Statistica Sinica **30** (2020), 903-924
doi:https://doi.org/10.5705/ss.202018.0041

AN ALGEBRA FOR THE CONDITIONAL MAIN EFFECT PARAMETERIZATION

Arman Sabbaghi

Purdue University

Abstract: The conditional main effect (CME) parameterization system resolves the long-standing aliasing dilemma associated with the traditional orthogonal components system for two-level regular fractional factorial designs. However, the algebra of the CME system is not yet fully understood, which impedes the development of general results on this system that possess a broad scope of application across designs. Therefore, we establish a comprehensive algebra for the CME system based on indicator functions. Our algebra facilitates derivations of general partial aliasing relations for a wide variety of two-level designs. Using the proposed algebra, we examine the implications for resolution IV designs of traditional design criteria under the CME system. A novel feature of our algebra is that it enables immediate and simple D-efficiency calculations for two-level regular designs and for models consisting of multiple conditional and traditional effects.

Key words and phrases: Complex aliasing, experimental design, regular fractional factorial design.

1. Introduction

Two-level regular fractional factorials are convenient designs for inferences on main effects and interactions in experiments with many factors and run size constraints. The traditional method of analysis is based on the orthogonal components parameterization of the factorial effects (Wu and Hamada (2009, p. 274)). Ever since the work of Finney (1945), the major disadvantage of regular designs was thought to be that, under this traditional system, no two aliased effects can be disentangled without follow-up runs (Wu (2015)). Su and Wu (2017) recently resolved this long-standing dilemma by reparameterizing the traditional main effects and two-factor interactions into main effects and conditional main effects (CMEs). In contrast to the orthogonal components system, analyses of regular designs under the CME system are similar to those of nonregular designs, due to the partial aliasing among the conditional and traditional effects. From the work of Hamada and Wu (1992) on the analysis of nonregular designs and partial aliasing, this feature of the CME system can be used to eliminate the need to employ follow-up runs to perform conclusive analyses on regular designs. Wu (2015, p. 615) first noted this advantage of the CME system and its utility for real-life applications. Su and Wu (2017) then proposed an analysis strategy for resolution III and IV designs under this system, based on partial aliasing relations among conditional and traditional effects. They leveraged the structure of their groupings of CMEs to develop simple rules for selecting parsimonious and interpretable models that yield unambiguous inferences in two-level regular fractional factorials.

The innovation of the CME system created several novel avenues of research for both experiments and observational studies. Mukerjee, Wu and Chang (2017) proposed an effect hierarchy for this system. Furthermore, they developed a design strategy incorporating a minimum aberration criterion to sequentially minimize the bias in estimations of main effects by successive iterations in the effect hierarchy. Mak and Wu (2019) proposed an analysis strategy for general observational data under this system that can perform bi-level variable selection and separate active effects from correlated groups of inert effects. A unified and insightful review on this system, the above recent advances, and related topics are provided by Wu (2018).

However, a significant feature of the CME system that has yet to be addressed is its algebra, which is not as transparent as the Galois field theory of the traditional orthogonal components system. The lack of an accessible algebra impedes the development of general results on CMEs that can be applied across the spectrum of two-level designs. For example, Su and Wu (2017, p. 10) noted that, although they could determine partial aliasing relations between conditional and traditional effects in small regular designs, their approach would not be feasible for deriving general partial aliasing relations in large designs. In addition, Mukerjee, Wu and Chang (2017) considered simple models consisting of traditional effects and just one CME, because more general models involving any number of conditional and traditional effects would incur heavy algebra under their framework. These examples highlight the need for an algebra that can facilitate the derivation of general results and properties under the CME system for broad types of two-level designs.

We establish an algebra for the CME system based on indicator functions for two-level designs. Fontana, Pistone and Rogantin (2000) introduced the indicator function based on the algebraic perspective of Pistone and Wynn (1996). They applied indicator functions to address multiple aspects of the classification of unreplicated two-level regular fractional factorials. Ye (2003) extended indicator functions to two-level nonregular fractions with replicate runs for ranking designs. Ye (2004) then used indicator functions to prove that a two-level design with no partial aliasing under the orthogonal components system must be a two-level regular fractional factorial, potentially with replicate runs. The orthogonal components system was always considered in these and other investigations on two-level designs that involved indicator functions. In contrast, we use an orthogonal basis of functions, the span of which contains indicator functions, to explicitly represent both conditional and traditional effects, and we define an inner product of these representations using the indicator function for a two-level design, the properties of which are studied here under the CME system. The contributions of our algebra are threefold. First, in contrast to the work of Su and Wu (2017), it facilitates general derivations of partial aliasing relations among conditional and traditional effects for broad classes of large designs. For example, Properties 2–5 of Su and Wu (2017, pp. 3–6) follow as simple calculations under our inner product, with no, or rather weak conditions. Second, our algebra reveals the implications for CME analyses of resolution IV designs of the maximum clear two-factor interactions (Wu and Hamada (2009, p. 217)) and minimum aberration (Fries and Hunter (1980)) design criteria. Third, it enables immediate and simple D-efficiency calculations for two-level regular designs and models consisting of multiple CMEs, main effects, and two-factor interactions. This particular contribution distinguishes our work from that of Mukerjee, Wu and Chang (2017).

We begin in Section 2 with a review of the CME system, its connections with traditional effects and nonregular designs, and its groupings of effects. Our algebra is defined in Section 3. We apply our algebra in Section 4 to derive general partial aliasing relations among conditional and traditional effects. In Section 5, we discuss the implications of traditional design criteria for CME analyses of resolution IV designs. We apply the proposed algebra to D-efficiency calculations in Section 6. Illustrative examples of our results are provided throughout the latter three sections, and the proofs are provided in the Supplementary Material. A practical application that demonstrates the importance of our results for real-world CME analyses is presented in Section 7. Section 8 concludes the paper.

2. Review of the CME System

Let $\mathcal{D}_r$ denote the 2^r full factorial for $r \geq 2$ factors, with the levels of factors $A_1, \ldots, A_r$ denoted by $-$ and $+$. A fraction of $\mathcal{D}_r$ is denoted by $\mathcal{F} \subseteq \mathcal{D}_r$.

As described by Cheng (2014, pp. 71–75), the main effects and interactions for a two-level design $\mathcal{D}_r$ are defined as contrasts of all of its 2^r treatment effects $\alpha(s_1, \ldots, s_r)$, where $s_1, \ldots, s_r \in \{-, +\}$ denote the factors' levels. The $\alpha(s_1, \ldots, s_r)$ are unknown, and, in general, the factorial effects are estimated using a least squares linear regression (Cheng (2014, pp. 81–83)). Similarly, a CME is defined as a contrast of the treatment effects that captures the main effect of one factor conditional on the level of a second. The estimation of a CME is also performed using a regression. A general description of the CME is given by Wu and Hamada (2009, p. 164) and Cheng (2014, pp. 71–72).

To illustrate the traditional and conditional effects, consider $\mathcal{D}_2$, and let $\alpha = (\alpha(-,-), \alpha(-,+), \alpha(+,-), \alpha(+,+))^\mathsf{T}$, where $\alpha(s_1, s_2)$ is the treatment effect for $(s_1, s_2) \in \{-,+\}^2$. The main effects of A_1 and A_2 are $\mathrm{ME}(A_1) = 2^{-1}(-1,-1,1,1)\alpha$ and $\mathrm{ME}(A_2) = 2^{-1}(-1,1,-1,1)\alpha$, respectively, and their interaction is $\mathrm{INT}(A_1, A_2) = 2^{-1}\{(-1,-1,1,1) \odot (-1,1,-1,1)\}\alpha$, where $\odot$ is the Hadamard product. The CMEs of A_1 given A_2 are

$$\mathrm{CME}(A_1 \mid A_2+) = \{(-1,-1,1,1) \odot (0,1,0,1)\}\alpha,$$

$$\mathrm{CME}(A_1 \mid A_2-) = \{(-1,-1,1,1) \odot (1,0,1,0)\}\alpha.$$

The sum and difference of $\mathrm{CME}(A_1 \mid A_2+)$ and $\mathrm{CME}(A_1 \mid A_2-)$ effectively define $\mathrm{ME}(A_1)$ and $\mathrm{INT}(A_1, A_2)$, respectively (Wu and Hamada (2009, p. 164)), and thus reparameterize them. Wu (2018, p. 252) provides physical interpretations of these effects, as well as the connections between them. If we let $y = (y(-,-), y(-,+), y(+,-), y(+,+))^\mathsf{T}$ denote the observed outcomes, where $y(s_1, s_2)$ is the response for the experimental unit assigned $(s_1, s_2) \in \{-,+\}^2$, then the estimators of these effects are $\widehat{\mathrm{ME}(A_1)} = 2^{-1}(-1,-1,1,1)y$, $\widehat{\mathrm{ME}(A_2)} = 2^{-1}(-1,1,-1,1)y$, $\widehat{\mathrm{INT}(A_1, A_2)} = 2^{-1}(1,-1,-1,1)y$, $\widehat{\mathrm{CME}(A_1 \mid A_2+)} = (0,-1,0,1)y$, and $\widehat{\mathrm{CME}(A_1 \mid A_2-)} = (-1,0,1,0)y$. The correlations between the estimators of the traditional effects and the CMEs, and between estimators of the distinct CMEs themselves, are strictly less than one in absolute value. Consequently, including CMEs with traditional effects in the analysis of a regular design introduces partial aliasing relations, and results in the design becoming of the nonregular type in its analysis (Wu (2018, p. 251)).

To simplify the exposition in this paper, our references to partial aliasing relations or to correlations between effects in a design signify the aliasing relations or correlations between the corresponding estimators. Formal definitions of different groups of conditional and traditional effects under the CME system follow below.

Definition 1 (Twin CMEs (Su and Wu (2017))). *For distinct factors* A_1 *and* A_2, $\mathrm{CME}(A_1 \mid A_2+)$ *and* $\mathrm{CME}(A_1 \mid A_2-)$ *are twins, with* A_1 *the parent effect,* A_2 *the conditioned effect, and conditioned levels* $+$ *and* $-$, *respectively.*

Definition 2 (Parent-child pair (Mak and Wu (2019))). *For distinct factors* A_1 *and* A_2, $\mathrm{CME}(A_1 \mid A_2 s)$, *with* $s \in \{-,+\}$, *and its corresponding parent main effect* $\mathrm{ME}(A_1)$ *constitute a parent-child pair.*

Definition 3 (Uncle-nephew pair (Mak and Wu (2019))). *For distinct factors* A_1 *and* A_2, $\mathrm{CME}(A_1 \mid A_2 s)$, *with* $s \in \{-,+\}$, *and its corresponding conditioned main effect* $\mathrm{ME}(A_2)$ *constitute an uncle-nephew pair.*

Definition 4 (Sibling CMEs (Su and Wu (2017))). *For distinct factors* $A_1, A_2,$ *and* A_3, $\mathrm{CME}(A_1 \mid A_2 s)$ *and* $\mathrm{CME}(A_1 \mid A_3 s')$, *with* $s, s' \in \{-,+\}$, *are siblings.*

Definition 5 (Cousin CMEs (Mak and Wu (2019))). *For distinct factors* $A_1, A_2,$ *and* A_3, $\mathrm{CME}(A_1 \mid A_2 s)$ *and* $\mathrm{CME}(A_3 \mid A_2 s)$, *with* $s \in \{-,+\}$, *are cousins.*

Definition 6 (Family of CMEs (Su and Wu (2017))). *For a fraction* $\mathcal{F} \subseteq \mathcal{D}_r$, *any two* $\mathrm{CME}(A_i \mid A_j s)$ *and* $\mathrm{CME}(A_l \mid A_k s')$, *with* $i, j, l, k \in \{1, \ldots, r\}$ *and* $s, s' \in \{-,+\}$, *whose corresponding two-factor interactions* $\mathrm{INT}(A_i, A_j)$ *and* $\mathrm{INT}(A_l, A_k)$ *are fully aliased in* $\mathcal{F}$, *belong to one family of CMEs of* $\mathcal{F}$, *and are referred to as family members.*

For a regular fractional factorial, any two of its distinct families must be disjoint, by virtue of the Galois field theory construction of regular designs. In addition, by inspection, a two-factor interaction $\mathrm{INT}(A_1, A_2)$ that is orthogonal to all other main effects and two-factor interactions in a regular fraction (i.e., a clear two-factor interaction) corresponds to the trivial family of CMEs $\{\mathrm{CME}(A_1 \mid A_2+), \mathrm{CME}(A_1 \mid A_2-), \mathrm{CME}(A_2 \mid A_1+), \mathrm{CME}(A_2 \mid A_1-)\}$. The number of nontrivial families, each of which contain distinct pairs of factors in their CMEs, in a regular fraction is equal to the number of aliasing relations that contain more than one two-factor interaction.

3. Indicator Functions and the Inner Product for the CME System

A design $\mathcal{F} \subseteq \mathcal{D}_r$ with distinct runs is specified completely by its indicator function $F_{\mathcal{F}} : \{-,+\}^r \to \mathbb{R}$, defined by Fontana, Pistone and Rogantin (2000, p. 153) as

$$F_{\mathcal{F}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{F}, \\ 0 & \text{otherwise.} \end{cases}$$

This function generalizes traditional design descriptions, for example, those based on defining relations, via the concept of algebraic variety (Fontana, Pistone and Rogantin (2000, p. 150)). Indicator functions also exist for designs with replicate runs (Ye (2003)), but we do not consider those here.

From Fontana, Pistone and Rogantin (2000, pp. 152–153) and Ye (2003, p. 985), $F_{\mathcal{F}}$ is expressed as a unique linear combination of the following set of orthogonal functions over $\{-,+\}^r$. Let $\mathcal{P}_r$ denote the power set of $\{1,\ldots,r\}$. For each $I \in \mathcal{P}_r$, define $X_I : \{-,+\}^r \to \mathbb{R}$ as $X_I(x) = \prod_{i\in I} x_i$, with $X_\phi \equiv 1$ being a constant function. Then, $\{X_I : I \in \mathcal{P}_r\}$ is an orthogonal basis of functions over $\{-,+\}^r$, and

$$F_{\mathcal{F}}(x) = \sum_{I\in\mathcal{P}_r} b_{\mathcal{F},I} X_I(x),$$

for unique $b_{\mathcal{F},I} \in \mathbb{R}$. Each X_I in this basis represents a traditional effect. For example, $X_{\{i\}}$ represents $\mathrm{ME}(A_i)$, and $X_{\{i,j\}}$ represents $\mathrm{INT}(A_i, A_j)$, for distinct $i, j \in \{1,\ldots,r\}$. For any fraction $\mathcal{F} \subseteq \mathcal{D}_r$, the work of Fontana, Pistone and Rogantin (2000, p. 154) yields that $b_{\mathcal{F},\phi} = 2^{-r}|\mathcal{F}|$, and $b_{\mathcal{F},I} = 2^{-r}\sum_{x\in\mathcal{F}} X_I(x)$ for $I \in \mathcal{P}_r$. The indicator function coefficients $b_{\mathcal{F},I}$ encode information on the correlations between the effects in $\mathcal{F}$. This is illustrated for the case of regular designs and traditional effects in the following proposition of Fontana, Pistone and Rogantin (2000, p. 154).

Definition 7. *The symmetric difference of $I, J \in \mathcal{P}_r$ is $I \triangle J = (I \cup J) - (I \cap J)$.*

Proposition 1 (Fontana, Pistone and Rogantin (2000))**.** *The correlation between any two traditional effects in a regular design $\mathcal{F} \subseteq \mathcal{D}_r$, corresponding to $I, J \in \mathcal{P}_r$ and not belonging to the defining contrast subgroup of $\mathcal{F}$, is $b_{\mathcal{F},I\triangle J}/b_{\mathcal{F},\phi}$.*

The above orthogonal basis of functions underlies our algebra for the CME system. Specifically, CMEs are easily expressed using this orthogonal basis, with $\mathrm{CME}(A_i \mid A_j+)$ and $\mathrm{CME}(A_i \mid A_j-)$, for distinct $i, j \in \{1,\ldots,r\}$, represented by $X^+_{i|j} \equiv 2^{-1}(X_{\{i\}} + X_{\{i,j\}})$ and $X^-_{i|j} \equiv 2^{-1}(X_{\{i\}} - X_{\{i,j\}})$, respectively. In these expressions, CMEs are again seen to be functions of traditional effects, and can be considered additional factors of interest in a two-level design. We then apply the indicator function of a fraction to define the following inner product of the functions over $\{-,+\}^r$ that correspond to conditional and traditional effects, thus establishing our algebra for the CME system.

Definition 8. *For a fractional factorial design $\mathcal{F} \subseteq \mathcal{D}_r$, $i, j, l, k \in \{1, \ldots, r\}$, $s, s' \in \{-, +\}$, and $I, J \in \mathcal{P}_r$,*

$$\langle X_I, X_J \mid \mathcal{F} \rangle = 2^{-r} \sum_{x \in \mathcal{D}_r} F_{\mathcal{F}}(x) X_I(x) X_J(x),$$

$$\left\langle X_{i|j}^s, X_I \mid \mathcal{F} \right\rangle = 2^{-r} \sum_{x \in \mathcal{D}_r} F_{\mathcal{F}}(x) X_{i|j}^s(x) X_I(x),$$

$$\left\langle X_{i|j}^s, X_{l|k}^{s'} \mid \mathcal{F} \right\rangle = 2^{-r} \sum_{x \in \mathcal{D}_r} F_{\mathcal{F}}(x) X_{i|j}^s(x) X_{l|k}^{s'}(x).$$

As we show in the following sections, the partial aliasing relations and other properties of a two-level design under the CME system can be derived in a simple and unrestricted manner using this inner product of coordinate-free representations of the conditional effects, traditional effects, and the design's indicator function.

Note that if we wish to use a different orthogonal basis containing functions corresponding to CMEs, then we must necessarily select a set of conditional and traditional effects that are orthogonal in $\mathcal{D}_r$. However, such selections may fail to permit a coordinate-free presentation. They may also unduly restrict the CMEs that can be studied under the corresponding algebra, thus hindering our ability to understand the CME system for broad types of two-level designs.

4. Partial Aliasing Relations Under the CME System

The inner product in Definition 8 facilitates derivations of the partial aliasing relations among conditional and traditional effects.

Lemma 1 (Fontana, Pistone and Rogantin (2000); Ye (2003))**.** *For $\mathcal{F} \subseteq \mathcal{D}_r$ and $I, J \in \mathcal{P}_r$,*

$$\langle X_I, X_J \mid \mathcal{F} \rangle = \langle X_\phi, X_{I \triangle J} \mid \mathcal{F} \rangle = b_{\mathcal{F}, I \triangle J}.$$

Proposition 2. *For $\mathcal{F} \subseteq \mathcal{D}_r$ and any $i, j, l, k \in \{1, \ldots, r\}$, with $i \neq j, l \neq k$, and $s, s' \in \{-, +\}$,*

$$\left\langle X_{i|j}^s, X_{l|k}^{s'} \mid \mathcal{F} \right\rangle = 2^{-2} \left(b_{\mathcal{F}, \{i\} \triangle \{l\}} + s' b_{\mathcal{F}, \{i\} \triangle \{l,k\}} \right.$$
$$\left. + s b_{\mathcal{F}, \{i,j\} \triangle \{l\}} + s s' b_{\mathcal{F}, \{i,j\} \triangle \{l,k\}} \right).$$

Corollary 1. *The correlation between* $\mathrm{CME}(A_i \mid A_j s)$ *and* $\mathrm{CME}(A_l \mid A_k s')$ *in a regular design $\mathcal{F} \subseteq \mathcal{D}_r$ of resolution at least* III*, for any $i, j, l, k \in \{1, \ldots, r\}$, with $i \neq j, l \neq k$, and $s, s' \in \{-, +\}$, is*

$$2^{-1} b_{\mathcal{F},\phi}^{-1} \left(b_{\mathcal{F}, \{i\} \triangle \{l\}} + s' b_{\mathcal{F}, \{i\} \triangle \{l,k\}} + s b_{\mathcal{F}, \{i,j\} \triangle \{l\}} + s s' b_{\mathcal{F}, \{i,j\} \triangle \{l,k\}} \right).$$

Proposition 3. *For $\mathcal{F} \subseteq \mathcal{D}_r$ and any $i, j \in \{1, \ldots, r\}$, with $i \neq j$, $I \in \mathcal{P}_r$, and $s \in \{-, +\}$,*

$$\left\langle X^s_{i|j}, X_I \mid \mathcal{F} \right\rangle = 2^{-1}(b_{\mathcal{F},\{i\}\triangle I} + s b_{\mathcal{F},\{i,j\}\triangle I}).$$

Corollary 2. *For a regular design $\mathcal{F} \subseteq \mathcal{D}_r$ of resolution at least* III *and any $i, j \in \{1, \ldots, r\}$, with $i \neq j$, $I \in \mathcal{P}_r$, and $s \in \{-, +\}$, the correlation between* $\mathrm{CME}(A_i \mid A_j s)$ *and the traditional effect corresponding to I is*

$$2^{-1/2} b^{-1}_{\mathcal{F},\phi} \left(b_{\mathcal{F},\{i\}\triangle I} + s b_{\mathcal{F},\{i,j\}\triangle I} \right).$$

Thus:

(a) If $\mathrm{ME}(A_i)$ *is aliased with the traditional effect corresponding to I in $\mathcal{F}$, then the correlations of* $\mathrm{CME}(A_i \mid A_j+)$ *and* $\mathrm{CME}(A_i \mid A_j-)$ *with the latter effect are both $2^{-1/2}$.*

(b) If $\mathrm{INT}(A_i, A_j)$ *is aliased with the traditional effect corresponding to I in $\mathcal{F}$, then the correlations of* $\mathrm{CME}(A_i \mid A_j+)$ *and* $\mathrm{CME}(A_i \mid A_j-)$ *with the latter effect are $2^{-1/2}$ and $-2^{-1/2}$, respectively.*

(c) If neither $\mathrm{ME}(A_i)$ *nor* $\mathrm{INT}(A_i, A_j)$ *are aliased with the traditional effect corresponding to I in $\mathcal{F}$, then the correlations of* $\mathrm{CME}(A_i \mid A_j+)$ *and* $\mathrm{CME}(A_i \mid A_j-)$ *with the latter effect are both zero.*

These results clearly demonstrate that the partial aliasing relations among the conditional and traditional effects in a design follow immediately from its indicator function coefficients.

Example 1. For our first, simple illustration of these results, consider the 2^{3-1}_{III} design $\mathcal{F}$ in Table 1 with indicator function $F_{\mathcal{F}}(x) = 1/2 + X_{\{1,2,3\}}(x)/2$. Suppose we wish to calculate the correlation between the siblings $\mathrm{CME}(A_1 \mid A_2+)$ and $\mathrm{CME}(A_1 \mid A_3-)$ in $\mathcal{F}$. From Corollary 1 and the indicator function coefficients, we immediately have that this correlation is $1/2$. Now, suppose we wish to calculate the correlations of $\mathrm{CME}(A_1 \mid A_2+)$ and $\mathrm{CME}(A_1 \mid A_2-)$ with $\mathrm{ME}(A_3)$ in $\mathcal{F}$. In this case, from Corollary 2(b), their respective correlations are $2^{-1/2}$ and $-2^{-1/2}$.

Example 2. To illustrate the utility of these results for larger designs of practical interest, let $\mathcal{F}_1$ denote the minimum aberration 2^{9-4}_{IV} design, and let $\mathcal{F}_2$ denote the 2^{9-4}_{IV} design that maximizes the number of clear two-factor interactions, as provided by Wu and Hamada (2009, p. 254). Following the notation of Wu and Hamada (2009, p. 215), the defining contrast subgroups of $\mathcal{F}_1$ and $\mathcal{F}_2$ are

Table 1. The 2^{3-1}_{III} design defined by $A_3 = A_1A_2$.

A_1	A_2	A_3
$-$	$-$	$+$
$-$	$+$	$-$
$+$	$-$	$-$
$+$	$+$	$+$

$$\{1236, 1247, 1258, 13459, 3467, 3568, 24569, 4578, 23579, 23489, 12345678, 15679, 14689, 13789, 26789\} \text{ and}$$

$$\{1236, 1247, 1348, 23459, 3467, 2468, 14569, 2378, 13579, 12589, 1678, 25679, 35689, 45789, 123456789\},$$

respectively, with the identity elements excluded from these subgroups, without loss of essential information. For this example, suppose we wish to conduct inferences on CME($A_1 \mid A_2+$) and CME($A_1 \mid A_2-$). Corollary 2 immediately yields that, in $\mathcal{F}_1$, these CMEs are correlated with INT(A_3, A_6), INT(A_4, A_7), and INT(A_5, A_8). However, in $\mathcal{F}_2$, they are correlated only with INT(A_3, A_6) and INT(A_4, A_7). The absolute magnitudes of these correlations are all equal to $2^{-1/2}$. Accordingly, we may choose design $\mathcal{F}_2$ over $\mathcal{F}_1$ to be able to obtain more conclusive inferences on these selected CMEs. Another immediate, and related, result is that $\mathcal{F}_2$ has fewer CMEs that are aliased with at least one main effect or two-factor interaction (excluding the corresponding parent main effect and two-factor interaction) than $\mathcal{F}_1$. Note that our orthogonal basis of functions enables us to derive these properties of $\mathcal{F}_1$ and $\mathcal{F}_2$ in a coordinate-free manner. As such, we can easily consider the conditional and traditional effects for large designs. We continue to explore the properties of $\mathcal{F}_1$ and $\mathcal{F}_2$ in later examples.

The properties of CMEs derived by Su and Wu (2017, pp. 4–6) follow as simple corollaries of our Propositions 2 and 3, with the formal proofs provided in the online Supplementary Material.

Corollary 3 (Su and Wu (2017)). *Twin CMEs are orthogonal.*

Corollary 4 (Property 2 of Su and Wu (2017)). *In regular designs, a CME is orthogonal to all traditional effects, except those fully aliased with its parent main effect or corresponding two-factor interaction.*

Corollary 5 (Property 3 of Su and Wu (2017)). *Sibling CMEs are correlated in regular designs of resolution at least* III.

Corollary 6 (Properties 4 and 5 of Su and Wu (2017))**.** *In regular designs of resolution at least* IV*, nontwin CMEs in a family are correlated, and CMEs with different parents and non-aliased corresponding two-factor interactions are orthogonal.*

Example 3. Consider designs $\mathcal{F}_1$ and $\mathcal{F}_2$ from Example 2. Recall that $\mathcal{F}_2$ has fewer CMEs that are aliased with at least one main effect or two-factor interaction (excluding the corresponding parent main effect and two-factor interaction) than $\mathcal{F}_1$. Furthermore, from Corollary 6 and the fact that $\mathcal{F}_1$ has less aberration than $\mathcal{F}_2$, $\mathcal{F}_1$ has fewer nontwin CME family members with different parent effects but with the same interaction effect than $\mathcal{F}_2$. In general, the makeup of CME families can be an important consideration when choosing between several candidate designs for a robust type of CME analysis. This is illustrated in the practical application discussed in Section 7, in which we consider four 2^{8-3}_{IV} designs that have different CME family compositions with respect to three distinct temperature factors.

Two additional properties related to uncle–nephew effect pairs and cousin CMEs follow immediately from Propositions 2 and 3.

Corollary 7. *Uncle–nephew effect pairs are orthogonal in regular designs of resolution at least* III*.*

Corollary 8. *Cousin CMEs are orthogonal in regular designs of resolution at least* IV*.*

These orthogonalities are useful when employing models with CMEs and their conditioned main effects.

5. Traditional Design Criteria and CMEs in Resolution IV Designs

In this section, we apply our derived partial aliasing relations to examine the implications of the maximum clear two-factor interactions and minimum aberration criteria for CME analyses of resolution IV regular designs. Our focus on such designs corresponds to an original motivation for the maximum clear two-factor interactions criterion, namely, comparing and rank-ordering regular designs that have the same number of clear main effects, but different numbers of clear two-factor interactions (Mukerjee and Wu (2006, p. 64)).

Definition 9. *A CME is clear in a design if it is orthogonal to all main effects, excluding its parent main effect, and two-factor interactions, excluding its* *corresponding two-factor interaction.*

Proposition 4. *For the class of 2_{IV}^{r-p} fractional factorials, a design has the maximum number of clear two-factor interactions if and only if it has the maximum number of clear CMEs.*

Corollary 9. *A fractional factorial with the maximum number of clear two-factor interactions among 2_{IV}^{r-p} designs minimizes the total number of CMEs across families containing more than four members for the class of 2_{IV}^{r-p} designs.*

Example 4. We illustrate these implications of the maximum clear two-factor interactions criterion using the 2_{IV}^{7-2} designs $\mathcal{F}_3$ and $\mathcal{F}_4$, with respective defining contrast subgroups {1236, 12457, 34567} and {1236, 3457, 124567}. The identity elements in each are excluded, without loss of essential information. From Wu and Hamada (2009, p. 254), $\mathcal{F}_3$ has the maximum number of clear two-factor interactions among the 2_{IV}^{7-2} designs. Thus, from Proposition 4, $\mathcal{F}_3$ has more clear CMEs than $\mathcal{F}_4$. Furthermore, $\mathcal{F}_3$ has three families with more than four members in each, because it has three aliasing relations containing more than one two-factor interaction. Similarly, $\mathcal{F}_4$ has six families with more than four members in each, because it has six aliasing relations containing more than one two-factor interaction. Following the notation of Wu and Hamada (2009, p. 215), the aliasing relations for $\mathcal{F}_3$ are

$$12 = 36 = 457 = 1234567,$$

$$13 = 26 = 23457 = 14567,$$

$$16 = 23 = 24567 = 13457,$$

and the aliasing relations for $\mathcal{F}_4$ are

$$12 = 36 = 123457 = 4567,$$

$$13 = 26 = 1457 = 234567,$$

$$16 = 23 = 134567 = 2457,$$

$$34 = 1246 = 57 = 123567,$$

$$35 = 1256 = 47 = 123467,$$

$$37 = 1267 = 45 = 123456.$$

Note that each of these families has exactly eight members. Hence, in comparison with $\mathcal{F}_4$, $\mathcal{F}_3$ has a smaller total number of CMEs across its families that contain more than four members in each, which follows from Corollary 9.

To present the implications of the minimum aberration criterion for CME analyses, we introduce notation to denote the number of distinct factor pairs among the CMEs in a design's family.

Definition 10. *For a design $\mathcal{F} \subseteq \mathcal{D}_r$, with $T_{\mathcal{F}}$ families, let $N_t(\mathcal{F})$ denote the number of distinct factor pairs among the CMEs in its family t, for $t = 1, \ldots, T_{\mathcal{F}}$.*

Example 5. The distinct factor pairs in the three families of $\mathcal{F}_3$ that have more than four members are $\{(A_1, A_2), (A_3, A_6)\}$, $\{(A_1, A_3), (A_2, A_6)\}$, and $\{(A_1, A_6), (A_2, A_3)\}$. Thus, $N_t(\mathcal{F}_3) = 2$ for all of these families, for $t = 1, 2, 3$. In addition, $N_t(\mathcal{F}_4) = 2$ for all families $t = 1, \ldots, 6$ of $\mathcal{F}_4$ that have more than four members.

Cheng, Steinberg and Sun (1999) and Cheng (2014, p. 172) provide an expression for a regular design's count of defining words of length four, in terms of the numbers of two-factor interactions in its aliasing sets. Using Definition 10, Lemma 2 reformulates this into a corresponding expression for minimum aberration designs under the CME system. We then combine it with Corollary 1 to show in Proposition 5 how minimum aberration designs minimize aggregate measures of correlations among CMEs.

Lemma 2. *For the class of 2_{IV}^{r-p} fractional factorials, a design $\mathcal{F}^*$ has the minimum aberration if and only if*

$$\sum_{t=1}^{T_{\mathcal{F}^*}} N_t(\mathcal{F}^*)\{N_t(\mathcal{F}^*) - 1\} \leq \sum_{t=1}^{T_{\mathcal{F}}} N_t(\mathcal{F})\{N_t(\mathcal{F}) - 1\},$$

for all 2_{IV}^{r-p} designs $\mathcal{F}$.

Proposition 5. *A fractional factorial with minimum aberration among 2_{IV}^{r-p} designs minimizes, for each exhaustive selection of CMEs such that no two involve the same pair of factors, both the sum of the absolute correlations and the sum of the squared correlations between nonsibling effects for the class of 2_{IV}^{r-p} designs.*

Example 6. For each exhaustive selection of CMEs in $\mathcal{F}_3$, such that no two involve the same pair of factors, the sum of the absolute correlations and the sum of the squared correlations between nonsibling effects are 1.5 and 0.75, respectively. The corresponding sums for $\mathcal{F}_4$ are 3 and 1.5. The inequalities $1.5 < 3$ and $0.75 < 1.5$ in these two respective sums correspond to Proposition 5 and the fact that $\mathcal{F}_3$ is the minimum aberration 2_{IV}^{7-2} design.

From Proposition 4 and Corollary 9, a resolution IV design with the maximum number of clear two-factor interactions among its peer class of 2_{IV}^{r-p} designs could be useful when employing models composed of main effects, nonsibling CMEs, and two-factor interactions. Proposition 5 demonstrates that the minimum aberration 2_{IV}^{r-p} design could be useful when it is desired to conduct an experiment with minimum aggregate correlations among distinct types of CMEs. These implications facilitate immediate comparisons of large designs under the CME system for practical applications.

Example 7. We illustrate the immediate applicability of this section's results for the larger designs $\mathcal{F}_1$ and $\mathcal{F}_2$ from Example 2. By inspection, $\mathcal{F}_1$ has 13 families that contain more than four members in each. One family has 16 CMEs, and the remainder have eight CMEs. In addition, $\mathcal{F}_2$ has seven families that contain more than four members, each with 12 CMEs. The total number of CMEs across the above families of $\mathcal{F}_1$ is 112, whereas that of $\mathcal{F}_2$ is 84. The smaller number for $\mathcal{F}_2$ corresponds to Corollary 9 and the fact that $\mathcal{F}_2$ has the maximum number of clear two-factor interactions among the 2_{IV}^{9-4} designs. Now, for each exhaustive selection of CMEs in $\mathcal{F}_1$, such that no two involve the same pair of factors, the sum of the absolute correlations and the sum of the squared correlations among nonsibling effects are 9 and 4.5, respectively. The corresponding sums for $\mathcal{F}_2$ are 10.5 and 5.25. The inequalities $9 < 10.5$ and $4.5 < 5.25$ in the two sums corresponds to Proposition 5 and the fact that $\mathcal{F}_1$ is the minimum aberration 2_{IV}^{9-4} design. In addition to demonstrating the applicability of our results for large designs, this example also illustrates that, as in the case of the orthogonal components system, the maximum clear two-factor interactions and minimum aberration criteria may disagree on the choice of design for a CME analysis.

6. D-Efficiency Under the CME System

Our algebra reduces D-efficiency calculations for general classes of designs and models under the CME system. We demonstrate this result for resolution III and IV regular designs and models consisting of multiple main effects, two-factor interactions, and CMEs. Negligible additions of notation will be introduced when extending these calculations to other designs and models.

We first describe the assumptions and notation utilized in this section. We assume that the factors for a regular design $\mathcal{F} \subseteq \mathcal{D}_r$ are partitioned into two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, with a selection of conditional and traditional effects involving the factors in $\mathcal{S}_1$ being of interest, and a selection of only traditional effects involving the factors in $\mathcal{S}_2$ being of interest. For $i \in \{1,2\}$, we let $\mathcal{S}_i^{\text{Trad}}$ denote the set of functions $Z_I \equiv 2|\mathcal{F}|^{-1}X_I$ that correspond to the selected traditional effects involving the factors in $\mathcal{S}_i$. Similarly, we let $\mathcal{S}_1^{\text{CME}}$ denote the set of functions $Z_{i|j}^s \equiv 2^2(s2^r b_{\mathcal{F},\{j\}} + |\mathcal{F}|)^{-1}X_{i|j}^s$, where $2^{-1}(s2^r b_{\mathcal{F},\{j\}} + |\mathcal{F}|)$ is the number of runs in $\mathcal{F}$ in which A_j is at level $s \in \{-,+\}$, that correspond to the selected CMEs involving the factors in $\mathcal{S}_1$. We specify the model matrix M for this selection of effects in $\mathcal{F}$ as follows:

$$M = \left(\mathbf{1}_{|\mathcal{F}|}\ S_1^{\text{Trad}}\ S_2^{\text{Trad}}\ S_1^{\text{CME}}\right), \tag{6.1}$$

where $\mathbf{1}_{|\mathcal{F}|}$ is an $|\mathcal{F}|\times 1$ vector, with all entries equal to one; S_i^{Trad} is an $|\mathcal{F}|\times|\mathcal{S}_i^{\text{Trad}}|$ matrix, the columns of which are the contrast vectors for the effects in $\mathcal{S}_i^{\text{Trad}}$, for $i \in \{1,2\}$; and S_1^{CME} is an $|\mathcal{F}| \times |\mathcal{S}_1^{\text{CME}}|$ matrix, the columns of which are the contrast vectors for the CMEs in $\mathcal{S}_1^{\text{CME}}$. We let $q = 1 + |\mathcal{S}_1^{\text{Trad}}| + |\mathcal{S}_2^{\text{Trad}}| + |\mathcal{S}_1^{\text{CME}}|$ denote the number of columns in M.

Example 8. To illustrate the above notation, consider the 2_{III}^{3-1} design $\mathcal{F}$ from Example 1. Suppose $\mathcal{S}_1 = \{A_1, A_2\}$ and $\mathcal{S}_2 = \{A_3\}$, with $\mathcal{S}_1^{\text{Trad}} = \{Z_{\{2\}}\}$, $\mathcal{S}_1^{\text{CME}} = \{Z_{1|2}^+\}$, and $\mathcal{S}_2^{\text{Trad}} = \{Z_{\{3\}}\}$. Then, the model matrix, as specified in equation (6.1), for this design and model is

$$M = \begin{pmatrix} 1 & -0.5 & 0.5 & 0 \\ 1 & 0.5 & -0.5 & -1 \\ 1 & -0.5 & -0.5 & 0 \\ 1 & 0.5 & 0.5 & 1 \end{pmatrix}.$$

Definition 11 (Montgomery (2013)). *Let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{D}_r$, with $|\mathcal{F}_1| = |\mathcal{F}_2|$, and suppose that the same sets of effects $\mathcal{S}_1^{\text{Trad}}$, $\mathcal{S}_2^{\text{Trad}}$, and $\mathcal{S}_1^{\text{CME}}$ are of interest for estimation under them. Let M_i denote the model matrix under $\mathcal{F}_i$ for $i \in \{1,2\}$. The relative D-efficiency of $\mathcal{F}_1$ to $\mathcal{F}_2$ is*

$$\left\{\frac{\det(M_1^{\mathsf{T}}M_1)}{\det(M_2^{\mathsf{T}}M_2)}\right\}^{1/q}.$$

From Definition 11, the D-efficiency calculation for a model matrix M revolves around $\det(M^{\mathsf{T}}M)$. The following lemma formally presents the derivation of $\det(M^{\mathsf{T}}M)$ under our algebra.

Lemma 3. *Consider model matrix M in equation (6.1). For $c, d \in \{1,\ldots,q\}$, let $Z_{(c)}$ and $Z_{(d)}$ denote the functions in $\{X_\phi\} \cup \mathcal{S}_1^{\text{Trad}} \cup \mathcal{S}_2^{\text{Trad}} \cup \mathcal{S}_1^{\text{CME}}$ that correspond to columns c and d, respectively, of M, with $Z_{(1)} = X_\phi$. Then, entry (c,d) of $M^{\mathsf{T}}M$ is $2^r\left\langle Z_{(c)}, Z_{(d)} \mid \mathcal{F}\right\rangle$.*

Using this lemma and our partial aliasing relations under the CME system, the entries of $M^{\mathsf{T}}M$ for a model matrix M containing both conditional and traditional effects can be described in a simple and general manner using indicator function coefficients for the fraction $\mathcal{F}$. We proceed to formally reduce D-efficiency calculations in this manner for resolution III and IV regular fractions, and for models in which $\mathcal{S}_1^{\text{Trad}}$ consists of the main effects for all factors in $\mathcal{S}_1$, and $\mathcal{S}_2^{\text{Trad}}$ consists of all main effects and a selection of non-aliased two-factor interactions involving the factors in $\mathcal{S}_2$.

Proposition 6. *Consider a 2^{r-p}_{III} design $\mathcal{F} \subseteq \mathcal{D}_r$, and let its model matrix M be structured as*

$$M = \left(\mathbf{1}_{|\mathcal{F}|}\ S_1^{\text{Trad}}\ S_2^{\text{ME}}\ S_2^{\text{INT}}\ S_1^{\text{CME}}\right),$$

where the columns of matrix S_2^{ME} are the main effect contrast vectors in S_2^{Trad}, and the columns of matrix S_2^{INT} are the two-factor interaction contrast vectors in S_2^{Trad}. Let n_1 and n_2 denote the number of columns in S_1^{CME} and S_2^{INT}, respectively. Then, $M^{\mathsf{T}}M$ is of the form

$$M^{\mathsf{T}}M = \begin{pmatrix} D_1 & C_1 & C_2 \\ C_1^{\mathsf{T}} & D_2 & C_3 \\ C_2^{\mathsf{T}} & C_3^{\mathsf{T}} & W \end{pmatrix}, \tag{6.2}$$

where

- *D_1 is a $(1+|\mathcal{S}_1|+|\mathcal{S}_2|) \times (1+|\mathcal{S}_1|+|\mathcal{S}_2|)$ diagonal matrix, with entry $(1,1)$ equal to $|\mathcal{F}|$, and entry (c,c), for $c \in \{2,\ldots,(1+|\mathcal{S}_1|+|\mathcal{S}_2|)\}$, equal to 2^{p-r+2},*

- *D_2 is an $n_2 \times n_2$ diagonal matrix, with entry (c,c), for $c \in \{1,\ldots,n_2\}$, equal to 2^{p-r+2},*

- *W is an $n_1 \times n_1$ matrix, with entry (c,d), for $c,d \in \{1,\ldots,n_1\}$ and in which $Z^s_{i|j}$ and $Z^{s'}_{l|k}$ correspond to the contrast vectors in columns c and d, respectively, of S_1^{CME}, equal to $2^{p-r+2}b^{-1}_{\mathcal{F},\phi}(b_{\mathcal{F},\{i\}\triangle\{l\}} + s'b_{\mathcal{F},\{i\}\triangle\{l,k\}} + sb_{\mathcal{F},\{i,j\}\triangle\{l\}} + ss'b_{\mathcal{F},\{i,j\}\triangle\{l,k\}})$,*

- *C_1 is an $(1+|\mathcal{S}_1|+|\mathcal{S}_2|) \times n_2$ matrix, with entry $(1,d)$ equal to zero for all $d \in \{1,\ldots,n_2\}$, and entry (c,d), for $c \in \{2,\ldots,(1+|\mathcal{S}_1|+|\mathcal{S}_2|)\}$, $d \in \{1,\ldots,n_2\}$, and in which Z_I and Z_J correspond to the contrast vectors in column c of $(\mathbf{1}_{|\mathcal{F}|}\ S_1^{\text{Trad}}\ S_2^{\text{ME}})$ and column d of S_2^{INT}, respectively, equal to $2^{p-r+2}b^{-1}_{\mathcal{F},\phi}b_{\mathcal{F},I\triangle J}$,*

- *C_2 is an $(1+|\mathcal{S}_1|+|\mathcal{S}_2|)\times n_1$ matrix, with entry $(1,d)$ equal to zero for all $d\in\{1,\dots,n_1\}$, and with entry (c,d), for $c\in\{2,\dots,(1+|\mathcal{S}_1|+|\mathcal{S}_2|)\}$, $d\in\{1,\dots,n_1\}$, and in which Z_I and $Z^s_{i|j}$ correspond to the contrast vectors in column c of $(\mathbf{1}_{|\mathcal{F}|}\ S_1^{\mathrm{Trad}}\ S_2^{\mathrm{ME}})$ and column d of S_1^{CME}, respectively, equal to $2^{p-r+2}b^{-1}_{\mathcal{F},\phi}(b_{\mathcal{F},\{i\}\triangle I}+sb_{\mathcal{F},\{i,j\}\triangle I})$, and*

- *C_3 is an $n_2\times n_1$ matrix, with entry (c,d), for $c\in\{1,\dots,n_2\}$, $d\in\{1,\dots,n_1\}$, and in which Z_I and $Z^s_{i|j}$ correspond to the contrast vectors in column c of S_2^{INT} and column d of S_1^{CME}, respectively, equal to $2^{p-r+2}b^{-1}_{\mathcal{F},\phi}(b_{\mathcal{F},\{i\}\triangle I}+sb_{\mathcal{F},\{i,j\}\triangle I})$.*

In addition,

$$\det\left(M^{\mathsf{T}}M\right)=|\mathcal{F}|2^{(p-r+2)(|\mathcal{S}_1|+|\mathcal{S}_2|)}\det\left\{\begin{pmatrix}D_2 & C_3\\ C_3^{\mathsf{T}} & W\end{pmatrix}-\begin{pmatrix}C_1^{\mathsf{T}}\\ C_2^{\mathsf{T}}\end{pmatrix}D_1^{-1}\begin{pmatrix}C_1 & C_2\end{pmatrix}\right\}.$$

Corollary 10. *Consider a 2^{r-p}_{IV} design $\mathcal{F}\subseteq\mathcal{D}_r$, and let its model matrix M be structured as in Proposition 6. Then, the entries of matrix C_1 in equation (6.2) are all equal to zero. Furthermore,*

$$\det\left(M^{\mathsf{T}}M\right)$$
$$=|\mathcal{F}|2^{(p-r+2)(|\mathcal{S}_1|+|\mathcal{S}_2|+n_2)}\det\left\{W-\left(C_2^{\mathsf{T}}\ C_3^{\mathsf{T}}\right)\begin{pmatrix}D_1 & \mathbf{0}\\ \mathbf{0} & D_2\end{pmatrix}^{-1}\begin{pmatrix}C_2\\ C_3\end{pmatrix}\right\}.$$

In Proposition 6, the entries of C_1 correspond to correlations between main effects and two-factor interactions, the entries of C_2 correspond to correlations between main effects and CMEs, and the entries of C_3 correspond to correlations between two-factor interactions and CMEs. The off-diagonal entries of W correspond to correlations between CMEs. Proposition 6 and Corollary 10 reduce D-efficiency calculations to the determinant of a $(n_1+n_2)\times(n_1+n_2)$ matrix for resolution III designs, and to the determinant of a $n_1\times n_1$ matrix for resolution IV designs. They also facilitate immediate characterizations of the D-efficiencies for several candidate designs under broad classes of models that involve different selections of main effects, two-factor interactions, and CMEs. These features of our results are illustrated in the case study in Section 7.

Example 9. For designs $\mathcal{F}_1$ and $\mathcal{F}_2$ from Example 2, let $\mathcal{S}_1=\{A_1, A_2, A_3, A_4, A_5\}$, $\mathcal{S}_1^{\mathrm{CME}}=\{Z^+_{1|4}, Z^-_{1|5}, Z^+_{2|3}, Z^-_{2|4}\}$, and $\mathcal{S}_2^{\mathrm{Trad}}$ contain the functions Z_I that correspond to the main effects and two-factor interactions from $\mathcal{S}_2=\{A_6, A_7, A_8, A_9\}$. From Corollary 10, the D-efficiencies of these designs with respect to this model, which contains a large selection of conditional and traditional effects, are reduced to determinants of simple 4×4 matrices. For design $\mathcal{F}_1$,

$$W - \begin{pmatrix} C_2^{\mathsf{T}} & C_3^{\mathsf{T}} \end{pmatrix} \begin{pmatrix} D_1 & \mathbf{0} \\ \mathbf{0} & D_2 \end{pmatrix}^{-1} \begin{pmatrix} C_2 \\ C_3 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and thus $\det\left(M_1^{\mathsf{T}} M_1\right)$ is nonzero. For design $\mathcal{F}_2$,

$$W - \begin{pmatrix} C_2^{\mathsf{T}} & C_3^{\mathsf{T}} \end{pmatrix} \begin{pmatrix} D_1 & \mathbf{0} \\ \mathbf{0} & D_2 \end{pmatrix}^{-1} \begin{pmatrix} C_2 \\ C_3 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and thus $\det\left(M_2^{\mathsf{T}} M_2\right)$ is zero, which means this model is not estimable by $\mathcal{F}_2$.

7. Practical Application of a CME Analysis

We demonstrate the utility of our algebra for real-world CME analyses by considering the painted panel experiment of Lorenzen and Anderson (1993, pp. 242–249). The data for this experiment are provided in the Supplementary Material. This case study illustrates how our results can reveal the possible scope of analyses, and the broad equivalencies and subtle differences, for several candidate designs under the CME system.

The experimenters' objective was to study the effects of the factors in Table 2 on painted panel film build. The effects thought *a priori* to be active were all of the main effects, $\mathrm{INT}(A_7, A_8)$, and $\mathrm{INT}(A_1, A_5)$, with higher-order interactions assumed to be inert. They selected a 2^{8-3}_{IV} design that had $\mathrm{INT}(A_7, A_8)$ and $\mathrm{INT}(A_1, A_5)$ clear, with defining contrast subgroup $\{3456, 12457, 2358, 12367, 2468, 13478, 15678\}$. Three other such designs exist, with defining contrast subgroups {1236, 1247, 13458, 3467, 24568, 23578, 15678}, {2467, 2357, 15678, 3456, 12458, 12368, 13478}, and {3468, 1248, 23578, 1236, 24567, 13457, 15678}, respectively (Wu and Hamada (2009, p. 254)). These four designs are denoted by $\mathcal{F}_1^{\mathrm{LA}}$, $\mathcal{F}_2^{\mathrm{LA}}$, $\mathcal{F}_3^{\mathrm{LA}}$, and $\mathcal{F}_4^{\mathrm{LA}}$, respectively. We use our algebra to evaluate their properties under the CME system. This evaluation is important in practice because additional interactions are typically active, but fully aliased in candidate designs. Thus, CMEs should be considered to obtain interpretable and conclusive inferences, and the designs' properties under the CME system should be assessed to better inform the final design selection. Here, our algebra explains how the

Table 2. The factors and their levels in the experiment of Lorenzen and Anderson (1993, pp. 242, 246).

Booth Humidity A_1	Substrate Temperature A_2	Fluid Flow Rate A_3	Target Distance A_4	Booth Temperature A_5	Base Coat Temperature A_6	Atomizing Air Pressure A_7	Fan Air Pressure A_8
70 (+)	100 (+)	20 (+)	15 (+)	90 (+)	85 (+)	50 (+)	50 (+)
50 (−)	70 (−)	0 (−)	12 (−)	70 (−)	65 (−)	40 (−)	40 (−)

chosen design $\mathcal{F}_1^{\mathrm{LA}}$ can yield inferences on CMEs corresponding to potentially active interactions involving the distinct temperature factors (A_2, A_5, A_6) that are more ambiguous than those of $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$.

Our results in Section 4 enable immediate comparisons of the partial aliasing relations for the CMEs in the four designs. Consider the CMEs involving the temperature factors. Proposition 2 and Corollary 1 yield, in a simple manner, all correlated CMEs that involve these factors for any of the designs, and that their absolute correlations are always 1/2, with the sign equal to the product of their conditioned levels. In $\mathcal{F}_1^{\mathrm{LA}}$ and $\mathcal{F}_3^{\mathrm{LA}}$, triples of CMEs involving A_2, A_5, and A_6 exist that are correlated. Examples in $\mathcal{F}_1^{\mathrm{LA}}$ are $\mathrm{CME}(A_2 \mid A_8 s_8)$, $\mathrm{CME}(A_5 \mid A_3 s_3)$, and $\mathrm{CME}(A_6 \mid A_4 s_4)$, and examples in $\mathcal{F}_3^{\mathrm{LA}}$ are $\mathrm{CME}(A_2 \mid A_7 s_7)$, $\mathrm{CME}(A_5 \mid A_3 s_3)$, and $\mathrm{CME}(A_6 \mid A_4 s_4)$, for $s_3, s_4, s_7, s_8 \in \{-,+\}$. In contrast, for $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$, only pairs of CMEs involving these factors exist that are correlated. Examples in $\mathcal{F}_2^{\mathrm{LA}}$ are $\mathrm{CME}(A_2 \mid A_1 s_1)$ and $\mathrm{CME}(A_6 \mid A_3 s_3)$, and examples in $\mathcal{F}_4^{\mathrm{LA}}$ are $\mathrm{CME}(A_2 \mid A_3 s_3)$ and $\mathrm{CME}(A_6 \mid A_1 s_1)$, for $s_1, s_3 \in \{-,+\}$. A practical consequence of this difference in partial aliasing relations is that $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$ can yield CME analyses that are more conclusive than those of $\mathcal{F}_1^{\mathrm{LA}}$ and $\mathcal{F}_3^{\mathrm{LA}}$ when more than one of the temperature factors have active two-factor interactions.

The combination of our results in Section 5 with the previously identified partial aliasing relations reveal several properties of the CME families in these designs. First, from Corollary 9 and the fact that each design has the maximum number of clear two-factor interactions among 2_{IV}^{8-3} designs, the designs all have the same (and minimum) number of CMEs across their nontrivial families. In fact, each design has one nontrivial family that contains 12 members, and six nontrivial families that each contain eight members. Each design also has 13 trivial families (with four members in each). However, $\mathcal{F}_1^{\mathrm{LA}}$ and $\mathcal{F}_3^{\mathrm{LA}}$ differ from $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$ in the composition of the families involving temperature factors. Specifically, $\mathcal{F}_1^{\mathrm{LA}}$ and $\mathcal{F}_3^{\mathrm{LA}}$ have CMEs involving A_2, A_5, and A_6 in their respective families of size 12, whereas $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$ have CMEs involving A_2

and A_6 just in their families of size eight, and CMEs involving A_5 just in their families of size four. Second, from Proposition 5 and the fact that each design has minimum aberration among 2^{8-3}_{IV} designs, for any exhaustive selection of CMEs, such that no two involve the same pair of factors, the sums of their respective absolute correlations and squared correlations among nonsibling CMEs are equal (and the minimum possible respective values). These sums are $9/2$ and $9/4$, respectively. However, $\mathcal{F}_1^{\mathrm{LA}}$ and $\mathcal{F}_3^{\mathrm{LA}}$ again differ from $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$ in that, for any such selection, the former two have larger sums of absolute correlations and squared correlations among nonsibling CMEs that involve the temperature factors. The sums for $\mathcal{F}_1^{\mathrm{LA}}$ and $\mathcal{F}_3^{\mathrm{LA}}$ are 3 and $3/2$, respectively, whereas the sums for $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$ are 1 and $1/2$, respectively. These results demonstrate that, although the designs are broadly equivalent in terms of their CME family structures and aggregate correlations among nonsibling CMEs, they also have subtle differences related to the CMEs that involve the temperature factors due to their distinct partial aliasing relations, which are easily derived from our algebra. These differences again play a role in the CME analysis for $\mathcal{F}_1^{\mathrm{LA}}$ in terms of its robustness to a temperature factor other than A_5 having an active two-factor interaction, and the degree to which conclusions can be drawn on the CMEs of the other temperature factors.

The results in Section 6 facilitate our understanding of the designs' D-efficiencies for models involving CMEs, main effects, and the previously identified $\mathrm{INT}(A_7, A_8)$ and $\mathrm{INT}(A_1, A_5)$. Let $\mathcal{S}_1 = \{A_2, A_3, A_4, A_6\}$, $\mathcal{S}_2 = \{A_1, A_5, A_7, A_8\}$, $\mathcal{S}_1^{\mathrm{Trad}} = \{Z_{\{2\}}, Z_{\{3\}}, Z_{\{4\}}, Z_{\{6\}}\}$, and $\mathcal{S}_2^{\mathrm{Trad}} = \{Z_{\{1\}}, Z_{\{5\}}, Z_{\{7\}}, Z_{\{8\}}, Z_{\{1,5\}}, Z_{\{7,8\}}\}$. Then, for any of the designs and choice of $\mathcal{S}_1^{\mathrm{CME}}$, matrix C_3 in Proposition 6 has all of its entries equal to zero. Thus, $\det\left(M^{\mathsf{T}}M\right) = 2^{-25}\det\{W - C_2^{\mathsf{T}} D_1^{-1} C_2\}$ for the model matrix M, by Corollary 10. This expression can be readily evaluated to characterize these designs' D-efficiencies for broad classes of models. For the first example, suppose $\mathcal{S}_1^{\mathrm{CME}} = \{Z^{s}_{2|3}, Z^{s'}_{2|4}, Z^{s''}_{6|3}, Z^{s'''}_{6|4}\}$, for $s, s', s'', s''' \in \{-, +\}$, which corresponds to CMEs that involve the two temperature factors other than A_5. The D-efficiencies are equal for all of the designs and choices of the conditioned levels, thus reducing to the single calculation

$$\det\left(M^{\mathsf{T}}M\right) = 2^{-25}\det\left\{\frac{1}{8}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\right\}.$$

For the second example, consider $\mathcal{S}_1^{\mathrm{CME}} = \{Z^{s}_{i|j}, Z^{s'}_{i|k}, Z^{s''}_{i|l}\}$, for distinct A_i, A_j,

A_k, $A_l \in \mathcal{S}_1$ and $s, s', s'' \in \{-, +\}$. The designs' D-efficiencies are again equal, and immediately reduce to the determinants of 3×3 matrices with the same structure for each such selection of CMEs. To illustrate, if $i = 2$, $j = 3$, $k = 4$, and $l = 6$, then $\det\left(M^{\mathsf{T}} M\right) = 2^{-34}$ for any of the designs and conditioned levels. The ease with which these broad D-efficiency characterizations were obtained further highlights the significance of our algebra for practical applications.

We now perform a CME analysis of the chosen design $\mathcal{F}_1^{\mathrm{LA}}$. The experimenters conducted an ANOVA test and concluded that the following were active: $\mathrm{ME}(A_1)$, $\mathrm{ME}(A_2)$, $\mathrm{ME}(A_3)$, $\mathrm{ME}(A_4)$, $\mathrm{ME}(A_5)$, $\mathrm{ME}(A_8)$, $\mathrm{INT}(A_4, A_7)$, and one or more of $\mathrm{INT}(A_2, A_8)$, $\mathrm{INT}(A_3, A_5)$, and $\mathrm{INT}(A_4, A_6)$ (Lorenzen and Anderson (1993, pp. 246–248)). More conclusive inferences on the last set of two-factor interactions cannot be obtained from the traditional analysis because they are all fully aliased in the chosen design. For the corresponding CMEs, we have from our previous results that $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$ are the preferable designs, because they can more easily resolve the ambiguity related to which temperature factors have significant effects beyond the main effects. To complete the CME analysis of this experiment, we use the three rules in the method of Su and Wu (2017, pp. 5–6), concluding that $\mathrm{CME}(A_2 \mid A_8-)$ is significant. Details of the analysis are provided in the Supplementary Material. Thus, we obtain interpretable, final conclusions from the CME analysis that substrate temperature affects film build at low air pressure, but not high air pressure, and that, contrary to the experimenters' prior knowledge, booth temperature has no significant effects beyond its main effect.

8. Conclusion

As recognized by Wu (2018), an important theme for modern experimental design is the consideration of parameterizations for factorial effects that better address real-life problems than more traditional systems. This study underscores that theme. We developed an accessible algebra for the CME system that facilitates the derivation of general results and properties for broad types of two-level designs and models consisting of multiple conditional and traditional effects. The framework for our algebra is based on indicator functions. Our work is distinct from previous studies on indicator functions, such as those of Fontana, Pistone and Rogantin (2000), Ye (2003), and Ye (2004), because they only consider applying indicator functions to derive design properties under traditional effects, whereas we consider their applications using our inner product in Definition 8, under both conditional and traditional effects. Our study of partial aliasing relations, design criteria, and D-efficiency calculations using our algebra conclusively demonstrates its advantages. It enables an unrestricted approach to understanding two-level designs under the CME system, with no limits on the designs or on their effects. The algebra provides concise, simple calculations of design characteristics, based on a small selection of indicator function coefficients. These points are further supported by our case study, which highlights both the usefulness of our algebra and a key advantage of CMEs as interpretable effects in many applications. A more general lesson of the case study is that our algebra enables easier comparisons of several large candidate designs under the CME system, and thus choices that are better informed.

Supplementary Material

The online Supplementary Material contains detailed proofs for the results in Sections 4, 5, and 6, and extended data analyses for the painted panel experiment in Section 7.

Acknowledgments

We thank Chien-Fu Jeff Wu and two reviewers for their insightful comments.

References

Cheng, C.-S. (2014). *Theory of Factorial Design: Single- and Multi-Stratum Experiments*, 1st Edition. Monographs on Statistics and Applied Probability. CRC Press.

Cheng, C.-S., Steinberg, D. M. and Sun, D. (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *Journal of the Royal Statistical Society. Series B. (Statistical Methodology)* **61**, 85–93.

Finney, D. J. (1945). The fractional replication of factorial experiments. *Ann. Eugenics* **12**, 291–301.

Fontana, R., Pistone, G. and Rogantin, M. P. (2000). Classification of two-level factorial fractions. *Journal of Statistical Planning and Inference* **87**, 149–172.

Fries, A. and Hunter, W. G. (1980). Minimum aberration 2^{k-p} designs. *Technometrics* **2**, 601–608.

Hamada, M. S. and Wu, C. F. J. (1992). Analysis of designed experiments with complex aliasing. *Journal of Quality Technology* **24**, 130–137.

Lorenzen, T. J. and Anderson, V. L. (1993). *Design of Experiments: A No-Name Approach*, 1st Edition. Statistics: Textbooks and Monographs. CRC Press.

Mak, S. and Wu, C. F. J. (2019). `cmenet`: A new method for bi-level variable selection of conditional main effects. *Journal of the American Statistical Association* **114**, 844–856.

Montgomery, D. C. (2013). *Design and Analysis of Experiments.* John Wiley & Sons, Inc.

Mukerjee, R. and Wu, C. F. J. (2006). *A Modern Theory of Factorial Designs*, 1st Edition. Springer.

Mukerjee, R., Wu, C.-F. J. and Chang, M.-C. (2017). Two-level minimum aberration designs under a conditional model with a pair of conditional and conditioning factors. *Statistica Sinica* **27**, 997–1016.

Pistone, G. and Wynn, H. P. (1996). Generalised confounding with Gröbner bases. *Biometrika* **83**, 653–666.

Su, H. and Wu, C. F. J. (2017). CME Analysis: a new method for unraveling aliased effects in two-level fractional factorial experiments. *Journal of Quality Technology* **49**, 1–10.

Wu, C.-F. J. (2015). Post-Fisherian experimentation: from physical to virtual. *Journal of the American Statistical Association* **110**, 612–620.

Wu, C.-F. J. (2018). A fresh look at effect aliasing and interactions: some new wine in old bottles. *Annals of the Institute of Statistical Mathematics* **70**, 249–268.

Wu, C.-F. J. and Hamada, M. S. (2009). *Experiments: Planning, Analysis, and Optimization*, 2nd Edition. Wiley Series in Probability and Statistics. Wiley.

Ye, K. Q. (2003). Indicator function and its application in two-level factorial designs. *The Annals of Statistics* **31**, 984–994.

Ye, K. Q. (2004). A note on regular fractional factorial designs. *Statistica Sinica* **14**, 1069–1074.

Department of Statistics, Purdue University, West Lafayette, IN 47907, USA.

E-mail: sabbaghi@purdue.edu

(Received January 2018; accepted June 2018)