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 of 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. We establish a comprehensive algebra for the CME system based on indicator functions. Our algebra facilitates the derivations of general partial aliasing relations for a wide variety of two-level designs. By means of our algebra, we illuminate the implications of traditional design criteria under the CME system for resolution IV designs. A novel feature of our algebra is that it enables immediate and simple D-efficiency calculations for two-level regular designs and 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 inference on main effects and interactions in experiments with many factors and run size constraints. Their traditional method of analysis is based on the orthogonal components parameterization of factorial effects (Wu and Hamada, 2009, p. 274). From the time of Finney (1945), the major disadvantage of regular designs was thought to be that, under this traditional system, any two aliased effects could not be disentangled without follow-up runs (Wu, 2015). Su and Wu (2017) recently resolved this long-standing dilemma by employing a reparameterization of the traditional main effects and two-factor interactions into main effects and conditional main effects (CMEs). In contrast to the orthogonal components system, the analysis of regular designs under the CME system parallels the analysis of nonregular designs due to the existence of partial aliasing among 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 for follow-up runs to perform conclusive analyses on regular designs. Wu (2015, p. 615) first noted this saving grace 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 can yield unambiguous inferences in two-level regular fractional factorials.

The innovation of the CME system created several novel avenues of study for both experiments and observational studies. Mukerjee, Wu, and Chang (2017) introduced a new effect hierarchy for this system, and developed a design strategy with a minimum aberration criteria to sequentially minimize the bias in estimation of main effects by successive iterations in the effect hierarchy. Mak and Wu (2018) 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 is provided by Wu (2018).

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 for the CME system impedes the development of general results on CMEs that can be broadly applied across the vast 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. Also, 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 whose span 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 whose properties under the CME system are to be studied. The contributions of our algebra are three-fold. First, as opposed 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, p. 3–6) follow as simple calculations under our inner product, with no or rather weak conditions. Second, our algebra illuminates the implications of the maximum clear two-factor interactions (Wu and Hamada, 2009, p. 217) and minimum aberration (Fries and Hunter, 1980) design criteria for CME analysis of resolution IV designs. 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. Our study of the implications of traditional design criteria for the CME analysis of resolution IV designs is in Section 5. The application of our algebra for D-efficiency calculation is described in Section 6. Illustrative examples of our results are provided throughout the latter three sections, and their proofs are in the supplement. A practical application that demonstrates the importance of our results for real-world CME analyses is in Section 7. Our concluding remarks are in Section 8.

2. Review of the Conditional Main Effect 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, p. 71–75), 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 factorial effects are generally estimated by least squares linear regression (Cheng, 2014, p. 81–83). A CME is similarly defined as a contrast of the treatment effects that captures the main effect of one factor conditional on the level of a second, and its estimation is also performed using regression. A general description of the CME contrast is given by Wu and Hamada (2009, p. 164) and Cheng (2014, p. 71–72).

To illustrate 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 $\mathrm{ME}(A_1) = 2^{-1}(-1,-1,1,1)^\mathsf{T}\alpha$ and $\mathrm{ME}(A_2) = 2^{-1}(-1,1,-1,1)^\mathsf{T}\alpha$ respectively, and their interaction is $\mathrm{INT}(A_1, A_2) = 2^{-1}\{(-1,-1,1,1) \odot (-1,1,-1,1)\}^\mathsf{T}\alpha$, where $\odot$ is the Hadamard product. The CMEs of A_1 given A_2 are

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

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

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 thereby reparameterize them. Wu (2018, p. 252) provides physical interpretations of these effects, and 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 $\widehat{\mathrm{ME}(A_1)} = 2^{-1}(-1,-1,1,1)^\mathsf{T}y$, $\widehat{\mathrm{ME}(A_2)} = 2^{-1}(-1,1,-1,1)^\mathsf{T}y$, $\widehat{\mathrm{INT}(A_1, A_2)} = 2^{-1}(1,-1,-1,1)^\mathsf{T}y$, $\widehat{\mathrm{CME}(A_1 \mid A_2+)} = (0,-1,0,1)^\mathsf{T}y$, $\widehat{\mathrm{CME}(A_1 \mid A_2-)} = (-1,0,1,0)^\mathsf{T}y$, respectively. Correlations between the estimators of traditional effects and CMEs, and among estimators of distinct CMEs themselves, are strictly less than one in absolute value. Consequently, the inclusion of CMEs with traditional effects in the analysis of a regular design will introduce partial aliasing relations, and result in the design becoming of the nonregular type in its analysis (Wu, 2018, p. 251).

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

Definition 1 (Twin CMEs (Su and Wu, 2017)). For distinct factors A_1 and A_2, CME($A_1 \mid A_2+$) and 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, 2018)). For distinct factors A_1 and A_2, CME($A_1 \mid A_2 s$), with $s \in \{-,+\}$, and its corresponding parent main effect ME(A_1) constitute a parent-child pair.

Definition 3 (Uncle-nephew pair (Mak and Wu, 2018)). For distinct factors A_1 and A_2, CME($A_1 \mid A_2 s$), with $s \in \{-,+\}$, and its corresponding conditioned main effect ME(A_2) constitute a uncle-nephew pair.

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

Definition 5 (Cousin CMEs (Mak and Wu, 2018))**.** For distinct factors A_1, A_2, and A_3, CME$(A_1 \mid A_2 s)$ and 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 CME$(A_i \mid A_j s)$ and 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 INT(A_i, A_j) and 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. Also, by inspection, a two-factor interaction 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 $\{\text{CME}(A_1 \mid A_2+), \text{CME}(A_1 \mid A_2-), \text{CME}(A_2 \mid A_1+), \text{CME}(A_2 \mid A_1-)\}$. The number of non-trivial families, each of which contain distinct pairs of factors in their CMEs, in a regular fraction is equal to the number of the design's aliasing relations that contain more than one two-factor interaction.

3. Indicator Functions and the Inner Product for the Conditional Main Effect System

A design $\mathcal{F} \subseteq \mathcal{D}_r$ with distinct runs is completely specified 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, e.g., those based on defining relations, via the concept of an algebraic variety (Fontana et al., 2000, p. 150). Indicator functions also exist for designs with replicate runs (Ye, 2003), but they are not considered in this paper.

From Fontana, Pistone, and Rogantin (2000, p. 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 ME(A_i), and $X_{\{i,j\}}$ represents 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 correlations between effects in $\mathcal{F}$. This fact 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 et al., 2000))**.** *The correlation between any two traditional effects in a regular design $\mathcal{F} \subseteq \mathcal{D}_r$ that correspond to $I, J \in \mathcal{P}_r$ and do not belong in the defining contrast subgroup of $\mathcal{F}$ is $b_{\mathcal{F},I\triangle J}/b_{\mathcal{F},\phi}$.*

The orthogonal basis of functions above underlies our algebra for the CME system. Specifically, CMEs are easily expressed using this orthogonal basis, with CME($A_i \mid A_j+$) and 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 as additional factors of interest in the study of 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, and thereby establish our algebra for the CME system.

Definition 8. For a fractional factorial design $\mathcal{F} \subseteq \mathcal{D}_r$, and $i, j, l, k \in \{1, \ldots, r\}$, $s, s' \in \{-, +\}$, 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 will demonstrate in the following sections, partial aliasing relations and other properties for a two-level design under the CME system can be derived in a simple and unrestricted manner by this inner product involving coordinate-free representations of conditional effects, traditional effects, and the design's indicator function.

It is important to note that if one wishes to utilize a different orthogonal basis that contains functions corresponding to CMEs, then one 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, and unduly restrict the CMEs that could be studied under the corresponding algebra, thereby frustrating one's ability to comprehensively understand the CME system for broad types of two-level designs.

4. Partial Aliasing Relations Under the Conditional Main Effect System

The inner product in Definition 8 facilitates derivations of partial aliasing relations among conditional and traditional effects. Our formal descriptions, and illustrative examples, of them follow below.

Lemma 1 ((Fontana et al., 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 $s, s' \in \{-, +\}$,*

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

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 $i \neq j$, and $I \in \mathcal{P}_r$, $s \in \{-, +\}$,*

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

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* $i \neq j$, *and* $I \in \mathcal{P}_r$, $s \in \{-, +\}$, *the correlation between* $\mathrm{CME}(A_i \mid A_j s)$ *and the traditional effect corresponding to* I *is*

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

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* $2^{-1/2}$.

(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 zero.*

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

Table 1: The 2^{3-1}_{III} design defined by $A_3 = A_1 A_2$.

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

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 $\text{CME}(A_1 \mid A_2+)$ and $\text{CME}(A_1 \mid A_3-)$ in $\mathcal{F}$. We immediately have from Corollary 1 and the indicator function coefficients that this correlation is $1/2$. Now suppose we wish to calculate the correlations of $\text{CME}(A_1 \mid A_2+)$ and $\text{CME}(A_1 \mid A_2-)$ with $\text{ME}(A_3)$ in $\mathcal{F}$. We immediately have from Corollary 2(b) that their respective correlations are $2^{-1/2}$ and $-2^{-1/2}$.

Example 2. To illustrate the utility of these results for larger designs that are of interest in practice, let $\mathcal{F}_1$ denote the minimum aberration 2^{9-4}_{IV} design, and $\mathcal{F}_2$ the 2^{9-4}_{IV} design that maximizes the number of clear two-factor interactions, which are 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

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

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

respectively, with the identity elements excluded from these subgroups without essential loss of information. For this example, suppose CME$(A_1 \mid A_2+)$ and CME$(A_1 \mid A_2-)$ are of substantive interest for inference. 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), whereas in $\mathcal{F}_2$ they are only correlated 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, so that we can easily consider any conditional and traditional effects for such large designs.

We will continue to explore the properties of $\mathcal{F}_1$ and $\mathcal{F}_2$ in later examples.

The properties of CMEs that were derived by Su and Wu (2017, p. 4–6) follow as simple corollaries of our Propositions 2 and 3, with the formal proofs provided in the supplement.

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 for 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*, non-twin 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 one difference between them 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 $\mathcal{F}_1$. Another difference that is obtained from Corollary 6, and the fact that $\mathcal{F}_1$ has less aberration than $\mathcal{F}_2$, is that $\mathcal{F}_1$ has fewer non-twin CME family members with different parent effects but 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. An illustration of this is provided in the practical application of Section 7, in which four 2^{8-3}_{IV} designs are considered that have different CME family compositions with respect to three distinct temperature factors.

Two additional properties relating to uncle-nephew effect pairs and cousin CMEs are also immediate 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 can be useful to consider when it is of interest to entertain models involving CMEs and their conditioned main effects.

5. Traditional Design Criteria and Conditional Main Effects in Resolution IV Designs

In this section, we apply our derived partial aliasing relations to illuminate the implications of the maximum clear two-factor interactions and minimum aberration criteria for the CME analysis of resolution IV regular designs. Our focus on resolution IV regular designs corresponds to an original motivation for the maximum clear two-factor interactions criterion, namely, the comparison and rank-ordering of 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^{r-p}_{IV} 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^{r-p}_{IV} 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 essential loss of information. From Wu and Hamada (2009, p. 254), $\mathcal{F}_3$ has the maximum number of clear two-factor interactions among 2_{IV}^{7-2} designs. Thus, we can immediately conclude from Proposition 4 that $\mathcal{F}_3$ has more clear CMEs than $\mathcal{F}_4$. We also observe that $\mathcal{F}_3$ has three families containing more than four members in each, because it has three aliasing relations containing more than one two-factor interaction, and $\mathcal{F}_4$ has six families containing more than four members in each, because it has six aliasing relations containing more than one two-factor interaction. In the notation of Wu and Hamada (2009, p. 215), these aliasing relations for $\mathcal{F}_3$ are

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

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

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

and these 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 to $\mathcal{F}_4$, $\mathcal{F}_3$ has a smaller total number of CMEs across its families that contain more than four members in each, which corresponds to Corollary 9.

To present the implications of the minimum aberration criterion for CME analysis, we introduce notation for 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 $t = 1, \ldots, T_{\mathcal{F}}$.

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

Using Definition 10, we reformulate in Lemma 2 the expression of Cheng, Steinberg, and Sun (1999) and Cheng (2014, p. 172) 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 into a corresponding expression for minimum aberration designs under the CME system. We then combine it with Corollary 1 to characterize in Proposition 5 how minimum aberration designs minimize aggregate measures of correlations among CMEs.

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

Proposition 5. *A fractional factorial with minimum aberration among 2^{r-p}_{IV} designs minimizes, for each exhaustive selection of CMEs such that no two involve the same pair of factors, both the sum of absolute correlations and the sum of squared correlations among non-sibling effects for the class of 2^{r-p}_{IV} 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 absolute correlations and the sum of squared correlations among non-sibling 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^{7-2}_{IV} 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^{r-p}_{IV} designs could be useful for entertaining models composed of main effects, non-sibling CMEs, and two-factor interactions. Proposition 5 demonstrates that the minimum aberration 2^{r-p}_{IV} design could be useful when it is desired to conduct an experiment with minimum aggregate correlations among distinct types of CMEs. Ultimately, 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 such family has 16 CMEs, and the remainder have 8 CMEs. Also, $\mathcal{F}_2$ has 7 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 the corresponding number for $\mathcal{F}_2$ 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 2^{9-4}_{IV} 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 absolute correlations and the sum of squared correlations among non-sibling 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 these two respective sums corresponds to Proposition 5 and the fact that $\mathcal{F}_1$ is the minimum aberration 2^{9-4}_{IV} design. Besides demonstrating the applicability of our results for large designs, this example also illustrates that, as in the case for 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 Conditional Main Effect 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 notations 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 $\mathcal{S}_1$ of interest, and a selection of only traditional effects involving the factors in $\mathcal{S}_2$ 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

$$M = \begin{pmatrix} \mathbf{1}_{|\mathcal{F}|} & S_1^{\text{Trad}} & S_2^{\text{Trad}} & S_1^{\text{CME}} \end{pmatrix}, \tag{6.1}$$

where $\mathbf{1}_{|\mathcal{F}|}$ is the $|\mathcal{F}| \times 1$ vector whose entries are all 1, S_i^{Trad} is the $|\mathcal{F}| \times |\mathcal{S}_i^{\text{Trad}}|$ matrix whose columns are the contrast vectors for the effects in $\mathcal{S}_i^{\text{Trad}}$ for $i \in \{1, 2\}$, and S_1^{CME} is the $|\mathcal{F}| \times |\mathcal{S}_1^{\text{CME}}|$ matrix whose columns 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 these notations, 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 $\mathcal{S}_2^{\text{Trad}} = \{Z_{\{3\}}\}$. The model matrix as specified in equation (6.1) for this design and model is then

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

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

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

From Definition 11, the D-efficiency calculation for a model matrix M revolves around $\det(M^\top M)$. The following lemma formally presents the derivation of $\det(M^\top M)$ 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 of M, respectively, with $Z_{(1)} = X_\phi$. Then entry (c, d) of $M^\top M$ is $2^r \left\langle Z_{(c)}, Z_{(d)} \mid \mathcal{F} \right\rangle$.*

By means of this lemma and our partial aliasing relations under the CME system, the entries of $M^\top M$ 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 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*

$$M = \begin{pmatrix} \mathbf{1}_{|\mathcal{F}|} & S_1^{\text{Trad}} & S_2^{\text{ME}} & S_2^{\text{INT}} & S_1^{\text{CME}} \end{pmatrix},$$

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

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

- W *is the* $n_1 \times n_1$ *matrix whose* (c, d) *entry for* $c, d \in \{1, \ldots, n_1\}$*, in which* $Z^s_{i|j}$ *and* $Z^{s'}_{l|k}$ *respectively correspond to the contrast vectors in columns* c *and* d *of* S_1^{CME}*, is*

$$2^{p-r+2} b^{-1}_{\mathcal{F},\phi}(b_{\mathcal{F},\{i\}\triangle\{l\}} + s' b_{\mathcal{F},\{i\}\triangle\{l,k\}} + s b_{\mathcal{F},\{i,j\}\triangle\{l\}} + ss' b_{\mathcal{F},\{i,j\}\triangle\{l,k\}}),$$

- C_1 *is the* $(1 + |\mathcal{S}_1| + |\mathcal{S}_2|) \times n_2$ *matrix whose* $(1, d)$ *entry is* 0 *for all* $d \in \{1, \ldots, n_2\}$*, and whose* (c, d) *entry for* $c \in \{2, \ldots, (1+|\mathcal{S}_1|+|\mathcal{S}_2|)\}$ *and* $d \in \{1, \ldots, n_2\}$*, in which* Z_I *and* Z_J *respectively correspond to the contrast vectors in column* c *of* $\begin{pmatrix} \mathbf{1}_{|\mathcal{F}|} & S_1^{\mathrm{Trad}} & S_2^{\mathrm{ME}} \end{pmatrix}$ *and column* d *of* S_2^{INT}*, is* $2^{p-r+2} b^{-1}_{\mathcal{F},\phi} b_{\mathcal{F},I\triangle J}$*,*

- C_2 *is the* $(1 + |\mathcal{S}_1| + |\mathcal{S}_2|) \times n_1$ *matrix whose* $(1, d)$ *entry is* 0 *for all* $d \in \{1, \ldots, n_1\}$*, and whose* (c, d) *entry for* $c \in \{2, \ldots, (1+|\mathcal{S}_1|+|\mathcal{S}_2|)\}$ *and* $d \in \{1, \ldots, n_1\}$*, in which* Z_I *and* $Z^s_{i|j}$ *respectively correspond to the contrast vectors in column* c *of* $\begin{pmatrix} \mathbf{1}_{|\mathcal{F}|} & S_1^{\mathrm{Trad}} & S_2^{\mathrm{ME}} \end{pmatrix}$ *and column* d *of* S_1^{CME}*, is* $2^{p-r+2} b^{-1}_{\mathcal{F},\phi}(b_{\mathcal{F},\{i\}\triangle I} + s b_{\mathcal{F},\{i,j\}\triangle I})$*, and*

- C_3 *is the* $n_2 \times n_1$ *matrix whose* (c, d) *entry for* $c \in \{1, \ldots, n_2\}$ *and* $d \in \{1, \ldots, n_1\}$*, in which* Z_I *and* $Z^s_{i|j}$ *respectively correspond to the contrast vectors in column* c *of* S_2^{INT} *and column* d *of* S_1^{CME}*, is*

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

Also,

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

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. Also,*

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

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 the determinant of a $n_1 \times n_1$ matrix for resolution IV designs, respectively. 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 $\mathcal{S}_2 = \{A_6, A_7, A_8, A_9\}$. By virtue of Corollary 10, the D-efficiencies of these designs with respect to this model containing a large selection of conditional and traditional effects are reduced to determinants of simple 4×4 matrices. For design $\mathcal{F}_1$,

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

and thus $\det\left(M_1^\intercal M_1\right)$ is non-zero. For design $\mathcal{F}_2$,

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

and thus $\det\left(M_2^\intercal M_2\right)$ is 0, so that this model is not estimable by $\mathcal{F}_2$.

7. Practical Application of Conditional Main Effect Analysis

We demonstrate the utility of our algebra for real-world CME analyses by considering the painted panel experiment of Lorenzen and Anderson (1993, p. 242–249). The data for this experiment are in the supplement. This case study illustrates how our results can effectively shed light on 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. 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 the higher-order interactions assumed inert. They selected a 2^{8-3}_{IV} design which 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 $\{3468, 1248, 23578, 1236, 24567, 13457, 15678\}$ (Wu and Hamada, 2009, p. 254). These designs are denoted in order by $\mathcal{F}_1^{\mathrm{LA}}$, $\mathcal{F}_2^{\mathrm{LA}}$, $\mathcal{F}_3^{\mathrm{LA}}$, and $\mathcal{F}_4^{\mathrm{LA}}$. 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, so that CMEs should be considered to

Table 2: The factors and their levels in the experiment of Lorenzen and Anderson (1993, p. 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 (−)

yield interpretable and conclusive inferences, and the designs' properties under the CME system should be assessed to better inform the final design selection. For this particular case, our algebra facilitates the understanding of how the chosen design $\mathcal{F}_1^{\mathrm{LA}}$ can yield more ambiguous inferences on CMEs that correspond to potentially active interactions involving the distinct temperature factors (A_2, A_5, A_6) compared to $\mathcal{F}_2^{\mathrm{LA}}$ and $\mathcal{F}_4^{\mathrm{LA}}$.

Our results in Section 4 enable immediate comparisons of the partial aliasing relations for CMEs in the four designs. Consider the CMEs involving the temperature factors. Proposition 2 and Corollary 1 yield in a simple manner all of the correlated CMEs that involve them 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 more conclusive CME analyses than $\mathcal{F}_1^{\mathrm{LA}}$ 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 illuminate several properties of the CME families in these designs. First, Corollary 9 and the fact that each design has the maximum number of clear two-factor interactions among 2_{IV}^{8-3} designs enable us to immediately conclude that they all have the same (and minimum) number of CMEs across their non-trivial families. In fact, each design has 1 non-trivial family that contains 12 members, and 6 non-trivial families, with each of these families in any one of the designs containing 8 members. Each design also has 13 trivial families (with 4 members in each such trivial family). Having said that, $\mathcal{F}_1^{\mathrm{LA}}$ 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 8 and CMEs involving A_5 just in their families of size 4, respectively. Second, we have from Proposition 5 and the fact that each design has minimum aberration among 2^{8-3}_{IV} designs that for any exhaustive selection of CMEs such that no two involve the same pair of factors, their respective sums of absolute correlations and squared correlations among non-sibling CMEs will be 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 will have larger sums of absolute correlations and squared correlations among non-sibling 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, 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 non-sibling CMEs, they also have subtle differences for CMEs involving 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 the degree of its robustness to the case that another temperature factor besides A_5 has an active two-factor interaction, and the degree to which conclusions can be obtained 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, so that $\det\left(M^{\mathsf{T}}M\right) = 2^{-25}\det\left\{W - C_2^{\mathsf{T}}D_1^{-1}C_2\right\}$ 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 $\mathcal{S}_1^{\mathrm{CME}} = \{Z_{2|3}^{s}, Z_{2|4}^{s'}, Z_{6|3}^{s''}, Z_{6|4}^{s'''}\}$ for $s, s', s'', s''' \in \{-,+\}$, which corresponds to CMEs that involve the two temperature factors besides A_5. The D-efficiencies for all of the designs and choices of conditioned levels in this case are equal, and immediately reduced 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 $\mathcal{S}_1^{\text{CME}} = \{Z_{i|j}^{s}, Z_{i|k}^{s'}, Z_{i|l}^{s''}\}$ 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 in this case, 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 the CME analysis of the chosen design $\mathcal{F}_1^{\text{LA}}$. The experimenters concluded by ANOVA that the following were active: ME(A_1), ME(A_2), ME(A_3), ME(A_4), ME(A_5), ME(A_8), INT(A_4, A_7), and one or more of INT(A_2, A_8), INT(A_3, A_5), and INT(A_4, A_6) (Lorenzen and Anderson, 1993, p. 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. When considering the corresponding CMEs, we have from our previous results that $\mathcal{F}_2^{\text{LA}}$ and $\mathcal{F}_4^{\text{LA}}$ were preferable designs in that they could have more easily resolved the ambiguity of which temperature factors have significant effects beyond main effects. To complete the CME analysis of this experiment, we use the three rules in the method of Su and Wu (2017, p. 5–6), and conclude that

CME($A_2 \mid A_8-$) is significant. Details of the analysis are in the supplement. Thus, we obtain the interpretable, final conclusions from the CME analysis that substrate temperature has an effect on film build at high air pressure but not low air pressure, and that, contrary to the experimenters' prior knowledge, booth temperature does not have any 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 compared to more traditional systems. The work in this paper 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 the applications of indicator functions for deriving design properties under traditional effects, whereas we consider their applications via our inner product in Definition 8 under both conditional and traditional effects. Our studies of partial aliasing relations, design criteria, and D-efficiency calculations via our algebra conclusively demonstrate its advantages. Specifically, it enables both an unrestricted approach to understanding two-level designs under the CME system, with no limits to the designs or effects that could be considered, and concise, simple calculations of design characteristics based on a small selection of indicator function coefficients. This is 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 thereby facilitates more informed choices among them.

Acknowledgements

We thank Chien-Fu Jeff Wu and two reviewers for their insightful comments that improved this paper.

References

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

Cheng, C.-S., D. M. Steinberg, and D. Sun (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *J. R. Statist. Soc. Ser. B 61*(1), 85–93.

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

Fontana, R., G. Pistone, and M. P. Rogantin (2000). Classification of two-level factorial fractions. *Journal of Statistical Planning and Inference 87*(1), 149–172.

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

Hamada, M. S. and C. F. J. Wu (1992). Analysis of designed experiments with complex aliasing. *Journal of Quality Technology 24*(3), 130–137.

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

Mak, S. and C. F. J. Wu (2018). `cmenet`: a new method for bi-level variable selection of conditional main effects. *Journal of the American Statistical Association* (in press), 1–36.

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

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

Mukerjee, R., C.-F. J. Wu, and M.-C. Chang (2017). Two-level minimum aberration designs under a conditional model with a pair of conditional and conditioning factors. *Statistica Sinica 27*(3), 1–20.

Pistone, G. and H. P. Wynn (1996). Generalised confounding with Gröbner bases.

Biometrika 83(3), 653–666.

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

Wu, C.-F. J. (2015). Post-Fisherian experimentation: from physical to virtual. *Journal of the American Statistical Association 110*(510), 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*(2), 249–268.

Wu, C.-F. J. and M. S. Hamada (2009). *Experiments: Planning, Analysis, and Optimization* (2 ed.). 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*(3), 984–994.

Ye, K. Q. (2004). A note on regular fractional factorial designs. *Statistica Sinica 14*(4), 1069–1074.

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

E-mail: sabbaghi@purdue.edu

Phone: 765-496-0234

Fax: 765-494-0558