the data. If the model containing the two terms $A|C+$ and *D* were best, it would indicate that the *AC* interaction was causing the two confounded interactions to appear significant. On the other hand, if the model containing $D|B-$ and *A* were the best, it would indicate that the the *BD* interaction was causing the two confounded interactions to appear to be significant.

In the model search defined in steps 1 and 2, orthogonality is preserved by including only one CME from the same family or only one CME from a set of siblings.

3 | ANALYTICAL LABORATORY EXPERIMENT

Snee[17] describes an experiment performed in an analytical laboratory. The study was initiated because of the perceived variability in the viscosity measurements of a high-volume product. Viscosity was a key quality characteristic of the product, and it was decided to conduct a ruggedness test[18] using a 2^{7-3}_{IV} design to determine which factors were influencing variability in the viscosity measurements. The factors and levels used in the fractional factorial experiment are shown in Table 2.

3.1 | Analysis based on original 20 experiments

The experimental runs and response data from the experiment are shown in Table 3. The generators for this $\frac{1}{8}$th fraction were $E = BCD$, $F = ACD$, and $G = ABC$. The order of the experiments was completely randomized, but the table shows

TABLE 3 2_{IV}^{7-3} Design for analytical laboratory experiment

Standard Order	*A*	*B*	*C*	*D*	*E*	*F*	*G*	Viscosity *Y*
1	-	-	-	-	-	-	-	2796
1r	-	-	-	-	-	-	-	2788
2	+	-	-	-	-	+	+	2460
3	-	+	-	-	+	-	+	2904
4	+	+	-	-	+	+	-	2320
5	-	-	+	-	+	+	+	2800
5r	-	-	+	-	+	+	+	2700
6	+	-	+	-	+	-	-	3772
7	-	+	+	-	-	+	-	2420
8	+	+	+	-	-	-	+	3376
9	-	-	-	+	+	+	-	2220
10	+	-	-	+	+	-	+	2548
11	-	+	-	+	-	+	+	2080
12	+	+	-	+	-	-	-	2464
12r	+	+	-	+	-	-	-	2348
13	-	-	+	+	-	-	+	3216
14	+	-	+	+	-	+	-	2380
15	-	+	+	+	+	-	-	3196
16	+	+	+	+	+	+	+	2340
16r	+	+	+	+	+	+	+	2380

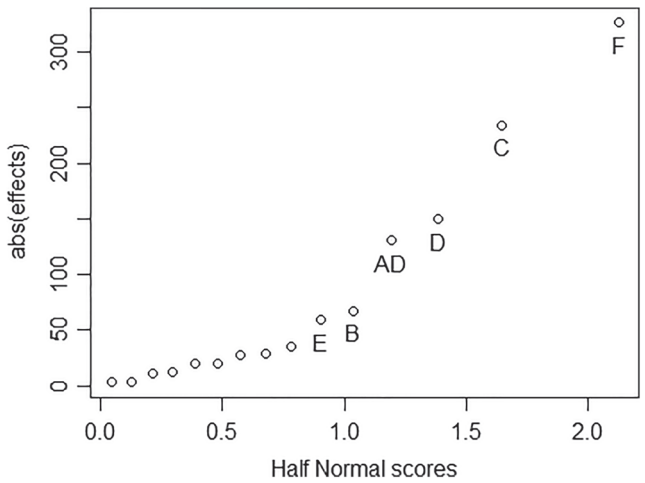

FIGURE 2 Half normal plot of coefficients

the runs in standard order. There were four replicate runs labeled 1r, 5r, 12r, and 16r. These additional runs were included to estimate the pure error and test for lack of fit.

A half normal plot of the coefficients or half effects from a saturated model are shown in Figure 2. There it can be seen that the large (in absolute value) and apparently significant effects are *F*, *C*, *D*, *AD*, *B*, and *E*.

Table 4 shows the estimated coefficients, or half effects, along with their standard errors (based on the pure error estimated from the replicates), *t* values, and *P* values. The significant effects at the $\alpha = .05$ level are the same ones identified in Figure 2. There was no lack of fit in the reduced model, containing only the significant effects, compared with the pure error (*P* value = .194). All six effects in the reduced model were significant with *P* values < .005, and the percentage of variation in the viscosity measurements explained by the model was $R^2 = .9788$.

The significant two-factor interaction *AD* is completely confounded with two other two-factor interactions, namely *CF* and *EG*. These three interactions cannot be separated. The strong effect heredity principle implies that the *CF* interaction is causing the significance. However, since the main effects *D* and *E* are also significant, the weak heredity principle could indicate that either the *AD* or *EG* interactions could be causing the significance. Figure 3 shows the plots of these three interactions.

If the interaction causing the significance is *CF*, then changes in only five factors (namely *F*, spindle number; *C*, mixing speed; *D*, mixing time; *E*, healing time, and *B*, moisture measurement) cause changes in the viscosity measurements.

TABLE 4 Coefficients from saturated model with replicates

Coefficients	Estimate	Standard Error	*t* value	Pr(>|t|)	
Intercept	2700.00	13.10	206.042	3.33e-09	
A	2.75	13.10	0.210	0.844035	
B	-67.25	13.10	-5.132	0.006829	**
C	233.75	13.10	17.838	5.80e-05	**
D	-149.25	13.10	-11.390	0.000339	**
E	58.75	13.10	4.483	0.010962	*
F	-326.25	13.10	-24.897	1.54e-05	**
G	11.75	13.10	0.897	0.420587	
AB	-20.00	13.10	-1.526	0.201653	
AC	35.50	13.10	2.709	0.053589	
AD	-130.00	13.10	-9.921	0.000580	**
AE	-11.50	13.10	-0.878	0.429719	
AF	3.50	13.10	0.267	0.802604	
AG	-28.50	13.10	-2.175	0.095288	
BD	27.00	13.10	2.060	0.108401	
ABD	19.75	13.10	1.507 0	.206249	

TABLE 5 Follow-up experiments for analytical laboratory

Order	*A*	*B*	*C*	*D*	*E*	*F*	*G*	Viscosity *Y*
21	-	+	-	+	-	-	-	2384
22	-	+	+	+	-	-	-	2976
23	-	+	-	+	-	+	-	2180
24	-	+	+	+	-	+	-	2300

Therefore, by tightly controlling these five factors, the variability in the viscosity measurement process should be reduced. On the other hand, if either the *AD* interaction or the *EG* interaction is causing the significance, then changes in six factors are causing the changes in viscosity measurements (namely *F*, spindle number; *C*, mixing speed; *D*, mixing time; *E*, healing time; *B*, moisture measurement; and *A*, sample preparation) if *AD* is active, or (*F*, spindle number; *C*, mixing speed; *D*, mixing time; *E*, healing time; *B*, moisture measurement; and *G*, protective lid) if *EG* is active. In either of these two cases, six factors in the measurement process would have to be tightly controlled to reduce variability in the viscosity measurement process but the true scenario cannot be determined by traditional methods of analysis without follow-up experiments.

3.2 | Analysis and conclusions after follow-up experiments

The experimenters decided to resolve the source of the significant confounded set of interactions by running four follow-up experiments. They ran the additional experiments shown in Table 5. These were obtained using the procedure of Box, Hunter, and Hunter[15] (pages 413-416).

A model was fit to the combined data set including both the follow-up experiments and the original 20. All the main effects were included in the model, along with each of the interactions in the significant set of confounded interactions found in the analysis of the original 20 runs. The results of this model are shown in Table 6. Here it can be seen that the same main effects found in the original analysis are significant along with the *CF* interaction. Neither the *AD* nor *EG* interactions were significant. This analysis clearly shows that the *CF* interaction was the one causing significance of the confounded set of interactions $AD = CF = EG$.

Based on this analysis, the experimenters concluded that they only needed to tightly control the variables $F =$ spindle number, $C =$ mixing speed, $D =$ mixing time, $E =$ healing time, and $B =$ moisture measurement in order to reduce the variability in viscosity measurements.

3.3 | Conclusions from the original 20 experiments based on CME analysis

Now consider the conclusions that could be drawn after a CME analysis with just the original 20 experiments, excluding the follow-ups. One way to perform a CME analysis is using a forward selection as suggested by Wu.[13] Following his example, a forward stepwise regression using the minimum AIC[19] stopping rule was used with the original 20 experiments

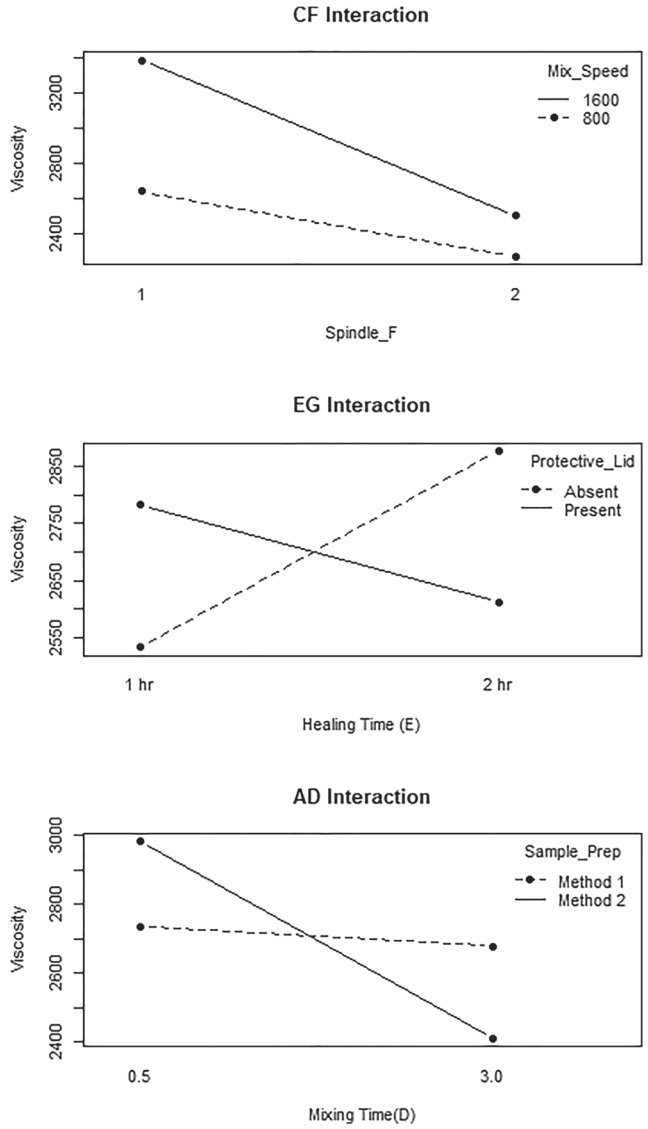

FIGURE 3 Confounded interaction plots

TABLE 6 Coefficients from model fit to combined data

Coefficients	Estimate	Standard Error	*t* value	Pr(>|t|)	
Intercept	2701.167	21.894	123.373	< 2e-16	
A	5.571	22.205	0.251	0.8058	
B	-64.429	22.205	-2.902	0.0124	*
C	221.537	20.292	10.918	6.44e-08	**
D	-146.429	22.205	-6.594	1.73e-05	**
E	57.421	22.351	2.569	0.0233	*
F	-306.063	20.292	-15.083	1.29e-09	**
G	10.421	22.351	0.466	0.6487	
A:D	-26.969	44.329	-0.608	0.5534	
C:F	-118.000	47.968	-2.460	0.0287	*
E:G	16.135	45.397	0.355	0.7280	
	R^2=.9711				

from the analytical laboratory. The candidate effects were the seven main effects A to G and the 12 twin CMEs associated with the confounded set of interactions $AD = CF = EG$. The CMEs used were $A|D+$, $A|D-$, $D|A+$, $D|A-$, $C|F+$, $C|F-$, $F|C+$, $F|C-$, $E|G+$, $E|G-$, $G|E+$, and $G|E-$.

Table 7 shows the results of the forward selection. A confirmatory model was fit in the variables selected. This five-term model had an R^2 of .9783 with all five terms significant with P values < .002. These statistics are similar to those for the model fit to the significant effects in Table 5, and it explains the variation in viscosity measurements equally well. However,

TABLE 7 Results of Forward Selection

Step	Variable entered	AIC
1	$F\|C+$	234.87
2	$C\|F+$	216.67
3	D	196.55
4	$E\|G-$	189.21
5	B	177.46

TABLE 8 Result of model search using rule 1

Model 1	$R^2 = .8358$	**Model 2**	$R^2 = .9781$
Term	P value	Term	P value
B	.116127	B	.00181
C	.000215	C	$1.98e^{-9}$
E	.225020	E	.00373
F	$8.36e^{-6}$	F	$2.67e^{-11}$
$A\|D+$	.04308	$D\|A+$	$2.0e^{-8}$
Model 3	$R^2 = .9310$	Model 4	$R^2 = .9665$
term	p-value	term	p-value
B	.055665	B	.00963
C	$2.47E^{-6}$	C	$3.75e^{-8}$
D	.000292	D	$7.75e^{-6}$
F	$4.14e^{-8}$	F	$3.36e^{-10}$
$G\|E-$	.003728	$E\|G-$	$1.95e^{-5}$
Model 5	$R^2 = .4337$	Model 6	$R^2 = .8783$
B	.4784	B	.1385
C	.0247	C	$9.74e^{-5}$
D	.1173	D	.0029
E	.9431	E	.2991
$C\|F+$	.1616	$F\|C+$	$1.76e^{-6}$

unlike the example presented by Wu,[13] it does not help in determining which of the confounded interactions caused the significance since CMEs that correspond to two of the aliased two-factor interactions (CF and EG) are significant. In addition, the results of this forward-selection CME analysis does not agree with the analysis of combined original data and follow-up experiments. That analysis clearly indicated the CF interaction caused the significance of the confounded string of two-factor interactions.

The other way of performing a CME analysis is to make a model search using rule 1 described by Su and Wu.[16] This rule says to modify the model identified in the half-normal plot by substituting a CME for one of the confounded two-factor interactions and its parent main effect that has a similar magnitude. For example, Figure 2 and Table 5 show that the magnitude of main effect D is similar to the magnitude of AD which is confounded with CF and EG. D is also a parent main effect of AD. Therefore, substitute either the CME $A|D+$ or $A|D-$ for AD and D in the model that contains the significant terms shown in Figure 2 and Table 5. According to Su and Wu,[16] $A|D+$ should be substituted since D and AD both have the same sign (ie, $-$). Alternatively, substitute $D|A+$ for D and AD. Another pair of models can be created by substituting $G|E-$ or $E|G-$ for EG and E since E and $EG = AD$ have similar magnitudes but opposite signs. A third pair of models can be created by substituting $C|F+$ or $F|C+$ for F and $CF = AD$ since F and $CF = AD$ have the same sign.

Table 8 shows the results of the model search over the six models. Model 2 with $R^2 = .9781$ explains about the same amount of variability in the viscosity measurements as the model containing the terms identified as significant in Figure 2 or Table 5. This model would indicate that the AD interaction was causing the significance of the confounded string $AD = CF = EG$. Again, this model search CME analysis does not agree with the results of the analysis of combined original data and follow-up experiments that showed CF caused significance of the confounded string.

4 | FERMENTATION PROCESS IMPROVEMENT EXPERIMENT

AlmeidaeSilva et al[20] used a 2^{8-4}_{IV} experimental design in an attempt to optimize the conditions for culturing *Paecilomyces variotti* in eucalyptus hemicellulosic hydrolysate in order to produce microbial protein.

The factors and levels for the fermentation experiment are shown in Table 9

TABLE 9 Factors and levels for fermentation experiment

Label	Factors	Levels -	Levels +
A	Inhibitors (furfural and acetic acid)	1.25g/L	7.8g/L
B	Rice bran	10.0g/L	30.0g/L
C	Urea	0.0g/L	2.0g/L
D	Magnesium sulfate	0.0g/L	1.5g/L
E	Ammonium sulfate	0.0g/L	2.0g/L
F	Potassium nitrate	0.0g/L	2.0g/L
G	Sodium phosphate	0.0g/L	2.0g/L
	Fermentation time	72 hrs.	96 hrs.

TABLE 10 Fermentation experiment

Standard Order	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	Biomass *Y*
1	-	-	-	-	-	-	-	-	5.75
2	+	-	-	-	-	+	+	+	6.70
3	-	+	-	-	+	-	+	+	11.12
4	+	+	-	-	+	+	-	-	10.67
5	-	-	+	-	+	+	+	-	4.92
6	+	-	+	-	+	-	-	+	5.35
7	-	+	+	-	-	+	-	+	2.81
8	+	+	+	-	-	-	+	-	10.83
9	-	-	-	+	+	+	-	+	6.08
10	+	-	-	+	+	-	+	-	7.27
11	-	+	-	+	-	+	+	-	9.68
12	+	+	-	+	-	-	-	+	4.20
13	-	-	+	+	-	-	+	+	3.90
14	+	-	+	+	-	+	-	-	3.78
15	-	+	+	+	+	-	-	-	11.57
16	+	+	+	+	+	+	+	+	7.39

The experimental runs in standard order are shown in Table 10, and the generators for the design were $E = BCD$, $F = ACD$, $G = ABC$, and $H = ABD$.

The biomass produced by this fungus during fermentation has all the amino acids necessary to feed humans and animals, and the goal of the experimentation was to find conditions to maximize the biomass.

4.1 | Analysis based on initial 16 experiments

Figure 4 shows a half-normal plot of the coefficients or half-effects calculated from the data. The largest values in absolute value are labeled, and a reduced model fit in these effects showed all terms were significant at the .05 level, except *G* which had a *P* value of .1123. A reduced model, without *G*, had and $R^2 = .7182$ and a residual standard error of 1.813.

The aliases for the two-factor interactions in the design are as follows:

$$CG + DH + AB + EF$$
$$AC + BG + DF + EH$$
$$CF + AD + EG + BH$$
$$CH + DG + AE + BF$$
$$CD + GH + AF + BE$$
$$BC + AG + DE + FH$$
$$CE + FG + AH + BD$$

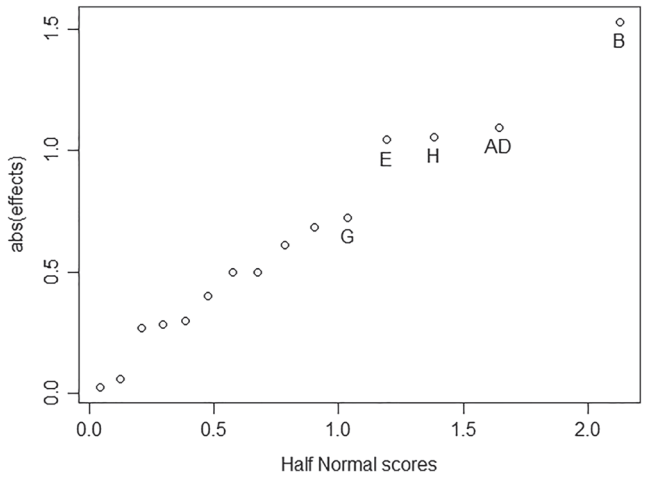

FIGURE 4 Half normal plot of coefficients

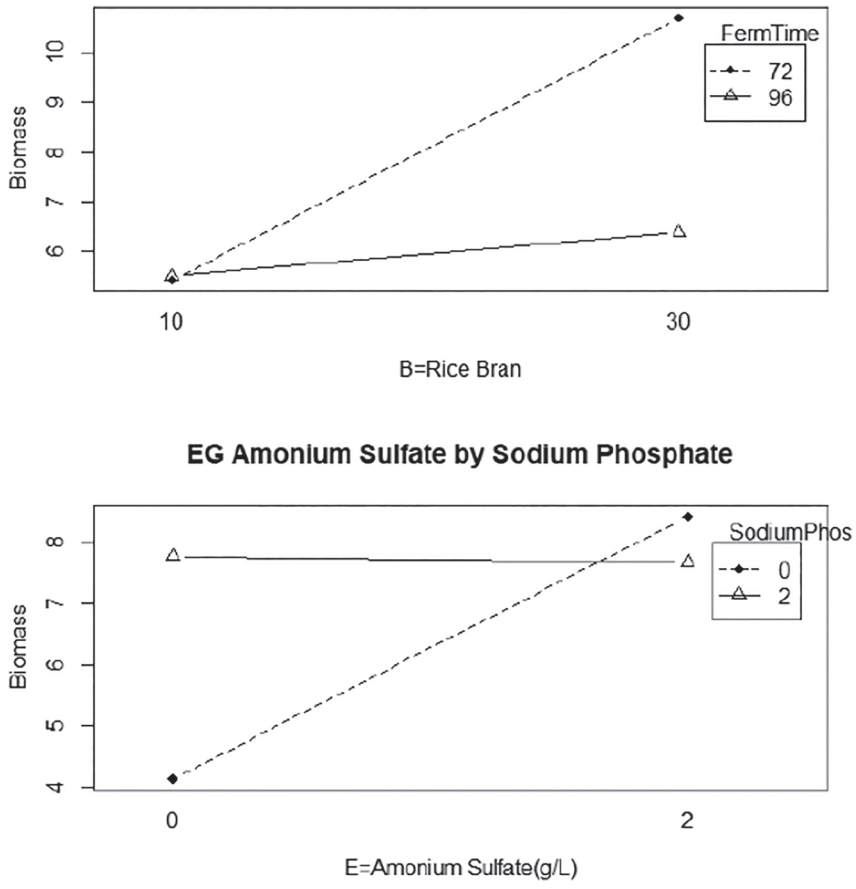

FIGURE 5 Confounded interaction plots

Based on the effect heredity principle, the apparently significant set of confounded two-factor interactions could represent *BH* since the two largest main effects were *B* and *H*. However, since main effects *E* and *G* are also large, the interaction could be *EG*. Figure 5 shows both of these interaction plots.

4.2 | Analysis and conclusions after follow-up experiments

The authors of the article felt that the experiment had given evidence that factor *A* (the inhibitor) had little effect and they confirmed this by citing other published reports. They also felt the experiment showed that main effects *D* (magnesium sulfate) and *F* (potassium nitrite) were insignificant. They decided to run the eight additional experiments shown in Table 11.

TABLE 11 Follow-up experiments

Run Number	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	Biomass *Y*
17	0	-	-	0	-	0	-	+	3.37
18	0	+	-	0	-	0	-	-	2.55
19	0	-	-	0	+	0	-	-	5.53
20	0	+	-	0	+	0	-	+	10.43
21	0	-	-	0	-	0	+	-	2.93
22	0	+	-	0	-	0	+	+	7.23
23	0	-	-	0	+	0	+	+	11.69
24	0	+	-	0	+	0	+	-	10.59

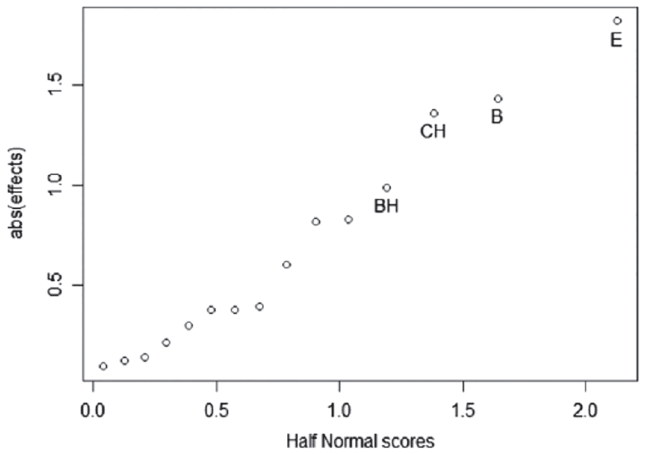

FIGURE 6 Half-normal plot of effects from 2^{5-1} Paecilomycesvariotii culture experiment

These experiments were run with factors they felt to be insignificant (ie, *A*, *D*, and *F*) held at the midpoint of the levels used in the initial 16 experiments. When combined with run numbers 6, 7, 8, 12, 13, 14, 15, and 16, from the initial 16 experiments, the eight follow-up experiments form a resolution V 2^{5-1} design with defining relation $I = BCEGH$ (ignoring factors *A*, *D*, and *F*). The resolution V design would allow fitting a model involving the five main effects and all two-factor interactions involving these main effects.

Figure 6 shows the half-normal plot of effects calculated from the resolution V design. The results of this set of follow-up experiments combined with eight runs from the original set of experiments suggest that main effects, *B*-rice bran, and *E*-ammonium sulfate along with interactions *CH* (urea × fermentation time) and *BH* (rice bran × fermentation time) appear to be significant. Since interactions exist, the main effects should not be interpreted in isolation.

Figure 7 shows the *BH* interaction plot. It shows how the effect of fermentation time depends upon the level of rice bran. When there is only 10 g/L of rice bran in the growth medium, it can be seen that increasing the fermentation time from 72 to 96 hours increases the biomass produced. However, if there are 30 g/L of rice bran in the growth medium, increasing the fermentation time actually decreases the biomass produced.

Figure 8 shows the interaction plot of fermentation time and urea, *CH*. On average, it can be seen that increasing the fermentation time seems to have little effect on biomass production. Also, on average, adding urea to the growth medium seems to have little effect on the biomass produced. However, as can be seen in the graph, the effect of fermentation time upon biomass depends upon whether urea is present and its effect appears to be exactly opposite depending on whether 2.0 g/L of urea is added to the growth medium.

While it is possible to have an interaction between two factors that do not have significant main effects (like the example shown in Figure 8), it is rare. As mentioned in the introduction, in the study of Li et al[5] of 113 published factorial experiments, this happened less than 1% of the time. Usually, interactions occur between factors where at least one of the two main effects are significant. Effect heredity describes this principle. In this experiment, since the two-factor interaction between fermentation time and urea is confounded with the three-factor interaction between rice bran, ammonium sulfate, and sodium phosphate (ie, $CH = BEG$), and all three of the latter factors have significant main effects, it is possible that the large effect labeled as *CH* on the normal plot is actually the three-factor interaction. In this case, the effect heredity principle (that would imply *BEG*) may provide a better interpretation of the data than the hierarchical ordering principle (that would imply *CH*).

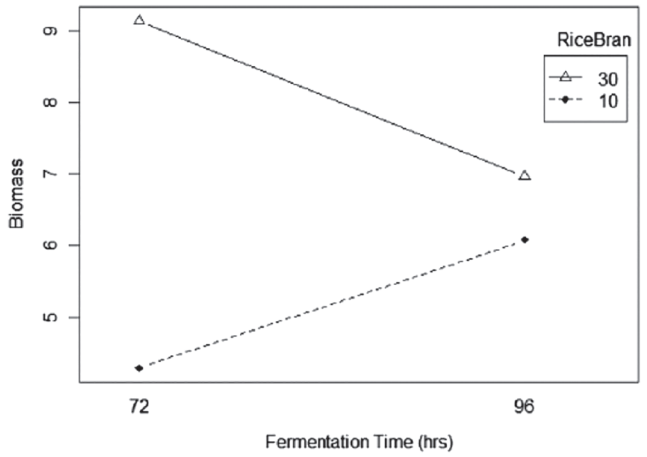

FIGURE 7 Interaction plot for fermentation time and level of rice bran

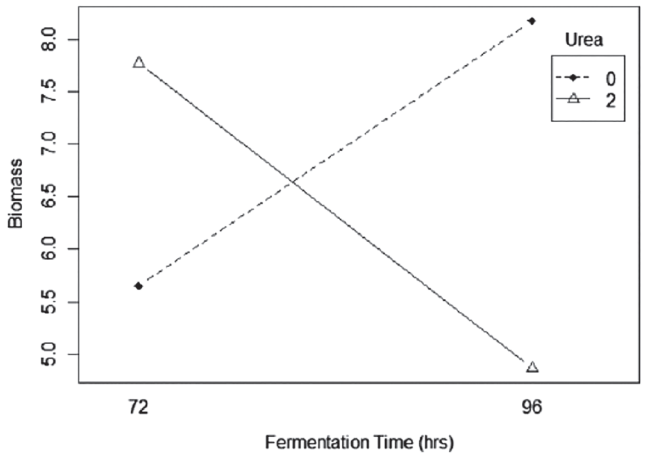

FIGURE 8 Interaction plot for urea and fermentation time

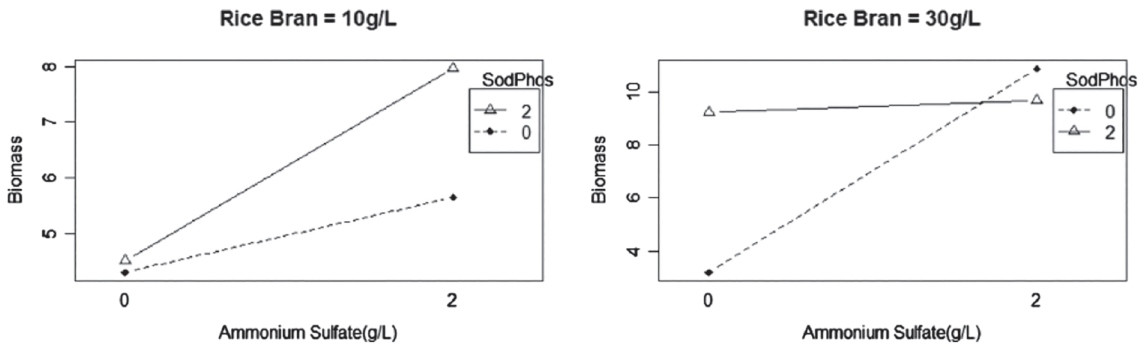

FIGURE 9 Interaction plots to interpret three-factor ammonium sulfate by rice bran by sodium phosphate interaction

A three-factor interaction means that the effect of one factor depends upon the combination of levels of two other factors. Figure 9 shows the effect of ammonium sulfate upon biomass at the four combinations of rice bran and sodium phosphate.

If this three-factor interaction is assumed to be important, the two-factor interactions do not tell the whole story. In Figure 9, it can be seen that adding 2.0 g/L of ammonium sulfate to the growth medium increases the biomass produced in general. However, this effect is greatest when there is 30 g/L of rice bran and no sodium phosphate in the growth medium. The optimum result appears to be with 30 g/L of rice bran, 2.0 g/L of ammonium sulfate, and 0.0 g/L of sodium phosphate in the growth medium. There the biomass yield was predicted to be 11.57 g/L.

The analysis of the combined set of all 24 experiments (the original 16 experiments in Table 10 plus the eight additional experiments in Table 11) showed that main effects *B*, *E*, and *G* were the most important (when excluding the factors *A*, *D*, and *F* that the authors felt were unimportant). Factor *H* (fermentaion time) and all its interactions were insignificant

TABLE 12 Coefficients from model fit to combined data

Coefficients	Estimate	Standard Error	*t* value	Pr(>|t|)	
Intercept	6.9308	0.3681	18.830	2.42e-12	**
B	1.3250	0.3681	3.600	0.002400	**
E	1.6200	0.3681	4.401	0.000446	**
G	0.9233	0.3681	2.509	0.023272	*
BE	0.4192	0.3681	1.139	0.271552	
BG	0.2942	0.3681	0.799	0.435881	
EG	-0.6442	0.3681	-1.750	0.099254	.
BEG	-1.1683	0.3681	-3.174	0.005889	**
R^2=.7704	s_e=1.803	df=16			

after combining the original data with the follow-up data. Therefore, a full factorial model containing just the factors *B*, *E*, and *G* and all their interactions was fit to the combined set of data. The results of this analysis are shown in Table 12.

The results in Table 12 show the three main effects *B*, *E*, and *G* were significant and the three-factor interaction *BEG* was significant at less than the .01 level. The three-factor interaction implies that the effect of factor *E* = ammonium sulfate depends on the combination of levels of both *B* = amount of rice bran and *G* = the amount of sodium phosphate. In the initial 16 experiments, the main effect for factor *H* was completely confounded with the *BEG* interaction, and this could be the reason it appeared to be significant in the analysis involving only main effects and two-factor interactions. The results of the analysis including the follow-up experiments reveals something much different than was found after the original 16 experiments.

The goal of the experimentation was to identify fermentation conditions that would result in the maximum amount of biomass. The model predictions shown in Figure 9 indicate that the high level of factor *E* (ammonium sulfate = 2.0 g/L), the low level of factor *G* (sodium phosphate = 0 g/L), and the high level of factor *B* (rice bran = 30.0 g/L) are predicted to result in the highest yield (biomass greater than 10.5 g/L). The three experimental runs at those conditions had an average biomass of 10.89 g/L with a standard deviation of .60 g/L.

4.3 | Conclusions from initial 16 experiments based on analysis of conditional main effects

The analysis of the initial 16 experiments showed that main effects *B*, *E*, and *H*, in addition to the confounded set of interactions $AD = BH = CF = EG$, were the significant effects. Using Wu's[13] suggested CME analysis, a forward stepwise regression using the minimum AIC stopping rule was used. The candidate terms for the forward selection were all eight main effects *A* to *H* and the 16 twin CMEs associated with the confounded string of interactions $AD = CF + EG + BH$. The first three terms entering the forward regression were $B|H-$, *H*, and *E*. A model involving these three terms had an $R^2 = .7064$. All terms were significant at the $\alpha = .05$ level and the model explained almost the same of the variation in biomass as the model that contained the four terms labeled on the half-normal plot in Figure 6 ($R^2 = .7182$). This model found by forward selection would imply that the *BH* interaction was causing the significance of the confounded set of interactions $AD = CF = EG + BH$.

The CME model search (using rule 1 described by Su and Wu.[16]) was performed by making the following substitutions of terms identified in the half-normal plot:

1. Replace *B*, and *BH* with either $B|H-$ or $H|B-$
2. Replace *H* and *BH* with either $H|B+$ or $B|H+$
3. Replace *E* and *EG* with either $G|E-$ or $E|G-$

The best of the six models was the one containing the terms *B*, *E*, and $H|B+$. This model had the same $R^2 = .7182$ as the model involving the labeled terms on the half-normal plot, and it again seems to indicate that the *BH* interaction is causing significance of the confounded string $AD = CF = EG = BH$.

Neither CME analysis of the original data in this experiment gives the same results as the analysis of follow-up experiments. The analysis of follow-up experiments seems to indicate a significant three-way interaction which helps to identify the optimum fermentation conditions. The proposed CME analysis does not consider three-factor interactions and would never work in identifying one from a resolution IV design. However, the experience supporting the *hierarchical ordering principle* would assert that this is a rare occurrence and this example should not deter anyone from trying a CME analysis when follow-up experiments cannot be obtained.

Although they don't agree with the analysis of follow-up experiments, the two proposed methods of CME analysis do give the same results for this experiment. They both indicate that the *BH* interaction is causing the significance of the

TABLE 13 Factors for injection molding experiment

Label	Factors
A	Mold temperature
B	Moisture content
C	Holding pressure
D	Cavity thickness
E	Booster pressure
F	Cycle time
G	Gate size
H	Screw speed

confounded string of interactions. The predicted optimum using the model including the terms B, E, H, and BH and the original data indicates that a maximum biomass can be produced at the high level of B = rice bran, the low level of H = fermentation time, and the high level of E = ammonium sulfate. These optimal conditions are somewhat similar to the optimal conditions predicted from the model involving B, E, G, and BEG obtained after follow-up experiments, but they do not specify that factor G = sodium phosphate should be set at its low level of 0 g/L, and they do specify that H = fermentation time (which was later found to have no significant effect after the follow-up experiments) should be set at its low level.

5 | INJECTION MOLDING EXPERIMENT

Box, Hunter, and Hunter[15] describe a 2^{8-4}_{IV} resolution IV fractional factorial experiment to study the shrinkage of parts in an injection molding process. The design was created as a saturated 2^{7-4}_{III} design augmented by its mirror image. The factors in the design are shown in Table 13, and each factor was studied at a low(−) and high(+) level.

5.1 | Analysis based on original 16 experiments

Based on a normal plot of the 15 effects calculated from the 16 runs, the effects that appeared to be significant were C = holding pressure, E = booster pressure, and the confounded string of two-factor interactions $AE = BF = CH = DG$. A model fit using the three terms C, E, and $AE+BF+CH+DG$ explained 94% of the variation in part shrinkage ($R^2 = .9398$) and all three terms in the model were significant.

5.2 | Analysis and conclusions after follow-up experiments

To get separate estimates of the interactions AE, BF, CH, and DG the four additional experiments listed in Appendix 12B of Box et al[15] were conducted. After incorporating the additional data, the separate estimates ($\hat{AE} = 0.6$, $\hat{BF} = 0.7$, $\hat{DG} = -1.6$, and $\hat{CH} = 4.9$) were obtained. These estimates showed that the CH (holding pressure × screw speed) interaction was responsible for the significance of the confounded string of interactions.

5.3 | Analysis of the original 16 experiments using CME analysis

Following Wu's[13] example, a forward selection procedure was run. The candidate variables were all eight main effects, A, B, C, D, E, F, G, and H and all 16 CMEs associated with each of the interactions AE, BH, CF, and EG (ie, $A|E+$, $A|E-$, $E|A+$, $E|A-$, $B|F+$, $B|F-$, $F|B+$, $F|B-$, $C|H+$, $C|H-$, $H|C+$, $H|C-$, $D|G+$, $D|G-$, $G|D+$, and $G|D-$).

The first two terms to enter the forward selection were $C|H+$ and E. This two-variable model had an $R^2 = .934$ with both terms highly significant. The CME $C|H+$ would indicate that the CH interaction was responsible for the significance of the confounded string of interactions $AE = BF = CH = DG$ and agrees with the results obtained after follow up experiments on p. 415 of box et al[15] In the table on that page, it can be seen that factor C = holding pressure had a strong positive effect upon shrinkage when H = screw speed is at its high level but a negligible effect when H = screw speed is at its low level.

Following the model search proposed by Su and Wu,[16] four orthogonal two-variable models were formed. In model 1, the CME $A|E-$ was substituted for parent main effect A and the interaction AE. It had an $R^2 = .6322$. In model 2, the CME $E|A-$ was substituted for parent main effect E and the interaction AE. It had an $R^2 = .9352$. In model 3, the CME $H|C+$ was substituted for parent main effect H and the interaction CH. It had an $R^2 = .4461$. Finally, in model 4, the CME $C|H+$ was substituted for parent main effect C and the interaction CH. It had an $R^2 = .934$.

TABLE 14 Summary of published studies

	Published Study				
Factors	**Analytical Lab**	**Fermentation**	**BHH Inject. Mld.**	**Wu[13] GM Canada**	**Su and Wu[16] Filtration**
No. Sig. Effects on Half Normal Plt.	6	4	3	4	4
No. Aliases for Sig. two-factor Interaction	3	4	4	2	2
Effect Heredity	Strong	Strong	Weak	Weak	Strong
Results	Analytical Lab	Fermentation	BHH Inject. Mld.	GM Canada	Filtration
Interact. Identified in Forward Sel.	CF or EG	BH	CH	CF	AC & BC
Interact. Identified in Model Search	AD	BH	AE	CF	AD & DB
Interact. Identified in follow-up Expts.	CF	BEG	CH	N.A.	N.A.

Models 2 and 4 were close with regard to the amount of variability they explained in the part shrinkage, but without having seen the results of the analysis results based on the follow-up experiments, it would appear that model 2 fit the data best. Because it included the terms $E|A-$ and C, it would imply that the AE interaction is causing the significance of the confounded set of two-factor interactions. This conclusion does not agree with either the analysis including follow-up experiments or the CME analysis based on forward selection.

6 | ASSESSMENT OF CME ANALYSIS BASED ON THREE PUBLISHED STUDIES THAT USED FOLLOW-UP EXPERIMENTS

Two methods of CME analysis were described in Section 2.1. The first method recommended by Wu[13] used a forward selection procedure that included as candidates all main effects and all CME's associated with each of the interactions in every confounded string of interactions that appeared to be significant on a half normal plot. The second method proposed by Su and Wu[16] used an orthogonal model search, where each model was obtained by substituting one CME for its parent main effect and corresponding interaction in the model identified with the half-normal plot. However, in the four examples presented (one by Wu,[13] and the other three by Su and Wu[16]), the two methods only give the same result in three of the four cases.

Table 14 compares the results of the two methods of CME analysis and the analysis of follow-up experiments for the three examples presented in this paper in addition to one example from Wu[13] and one example from Su and Wu[16]. The last two examples did not include follow-up experiments.

From the table, it can be seen that the results of the forward selection method of CME analysis matched the results of analysis of the data combined with follow-up experiments for only one of the three examples where follow-up experiments were included. The model search method of CME analysis did not give the same results as the analysis of analysis of the data combined with follow-up experiments for any of the examples where follow-up experiments were included. Additionally, the results of the forward selection method of CME analysis matched the results of the model search method for two of the five examples presented in Table 14.

The results of the CME analysis are not consistent across the five examples presented in the table. However, there are other differences between the examples with regard to the the factors listed in the left column in the upper half of Table 14.

7 | SIMULATION STUDY

In attempt to determine if the accuracy of CME analysis is dependent on the levels of the factors shown in Table 14, a modest simulation study was conducted varying the three factors shown in Table 15.

Half the simulations were made using a 2^{6-2}_{IV} design with generators ABC and BCD, which resulted in each two factor interaction being confounded with one other two-factor interaction (ie, two aliases for each two-factor interaction). The

TABLE 15 Factors and levels for simulation study

Label	Factors	Levels -	Levels +
A	Number of significant terms on half normal plot	3	5
B	Number of aliases for significant two-factor interaction	2	4
C	Heredity	Weak	Strong

TABLE 16 Results of simulation study

Standard Order	*A*	*B*	*C*	Forward Selection Correct	Forward Selection Incorrect	Model Search Correct	Model Search Incorrect
1	-	-	-	16	4	6	14
2	+	-	-	16	4	7	13
3	-	+	-	20	0	8	12
4	+	+	-	16	4	5	15
5	-	-	+	15	5	5	15
6	+	-	+	11	9	1	19
7	-	+	+	20	0	20	0
8	+	+	+	8	12	6	14

other half of the simulations were made using a 2^{8-4}_{IV} design with generators *BCD*, *ACD*, *ABC*, and *ABD*. Each two-factor interaction in this design was confounded with three other two-factor interactions (ie, four aliases for each two-factor interaction). The number of active effects in each simulation varied between three and five, and one of the active effects was a two-factor interaction. The rest were main effects. Strong heredity was created by choosing the active two-factor interaction to be one that involved two of the active main effects. Weak heredity was created by choosing the active two-factor interaction to be one that involved only one of the active main effects. Twenty simulations were performed at each of the 2^3 combinations of factor levels shown in Table 16.

From the results shown in Table 16, it can be seen that the forward selection method of CME analysis is more effective. It finds the correct active interaction in the significant but confounded set 76.0% of the time while the CME analysis based on the orthogonal model search only identified the correct interaction 29.3% of the time. This confirms the analysis of published studies with follow-up experiments where the orthogonal model search never identified the active interaction.

A logistic regression model was fit to the results in order to determine if the accuracy of the two methods was influenced by the factors studied. This analysis showed that factor *A* (the number of significant terms or active terms), factor *B* (the heredity of the interaction), and the *AB* interaction had significant effects on the accuracy of the forward selection method of analysis. The number of active terms had a negative effect on the accuracy of the forward selection, and this method was most accurate with only three active terms. This finding is similar to the analysis of published studies where the forward selection and analysis including follow-up experiments matched only for Box-Hunter and Hunter's injection molding experiment where there were only three large effects on the half-normal plot. Due to the interaction *AB* found in the simulations, the forward selection method was found to be most accurate when there were only three active effects and strong heredity. For this combination of factor levels, the forward selection identified the correct interaction in all 40 cases (ie, two replicates of these conditions).

None of the factors in the simulation study had significant effects on the accuracy of the model search method of CME analysis. Therefore, this method is only expected to identify the correct interaction in a confounded string of interactions 29.3% of the time on the average. This result again ratifies the result of the analysis of published studies where the model search never identified the interaction determined to be significant after follow-up experiments.

8 | CONCLUSIONS AND RECOMMENDATIONS

The results shown in this paper indicate that the forward selection method of CME analysis is the most accurate, yet it should not be trusted if there are five or more large effects on the half-normal plot and only weak heredity (ie, only one of the large main effects is involved in each alias of the significant string of two-factor interactions). In simulations, the only conditions where both methods of CME analysis works well is when there are only three large effects on the half-normal plot and strong heredity. However, this is the situation where the strong heredity principle would normally lead an analyst to choose the correct interaction without any further analysis.

Although CME analysis was conjectured by Wu[13] and Su and Wu[16] to be a useful way to determine which two-factor interaction was causing the significance of a string of confounded two-factor interactions, the results of its use on exper-

iments that included follow-up experiments and the results of simulations showed that it is not reliable. Follow-up experiments are the most reliable way of separating confounded two-factor interactions.

Based on the simulations in this paper, it is concluded that the model search method of CME analysis is not effective. Additionally, it is recommended that if a forward selection method of CME analysis is employed, then follow-up experiments should be performed to confirm the results of analysis.

One situation where CME analysis has been proven to be effective is when there is a nested factor whose effect is likely to be different depending on the level of the factor within which it is nested (see Goos and Jones[21]). In that case, again the forward selection of CME analysis is recommended.

ORCID

John Lawson https://orcid.org/0000-0003-2563-9233

REFERENCES

1. Lawson J. *Design and Analysis of Experiments with R*. Boca Raton, FL: CRC Press; 2015. ISBN 978-1-4398-3.
2. Lin DKJ, Draper NR. Projection properties of Plackett-Burman designs. *Technometrics*. 1992;34:423-428.
3. Wang JC, Wu CFJ. A hidden projection property of Plackett-Burman and related designs. *Statistica Sinica*. 1995;5:235-250.
4. Wu GFJ, Hamada M. *Experiments Planning, Analysis and Parameter Design Optimization*. New York, NY: John Wiley and Sons, Inc.; 2000. ISBN 0-471-25511-4.
5. Li X, Sudarsanam N, Frey D. Regularities in data from factorial experiments. *Complexity*. 2006;11:32-45.
6. Plackett RL, Burman JP. The design of optimum multifactorial experiments. *Biometrika*. 1946;33:305-325.
7. Li W, Nachtsheim C. Model robust factorial designs. *Technometrics*. 2000;42:345-352.
8. Jones B, Montgomery DC. Alternatives to resolution IV screening designs in 16 runs. *Int. J. Experimental Design and Process Optimisation*. 2010;1:285-295.
9. Lin DKJ. Spotlight interaction effects in main effect plans: a supersaturated design approach. *Quality Engineering*. 1999;11:133-139.
10. Jones B, Nachtsheim CJ. A class of three-level designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology*. 2011;4:1-15.
11. JMP©, Version 14. SAS Institute Inc., Cary NC, 1989-2019; 2019.
12. Lawson J. daewr: Contains data frames and functions used in the book–Design and Analysis of Experiments with R, CRC Press 2015. R package version 1.1.7, http://CRAN.R-project.org/package=daewr; 2016.
13. Wu CFJ. Post-Fisherian experimentation: from physical to virtual. *Journal of the American Statistical Association*. 2015;110:612-620.
14. Montgomery DC. *Design and Analysis of Experiments, Sixth Edition*. New York, NY: John Wiley and Sons, Inc.; 2005. ISBN 0-471-66159-7.
15. Box GEP, Hunter JS, Hunter WG. *Statistics for Experimenters - Design, An Introduction to Design, Data Analysis, and Model Building*. New York, NY: John Wiley and Sons, Inc.; 1978. ISBN 0-471-09315-7.
16. Su H, Wu CFJ. CME Analysis: a new method for unraveling aliased effects in two-level fractional factorial experiments. *Journal of Quality Technology*. 2017;49:1-10.
17. Snee RD. Experimenting with a large number of variables. In: Snee RD, Hare LB, Trout JR, eds. *Experiments in Industry: Design, Analysis, and Interpretation of Results*. American Society for Quality Control: Chemical and Process Industries Division Technical Supplement; 1985:25-35.
18. Wernimont G. Ruggedness evaluation of test procedures. *ASTM Standardization News*. 1977;5:61-64.
19. Sugiura N. Further analysis of the data by Akaike information criterion and the finite corrections. *Communication in Statistics-Theory and Methods*. 1978;7:13-26.
20. AlmeidaeSilva JB, Lima UA, Taqueda MES, Guaragna FG. Use of fractional factorial designs for selection of nutrients for culturing *Paecilomyces variotti* in Eucalyptus Hemicellulosic Hydrolysate. *Braziian Journal of Chemical Engineering*. 1998;15:273-279.
21. Goos P, Jones B. Optimal Experimental Designs in the Presence of Nested Factors. published online April 25, 2019; 2019.

AUTHOR BIOGRAPHY

John Lawson is Professor Emeritus at Brigham Young University. He received a MSc Degree in Mathematical Statistics from Rutgers University in 1977, and a PhD Degree in Applied Statistics from the Polytechnic Institute of N.Y in 1984. He was a Statistician at Johnson and Johnson Corporation, and the Manager of Statistical Services at FMC Chemicals Division prior to becoming a Professor of Statistics.

How to cite this article: Lawson J. Comparison of conditional main effects analysis to the analysis of follow-up experiments for separating confounded two-factor interaction effects in 2^{k-p}_{IV} fractional factorial experiments. *Qual Reliab Engng Int*. 2020;36:1454–1472. https://doi.org/10.1002/qre.2638