The Annals of Statistics
2019, Vol. 47, No. 2, 909–935
https://doi.org/10.1214/18-AOS1682

SIGNAL ALIASING IN GAUSSIAN RANDOM FIELDS FOR EXPERIMENTS WITH QUALITATIVE FACTORS

BY MING-CHUNG CHANG[*], SHAO-WEI CHENG[†,1] AND CHING-SHUI CHENG[‡,§]

National Central University[*], *National Tsing Hua University*[†], *Academia Sinica*[‡] *and University of California, Berkeley*[§]

Signal aliasing is an inevitable consequence of using fractional factorial designs. Unlike linear models with fixed factorial effects, for Gaussian random field models advocated in some Bayesian design and computer experiment literature, the issue of signal aliasing has not received comparable attention. In the present article, this issue is tackled for experiments with qualitative factors. The signals in a Gaussian random field can be characterized by the random effects identified from the covariance function. The aliasing severity of the signals is determined by two key elements: (i) the aliasing pattern, which depends only on the chosen design, and (ii) the effect priority, which is related to the variances of the random effects and depends on the model parameters. We first apply this framework to study the signal-aliasing problem for regular fractional factorial designs. For general factorial designs including nonregular ones, we propose an aliasing severity index to quantify the severity of signal aliasing. We also observe that the aliasing severity index is highly correlated with the prediction variance.

1. Introduction. In science, engineering and many other areas, experimentation is commonly used to explore the systematic patterns in the relationship between the factors and the response in a phenomenon. The systematic patterns are called *signals* in this paper. When an experiment is performed with limited resources, it is possible that the collected data do not contain enough information to distinguish and/or identify some signals. We refer to this situation as *signal aliasing*.

The various (fixed) factorial effects (such as main effects and interactions) that characterize how the factors affect the response are the signals of interest in a factorial experiment. Full factorial designs can be used to gather complete information about the factorial effects. Such designs, however, are rarely used when the number of factors is large. Alternatively, fractional factorial designs (FFDs) are often used in practice. An advantage of using an FFD is the cost saving due to the reduction in run size, but a consequence is that some factorial effects are aliased.

Received September 2017; revised December 2017.

[1]Supported by the Ministry of Science and Technology of Taiwan.

MSC2010 subject classifications. 62K15.

Key words and phrases. Bayesian design, computer experiment, eigen-decomposition, fixed-effect model, fractional factorial, kriging model, random-effect model.

Under fixed-effect models (FEM) such as regression or ANOVA models generally adopted in physical experiments, the signal aliasing caused by using an FFD can be characterized by the collinear relationship among the factorial effects. Any two factorial effects can be fully aliased, partially aliased, or mutually orthogonal. Readers are referred to Wu and Hamada (2009) and Cheng (2014) for more details about the aliasing of factorial effects in an FFD under an FEM.

An alternative modeling option for data from factorial experiments is the stochastic process approach that appeared in some Bayesian design literature [Mitchell, Morris and Ylvisaker (1995), Kerr (2001), Joseph (2006), Joseph and Delaney (2007)]. In this approach, the response function is regarded as a realization of a Gaussian random field. Suppose there are n factors $F_1, \dots, F_n$, each with p_i levels, $i = 1, \dots, n$. Denote by χ the set of all the factor-level combinations $\boldsymbol{x} = (x_1, \dots, x_n)^T$, also called treatments or design points, where $1 \le x_i \le p_i$ is the level of the ith factor. Consider the model

$$Y(\boldsymbol{x}) = Z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \tag{1.1}$$

where $Y(\boldsymbol{x})$ is the response observed at the design point $\boldsymbol{x}$, Z is a zero-mean stationary Gaussian random field with $\mathrm{Var}(Z(\boldsymbol{x})) = \sigma_Z^2$ and a covariance function $C(\boldsymbol{x}, \boldsymbol{x}')$ defined for any two design points $\boldsymbol{x}$ and $\boldsymbol{x}'$, and $\varepsilon(\boldsymbol{x})$, assumed to be independent of $Z(\boldsymbol{x})$, is a zero-mean random error with constant variance and $E[\varepsilon(\boldsymbol{x})\varepsilon(\boldsymbol{x}')] = 0$ for $\boldsymbol{x} \neq \boldsymbol{x}'$. Here, $Z(\boldsymbol{x})$ represents the systematic patterns in $Y(\boldsymbol{x})$. In *computer experiments*, such models are more popular than regression or ANOVA models, even when the experiments involve some qualitative factors [see Qian, Wu and Wu (2008), Han et al. (2009), Zhou, Qian and Zhou (2011), Deng, Hung and Lin (2015)]. Being able to evaluate the severity of signal aliasing is crucial for designing an experiment with a better ability to distinguish the most important signals. In Section 2, we present a motivating example to illustrate that the prediction variance of $Z(\boldsymbol{x})$ could be influenced by the severity level of aliasing.

Unlike the FEMs, it is not clear how to evaluate the "severity of effect aliasing" under a Gaussian random field model. While for FEMs the "signals" are explicitly defined in the mean response via some pre-determined fixed factorial effects, for $Z(\boldsymbol{x})$ they are hidden in the covariance function. For two-level designs, Mitchell, Morris and Ylvisaker (1995) defined random factorial effects directly as contrasts of $Z(\boldsymbol{x}), \boldsymbol{x} \in \chi$, in the same way as the fixed factorial effects are defined in FEMs, and showed that under certain covariance functions of the Gaussian random field, the contrast vectors that define the factorial effects are eigenvectors of the covariance function. We extend this result to general multi-level FFDs for experiments with *qualitative* factors, and apply the result to tackle the signal-aliasing problem.

Throughout the rest of this paper, we assume that the covariance function of Z has the form

$$C(\boldsymbol{x}_r, \boldsymbol{x}_s) = \tau_{t_1 \cdots t_n}, \tag{1.2}$$

where for any two design points $\boldsymbol{x}_r = (x_{r1}, \dots, x_{rn})^T$ and $\boldsymbol{x}_s = (x_{s1}, \dots, x_{sn})^T$ in χ, $(t_1, \dots, t_n)^T$ is the componentwise Hamming distance vector of $\boldsymbol{x}_r$ and $\boldsymbol{x}_s$ (i.e., for $i = 1, 2, \dots, n$, $t_i = 1$ if $x_{ri} \neq x_{si}$ and zero otherwise), and the $\tau_{t_1 \cdots t_n}$'s are parameters depending only on $(t_1, \dots, t_n)^T$. The parameter space is the collection of $\tau_{t_1 \cdots t_n}$'s that produce positive definite $C(\cdot, \cdot)$ on χ. The positive-definite property induces some constraints on the $\tau_{t_1 \cdots t_n}$'s. An obvious one is $C(\boldsymbol{x}, \boldsymbol{x}) = \tau_{0 \cdots 0} > 0$. The existence of the $\tau_{t_1 \cdots t_n}$'s satisfying the constraints is guaranteed by Theorem 3.1 (presented later in Section 3). It is also implied by the theorem that the parameter space of the $\tau_{t_1 \cdots t_n}$'s is a subset of $\mathbb{R}^{2^n}$. In the following, the vector of all the parameters $\tau_{t_1 \cdots t_n}$'s is denoted by $\boldsymbol{\tau}$. Because $\tau_{0 \cdots 0} = C(\boldsymbol{x}, \boldsymbol{x})$ for any $\boldsymbol{x} \in \chi$, the correlation function of $Z(\boldsymbol{x})$ is

$$R(\boldsymbol{x}_r, \boldsymbol{x}_s) = \frac{C(\boldsymbol{x}_r, \boldsymbol{x}_s)}{\tau_{0 \cdots 0}} = \frac{\tau_{t_1 \cdots t_n}}{\tau_{0 \cdots 0}}.$$

Some other covariance functions have been proposed for qualitative factors in the literature [e.g., Joseph and Delaney (2007), Qian, Wu and Wu (2008)]. A brief comparison of these covariance functions can be found in Chang (2015). For the signal-aliasing problem, the key is that $Z(\boldsymbol{x})$ can be expressed as a linear combination of random factorial effects, through which some ideas that have been used to study aliasing in FFDs under FEMs are generalized to Gaussian random fields.

The covariance function considered in Joseph and Delaney (2007) is a special case of (1.2) with

$$\frac{\tau_{t_1 \cdots t_n}}{\tau_{0 \cdots 0}} = \prod_{i=1}^{n} \left(\frac{\tau_{\boldsymbol{T}_i}}{\tau_{0 \cdots 0}} \right)^{t_i}, \tag{1.3}$$

where, for $i = 1, \dots, n$, $\boldsymbol{T}_i$ is the ith unit vector. In this case, the number of parameters in the covariance function of Z is reduced from 2^n to $n + 1$ ($\tau_{\boldsymbol{T}_i}$'s and $\tau_{0 \cdots 0}$). We denote the vector of the n parameters $\tau_{\boldsymbol{T}_1}, \dots, \tau_{\boldsymbol{T}_n}$ by $\boldsymbol{\tau}_{\text{ex}}$.

The rest of this article is organized as follows. In addition to the motivating example referred to earlier, Section 2 contains a review of the treatment factorial structure. Hidden random factorial effects identified from the covariance function of a Gaussian random field are discussed in Section 3. Section 4 addresses the issue of effect priority. In Section 5 we study the signal aliasing in model (1.1) under regular FFDs with prime-number levels. A criterion for measuring the severity level of aliasing in an FFD (regular or nonregular), called aliasing severity index, is proposed in Section 6. Some discussions are given in Section 7. All the proofs of the theorems are provided in the Appendix.

2. A motivating example and treatment factorial structure. We use an example to demonstrate the need to study signal aliasing for Gaussian random fields. Consider two 2^{6-2} regular FFDs reported in Wu and Hamada [(2009), page 253] with the defining contrast subgroups (DCSGs):

$$d_1 : I = \mathbf{1235} = \mathbf{1246} = \mathbf{3456},$$

$$d_2 : I = \mathbf{125} = \mathbf{1346} = \mathbf{23456}.$$

Among all the 2^{6-2} regular FFDs, d_1 has minimum aberration [Fries and Hunter (1980)]; so under the *effect hierarchy principle* [Wu and Hamada (2009), page 172], d_2 has more severe aliasing than d_1. On the other hand, suppose factors **4** and **6** are more important than the other factors. Then d_2 would have less severe aliasing than d_1, because among the two-factor interactions involving factors **4** or **6**, d_1 produces four aliased pairs (**14** = **26**, **24** = **16**, **34** = **56** and **45** = **36**) while d_2 produces only two aliased pairs (**14** = **36** and **34** = **16**).

We use prediction variances to compare the performances of d_1 and d_2 under a Gaussian random field. By the same argument given in Santner, Williams and Notz [(2003), page 93], under an N-run design $d = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_N\}$, for any $\boldsymbol{x} \in \chi$, the prediction variance of $Z(\boldsymbol{x})$ given $Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_N)$ is

$$\sigma_Z^2\left(1 - \boldsymbol{r}_{\boldsymbol{x},d}^T \mathbf{R}_d^{-1} \boldsymbol{r}_{\boldsymbol{x},d}\right), \tag{2.1}$$

where $\mathbf{R}_d$ is the correlation matrix of Z evaluated on d, and $\boldsymbol{r}_{\boldsymbol{x},d}$ is the $N \times 1$ vector formed by the correlations between $Z(\boldsymbol{x})$ and the $Z(\boldsymbol{x}_i)$'s, $\boldsymbol{x}_i \in d$.

Consider a $Z(\boldsymbol{x})$ with the covariance function (1.3). Without loss of generality, we set $\tau_{0\cdots0} = 1$. Then each $\tau_{\boldsymbol{T}_i}$ is a correlation. For $n = 6$, consider the set of parameters

$$\boldsymbol{\Theta}_{\text{ex},1} = \{\boldsymbol{\tau}_{\text{ex}} = (\tau, \tau, \tau, \tau, \tau, \tau) : 0 \le \tau \le 1\}. \tag{2.2}$$

This set is a line segment in the parameter space of $\boldsymbol{\tau}_{\text{ex}}$. Because $\tau_{\boldsymbol{T}_1} = \cdots = \tau_{\boldsymbol{T}_6}$, every $\boldsymbol{\tau}_{\text{ex}}$ in $\boldsymbol{\Theta}_{\text{ex},1}$ assumes that all the six factors are *equally* important for $Z(\boldsymbol{x})$. By a discussion that will be given in Section 3, for τ close to 1, the six factors are (equally) unimportant, while for τ close to 0, they are (equally) very important. Under d_1 and d_2, the average prediction variances (APVs) of $Z(\boldsymbol{x})$ over all the points $\boldsymbol{x}$ in χ for $\boldsymbol{\tau}_{\text{ex}} \in \boldsymbol{\Theta}_{\text{ex},1}$ are calculated from (2.1) and plotted in the left panel of Figure 1. The difference of the APVs for d_1 and d_2 is given in the right panel. The figure shows that d_1 performs slightly better than d_2 on the APV uniformly for $\boldsymbol{\tau}_{\text{ex}} \in \boldsymbol{\Theta}_{\text{ex},1}$. This is expected since d_1 is a minimum aberration design favoring no particular factor, and for $\boldsymbol{\tau}_{\text{ex}} \in \boldsymbol{\Theta}_{\text{ex},1}$, all the six factors are of equal importance. We will see in Section 3 that when τ increases, the aliasing in both designs becomes less severe. Indeed we observe in the figure that the APVs decrease.

Now, consider another set of parameters, which is also a line segment in the parameter space of $\boldsymbol{\tau}_{\text{ex}}$:

$$\boldsymbol{\Theta}_{\text{ex},2} = \{\boldsymbol{\tau}_{\text{ex}} = (0.5, 0.5, 0.5, \tau, 0.5, \tau) : 0 \le \tau \le 1\}. \tag{2.3}$$

For τ closer to 0, factors **4** and **6** are more important than the others, while for τ closer to 1, the situation is reversed. Since d_2 causes less severe aliasing than

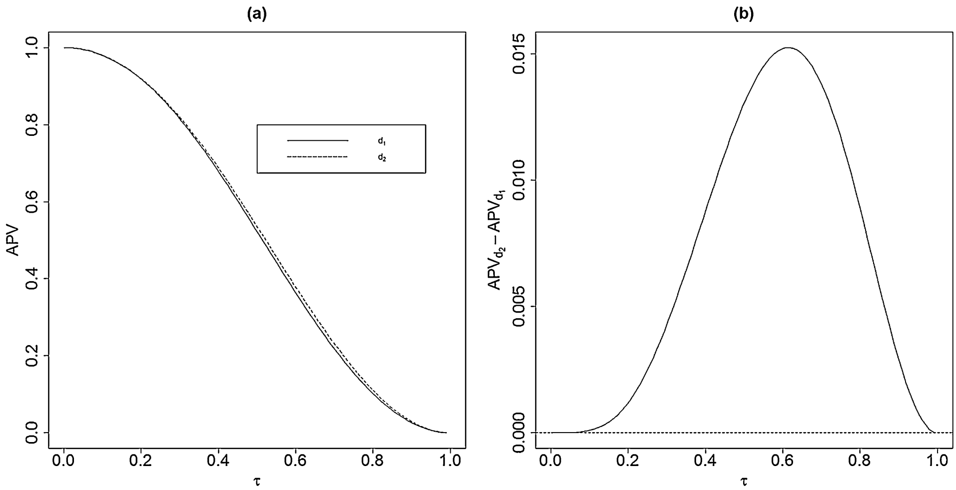

FIG. 1. (a) *Plots of the average prediction variances under* d_1 *and* d_2, *and* (b) *difference of the average prediction variances between* d_2 *and* d_1, *for* $\mathbf{\Theta}_{\mathrm{ex},1}$.

d_1 on the effects involving factors **4** or **6**, which are more important when τ is small, it is expected that d_2 produces smaller APV than d_1 in this case. A similar statement can be made for the case where τ is large. Indeed Figure 2 shows that d_2 performs better than d_1 on the APV when τ is small, and the two APV curves have a crossover point at $\tau = 0.39$.

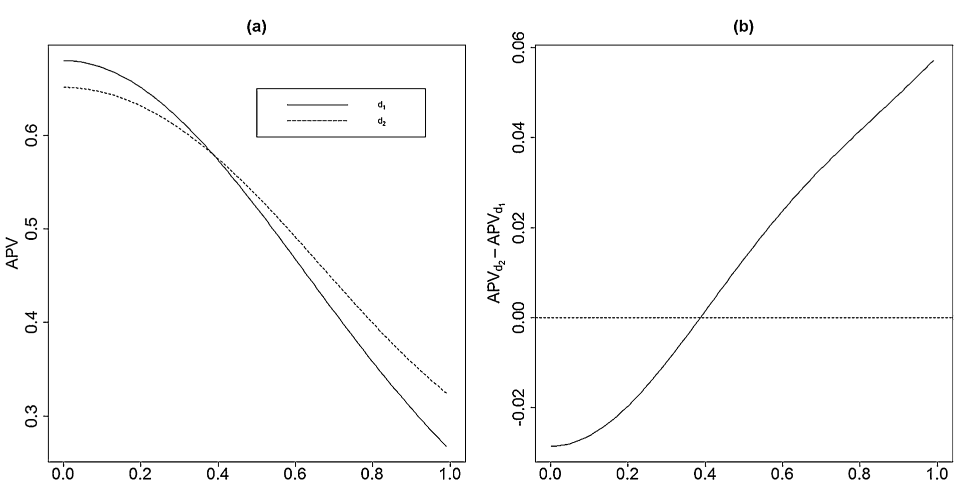

FIG. 2. (a) *Plots of the average prediction variances under* d_1 *and* d_2, *and* (b) *difference of the average prediction variances between* d_2 *and* d_1, *for* $\mathbf{\Theta}_{\mathrm{ex},2}$.

This example suggests that under Z, the power of an FFD in generating good predictions could depend on the severity level of signal aliasing produced by the FFD. The example will be revisited in Section 6. The two designs will be evaluated there using an aliasing severity index proposed in this article, and it will be shown that the APVs in Figures 1 and 2 are highly correlated with the aliasing severity indexes of d_1 and d_2.

The rest of this section is devoted to a review of factorial treatment structures. Let $\Xi = \prod_{i=1}^{n} p_i$ be the number of all the treatments (design points) and denote the Ξ design points in χ by $\boldsymbol{x}_1, \dots, \boldsymbol{x}_\Xi$. For each $S \subseteq \{1, \dots, n\}$, let V_S be the set of all the vectors $(v_1, \dots, v_\Xi)^T \in \mathbb{R}^\Xi$ such that $v_i = v_j$ if $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ have the same levels for all the factors in S. Then for $S = \{i_1, \dots, i_k\}$, V_S is a $(\prod_{j=1}^{k} p_{i_j})$-dimensional subspace of $\mathbb{R}^\Xi$. In particular, $V_\varnothing$ is the 1-dimensional subspace spanned by the vector of all ones and, for $S = \{1, \dots, n\}$, $V_S = \mathbb{R}^\Xi$. We further define 2^n spaces W_S, one for each $S \subseteq \{1, \dots, n\}$, as

$$W_S = V_S \cap \left(\bigcap_{S': S' \subset S} V_{S'}^{\perp} \right), \tag{2.4}$$

where $\perp$ is the orthogonal complement operator. For example, when $n = 2$ and $p_1 = p_2 = 2$, denote the four design points by $\boldsymbol{x}_1 = (-1, -1)$, $\boldsymbol{x}_2 = (-1, 1)$, $\boldsymbol{x}_3 = (1, -1)$ and $\boldsymbol{x}_4 = (1, 1)$. For $S = \{1\}$, because factor **1** has the same level in $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and the same level in $\boldsymbol{x}_3$ and $\boldsymbol{x}_4$, we have that $V_S = \{(v_1, v_2, v_3, v_4)^T : v_1 = v_2, v_3 = v_4\}$. Thus V_S is the 2-dimensional space spanned by $(1, 1, 1, 1)^T$ and $(-1, -1, 1, 1)^T$. Since $S' = \varnothing$ is the only proper subset of $\{1\}$, and $V_\varnothing$ is the 1-dimensional space spanned by $(1, 1, 1, 1)^T$, it follows from (2.4) that W_S is the 1-dimensional space spanned by $(-1, -1, 1, 1)^T$, which is a main-effect contrast of factor **1**. For $S = \{1, 2\}$, $V_S = \mathbb{R}^4$. Since $\varnothing$, $\{1\}$ and $\{2\}$ are the proper subsets of S, by (2.4), W_S consists of all the vectors that are orthogonal to $(1, 1, 1, 1)^T$, the main-effect contrasts of factor **1** and the main-effect contrasts of factor **2**. It follows that W_S is the 1-dimensional space spanned by $(1, -1, -1, 1)^T$, which is an interaction contrast of the two factors.

By Theorem 6.2 in Cheng (2014), the 2^n subspaces W_S's, $S \subseteq \{1, \dots, n\}$, form an orthogonal decomposition of $\mathbb{R}^\Xi$. The readers are referred to Cheng [(2014), Chapters 4 and 6] for more details about factorial treatment structures.

For each nonempty S, all the vectors in W_S are orthogonal to the vector of all 1's, and thus define treatment *contrasts*. They are main-effect contrasts of F_i if $S = \{i\}$ and are interaction contrasts of $F_{i_1}, \dots, F_{i_k}$ if $S = \{i_1, \dots, i_k\}$. In this paper, an orthogonal basis of W_S is called a set of *statistical orthogonal contrasts* (SOCs) of W_S. If $\{\boldsymbol{c}_1^i, \dots, \boldsymbol{c}_{p_i-1}^i\}$ is a set of SOCs for the main effects of F_i, $i = 1, \dots, n$, then

$$\{\boldsymbol{c}_{s_1}^{i_1} \odot \cdots \odot \boldsymbol{c}_{s_k}^{i_k} : 1 \le s_j \le p_{i_j} - 1, j = 1, \dots, k\} \tag{2.5}$$

is a set of SOCs for the interactions of factors $F_{i_1}, \dots, \mathcal{F}_{i_k}$, where $\odot$ is the componentwise product. A collection of SOCs of W_S for all $S \neq \varnothing$, $\Xi - 1$ in total, is called a *statistical orthogonal contrast basis*; see also Cheng and Ye (2004). For consistency, all the SOCs discussed in this paper are scaled to have a squared length Ξ.

When $p_1 = \cdots = p_n = p$ and p is a prime or a prime power, an alternative approach to defining SOCs is through the use of finite geometry. More details can be found in Cheng (2014), Section 6.6.

3. Hidden random effects. Motivated by Mitchell, Morris and Ylvisaker (1995), we generalize the notion of factorial effects in FEMs to Gaussian random field models by eigen-decomposing the covariance function of Z, where Z is the Gaussian random field in (1.1). Let $\boldsymbol{Z}$ be the vector $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_\Xi))^T$ and $\mathbf{C}_{\boldsymbol{\tau}}$ be the covariance matrix of $\boldsymbol{Z}$. Because $\mathbf{C}_{\boldsymbol{\tau}}$ is symmetric and positive definite, we have

$$\mathbf{C}_{\boldsymbol{\tau}} = \mathbf{E}_{\boldsymbol{\tau}} \boldsymbol{\Lambda}_{\boldsymbol{\tau}} \mathbf{E}_{\boldsymbol{\tau}}^T, \tag{3.1}$$

where the columns of $\mathbf{E}_{\boldsymbol{\tau}}$ are Ξ orthogonal normalized eigenvectors of $\mathbf{C}_{\boldsymbol{\tau}}$ and $\boldsymbol{\Lambda}_{\boldsymbol{\tau}}$ is a diagonal matrix with the associated positive eigenvalues as its diagonal entries. Denote the lth column of $\mathbf{E}_{\boldsymbol{\tau}}$ by $\boldsymbol{g}_{l,\boldsymbol{\tau}}$ and let the corresponding eigenvalue be $\xi_l(\boldsymbol{\tau})$. Then a covariance-function representation of (3.1) is

$$C(\boldsymbol{x}_r, \boldsymbol{x}_s; \boldsymbol{\tau}) = \sum_{l=1}^{\Xi} \frac{g_l(\boldsymbol{x}_r; \boldsymbol{\tau})}{\|\boldsymbol{g}_{l,\boldsymbol{\tau}}\|} \frac{g_l(\boldsymbol{x}_s; \boldsymbol{\tau})}{\|\boldsymbol{g}_{l,\boldsymbol{\tau}}\|} \xi_l(\boldsymbol{\tau}),$$

where $\boldsymbol{x}_r, \boldsymbol{x}_s \in \chi$, $g_l(\boldsymbol{x}; \boldsymbol{\tau})$ is the component of $\boldsymbol{g}_{l,\boldsymbol{\tau}}$ corresponding to $\boldsymbol{x}$, and $\|\cdot\|$ is the Euclidean norm of a vector. For consistency, all the $\boldsymbol{g}_{l,\boldsymbol{\tau}}$'s are scaled to have a squared length Ξ. The parameter $\boldsymbol{\tau}$ is put in the representation to emphasize that the eigenvectors and eigenvalues depend on $\boldsymbol{\tau}$ in general.

Because $\boldsymbol{Z}$ is a random vector in $\mathbb{R}^\Xi$ and the eigenvectors form an orthogonal basis of $\mathbb{R}^\Xi$, we can express $\boldsymbol{Z}$ as a linear combination of eigenvectors:

$$\boldsymbol{Z} = \sum_{l=1}^{\Xi} \beta_{l,\boldsymbol{\tau}} \boldsymbol{g}_{l,\boldsymbol{\tau}}, \tag{3.2}$$

where $\beta_{l,\boldsymbol{\tau}} = \boldsymbol{g}_{l,\boldsymbol{\tau}}^T \boldsymbol{Z} / \|\boldsymbol{g}_{l,\boldsymbol{\tau}}\|^2$ is the projection score of $\boldsymbol{Z}$ on $\boldsymbol{g}_{l,\boldsymbol{\tau}}$. The $\beta_{l,\boldsymbol{\tau}}$'s are independent normal random variables with zero means and variances $\sigma_l^2(\boldsymbol{\tau}) = \xi_l(\boldsymbol{\tau}) / \|\boldsymbol{g}_{l,\boldsymbol{\tau}}\|^2$, $l = 1, \dots, \Xi$.

We can write (3.2) as

$$Z(\boldsymbol{x}) = \sum_{l=1}^{\Xi} \beta_{l,\boldsymbol{\tau}} g_l(\boldsymbol{x}; \boldsymbol{\tau}). \tag{3.3}$$

Let H be the collection of all the 2^n vectors $\boldsymbol{t} = (t_1, \dots, t_n)^T$, where each t_i is 0 or 1. The following theorem, which extends Proposition 3.2 in Mitchell, Morris and Ylvisaker (1995), shows that under the covariance function (1.2), the $\boldsymbol{g}_{l,\boldsymbol{\tau}}$'s and $\beta_{l,\boldsymbol{\tau}}$'s in (3.2) do not depend on $\boldsymbol{\tau}$.

THEOREM 3.1. *Let* $\mathbf{C}_{\boldsymbol{\tau}}$ *be the covariance matrix of* $\mathbf{Z}$ *under the covariance function* (1.2). *Then* $\mathbf{C}_{\boldsymbol{\tau}}$ *has at most* 2^n *distinct eigenvalues. For each* $S \subseteq \{1, \dots, n\}$, *let* $\boldsymbol{u}_S = (u_{S1}, \dots, u_{Sn})^T$ *be a vector such that* u_{Si} *is equal to* 1 *if* $i \in S$, *and zero otherwise. Then all the vectors in* W_S *are eigenvectors of* $\mathbf{C}_{\boldsymbol{\tau}}$ *with the eigenvalue*

$$\xi_S(\boldsymbol{\tau}) = \sum_{\boldsymbol{t} \in H} (-1)^{\boldsymbol{t}^T \boldsymbol{u}_S} \left[\prod_{i=1}^{n} (p_i - 1)^{t_i(1-u_{Si})} \right] \tau_{\boldsymbol{t}}. \tag{3.4}$$

The linear equation (3.4) *is invertible and gives a one-to-one correspondence between the* ξ_S*'s and* $\tau_{\boldsymbol{t}}$*'s.*

Note that the W_S's do not depend on $\boldsymbol{\tau}$. By Theorem 3.1, a set of SOCs of W_S are mutually orthogonal eigenvectors of $\mathbf{C}_{\boldsymbol{\tau}}$. Suppose that $\{\boldsymbol{g}_l : 1 \le l \le \Xi\}$ is a statistical orthogonal contrast basis. Then (3.2) reduces to

$$\boldsymbol{Z} = \sum_{l=1}^{\Xi} \beta_l \boldsymbol{g}_l, \tag{3.5}$$

and for each S, all the β_l's associated with S have the identical variance

$$\sigma_l^2(\boldsymbol{\tau}) = \xi_S(\boldsymbol{\tau}) / \Xi. \tag{3.6}$$

For each $\boldsymbol{x} \in \Xi$, we can also write $Z(\boldsymbol{x})$ as

$$Z(\boldsymbol{x}) = \sum_{l=1}^{\Xi} \beta_l g_l(\boldsymbol{x}) \tag{3.7}$$

and regard (3.7) as a linear random-effect model with *independent* β_l's acting like random effects and g_l's like contrast codings in FEMs for factorial designs. The expression (3.7) is also given in Mitchell, Morris and Ylvisaker (1995) for two-level designs, where each β_l is a factorial effect. Throughout this paper, the random effects β_l's in (3.7), identified from the covariance function of Z, are called *hidden random effects* and are regarded as the signals of Z. The aliasing of these signals under FFDs will be discussed in Sections 5 and 6. We note that the proof in Mitchell, Morris and Ylvisaker (1995) cannot be extended to the multi-level case. Joseph and Delaney (2007) considered the multi-level case under the more restricted covariance structure in (1.3).

Since $\mathbf{C}_{\boldsymbol{\tau}}$ is positive definite if and only if all the ξ_S's in (3.4) are greater than zero, it follows immediately from the theorem that the $\tau_{\boldsymbol{t}}$'s can only take the values such that $\xi_S(\boldsymbol{\tau}) > 0$ for all $S \subseteq \{1, \dots, n\}$. This requirement defines the parameter space of $\boldsymbol{\tau}$, which, by the one-to-one correspondence between the $\tau_{\boldsymbol{t}}$'s and ξ_S's, is a subset of $\mathbb{R}^{2^n}$.

An illustrative example of Theorem 3.1 is given below.

EXAMPLE 3.1. Consider a 2^3 full factorial design and a Gaussian random field $\boldsymbol{Z}$ with the covariance function (1.2). The three factors are labelled by **1**, **2** and **3** and the eight runs are arranged in a Yates' order. A 1-dimensional W_S representing a factorial effect (including the grand mean) is associated with each of the eight subsets S of $\{1, 2, 3\}$. Denote the codings of the factorial effects by $g_{\mathbf{0}}, g_{\mathbf{1}}, g_{\mathbf{2}}, g_{\mathbf{3}}, \dots, g_{\mathbf{123}}$, respectively. By Theorem 3.1, these eight vectors produce an orthogonal eigenvector basis for the covariance matrix $\mathbf{C}_{\boldsymbol{\tau}}$ of $\boldsymbol{Z}$, and $\mathbf{C}_{\boldsymbol{\tau}}$ can be eigen-decomposed as $\mathbf{E}\boldsymbol{\Lambda}_{\boldsymbol{\tau}}\mathbf{E}^T$, where

$$\mathbf{E} = \frac{1}{\sqrt{8}} \begin{array}{c} \begin{array}{cccccccc} g_{\mathbf{0}} & g_{\mathbf{1}} & g_{\mathbf{2}} & g_{\mathbf{3}} & g_{\mathbf{12}} & g_{\mathbf{13}} & g_{\mathbf{23}} & g_{\mathbf{123}} \end{array} \\ \left(\begin{array}{rrrrrrrr} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{array} \right) \end{array}.$$

The eigenvalues in $\boldsymbol{\Lambda}_{\boldsymbol{\tau}}$ can be obtained from (3.4). Take $S = \{1, 2\}$ as an example. We have that $g_{\mathbf{12}}$ is an eigenvector with the eigenvalue

$$\xi_S(\boldsymbol{\tau}) = (\tau_{000} + \tau_{001} + \tau_{110} + \tau_{111}) - (\tau_{100} + \tau_{101} + \tau_{010} + \tau_{011}) = g_{\mathbf{12}}^T \boldsymbol{\tau},$$

which is the sum of the $\tau_{\boldsymbol{t}}$'s with $t_1 + t_2 = 0 \pmod 2$ minus the sum of the $\tau_{\boldsymbol{t}}$'s with $t_1 + t_2 = 1 \pmod 2$. The other eigenvalues have similar forms. The parameter space of $\boldsymbol{\tau}$ is $\{\boldsymbol{\tau} : \mathbf{E}^T \boldsymbol{\tau} > \mathbf{0}\}$. By (3.7), $\boldsymbol{Z} = \beta_0 g_{\mathbf{0}} + \beta_{\mathbf{1}} g_{\mathbf{1}} + \cdots + \beta_{\mathbf{123}} g_{\mathbf{123}}$, where the β's are independent normally distributed random variables with zero means. By (3.6), the variances of β's are the corresponding eigenvalues divided by 8.

Applying Theorem 3.1 to the covariance function (1.3), a special case of (1.2), we have

$$\begin{aligned} \xi_S(\boldsymbol{\tau}_{\text{ex}}) &= \tau_{0\cdots0} \sum_{\boldsymbol{t}\in H} \left[\prod_{i=1}^n (-1)^{t_i u_{Si}} (p_i - 1)^{t_i(1-u_{Si})} \left(\frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots0}} \right)^{t_i} \right] \\ &= \tau_{0\cdots0} \sum_{\boldsymbol{t}\in H} \left\{ \prod_{i=1}^n \left(-\frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots0}} \right)^{t_i u_{Si}} \left[(p_i - 1) \frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots0}} \right]^{t_i(1-u_{Si})} \right\} \\ &= \tau_{0\cdots0} \left[\prod_{i\in S} \left(1 - \frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots0}} \right) \right] \left\{ \prod_{i\notin S} \left[1 + (p_i - 1) \frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots0}} \right] \right\}, \end{aligned}$$

where the last equality holds because, for $1 \le i \le n$, we have $t_i u_{Si} = t_i$ if $i \in S$, and $t_i u_{Si} = 0$ if $i \notin S$. This result also appeared in Joseph and Delaney (2007) with a different parametrization, and the special case with $p_1 = \cdots = p_n = 2$ was obtained earlier by Mitchell, Morris and Ylvisaker (1995). The formula in (3.8) implies some interesting properties. First, when $\tau_{\boldsymbol{T}_i} > 0$ for $i = 1, \ldots, n$, we have

$$1 - \frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots 0}} < 1 + (p_i - 1)\frac{\tau_{\boldsymbol{T}_i}}{\tau_{0\cdots 0}},$$

which implies $\xi_S(\boldsymbol{\tau}_{\mathrm{ex}}) > \xi_{S'}(\boldsymbol{\tau}_{\mathrm{ex}})$ if $S \subset S'$. Second, when $\tau_{\boldsymbol{T}_i}/\tau_{0\cdots 0}$ decreases, $\xi_S(\boldsymbol{\tau}_{\mathrm{ex}})$ increases if S involves the ith factor and decreases otherwise. Thus as $\tau_{\boldsymbol{T}_i}/\tau_{0\cdots 0}$ decreases, all the β_l's associated with the S's involving the ith factor tend to produce larger values. This indicates that $\tau_{\boldsymbol{T}_i}$ can be used to assess the overall influence of the ith factor on Z. This observation supports the arguments given in Section 2 for $\boldsymbol{\Theta}_{\mathrm{ex},1}$ and $\boldsymbol{\Theta}_{\mathrm{ex},2}$.

The procedure used to identify hidden random effects has a strong theoretical connection with principal component analysis. Both methods utilize the eigen-decomposition of covariance matrices. Steinberg and Bursztyn (2004) also applied it to explore Bayesian regression models associated with random fields in the case of *quantitative* factors. There are, however, major differences between their study and ours. First, they performed eigen-decomposition on the fitted covariance matrix (with the parameters replaced by their estimates), while we derive the eigenvalues and eigenvectors from the covariance matrix with unknown parameters. The evaluation of aliasing severity of designs is usually conducted before experiments for the purpose of choosing a better design, and there is no data available to estimate the parameters then. Second, under their covariance function (for quantitative factors), the eigenvectors depend on the parameters, while our covariance function (for qualitative factors) has eigenvectors irrelevant to the parameters. As shown in Steinberg and Bursztyn (2004) and Chang (2015), a theorem such as Theorem 3.1 is not achievable for most commonly used covariance functions for quantitative factors. In this case, the eigen-decomposition leads to a model like (3.3) in which the contrast codings depend on the unknown parameters. For such contrast codings, a study of signal aliasing is difficult, or probably impossible, to implement before experiments. An alternative approach to studying signal aliasing for quantitative factors was proposed in Chang (2015) and Chang and Cheng (2018), in which they defined contrast codings directly through a set of pre-specified orthogonal functions of $\boldsymbol{x}$ that do not depend on the parameters, but paid the price that the resulting hidden random effects β_l's are correlated.

4. Priority of hidden random effects. It is impossible to evaluate the aliasing severity of FFDs without making an effect priority assumption on what effects are more likely to be important. Under FEMs, the commonly used *effect hierarchy principle* assumes that lower-order effects are more likely to be important than higher-order effects and effects of the same order are equally likely to be important. This principle justifies various optimality criteria such as minimum aberration in the development of FFD theory.

The effect priority principle for Gaussian random field models can be developed from the distributions of the hidden random effects β_l's in (3.7), which are independently normally distributed with zero means. A β_l with a large variance is more likely to produce a value away from 0. It is therefore reasonable to assign each β_l a priority proportional to its variance $\sigma_l^2(\boldsymbol{\tau})$. The variances provide a quantitative assessment of effect priority. For example, β_l is regarded as twice as important as $\beta_{l'}$ at some $\boldsymbol{\tau}$ if $\sigma_l^2(\boldsymbol{\tau}) = 2\sigma_{l'}^2(\boldsymbol{\tau})$. In contrast, an assumption such as effect hierarchy only offers a relative priority order of effects, which is essentially an ordinal (or qualitative) evaluation.

We call an order of β_l's ranked by their variances a *priority order* of β_l's. The priority orders depend on $\boldsymbol{\tau}$. Ideally, we can partition the parameter space of $\boldsymbol{\tau}$ into disjoint subsets, each of which contains the $\boldsymbol{\tau}$'s that generate an identical priority order. Some principles are defined below for choosing a subset of the parameter space with priority orders that are suitable, or at least reasonable, for common studies of aliasing severity in FFDs. For each β_l, denote the set of the associated factors by S_l and let $|S_l|$ be the cardinality of S_l.

DEFINITION 4.1. Let $\boldsymbol{\Theta}$ be a subset of the parameter space. A priority order of the hidden random effects is said to be consistent with the effect hierarchy principle in $\boldsymbol{\Theta}$ if, for any $\boldsymbol{\tau} \in \boldsymbol{\Theta}$ and $1 \le l \ne l' \le \Xi$, we have $\sigma_l^2(\boldsymbol{\tau}) = \sigma_{l'}^2(\boldsymbol{\tau})$ when $|S_l| = |S_{l'}|$ and $\sigma_l^2(\boldsymbol{\tau}) > \sigma_{l'}^2(\boldsymbol{\tau})$ when $|S_l| < |S_{l'}|$. A priority order of the hidden random effects is said to be consistent with the effect heredity principle [Wu and Hamada (2009), page 173] if for any $\boldsymbol{\tau} \in \boldsymbol{\Theta}$ and (l, l') such that S_l is a proper subset of $S_{l'}$, $1 \le l, l' \le \Xi$, we have $\sigma_l^2(\boldsymbol{\tau}) > \sigma_{l'}^2(\boldsymbol{\tau})$.

The latter property in Definition 4.1 is called *nested decreasing interaction variance* in Kerr (2001). For the covariance function (1.3), it follows from (3.8) that the priority orders of β_l's are consistent with effect heredity in $\boldsymbol{\Theta} = \{\boldsymbol{\tau}_{\mathrm{ex}} = (\tau_{\boldsymbol{T}_1}, \dots, \tau_{\boldsymbol{T}_n}) : \tau_{\boldsymbol{T}_i} > 0, i = 1, \dots, n\}$. Furthermore, when $p_1 = \cdots = p_n$ and $\tau_{\boldsymbol{T}_l} = \tau_{\boldsymbol{T}_2} = \cdots = \tau_{\boldsymbol{T}_n} > 0$, it is consistent with effect hierarchy. For the covariance function (1.2), the priority orders of β_l's are more diverse. We use an example to illustrate a way of using (3.4) to construct a $\boldsymbol{\Theta}$ with a priority order consistent with effect hierarchy.

EXAMPLE 4.1. Consider the case $n = 3$ and $p_1 = p_2 = p_3 = 2$. To determine a $\boldsymbol{\Theta}$ with a priority order consistent with effect hierarchy, we can construct a set $\boldsymbol{E} = \{\boldsymbol{e} = (e_0, e_1, e_2, e_3) : 0 < e_3 < e_2 < e_1 < e_0\}$, and assign $\sigma_l^2(\boldsymbol{\tau})$ to be e_0 if $|S_l| = 0$ and e_k if $|S_l| = k$, $k = 1, 2, 3$. By (3.4) and (3.6), $\boldsymbol{\tau}$ can be expressed as a function of $\boldsymbol{e}$ as follows: $\tau_{000} = e_0 + 3e_1 + 3e_2 + e_3$, $\tau_{100} = \tau_{010} = \tau_{001} = e_0 + e_1 - e_2 - e_3$, $\tau_{110} = \tau_{101} = \tau_{011} = e_0 - e_1 - e_2 + e_3$ and $\tau_{111} = e_0 - 3e_1 + 3e_2 - e_3$. Hence, $\boldsymbol{\Theta}$ can be chosen as the collection of the $\boldsymbol{\tau}$'s satisfying the above equations for some $\boldsymbol{e} \in \boldsymbol{E}$.

Assigning a $\boldsymbol{\Theta}$ (i.e., choosing a collection of different $\boldsymbol{\tau}$'s) with suitable priority orders is a critical step in the identification of FFDs with least severe aliasing. Different choices of $\boldsymbol{\Theta}$ could lead to different optimal designs. The chosen $\boldsymbol{\Theta}$ reflects the experimenters' belief in what signals deserve more attention, and provides a foundation for comparing the aliasing severity of FFDs. In our experience, when the priority orders of β_l's in $\boldsymbol{\Theta}$ are consistent with some well-accepted assumptions such as the effect heredity principle or the effect hierarchy principle, optimal or good designs for $\boldsymbol{\Theta}$ are generally good for most $\boldsymbol{\tau}$'s, either in $\boldsymbol{\Theta}$ or not in $\boldsymbol{\Theta}$, with priority orders consistent with these principles.

5. Signal aliasing under regular fractional factorial designs. In this section we study the signal-aliasing problem for regular p^{n-m} FFDs, where p is a prime number. In the FEM approach, the aliasing structure of factorial effects under a p^{n-m} design has been extensively studied. In the case of a p^n full factorial design, for any set S of k-factors, the $(p-1)^k$-dimensional space W_S represents the factorial effects of the k factors in S, and can be further decomposed into $(p-1)^{k-1}$ mutually orthogonal subspaces, each associated with an *effect component*, or more commonly referred to as a *word*. Each word carries $p-1$ degrees of freedom and the associated space can be spanned by $p-1$ SOCs, which are called SOCs of the word. For example, under a three-level design, the three-factor interaction $\mathbf{1}\times\mathbf{2}\times\mathbf{3}$ can be decomposed into four effect components represented by the words $\mathbf{123}$, $\mathbf{123}^2$, $\mathbf{12}^2\mathbf{3}$ and $\mathbf{12}^2\mathbf{3}^2$, respectively. Each word has 2 degrees of freedom and two SOCs can be constructed.

Under a p^{n-m} design d, any two words are either fully aliased or mutually orthogonal. The aliasing structure of d is determined by its DCSG, which contains $B_0 = (p^m - 1)/(p-1)$ words. An *alias set* of d is a collection of words that are aliased with one another. There are $B_1 = (p^{n-m} - 1)/(p-1)$ alias sets, each containing p^m words. Readers are referred to Wu and Hamada (2009) and Cheng (2014) for more details on the aliasing structure of regular FFDs.

Let $N = p^{n-m}$. For each alias set $\mathcal{A}$ of d, define $W_{\mathcal{A}}^d$ as the subspace of $\mathbb{R}^N$ that is spanned by $p-1$ $N\times 1$ vectors obtained by restricting a set of SOCs of $\boldsymbol{a}$ to the treatments in d, where $\boldsymbol{a}$ is an arbitrary word in $\mathcal{A}$. An orthogonal basis of $W_{\mathcal{A}}^d$ is called a set of SOCs of $\mathcal{A}$. We note that $W_{\mathcal{A}}^d$ is independent of the choice of $\boldsymbol{a}$; also, SOCs of words are $\Xi\times 1$ vectors, while SOCs of alias sets are $N\times 1$ vectors. For each $1\le i\le B_1$, let $\mathcal{A}_i$ be the ith alias set and, for each $1\le j\le p^m$, let $\boldsymbol{a}_{ij}$ be the jth word in $\mathcal{A}_i$. By Theorem 3.1, all the SOCs of $\boldsymbol{a}_{ij}$ are eigenvectors of $\mathbf{C}_{\boldsymbol{\tau}}$ with the same eigenvalue, say ξ_{ij}. Similarly, let $\boldsymbol{a}_{01},\dots,\boldsymbol{a}_{0B_0}$ be the words in the DCSG of d; then all the SOCs of each $\boldsymbol{a}_{0j}$ are eigenvectors of $\mathbf{C}_{\boldsymbol{\tau}}$ with the same eigenvalue, say ξ_{0j}.

Let $\boldsymbol{Z}_d$ be the $N\times 1$ vector obtained by restricting $\boldsymbol{Z}$ to the treatments in d. The following theorem, which extends Proposition 3.2 in Mitchell, Morris and Ylvisaker (1995), presents the signal aliasing pattern in $\boldsymbol{Z}_d$.

THEOREM 5.1. *Let $\mathbf{C}_{\boldsymbol{\tau}}^d$ be the covariance matrix of $\mathbf{Z}_d$ under* (1.2). *Then*

$$\mathbf{C}_{\boldsymbol{\tau}}^d = \left(\frac{\xi_\varnothing + (p-1)\sum_{j=1}^{B_0} \xi_{0j}}{p^m}\right)\mathbf{P}_0 + \sum_{i=1}^{B_1}\left(\sum_{j=1}^{p^m} \frac{\xi_{ij}}{p^m}\right)\mathbf{P}_i, \tag{5.1}$$

where $\mathbf{P}_0$ is the orthogonal projection matrix onto the space of $N \times 1$ vectors with constant entries, $\xi_\varnothing$ is the eigenvalue of $\mathbf{C}_{\boldsymbol{\tau}}$ associated with the eigenvector with constant entries, and, for each $1 \le i \le B_1$, $\mathbf{P}_i$ is the orthogonal projection matrix onto $W_{\mathcal{A}_i}^d$. Thus $\mathbf{C}_{\boldsymbol{\tau}}^d$ has at most $B_1 + 1$ eigenvalues. The $N \times 1$ vector of ones is an eigenvector whose associated eigenvalue is the average of the eigenvalues of $\mathbf{C}_{\boldsymbol{\tau}}$ associated with the $\Xi \times 1$ vector of ones and a set of SOCs of the words in the DCSG of d. For each $1 \le i \le B_1$, SOCs of $\mathcal{A}_i$ are eigenvectors with the same eigenvalue, which is equal to the average of all the eigenvalues of $\mathbf{C}_{\boldsymbol{\tau}}$ associated with SOCs of the words in $\mathcal{A}_i$.

The following example gives an illustration of the theorem.

EXAMPLE 5.1. Consider an experiment with three 3-level factors $\mathbf{1}$, $\mathbf{2}$ and $\mathbf{3}$. Under the 3^{3-1} design d defined by $I = \mathbf{123}$, the four alias sets are $\mathcal{A}_1 = \{\mathbf{1}, \mathbf{12^23^2}, \mathbf{23}\}$, $\mathcal{A}_2 = \{\mathbf{2}, \mathbf{12^23}, \mathbf{13}\}$, $\mathcal{A}_3 = \{\mathbf{3}, \mathbf{123^2}, \mathbf{12}\}$ and $\mathcal{A}_4 = \{\mathbf{12^2}, \mathbf{23^2}, \mathbf{13^2}\}$. The jth word in $\mathcal{A}_i$, $i = 1, \dots, 4$, is $\boldsymbol{a}_{ij}$; for instance, $\boldsymbol{a}_{11} = \mathbf{1}$, $\boldsymbol{a}_{12} = \mathbf{12^23^2}$, $\boldsymbol{a}_{13} = \mathbf{23}$. For each $S \subseteq \{1, 2, 3\}$, we denote ξ_S by $\xi_{\boldsymbol{i_1 \dots i_k}}$ if $S = \{i_1, \dots, i_k\}$, and $\xi_{\mathbf{0}}$ if $S = \varnothing$. Then by Theorem 5.1, $\mathbf{C}_{\boldsymbol{\tau}}^d$ has at most five distinct eigenvalues: $\frac{1}{3}(\xi_{\mathbf{0}}(\boldsymbol{\tau}) + 2\xi_{\mathbf{123}}(\boldsymbol{\tau}))$ (with the vector of all 1's as an eigenvector) and one eigenvalue for each alias set. For example, the eigenvalue corresponding to the first alias set is $\frac{1}{3}(\xi_{\mathbf{1}}(\boldsymbol{\tau}) + \xi_{\mathbf{123}}(\boldsymbol{\tau}) + \xi_{\mathbf{23}}(\boldsymbol{\tau}))$, with the eigenspace spanned by the vectors obtained by restricting the two SOCs of $\mathbf{1}$ (or $\mathbf{12^23^2}$ or $\mathbf{23}$) to d. The eigenvalues and eigenspaces of the other alias sets can be similarly obtained.

Theorem 5.1 can be interpreted from the random-effect expression of $\boldsymbol{Z}$. By Theorem 5.1 and the argument used to obtain (3.5), we can express $\boldsymbol{Z}_d$ as

$$\boldsymbol{Z}_d = \sum_{s=1}^{N} \beta_s^* \boldsymbol{g}_s^*,$$

where the β_s^*'s are independent normally distributed random variables with zero means and each $\boldsymbol{g}_s^*$ is a vector with constant entries or an SOC of an alias set of d. We normalize each $\boldsymbol{g}_s^*$ to have squared length N. Suppose $\boldsymbol{g}_s^*$ is an SOC of the alias set $\mathcal{A}$. Then $\operatorname{Var}(\beta_s^*) = \xi_{\mathcal{A}}^d(\boldsymbol{\tau})/N$, where $\xi_{\mathcal{A}}^d(\boldsymbol{\tau})$ is the average of the eigenvalues of $\mathbf{C}_{\boldsymbol{\tau}}$ associated with SOCs of the words in $\mathcal{A}$. To link the β_s^*'s to the β_l's in (3.5), we write

$$\mathbf{Z}_d = \mathbf{L}_d \mathbf{Z}, \tag{5.2}$$

where $\mathbf{L}_d$ is the $N \times \Xi$ matrix with the (i, j)th entry being 1 if the ith element of $\mathbf{Z}_d$ is the jth element of $\mathbf{Z}$ and zero otherwise. By (3.5) and (5.2),

$$\beta_s^* = \frac{1}{\|\boldsymbol{g}_s^*\|^2} \boldsymbol{g}_s^{*T} \mathbf{Z}_d = \frac{1}{N} \boldsymbol{g}_s^{*T} (\mathbf{L}_d \mathbf{Z}) = \frac{1}{N} \sum_{l=1}^{\Xi} \beta_l \boldsymbol{g}_s^{*T} (\mathbf{L}_d \boldsymbol{g}_l), \tag{5.3}$$

where $\mathbf{L}_d \boldsymbol{g}_l$ is the vector obtained by restricting $\boldsymbol{g}_l$ to d.

Because words from different alias sets are mutually orthogonal, we have $\boldsymbol{g}_s^{*T} \mathbf{L}_d \boldsymbol{g}_l = 0$ if the corresponding word of $\boldsymbol{g}_l$ is not in $\mathcal{A}$. Then it follows from (5.3) that

$$\beta_s^* = \frac{1}{N} \sum_{\boldsymbol{a} \in \mathcal{A}} \sum_{\boldsymbol{g}_l \widehat{\in} \boldsymbol{a}} (\boldsymbol{g}_s^{*T} \mathbf{L}_d \boldsymbol{g}_l) \beta_l, \tag{5.4}$$

where the notation $\boldsymbol{g}_l \widehat{\in} \boldsymbol{a}$ stands for "$\boldsymbol{g}_l$ is an SOC of the word $\boldsymbol{a}$". The β_s^*'s are called *alias interactions* in Kerr (2001).

Since $\boldsymbol{g}_s^*$ is in $W_{\mathcal{A}}^d$ and, for any $\boldsymbol{a} \in \mathcal{A}$, $\{\mathbf{L}_d \boldsymbol{g}_l : \boldsymbol{g}_l \widehat{\in} \boldsymbol{a}\}$ is an orthogonal basis of $W_{\mathcal{A}}^d$, we have

$$\begin{aligned} \operatorname{Var}(\beta_s^*) &= \frac{1}{N} \sum_{\boldsymbol{a} \in \mathcal{A}} \sum_{\boldsymbol{g}_l \widehat{\in} \boldsymbol{a}} \left(\boldsymbol{g}_s^{*T} \frac{\mathbf{L}_d \boldsymbol{g}_l}{\|\mathbf{L}_d \boldsymbol{g}_l\|} \right)^2 \operatorname{Var}(\beta_l) \\ &= \frac{1}{N} \sum_{\boldsymbol{a} \in \mathcal{A}} \frac{\xi_{\boldsymbol{a}}(\boldsymbol{\tau})}{\Xi} \|\boldsymbol{g}_s^*\|^2 = \sum_{\boldsymbol{a} \in \mathcal{A}} \frac{\xi_{\boldsymbol{a}}(\boldsymbol{\tau})}{\Xi}, \end{aligned} \tag{5.5}$$

where $\xi_{\boldsymbol{a}}(\boldsymbol{\tau})$ is the eigenvalue of $\mathbf{C}_{\boldsymbol{\tau}}$ corresponding to the SOCs of $\boldsymbol{a}$, and the second equality holds since $\operatorname{Var}(\beta_l) = \xi_{\boldsymbol{a}}(\boldsymbol{\tau})/\Xi$ for $\boldsymbol{g}_l \widehat{\in} \boldsymbol{a}$. We note that the conclusion on the eigenvalues in Theorem 5.1 also follows from (5.5) and a similar argument for SOCs of the words in the DCSG of d.

The discussion above provides some insight for understanding the signal aliasing phenomenon in Gaussian random fields. First, a regular FFD d can provide information only on the *joint effects* β_s^*'s, which are linear combinations of the hidden random effects β_l's in the same alias set as shown in (5.4). It is similar to what occurs in using FFDs under FEMs. A major difference is that the joint effects under FEMs are parameters that need to be estimated, while under Gaussian random fields they are random variables with variances depending on some parameters. Second, the aliasing pattern of β_l's in d is determined by the collinear structure of $\mathbf{L}_d \boldsymbol{g}_l$'s. Because d is regular, this structure can be similarly characterized via flats and pencils in a finite geometry as in the FFD theory developed for FEMs. Third, as shown in (5.5), the variance of β_s^* is an accumulation of p^m variances, each contributed by a word. This property is useful in the evaluation of aliasing severity. As discussed in Section 4, $\xi_{\boldsymbol{a}}(\boldsymbol{\tau})/\Xi$ represents the priority of the SOCs of $\boldsymbol{a}$. If an alias set contains too many words (or β_l's) with large variances, then the joint effects of the alias set would tend to have a variance much larger than those of the other alias sets, which indicates severe aliasing in the alias set. A good design is expected to distribute the words with high priorities as uniformly as possible over the alias sets; see more discussions in Section 6.

For a general FFD d (regular or nonregular), the study of signal aliasing in $\boldsymbol{Z}_d$ can be divided into two parts: to investigate the collinear structure of $\mathbf{L}_d \boldsymbol{g}_l$'s and to evaluate the aliasing severity of the collinear structure with the consideration of effect priority, for example, whether the β_l's with large variances are seriously aliased. Because $\mathrm{Var}(\beta_l)$'s are functions of $\boldsymbol{\tau}$, the aliasing severity depends on $\boldsymbol{\tau}$ and d, while the collinear structure only depends on d. A more detailed discussion on the evaluation of the aliasing severity in general FFDs is given in the next section.

6. Aliasing severity index. In this section we propose an aliasing severity index for FFDs under Gaussian random field models, motivated by the idea of minimum aberration. For FFDs with multi-level qualitative factors, the generalized minimum aberration (GMA) criterion proposed by Xu and Wu (2001) is a popular criterion for comparing and assessing aliasing severity of FFDs under FEMs. This criterion, based on the effect hierarchy principle, is to sequentially minimize the terms in a generalized word-length pattern (GWLP). Denote the GWLP of a design d by $\boldsymbol{w}_d = (w_1, \dots, w_n)$. For two n-factor designs d_1 and d_2 with GWLPs $\boldsymbol{w}_{d_1} = (w_{11}, \dots, w_{1n})$ and $\boldsymbol{w}_{d_2} = (w_{21}, \dots, w_{2n})$, respectively, d_1 has less aberration than d_2 if there exists some j, $1 \le j \le n$, such that $w_{1j} < w_{2j}$ and $w_{1i} = w_{2i}$ for $i < j$. This criterion can be implemented by a more general optimization procedure as follows. For a GWLP $\boldsymbol{w}_d$, let $U(\boldsymbol{w}_d) = \sum_{i=1}^n b_i w_i$, where b_i is a quantity assigned to reflect the effect priority of the ith-order effects, with a larger quantity indicating higher priority. The b_i's can be chosen so that d_1 has less aberration than d_2 if and only if $U(\boldsymbol{w}_{d_1}) < U(\boldsymbol{w}_{d_2})$. For example, for p^{n-m} designs, because $w_i \le B_0$, $1 \le i \le n$, minimizing $U(\boldsymbol{w}_d)$ with b_i's satisfying $b_i > B_0 \sum_{j=i+1}^n b_j$, for $i = 1, \dots, n-1$, is equivalent to the GMA criterion.

The quantity $U(\boldsymbol{w}_d)$ can be linked with SOCs of W_S's, $S \subseteq \{1, \dots, n\}$. It was mentioned in Cheng [(2014), page 340] that w_i is proportional to the sum of squares of the inner products between the vector of ones and the vectors obtained by evaluating the SOCs of W_S's with $|S| = i$ on d. In the two-level case, for each $S \subseteq \{1, \dots, n\}$, let $\boldsymbol{g}_S$ be the vector obtained by evaluating the SOC of W_S on the full factorial design. For an N-run design d, let $\mathbf{L}_d$ be as defined in (5.2). Then we have

$$w_i \propto \frac{1}{N^2} \sum_{\substack{S \subseteq \{1,\dots,n\} \\ |S|=i}} (\mathbf{1}_N^T \mathbf{L}_d \boldsymbol{g}_S)^2, \tag{6.1}$$

where $\mathbf{1}_N$ is the vector of N ones. For any pair S' and S'' such that $S' \triangle S'' = S$, where $S' \triangle S'' = (S' \setminus S'') \cup (S'' \setminus S')$ is the symmetric difference of S_1 and S_2, it is known that $\mathbf{1}_N^T \mathbf{L}_d \boldsymbol{g}_S = (\mathbf{L}_d \boldsymbol{g}_{S'})^T (\mathbf{L}_d \boldsymbol{g}_{S''})$. Because the number of such pairs is a constant for any $S \subseteq \{1, \dots, n\}$, we have

$$\sum_{\substack{S', S'' \subseteq \{1,\dots,n\} \\ S' \triangle S'' = S}} \big((\mathbf{L}_d \boldsymbol{g}_{S'})^T (\mathbf{L}_d \boldsymbol{g}_{S''})\big)^2 \propto \big(\mathbf{1}_N^T \mathbf{L}_d \boldsymbol{g}_S\big)^2. \tag{6.2}$$

It follows from (6.1) and (6.2) that $U(\boldsymbol{w}_d)$ is proportional to

$$V(\boldsymbol{w}_d) = \frac{1}{N^2} \sum_{S', S'' \subseteq \{1,\dots,n\}} b_{|S' \triangle S''|} \big((\mathbf{L}_d \boldsymbol{g}_{S'})^T (\mathbf{L}_d \boldsymbol{g}_{S''})\big)^2.$$

In $V(\boldsymbol{w}_d)$, the inner products are affected by the collinear structure of $\mathbf{L}_d \boldsymbol{g}_S$'s, and the b_i's are determined by the effect priority. The former measure *collinear strengths* between any pairs of $\mathbf{L}_d \boldsymbol{g}_{S'}$ and $\mathbf{L}_d \boldsymbol{g}_{S''}$, and the latter are *severity weights* assigned for the aliasing between S' and S''.

Our aliasing severity index is motivated by $V(\boldsymbol{w}_d)$. Consider the model in (1.1) and the random-effect expression $\sum_{l=1}^{\Xi} \beta_l g_l$ of Z in (3.7) with $\mathrm{Var}(\beta_l) = \sigma_l^2(\boldsymbol{\tau})$. Let $\boldsymbol{g}_l$'s and $\mathbf{L}_d \boldsymbol{g}_l$'s, $1 \le l \le \Xi$, be as defined in Section 5 for a general N-run FFD d. Because the priorities of β_l's vary with $\boldsymbol{\tau}$, a reasonable modification of $V(\boldsymbol{w}_d)$ for Z is

$$\frac{1}{N^2} \sum_{1 \le l, l' \le \Xi} b_{l,l'}(\boldsymbol{\tau}) \big((\mathbf{L}_d \boldsymbol{g}_l)^T (\mathbf{L}_d \boldsymbol{g}_{l'})\big)^2,$$

where $b_{l,l'}(\boldsymbol{\tau})$ is a nonnegative function of $\boldsymbol{\tau}$ used to assign a reasonable severity weight to the aliasing between β_l and $\beta_{l'}$ under $\boldsymbol{\tau}$. Because the $\mathrm{Var}(\beta_l)$'s are quantitative assessments of the priorities of the β_l's, we propose to use $b_{l,l'}(\boldsymbol{\tau}) = \sigma_l^2(\boldsymbol{\tau}) \sigma_{l'}^2(\boldsymbol{\tau})$, which assigns greater severity weights to the aliasing of the β_l's with higher priorities. The modification of $V(\boldsymbol{w}_d)$ for Z then becomes

$$\frac{1}{N^2} \sum_{1 \le l, l' \le \Xi} \sigma_l^2(\boldsymbol{\tau}) \sigma_{l'}^2(\boldsymbol{\tau}) \big((\mathbf{L}_d \boldsymbol{g}_l)^T (\mathbf{L}_d \boldsymbol{g}_{l'})\big)^2. \tag{6.3}$$

For each $S \subseteq \{1, \dots, n\}$, let $\mathbf{X}_S$ be the matrix whose columns are the $\mathbf{L}_d \boldsymbol{g}_l$'s associated with W_S. For $S, S' \subseteq \{1, \dots, n\}$, where $S = S'$ is allowed, denote the sum of squares of the entries in $\mathbf{X}_S^T \mathbf{X}_{S'}$ by $\|\mathbf{X}_S^T \mathbf{X}_{S'}\|^2$. Because all the β_l's associated with W_S have variance $\xi_S(\boldsymbol{\tau})/\Xi$, the quantity in (6.3) is proportional to

$$A^*(d; \boldsymbol{\tau}) = \frac{1}{N^2} \sum_{S, S' \subseteq \{1,\dots,n\}} \xi_S(\boldsymbol{\tau}) \xi_{S'}(\boldsymbol{\tau}) \big\|\mathbf{X}_S^T \mathbf{X}_{S'}\big\|^2. \tag{6.4}$$

In $A^*(d; \boldsymbol{\tau})$, $\|\mathbf{X}_S^T \mathbf{X}_{S'}\|^2$ evaluates the collinear strength between W_S and $W_{S'}$ under d, and $\xi_S(\boldsymbol{\tau}) \xi_{S'}(\boldsymbol{\tau})$ assigns a severity weight to the aliasing between W_S and $W_{S'}$. The former depends on d, while the latter depends on $\boldsymbol{\tau}$. A design that minimizes $A^*(d; \boldsymbol{\tau})$ would tend to make the collinear strengths between high-priority W_S's as small as possible, and tend to allow for larger collinear strengths between high-priority and low-priority W_S's and/or between low-priority W_S's.

Some properties of $A^*(d;\boldsymbol{\tau})$ are given below. A criterion for comparing designs with qualitative factors should not depend on the choices of SOCs of W_S's. The following theorem verifies such a property for $A^*(d;\boldsymbol{\tau})$.

THEOREM 6.1. *For any $S, S' \subseteq \{1,\dots,n\}$, where $S = S'$ is allowed, $\|\mathbf{X}_S^T\mathbf{X}_{S'}\|^2$ is invariant to the choices of SOCs of W_S's.*

We present a relationship between the $A^*(d;\boldsymbol{\tau})$ of FFDs and that of the full factorial design.

THEOREM 6.2. *Let D be the full factorial design. For any FFD d,*

$$A^*(d;\boldsymbol{\tau}) \ge A^*(D;\boldsymbol{\tau}).$$

Moreover, $A^*(D;\boldsymbol{\tau}) = \mathrm{tr}(\mathbf{C}_{\boldsymbol{\tau}}^2)$, where $\mathrm{tr}(\cdot)$ is the trace of a matrix.

This theorem states that D has the least severe aliasing for any $\boldsymbol{\tau}$. It allows us to use $A^*(D;\boldsymbol{\tau})$ as a benchmark for the aliasing severity index presented later.

When d is regular, $A^*(d;\boldsymbol{\tau})$ has a more structured form. In the following discussion, the notation $\boldsymbol{\tau}$ in $\xi_S(\boldsymbol{\tau})$ is suppressed for simplicity. For a regular p^{n-m} design, let $S(0, j)$ be the collection of the factors involved in the jth word in the DCSG, $j = 1,\dots,B_0$ and let $S(i, j)$ be the collection of the factors involved in the jth word in the ith alias set, $i = 1,\dots,B_1$, $j = 1,\dots,p^m$. Then we have the following result.

THEOREM 6.3. *For a regular p^{n-m} design d, where p is a prime number,*

$$\begin{aligned} A^*(d;\boldsymbol{\tau}) &= A^*(D;\boldsymbol{\tau}) \\ &\quad + (p-1)(p-2)\sum_{j=1}^{B_0}\xi_{S(0,j)}^2 + 2(p-1)\left[\sum_{j=1}^{B_0}\xi_{\varnothing}\xi_{S(0,j)}\right] \\ &\quad + 2(p-1)^2\left[\sum_{1\le s<t\le B_0}\xi_{S(0,s)}\xi_{S(0,t)}\right] \\ &\quad + 2(p-1)\left[\sum_{i=1}^{B_1}\sum_{1\le s<t\le p^m}\xi_{S(i,s)}\xi_{S(i,t)}\right]. \end{aligned} \tag{6.5}$$

Based on Theorem 6.3, it can be argued that for regular FFDs, minimizing $A^*(d;\boldsymbol{\tau})$ functions in a similar manner to the minimum aberration criterion. In (6.5), the first term $A^*(D;\boldsymbol{\tau})$ is a constant for any d, the second to fourth terms are contributed by the words in the DCSG of d, and the last term is contributed by the words in the alias sets. A regular FFD has small values on the second to fourth terms if the words in its DCSG have small variances. It suggests that the DCSG of a good design should be composed of low-priority words. Let $\xi_i^d = \sum_{j=1}^{p^m} \xi_{S(i,j)}$, $p_{i,j}^d = \xi_{S(i,j)}/\xi_i^d$, $i = 1, \dots, B_1$, $j = 1, \dots, p^m$ and denote the ith alias set of d by $\mathcal{A}_i$. While ξ_i^d is proportional to the variance of the combined effects β_s^*'s of $\mathcal{A}_i$ and can be regarded as a measure of the *joint priority* of $\mathcal{A}_i$, $p_{i,j}^d$ is the proportion of ξ_i^d contributed by the jth word in $\mathcal{A}_i$. The last term in (6.5) evaluates the aliasing severity of the alias sets and can be expressed in terms of ξ_i^d's and $p_{i,j}^d$'s as

$$2(p-1)T^d\left[\sum_{i=1}^{B_1}\left(\xi_i^d/\sqrt{T^d}\right)^2 P_i^d\right], \tag{6.6}$$

where $T^d = \sum_{i=1}^{B_1}(\xi_i^d)^2$ and $P_i^d = \sum_{1\le s<t\le p^m} p_{i,s}^d p_{i,t}^d$. This expression provides us with some clues on how to distribute high-priority words to the alias sets so as to generate a small value of the last term. Although the P_i^d's vary with ξ_i^d's, we can regard (6.6) as a diagonal quadratic form of the ξ_i^d's. Among designs with similar values of T^d, one with large ξ_i^d's paired with small P_i^d's would tend to generate a smaller value of (6.6). Because each P_i^d is a Schur concave function of the $p_{i,j}^d$'s and $\sum_{j=1}^{p^m} p_{i,j}^d = 1$, by the theory of majorization [Marshall and Olkin (1979)], P_i^d would be small if one of the $p_{i,j}^d$'s or a small number of them are much larger than the rest. The discussions above imply that a regular FFD has small $A^*(d;\boldsymbol{\tau})$ if it:

(i) does not assign high-priority words to its DCSG, and

(ii) distributes high-priority words to as many different alias sets as possible, and makes every alias set contain as few high-priority words as possible; this would be achieved, for example, if high-priority words are distributed as uniformly as possible over the alias sets.

Cheng, Steinberg and Sun (1999) pointed out that a two-level minimum aberration regular FFD with resolution at least three would assign the main effects to different alias sets and uniformly distribute as many two-factor interactions as possible over the alias sets not containing main effects. Thus minimizing $A^*(d;\boldsymbol{\tau})$ and minimum aberration share some common principles, with the priority of words (or effects) determined by their variances and word-lengths in the former and latter cases, respectively.

We propose an aliasing severity index for general FFDs under Gaussian random field models as follows.

DEFINITION 6.1. For an FFD d, the aliasing severity index of d for $\boldsymbol{\tau}$ is defined as

$$A(d;\boldsymbol{\tau}) = \frac{A^*(d;\boldsymbol{\tau}) - A^*(D;\boldsymbol{\tau})}{A^*(D;\boldsymbol{\tau})}.$$

A design d is said to have less severe aliasing for $\boldsymbol{\tau}$ than another design d' if $A(d; \boldsymbol{\tau}) < A(d'; \boldsymbol{\tau})$. An N-run design d is said to be optimal for $\boldsymbol{\tau}$ if for any other N-run design d', $A(d; \boldsymbol{\tau}) \leq A(d'; \boldsymbol{\tau})$.

It follows immediately from Theorem 6.2 that $A(d; \boldsymbol{\tau}) \geq 0$ for any d, and $A(D; \boldsymbol{\tau}) = 0$. Ranking designs by $A^*(d; \boldsymbol{\tau})$ is equivalent to ranking them by $A(d; \boldsymbol{\tau})$. Because the index depends on $\boldsymbol{\tau}$, for the covariance functions discussed in this article, a design that is optimal for all $\boldsymbol{\tau}$ might not exist for a given N. Under this circumstance, we can partition the parameter space of $\boldsymbol{\tau}$ into disjoint subsets such that each of them can individually produce a design that is optimal for all $\boldsymbol{\tau}$ in the subset. Another practically acceptable approach is to choose a $\boldsymbol{\tau}$ for which the priority order of β_l's is consistent with the experimental conditions or obeys some commonly acknowledged principles such as effect hierarchy or effect heredity in Definition 4.1, and identify locally optimal designs by minimizing $A(d; \boldsymbol{\tau})$ for the specified $\boldsymbol{\tau}$.

In the rest of this section we revisit the example in Section 2. For the two designs d_1 and d_2, we use $A(d; \boldsymbol{\tau})$ to evaluate their aliasing severity on $\boldsymbol{\Theta}_{\mathrm{ex},1}$ and $\boldsymbol{\Theta}_{\mathrm{ex},2}$ defined respectively in (2.2) and (2.3). The values of $A(d_1; \boldsymbol{\tau}_{\mathrm{ex}})$ and $A(d_2; \boldsymbol{\tau}_{\mathrm{ex}})$ for $\boldsymbol{\tau}_{\mathrm{ex}}$ in $\boldsymbol{\Theta}_{\mathrm{ex},1}$ and $\boldsymbol{\Theta}_{\mathrm{ex},2}$ are plotted in the left panels of Figures 3 and 4, respectively. The plots of $A(d_2; \boldsymbol{\tau}_{\mathrm{ex}}) - A(d_1; \boldsymbol{\tau}_{\mathrm{ex}})$ against τ are also given in the right panels. Figure 3 shows that the minimum aberration design d_1 has smaller $A(d; \boldsymbol{\tau})$ than d_2 for any $\boldsymbol{\tau}_{\mathrm{ex}} \in \boldsymbol{\Theta}_{\mathrm{ex},1}$. This is reasonable because, for any $\boldsymbol{\tau}_{\mathrm{ex}} \in \boldsymbol{\Theta}_{\mathrm{ex},1}$, the priority order of the hidden random effects is consistent with the effect hierarchy principle. In Figure 4, it can be observed that d_2 has smaller $A(d; \boldsymbol{\tau})$ than d_1 when τ is small, and the two curves have a crossover point at $\tau = 0.36$. According to the discussion

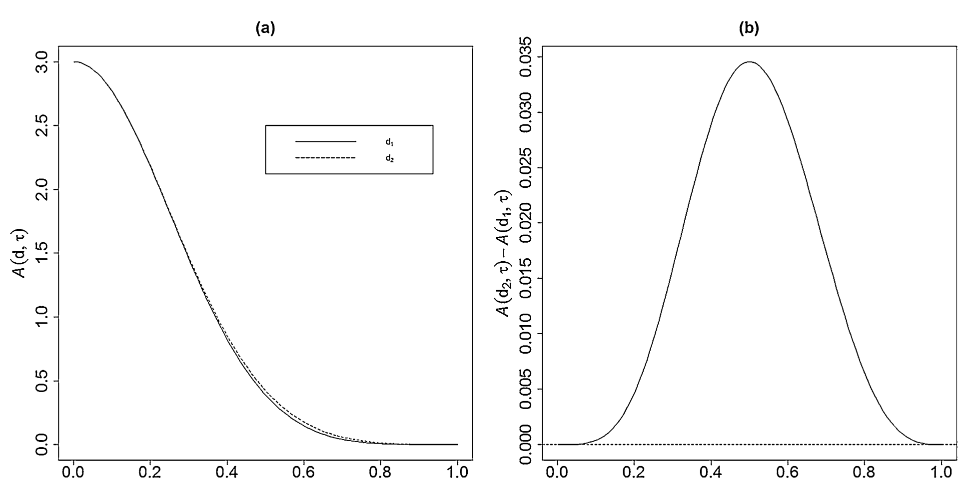

FIG. 3. (a) *Plots of* $A(d; \boldsymbol{\tau})$ *under* d_1 *and* d_2, *and* (b) *difference of* $A(d; \boldsymbol{\tau})$ *between* d_2 *and* d_1, *for* $\boldsymbol{\Theta}_{\mathrm{ex},1}$.

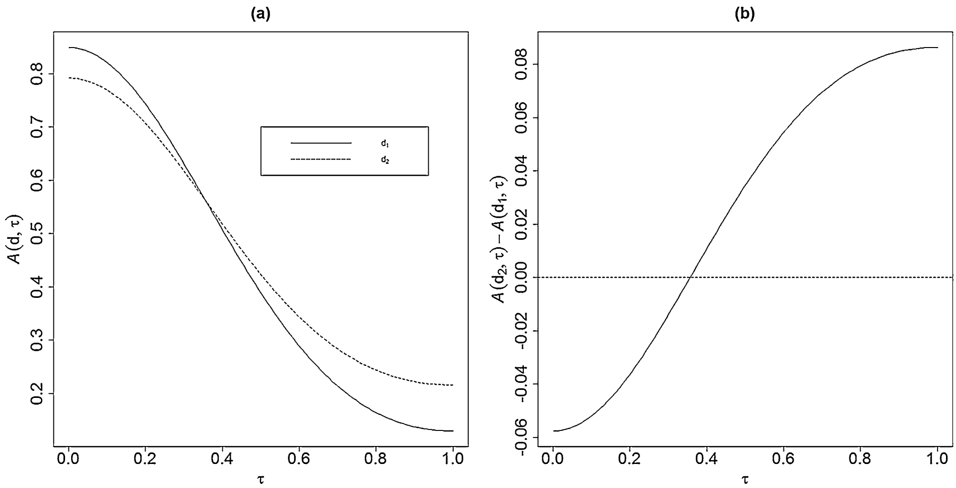

FIG. 4. (a) *Plots of* $A(d; \boldsymbol{\tau})$ *under* d_1 *and* d_2, *and* (b) *difference of* $A(d; \boldsymbol{\tau})$ *between* d_2 *and* d_1, *for* $\boldsymbol{\Theta}_{\mathrm{ex},2}$.

in Section 3, factors **4** and **6** are more important than the other factors when τ is small, and are less important when τ is large. Figure 4 shows that d_2 could have less severe aliasing than d_1 when factors **4** and **6** are more important than the other factors.

By comparing the APVs in Figures 1–2 and the $A(d; \boldsymbol{\tau})$'s in Figures 3–4, we find that APV is highly correlated with $A(d; \boldsymbol{\tau})$ in this case. It seems to suggest that the prediction power is strongly connected with the aliasing severity of signals. A detailed exploration of such connection would be a good direction for future research.

7. Concluding remarks. In this article, a framework is proposed to initiate a systematic investigation of signal-aliasing in Gaussian random fields with covariance functions suitable for experiments with qualitative factors. Via spectral decomposition of the covariance function, hidden random factorial effects are identified as projection scores of the Gaussian random field on the eigenvectors. The study of effect (signal) aliasing under fractional factorial designs is extended from the usual fixed-effects models to random field models.

We argue that the aliasing severity of the signals in a Gaussian random field is determined by (i) the aliasing pattern and (ii) priorities of the random factorial effects. The former is characterized by the collinear structure of the factorial-effect codings and the latter is determined by variances of the random factorial effects. The aliasing pattern depends on the design, while the effect priority depends on model parameters. An index is developed for evaluating the aliasing severity under general fractional factorial designs.

The study of signal-aliasing for Gaussian random fields is still in its early stage. Many problems remain unsolved. For example, although we present examples to demonstrate that the aliasing severity index introduced here is highly correlated with prediction variances, a detailed theoretical investigation is still lacking. Also, it is not clear how the identifiability and accuracy of parameter estimation are affected by the level of signal aliasing. Another interesting topic is the identification of optimal designs for an appropriately chosen subset of the parameters in the covariance structure. Tables of optimal designs for different choices of parameters will be valuable in practical applications.

Last but not the least, it is worth exploring how the concepts, methodologies and results for signal aliasing presented in this article can be generalized to Gaussian random field models for experiments with quantitative factors as well as those that contain some fixed factorial effects. Some difficulties in the former case was raised at the end of Section 3. For the latter, a general form of the Gaussian random field models frequently used in computer experiments [Santner, Williams and Notz (2003)] is

$$Y(\boldsymbol{x}) = \sum_{s=1}^{q} \alpha_s f_s(\boldsymbol{x}) + Z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \tag{7.1}$$

with an extra term compared with model (1.1), where the α_s's are unknown fixed factorial effects with effect codings $f_s(\boldsymbol{x})$. Model (7.1) is also referred to as a *universal kriging model* in the literature of spatial statistics [Cressie (1993)]. Welch et al. (1992) and Joseph, Hung and Sudjianto (2008) mentioned that a collection of wrongly specified fixed factorial effects f_s's in (7.1) can result in worse predictions with large prediction variances. We suspect that it happens because of, at least partially, a high resemblance of the functional surfaces of $Z(\boldsymbol{x})$ and some $f_s(\boldsymbol{x})$'s under the design.

APPENDIX: PROOFS

PROOF OF THEOREM 3.1. For simplicity, the notation $\boldsymbol{\tau}$ in $\xi_S(\boldsymbol{\tau})$ is suppressed. We first prove that the vectors in the space W_S defined in (2.4) are eigenvectors of $\mathbf{C}_{\boldsymbol{\tau}}$ with the same eigenvalue. This is proved in a fashion similar to that of Theorem 10.9 in Bailey (2008), page 196. For each $S \subseteq \{1, \dots, n\}$, let $\mathbf{A}_S$ be the $\Xi \times \Xi$ (0, 1)-matrix such that $[\mathbf{A}_S]_{i,j} = 1$ if and only if the ith and jth treatments have the same level of each factor in S, but different levels of every other factor, where $[\mathbf{A}_S]_{i,j}$ is the (i, j)th entry of $\mathbf{A}_S$. Then we have $\sum_{S\subseteq\{1,\dots,n\}} \mathbf{A}_S = \mathbf{J}$, where $\mathbf{J}$ is a matrix of ones. Let $\tau_S = \tau_{\mathbf{t}}$, where $t_i = 0$ if and only if $i \in S$. Then

$$\mathbf{C}_{\boldsymbol{\tau}} = \sum_{S\subseteq\{1,\dots,n\}} \tau_S \mathbf{A}_S. \tag{A.1}$$

Furthermore, let $\mathbf{J}_S$ be the $\Xi \times \Xi$ (0, 1)-matrix such that $[\mathbf{J}_S]_{i,j} = 1$ if and only if the ith and jth treatments have the same level of each factor in S. From the definitions of $\mathbf{A}_S$ and $\mathbf{J}_S$, it can be seen that $\mathbf{J}_S = \sum_{S^* \subseteq \{1,\dots,n\}: S \subseteq S^*} \mathbf{A}_{S^*}$. Inverting this equation using the Möbius function [Bailey (1996), page 61], we have

$$\mathbf{A}_S = \sum_{\substack{S^* \subseteq \{1,\dots,n\}: \\ S \subseteq S^*}} (-1)^{|S|+|S^*|} \mathbf{J}_{S^*}. \tag{A.2}$$

By substituting (A.2) into (A.1), we get

$$\mathbf{C}_{\boldsymbol{\tau}} = \sum_{S \subseteq \{1,\dots,n\}} \tau_S \Bigg[\sum_{\substack{S^* \subseteq \{1,\dots,n\}: \\ S \subseteq S^*}} (-1)^{|S|+|S^*|} \mathbf{J}_{S^*} \Bigg] = \sum_{S^* \subseteq \{1,\dots,n\}} \eta_{S^*} \mathbf{J}_{S^*}, \tag{A.3}$$

where

$$\eta_{S^*} = \sum_{\substack{S \subseteq \{1,\dots,n\}: \\ S \subseteq S^*}} (-1)^{|S|+|S^*|} \tau_S. \tag{A.4}$$

It is shown in Cheng [(2014), page 245] that $\mathbf{J}_S \boldsymbol{v} = m_S \boldsymbol{v}$ for $\boldsymbol{v} \in V_S$ and $\mathbf{J}_S \boldsymbol{v} = \mathbf{0}$ for $\boldsymbol{v} \in V_S^{\perp}$, where

$$m_S = \prod_{i \notin S} p_i. \tag{A.5}$$

For any $\boldsymbol{v} \in W_S$, by (2.4) and (A.3), we have

$$\mathbf{C}_{\boldsymbol{\tau}} \boldsymbol{v} = \Bigg(\sum_{\substack{S^* \subseteq \{1,\dots,n\}: \\ S \subseteq S^*}} m_{S^*} \eta_{S^*} \Bigg) \boldsymbol{v}. \tag{A.6}$$

So the vectors in W_S are eigenvectors of $\mathbf{C}_{\boldsymbol{\tau}}$ with the eigenvalue

$$\xi_S = \sum_{S^* : S \subseteq S^*} m_{S^*} \eta_{S^*}.$$

Next we prove (3.4). For an $S \subseteq \{1, \dots, n\}$ with $|S| = k$, from (A.4) and (A.6), we have

$$\xi_S = \sum_{\substack{S^* \subseteq \{1,\dots,n\}: \\ S \subseteq S^*}} m_{S^*} \Bigg[\sum_{\substack{T \subseteq \{1,\dots,n\}: \\ T \subseteq S^*}} (-1)^{|S^*|+|T|} \tau_T \Bigg] = \sum_{T \subseteq \{1,\dots,n\}} h_{S,T} \times \tau_T,$$

where

$$h_{S,T} = \sum_{\substack{S^* \subseteq \{1,\dots,n\}: \\ S \subseteq S^*, T \subseteq S^*}} m_{S^*} \times (-1)^{|S^*|+|T|}. \tag{A.7}$$

Without loss of generality, assume that $S = \{1, 2, \dots, k\}$ and T is the subset formed by the first r_1 elements of $\{1, \dots, k\}$ and the first r_2 elements of $\{k+1, \dots, n\}$, where $0 \le r_1 \le k$, $0 \le r_2 \le n-k$. To prove (3.4), it is enough to show that $h_{S,T} = (-1)^{k-r_1} \prod_{i \in \{k+r_2+1,\dots,n\}} (p_i - 1)$.

An S^* with $S \cup T \subseteq S^*$ must be a subset composed of $1, \ldots, k+r_2$, and t elements from $\{k+r_2+1, \ldots, n\}$, where $t = |S^*| - (k+r_2)$. Let D_t be the collection of all the t-element subsets of $\{k+r_2+1, \ldots, n\}$, $0 \le t \le n-k-r_2$. By (A.5) and (A.7), we have

$$\begin{aligned} h_{S,T} &= \sum_{S^* \subseteq \{1,\ldots,n\}: S \cup T \subseteq S^*} m_{S^*} \times (-1)^{|S^*|+|T|} \\ &= \sum_{t=0}^{n-k-r_2} \sum_{S' \in D_t} \left\{ \left[\prod_{i \in \{k+r_2+1,\ldots,n\} \setminus S'} p_i \right] (-1)^{(k+r_2+t)+(r_1+r_2)} \right\} \\ &= (-1)^{k+r_1} \sum_{t=0}^{n-k-r_2} \left\{ (-1)^t \left[\sum_{S' \in D_t} \left(\prod_{i \in \{k+r_2+1,\ldots,n\} \setminus S'} p_i \right) \right] \right\} \\ &= (-1)^{k-r_1} \prod_{i \in \{k+r_2+1,\ldots,n\}} (p_i - 1). \end{aligned}$$

Thus (3.4) is proved.

Finally, we show that there is a one-to-one correspondence between the ξ_S's and τ_t's. For each $S \subseteq \{1, \ldots, n\}$, choose a vector $\boldsymbol{q}_S$ in W_S with the first entry being one. It follows from $\mathbf{C}_{\boldsymbol{\tau}} \boldsymbol{q}_S = \xi_S \boldsymbol{q}_S$ that $\xi_S = \boldsymbol{q}_S^T \boldsymbol{c}_1$, where $\boldsymbol{c}_1^T$ is the first row of $\mathbf{C}_{\boldsymbol{\tau}}$. Because every entry of $\boldsymbol{c}_1$ is an element of $\boldsymbol{\tau}$ and every entry of $\boldsymbol{\tau}$ appears at least once in $\boldsymbol{c}_1$, there exists a matrix $\mathbf{B}$ of full column rank such that $\boldsymbol{c}_1 = \mathbf{B}\boldsymbol{\tau}$. Then $\xi_S = \boldsymbol{q}_S^T \mathbf{B}\boldsymbol{\tau}$. Denote the 2^n subsets of $\{1, \ldots, n\}$ by S_i, $i = 1, \ldots, 2^n$, and let $\boldsymbol{\xi} = (\xi_{S_1}, \ldots, \xi_{S_{2^n}})^T$ and $\mathbf{Q} = (\boldsymbol{q}_{S_1}, \ldots, \boldsymbol{q}_{S_{2^n}})^T$. Then $\boldsymbol{\xi} = \mathbf{QB}\boldsymbol{\tau}$. Because $\boldsymbol{q}_{+S_i} \in W_{S_i}$, $i = 1, \ldots, 2^n$, and the W_{S_i}'s are mutually orthogonal, $\mathbf{Q}$ has full row rank and the rank of $\mathbf{QB}$ equals the column rank of $\mathbf{B}$ [Seber (2008), page 37]. It implies that $\mathbf{QB}$ has full rank and $\boldsymbol{\tau} = (\mathbf{QB})^{-1}\boldsymbol{\xi}$, which assures the one-to-one correspondence. $\square$

PROOF OF THEOREM 5.1. By Theorem 3.1, $\mathbf{C}_{\boldsymbol{\tau}}$ can be eigen-decomposed as $\mathbf{C}_{\boldsymbol{\tau}} = \frac{1}{p^n} \mathbf{E} \boldsymbol{\Lambda}_{\boldsymbol{\tau}} \mathbf{E}^T$, where $\mathbf{E}$ is formed by SOCs of words. We use the alias sets to partition $\mathbf{E}$ into several submatrices as follows. For the jth word $\boldsymbol{a}_{ij}$ in $\mathcal{A}_i$, let $\mathbf{E}_{ij}$ be the matrix formed by the $p-1$ columns of $\mathbf{E}$ associated with $\boldsymbol{a}_{ij}$. Similarly, for the jth word $\boldsymbol{a}_{0j}$ in the DCSG of d, let $\mathbf{E}_{0j}$ be the matrix formed by the $p-1$ columns of $\mathbf{E}$ associated with $\boldsymbol{a}_{0j}$. Let $\mathbf{E}_{00}$ be the column of p^n ones. Then we have

$$\mathbf{C}_{\boldsymbol{\tau}} = \xi_{\varnothing} \mathbf{P}_{00} + \sum_{j=1}^{B_0} \xi_{0j} \mathbf{P}_{0j} + \sum_{i=1}^{B_1} \sum_{j=1}^{p^m} \xi_{ij} \mathbf{P}_{ij},$$

where $\mathbf{P}_{ij} = \frac{1}{p^n} \mathbf{E}_{ij} \mathbf{E}_{ij}^T$ for any $i, j \ge 0$. Because $\mathbf{E}_{00}^T \mathbf{E}_{00} = p^n$ and $\mathbf{E}_{ij}^T \mathbf{E}_{ij} = p^n \mathbf{I}_{p-1}$ if $(i, j) \ne (0, 0)$, $\mathbf{P}_{ij}$ is the orthogonal projection matrix onto the column space of $\mathbf{E}_{ij}$.

Let $\mathbf{L}_d$ be as defined in (5.2). Since $\mathbf{C}_{\boldsymbol{\tau}}^d = \mathbf{L}_d \mathbf{C}_{\boldsymbol{\tau}} \mathbf{L}_d^T$, we have

$$\begin{aligned}\mathbf{C}_{\boldsymbol{\tau}}^d = {} & \frac{\xi_\varnothing}{p^m}(p^m \mathbf{L}_d \mathbf{P}_{00} \mathbf{L}_d^T) + \sum_{j=1}^{B_0} \frac{\xi_{0j}}{p^m}(p^m \mathbf{L}_d \mathbf{P}_{0j} \mathbf{L}_d^T) \\ & + \sum_{i=1}^{B_1} \sum_{j=1}^{p^m} \frac{\xi_{ij}}{p^m}(p^m \mathbf{L}_d \mathbf{P}_{ij} \mathbf{L}_d^T).\end{aligned} \tag{A.8}$$

By the discussions above, $p^m \mathbf{L}_d \mathbf{P}_{00} \mathbf{L}_d^T$ in the first term and $p^m \mathbf{L}_d \mathbf{P}_{ij} \mathbf{L}_d^T$ in the third term are orthogonal projection matrices onto the column spaces of $\mathbf{L}_d \mathbf{E}_{00}$ and $\mathbf{L}_d \mathbf{E}_{ij}$, respectively. For the second term, let $[\mathbf{E}_{00}\ \mathbf{E}_{0j}]$ be the matrix obtained by adjoining $\mathbf{E}_{0j}$ to the right of $\mathbf{E}_{00}$. Because $\boldsymbol{a}_{0j}$ is a word in the DCSG and all the rows in $[\mathbf{E}_{00}\ \mathbf{E}_{0j}]$ have the same length $1/\sqrt{p}$, the following identity holds:

$$\big(\mathbf{L}_d[\mathbf{E}_{00}\ \mathbf{E}_{0j}]\big)\big(\mathbf{L}_d[\mathbf{E}_{00}\ \mathbf{E}_{0j}]\big)^T = p(\mathbf{L}_d \mathbf{E}_{00})(\mathbf{L}_d \mathbf{E}_{00})^T.$$

It follows that

$$p^m \mathbf{L}_d \mathbf{P}_{0j} \mathbf{L}_d^T = (p-1)\big(p^m \mathbf{L}_d \mathbf{P}_{00} \mathbf{L}_d^T\big). \tag{A.9}$$

We have $\mathbf{P}_0 = p^m \mathbf{L}_d \mathbf{P}_{00} \mathbf{L}_d^T$. For $\boldsymbol{a}_{ij}, \boldsymbol{a}_{ij'} \in \mathcal{A}_i$, $\mathbf{L}_d \mathbf{E}_{ij}$ and $\mathbf{L}_d \mathbf{E}_{ij'}$ have the same column space $W_{\mathcal{A}_i}^d$ and

$$p^m \mathbf{L}_d \mathbf{P}_{ij} \mathbf{L}_d^T = p^m \mathbf{L}_d \mathbf{P}_{ij'} \mathbf{L}_d^T. \tag{A.10}$$

Now (5.1) follows from (A.8)–(A.10). □

PROOF OF THEOREM 6.1. For each $S \subseteq \{1, \dots, n\}$, let $\mathbf{E}_S$ and $\mathbf{E}_S^*$ be the matrices whose columns are the vectors obtained by respectively evaluating two sets of SOCs on the full factorial design. Because the columns of $\mathbf{E}_S$ and $\mathbf{E}_S^*$ are orthogonal bases of the same space, there exists an orthogonal matrix $\boldsymbol{\Gamma}$ such that $\mathbf{E}_S = \mathbf{E}_S^* \boldsymbol{\Gamma}$. For a design d, let $\mathbf{X}_S = \mathbf{L}_d \mathbf{E}_S$ and $\mathbf{X}_S^* = \mathbf{L}_d \mathbf{E}_S^*$, where $\mathbf{L}_d$ is as defined in (5.2). Then $\mathbf{X}_S = \mathbf{X}_S^* \boldsymbol{\Gamma}$ and $\mathbf{X}_S \mathbf{X}_S^T = \mathbf{X}_S^* \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T \mathbf{X}_S^{*T} = \mathbf{X}_S^* \mathbf{X}_S^{*T}$. For any $S, S' \subseteq \{1, \dots, n\}$, we have $\|\mathbf{X}_S^T \mathbf{X}_{S'}\|^2 = \operatorname{tr}((\mathbf{X}_S^T \mathbf{X}_{S'})^T (\mathbf{X}_S^T \mathbf{X}_{S'})) = \operatorname{tr}(\mathbf{X}_{S'} \mathbf{X}_{S'}^T \mathbf{X}_S \mathbf{X}_S^T) = \operatorname{tr}(\mathbf{X}_{S'}^* \mathbf{X}_{S'}^{*T} \mathbf{X}_S^* \mathbf{X}_S^{*T}) = \|\mathbf{X}_S^{*T} \mathbf{X}_{S'}^*\|^2$. □

PROOF OF THEOREM 6.2. For $1 \le l, l' \le \Xi$, $\boldsymbol{g}_l^T \boldsymbol{g}_{l'} = \Xi$ if $l = l'$ and zero otherwise. By Theorem 3.1, we have $A^*(D; \boldsymbol{\tau}) = \sum_{l=1}^{\Xi} \xi_l(\boldsymbol{\tau})^2 = \operatorname{tr}(\mathbf{C}_{\boldsymbol{\tau}}^2)$, where ξ_l is the eigenvalue of $\mathbf{C}_{\boldsymbol{\tau}}$ corresponding to $\boldsymbol{g}_l$. The rest of the proof is done by considering a particular set of complex SOCs, introduced in Bailey (1982) and extensively used in Xu and Wu (2001). For the ith factor, define the codings of its complex SOCs on a design point $\boldsymbol{x} = (x_1, \dots, x_n)$ as $c_s^i(\boldsymbol{x}) = \zeta^{s x_i}$, $0 \le s \le p_i - 1$, where $\zeta = \exp(2\pi \mathrm{i}/p_i)$ and $\mathrm{i} = \sqrt{-1}$. It follows that $\overline{c_s^i(\boldsymbol{x})} c_t^i(\boldsymbol{x}) = c_{t-s}^i(\boldsymbol{x})$

and $\overline{c_s^i(\boldsymbol{x})}c_s^i(\boldsymbol{x}) = 1$. The codings of complex SOCs for a W_S, $S \subseteq \{1, \ldots, n\}$, are defined via c_s^i's as in (2.5), that is, for $\boldsymbol{s} = (s_1, \ldots, s_n) \in D$, the coding is $c_{\boldsymbol{s}}(\boldsymbol{x}) = \prod_{i=1}^n c_{s_i}^i(\boldsymbol{x})$. It follows that $\overline{c_{\boldsymbol{s}}(\boldsymbol{x})}c_{\boldsymbol{t}}(\boldsymbol{x}) = c_{\boldsymbol{t}-\boldsymbol{s}}(\boldsymbol{x})$ and

$$\overline{c_{\boldsymbol{s}}(\boldsymbol{x})}c_{\boldsymbol{s}}(\boldsymbol{x}) = 1. \tag{A.11}$$

We can define $\mathbf{X}_S$ and $\mathbf{L}_d \boldsymbol{g}_l$ for complex SOCs in the same way as before and, following the same arguments in the proof of Theorem 6.1, we can show that $\|\overline{\mathbf{X}_S}^T \mathbf{X}_{S'}\|^2$ does not depend on the choices of SOCs, where $\overline{\mathbf{X}_S}$ is the complex conjugate of $\mathbf{X}_S$ and $\|\cdot\|^2$ computes the sum of squares of the absolute values of the entries in a complex matrix. It suffices to show that $A^*(d;\boldsymbol{\tau}) \ge A^*(D;\boldsymbol{\tau})$ under complex SOCs for any FFD d. Since $(\overline{\mathbf{L}_d \boldsymbol{g}_l})^T(\mathbf{L}_d \boldsymbol{g}_l) = N$ by (A.11), $1 \le l \le \Xi$, we have

$$A^*(d;\boldsymbol{\tau}) = \sum_{l=1}^{\Xi} \xi_l(\boldsymbol{\tau})^2 + \frac{1}{N^2}\left\{\sum_{1\le l\ne l'\le \Xi} \left|(\overline{\mathbf{L}_d \boldsymbol{g}_l^*})^T(\mathbf{L}_d \boldsymbol{g}_{l'}^*)\right|^2 \xi_l(\boldsymbol{\tau})\xi_{l'}(\boldsymbol{\tau})\right\}.$$

The result follows from the fact that $A^*(D;\boldsymbol{\tau}) = \sum_{l=1}^{\Xi} \xi_l(\boldsymbol{\tau})^2$. $\square$

PROOF OF THEOREM 6.3. For a regular p^{n-m} design d, where p is a prime number, let $\boldsymbol{a}_{ij}$ $i, j \ge 0$, be as defined in the proof of Theorem 5.1, $\mathbf{E}_{ij}$ be as defined in the proof of Theorem 5.1, $\mathbf{L}_d$ be as defined in (5.2), and let $\mathbf{X}_{ij} = \mathbf{L}_d \mathbf{E}_{ij}$. Because $\boldsymbol{a}_{ij}$ and $\boldsymbol{a}_{i'j'}$ with $i \ne i'$ are orthogonal in d, we have $\|\mathbf{X}_{ij}^T \mathbf{X}_{i'j'}\|^2 = 0$ if $i \ne i'$. Then

$$\begin{aligned} A^*(d;\boldsymbol{\tau}) = \frac{1}{N^2}\Bigg(&\sum_{0\le s,t\le B_0} \|\mathbf{X}_{0s}^T \mathbf{X}_{0t}\|^2 \xi_{S(0,s)}\xi_{S(0,t)} \\ &+ \sum_{i=1}^{B_1} \sum_{1\le s,t\le p^m} \|\mathbf{X}_{is}^T \mathbf{X}_{it}\|^2 \xi_{S(i,s)}\xi_{S(i,t)}\Bigg). \end{aligned} \tag{A.12}$$

Following the same arguments used in the proof of Theorem 6.1, it can be shown that $\|\mathbf{X}_{ij}^T \mathbf{X}_{ij'}\|^2$ is also invariant to the choices of SOCs. Because $\mathbf{X}_{ij}$ and $\mathbf{X}_{ij'}$ have the same column space, we can arbitrarily choose a set of SOCs so that

$$\|\mathbf{X}_{ij}^T \mathbf{X}_{ij'}\|^2 = \begin{cases} N^2 & \text{if } i = j = j' = 0, \\ (p-1)N^2 & \text{if } i = 0 \text{ and only one of } j \text{ and } j' \text{ is } 0, \\ (p-1)^2 N^2 & \text{if } i = 0 \text{ and } j, j' > 0, \\ (p-1)N^2 & \text{if } i \ne 0. \end{cases}$$

Substituting these values into (A.12), we get

$$A^*(d;\boldsymbol{\tau}) = \xi_{\varnothing}^2 + \left[(p-1)^2 \sum_{j=1}^{B_0} \xi_{S(0,j)}^2\right] + \left[(p-1)\sum_{i=1}^{B_1}\sum_{j=1}^{p^m} \xi_{S(i,j)}^2\right]$$
$$+ 2(p-1)\left\{\left[\sum_{j=1}^{B_0} \xi_{\varnothing}\xi_{S(0,j)}\right] + \left[(p-1)\sum_{1\le s<t\le B_0} \xi_{S(0,s)}\xi_{S(0,t)}\right]\right.$$
$$\left.+\left[\sum_{i=1}^{B_1}\sum_{1\le s<t\le p^m} \xi_{S(i,s)}\xi_{S(i,t)}\right]\right\}.$$

By Theorem 6.2, $\xi_{\varnothing}^2 + (p-1)[\sum_{j=1}^{B_0} \xi_{S(0,j)}^2 + \sum_{i=1}^{B_1}\sum_{j=1}^{p^m} \xi_{S(i,j)}^2] = \mathrm{tr}(\mathbf{C}_{\boldsymbol{\tau}}^2) = A^*(D;\boldsymbol{\tau})$. Then Theorem 6.3 follows. $\square$

Acknowledgments. We are grateful to the referees for providing helpful comments and constructive suggestions, which led to an improved paper.

REFERENCES

BAILEY, R. A. (1982). The decomposition of treatment degrees of freedom in quantitative factorial experiments. *J. Roy. Statist. Soc. Ser. B* **44** 63–70. MR0655375

BAILEY, R. A. (1996). Orthogonal partitions in designed experiments. *Des. Codes Cryptogr.* **8** 45–77. MR1393974

BAILEY, R. A. (2008). *Design of Comparative Experiments*. *Cambridge Series in Statistical and Probabilistic Mathematics* **25**. Cambridge Univ. Press, Cambridge. MR2422352

CHANG, M. C. (2015). Effect aliasing in Gaussian random field models. Ph.D. dissertation, Institute of Statistics, National Tsing Hua Univ.

CHANG, M. C. and CHENG, S. W. (2018). Signal aliasing in Gaussian random fields for experiments with quantitative factors. *Ann. Statist.* To appear.

CHENG, C. S. (2014). *Theory of Factorial Design*: *Single- and Multi-Stratum Experiments*. CRC Press, Boca Raton, FL.

CHENG, C.-S., STEINBERG, D. M. and SUN, D. X. (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 85–93. MR1664104

CHENG, S.-W. and YE, K. Q. (2004). Geometric isomorphism and minimum aberration for factorial designs with quantitative factors. *Ann. Statist.* **32** 2168–2185. MR2102507

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. Wiley, New York. Revised reprint of the 1991 edition. MR1239641

DENG, X., HUNG, Y. and LIN, C. D. (2015). Design for computer experiments with qualitative and quantitative factors. *Statist. Sinica* **25** 1567–1581. MR3409081

FRIES, A. and HUNTER, W. G. (1980). Minimum aberration 2^{k-p} designs. *Technometrics* **22** 601–608. MR0596803

HAN, G., SANTNER, T. J., NOTZ, W. I. and BARTEL, D. L. (2009). Prediction for computer experiments having quantitative and qualitative input variables. *Technometrics* **51** 278–288. MR2751072

JOSEPH, V. R. (2006). A Bayesian approach to the design and analysis of fractionated experiments. *Technometrics* **48** 219–229. MR2277676

JOSEPH, V. R. and DELANEY, J. D. (2007). Functionally induced priors for the analysis of experiments. *Technometrics* **49** 1–11. MR2345447

JOSEPH, V. R., HUNG, Y. and SUDJIANTO, A. (2008). Blind kriging: A new method for developing metamodels. *J. Mech. Des.* **130** 1–8.

KERR, M. K. (2001). Bayesian optimal fractional factorials. *Statist. Sinica* **11** 605–630. MR1863153

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications. Mathematics in Science and Engineering* **143**. Academic Press, New York. MR0552278

MITCHELL, T. J., MORRIS, M. D. and YLVISAKER, D. (1995). Two-level fractional factorials and Bayesian prediction. *Statist. Sinica* **5** 559–573. MR1347607

QIAN, P. Z. G., WU, H. and WU, C. F. J. (2008). Gaussian process models for computer experiments with qualitative and quantitative factors. *Technometrics* **50** 383–396. MR2457574

SANTNER, T. J., WILLIAMS, B. J. and NOTZ, W. I. (2003). *The Design and Analysis of Computer Experiments*. Springer, New York. MR2160708

SEBER, G. A. F. (2008). *A Matrix Handbook for Statisticians*. Wiley, Hoboken, NJ. MR2365265

STEINBERG, D. M. and BURSZTYN, D. (2004). Data analytic tools for understanding random field regression models. *Technometrics* **46** 411–420. MR2101509

WELCH, W. J., BUCK, R. J., SACKS, J., WYNN, H. P., MITCHELL, T. J. and MORRIS, M. D. (1992). Screening, predicting and computer experiments. *Technometrics* **34** 15–25.

WU, C. F. J. and HAMADA, M. S. (2009). *Experiments: Planning, Analysis, and Optimization*, 2nd ed. Wiley, Hoboken, NJ. MR2583259

XU, H. and WU, C. F. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs [corrected republication of MR1863969]. *Ann. Statist.* **29** 1066–1077. MR1869240

ZHOU, Q., QIAN, P. Z. G. and ZHOU, S. (2011). A simple approach to emulation for computer models with qualitative and quantitative factors. *Technometrics* **53** 266–273. MR2857704

M.-C. CHANG
GRADUATE INSTITUTE OF STATISTICS
NATIONAL CENTRAL UNIVERSITY
TAIPEI
TAIWAN 11529
E-MAIL: mcchang0131@gmail.com

S.-W. CHENG
INSTITUTE OF STATISTICS
NATIONAL TSING HUA UNIVERSITY
HSINCHU
TAIWAN 30013
E-MAIL: swcheng@stat.nthu.edu.tw

C.-S. CHENG
INSTITUTE OF STATISTICS
ACADEMIA SINICA
TAIPEI
TAIWAN 11529
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELY, CALIFORNIA 94720
USA
E-MAIL: cheng@stat.sinica.edu.tw
cheng@stat.berkeley.edu