Adaptive One-Factor-at-a-Time Experimentation and Expected Value of Improvement

Daniel D. FREY

Department of Mechanical Engineering and
Engineering Systems Division
Massachusetts Institute of Technology
Cambridge, MA 02139
(*danfrey@mit.edu*)

Hungjen WANG

Engineering Systems Division
Massachusetts Institute of Technology
Cambridge, MA 02139
(*hjwang@mit.edu*)

This article concerns adaptive experimentation as a means for making improvements in design of engineering systems. A simple method for experimentation, called "adaptive one-factor-at-a-time," is described. A mathematical model is proposed and theorems are proven concerning the expected value of the improvement provided and the probability that factor effects will be exploited. It is shown that adaptive one-factor-at-a-time provides a large fraction of the potential improvements if experimental error is not large compared with the main effects and that this degree of improvement is more than that provided by resolution III fractional factorial designs if interactions are not small compared with main effects. The theorems also establish that the method exploits two-factor interactions when they are large and exploits main effects if interactions are small. A case study on design of electric-powered aircraft supports these results.

KEY WORDS: Adaptive experimentation; Design of experiments; Fractional factorial design; One factor at a time.

1. MOTIVATION

Engineering designers engage in experimentation for various reasons, including:

- To estimate the performance of a system at a single point in a design space (e.g., to test whether a design satisfies a requirement on performance)
- To learn about a design space (e.g., to estimate the effects of a set of design parameters on a set of design performance criteria)
- To create models (to build, calibrate, and/or test equations, simulations, and physical analogs)
- To seek improvements in the performance of the design.

Although all of these goals are important, this article places emphasis on the last one: seeking improvements in the design. Parameter estimation, learning, and modeling are all helpful in contributing to improvement, but it is ultimately the improvements in the systems that justify the investments, not the learning or models per se. Therefore, methods of experimentation should be assessed on the basis of how well they serve to improve the system in light of realistic uncertainties. This article seeks to develop theory to help designers choose among experimentation processes, especially when expected value of improvement is the primary evaluation criterion.

In seeking a higher expected value of improvement, this article emphasize's adaptive experimentation processes. Methodology for adaptive experimentation is not new. For example, response surface methodology (RSM) is a highly developed and widely practiced set of techniques. The key ideas behind RSM were first published by Box and Wilson (1951). The development of new concepts and methods proceeded rapidly, as documented in such review articles as those of Hill and Hunter (1966) and Myers, Khuri, and Carter (1989). RSM continues to be an active area of research. Its current role has been summarized by Wu and Hamada (2000): "If the input factors are quantitative and there are only a few of them, response surface methodology is an effective tool for studying this relationship." The reason that RSM is typically limited to a small number of variables is that many of the techniques used within RSM scale poorly. A motivation of this article is to explore a method that scales more favorably to a large number of variables.

Notwithstanding the substantial body of literature on RSM, there has not been much research on a specific aspect of adaptive experimentation that we seek to explore: how adaptation of an experimental plan based on data promotes the overall success of the investigation. Box (1999) made two key points on this subject: "There should be more studies of statistics from the dynamic point of view," and that "it is the investigation itself that must be regarded as the unit, and the success of the investigation that must be regarded as the objective." Inspired by these ideas, we set a goal of developing mathematical results regarding adaptive methods and their relationship to expected improvement. We chose to investigate adaptive one-factor-at-a-time (OFAT) experimentation principally because the immediate use of information to adjust the experimental plan provides insight into the benefits of adaptation. More complex adaptive experimentation processes may provide greater benefits, but may not illustrate the lessons as clearly.

2. BACKGROUND ON ONE–FACTOR–AT–A–TIME EXPERIMENTATION

In pursuing the study of adaptive experimentation, this article examines a simple process in which only one factor is changed from trial to trial. OFAT experimentation is generally discouraged in the literature on experimental design and quality

TECHNOMETRICS, AUGUST 2006, VOL. 48, NO. 3
DOI 10.1198/004017006000000075

improvement (Box, Hunter, and Hunter 1978; Logothetis and Wynn 1994; Czitrom 1999; Wu and Hamada 2000). Reasons cited for this include that (1) it requires more runs for the same precision in effect estimation, (2) it cannot estimate some interactions; (3) the conclusions from its analysis are not general, (4) it can miss optimal settings of factors, (5) OFAT can be susceptible to bias due to time trends.

Although these cautions are valid and should be taken into account in considering the use of OFAT, some researchers have articulated a role for OFAT and demonstrated that it has some advantages under some conditions. Friedman and Savage (1974) suggested that a OFAT approach might be preferred over balanced factorial plans when the experimenter seeks an optimum within a system likely to contain interactions. They suggested that OFAT might offer advantages because it concentrates observations in regions likely to contain the optimum. Daniel (1973) suggested that OFAT may be preferred when an experimenter wishes to react more quickly to data and can be safely used in those cases in which factor effects are three or four times the standard deviation due to pure experimental error. Koita (1994) showed that an OFAT method was effective in identifying selected interactions after running fractional factorial designs as part of an overall approach to sequential experimentation. McDaniel and Ankenman (2000) provided empirical evidence that for "small factor change problems," a strategy using OFAT and Box–Behnken designs often worked better than a comparable strategy using fractional factorial designs when there is no error in the response. Qu and Wu (2005) used OFAT techniques to construct resolution V designs with economical run size.

This article analyzes a specific process called "adaptive OFAT." The process is illustrated in Figure 1 for the case of three two-level factors labeled A, B, and C. First, an experiment is conducted at some baseline point in the design space. In Figure 1 this baseline point is $A = -1$, $B = +1$, $C = +1$, but any of the eight points in the space might have been chosen. Next, a factor is varied and another experiment is run. In Figure 1 factor A is varied first, but B or C also could have been chosen. If the experimental results suggest improvement in the response, then the change in factor A is retained. Thus all future experiments are affected by the results of the first two experiments. Next, another factor is changed; in Figure 1, C is toggled from $+1$ to -1. Another experiment is conducted and compared with the best result observed so far. If the most recent change does not seem favorable, then it is reversed before proceeding. The process ends when all of the factors have been changed. The final settings are determined by the best observation in the sequence.

Frey, Engelhardt, and Greitzer (2003) have shown by a meta-analysis of 66 datasets that adaptive OFAT can provide a large fraction of the potentially attainable improvement if the pure experimental error is not very large. This study also showed that under some conditions, adaptive OFAT, can be preferable to factorial designs that use the same number of experiments as adaptive OFAT, especially if interactions are not small. That study demonstrated the phenomenon that are explored from a theoretical perspective in this article.

It should be emphasized that although this article seeks to examine the potential benefits of adaptive OFAT, in many circumstances adaptive OFAT would be unsuitable, such as when the duration of a single experiment is long, when parallel experimentation is much more economical than sequential experimentation, when time trends exert a large influence making randomization essential, or when significantly more experiments are warranted than OFAT requires.

3. A MATHEMATICAL MODEL

This section presents a mathematical model useful for representing adaptive OFAT experimentation. A mathematical model of a response of an engineering system that captures effect hierarchy in a simple way is then proposed. Finally, the concept of "exploiting" factor effects is introduced.

3.1 A Model of Adaptive OFAT

We assume that there are n experimental factors, there is a response y that is a function of the experimental factors $(x_1, x_2, \ldots, x_n)$, and that the experimenter seeks to increase the response. Further, assume that the experimental factors have two levels each, coded as $x_i \in \{-1, +1\}$ for all $i \in 1, \ldots, n$. The adaptive OFAT process starts with values of the experimental factors denoted by $\tilde{x}_i$ and ends with values denoted by x_i^* as follows:

$$O_0 = y(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n), \tag{1}$$

$$O_1 = y(-\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n), \tag{2}$$

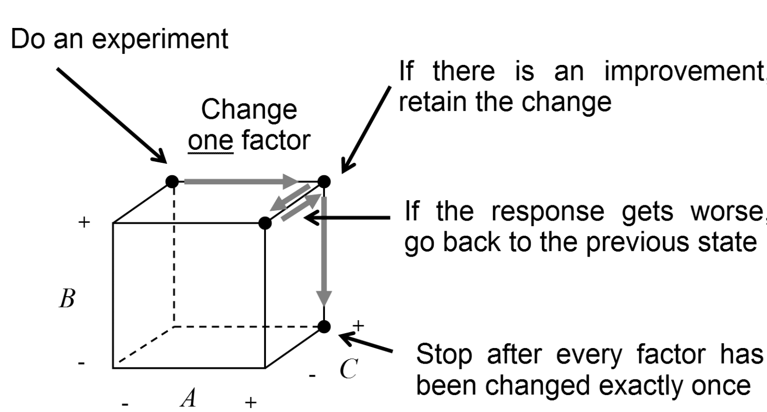

Figure 1. Adaptive OFAT as Applied to a System With Three Two-Level Factors (A, B, and C).

$$x_1^* = \tilde{x}_1 \operatorname{Sgn}(O_0 - O_1), \tag{3}$$

$$O_i = y(x_1^*, \ldots, x_{i-1}^*, -\tilde{x}_i, \tilde{x}_{i+1}, \ldots, \tilde{x}_n) \quad \text{for } i = 2, \ldots, n, \tag{4}$$

$$x_i^* = \tilde{x}_i \operatorname{Sgn}[\max(O_0, O_1, \ldots, O_{i-1}) - O_i] \quad \text{for } i = 2, \ldots, n. \tag{5}$$

The adaptive OFAT procedure begins with a baseline observation, O_0 of the response y at the starting set of values $(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$. Then the experimenter proceeds to toggle the coded level of the first variable and makes an observation O_1. The final value of the first factor, x_1^*, is determined by taking the difference between the observations O_0 and O_1 in such a way that the toggled value of the factor is retained if and only if it is associated with an apparent improvement in the response y. The final values of all subsequent variables are similarly set by taking one new observation and comparing it with a previously established baseline that is the highest response observed so far.

In (3) and (5) and throughout the article $\operatorname{Sgn}(x)$ is defined so that it equals -1 if $x \le 0$ and $+1$ otherwise. This is slightly different from a conventional definition, in which Sgn(0) is defined as 0 (Korn and Korn 1961); however, in adaptive OFAT the factor levels must be set to one of the two possible levels in all cases, making this modification necessary. This distinction does not affect the theorems presented here.

Adaptive OFAT requires sampling a very small fraction of the space of possible factor settings. For a system with n two-level factors, adaptive OFAT requires $n + 1$ experiments, and there are 2^n different design alternatives. Therefore, the method requires sampling about 20% of the design space for $n = 5$, and the fraction drops quickly as n rises.

3.2 A Mathematical Model of Factor Effects

We now define a mathematical model of a response y that includes main effects and two-factor interactions:

$$y(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j + \varepsilon_k, \tag{6}$$

$$\varepsilon_k \sim \mathrm{N}(0, \sigma_\varepsilon^2), \tag{7}$$

$$\beta_i \sim \mathrm{N}(0, \sigma_{\mathrm{ME}}^2), \tag{8}$$

$$\beta_{ij} \sim \mathrm{N}(0, \sigma_{\mathrm{INT}}^2), \tag{9}$$

$$y_{\max} \equiv \max\left(\sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j, x_i \in \{-1, +1\} \right). \tag{10}$$

In this model y is the response of the system and the coefficients β quantify the factor effects on the response. In (6) we omit the constant term β_0 that is conventionally included in such models, because this article is concerned with degree of improvement of the response, and our conclusions are unaffected by the baseline from which improvement begins.

In this model the effect coefficients β are modeled as random variables in which variations occur across multiple applications of adaptive OFAT, not across multiple observations within a single use of adaptive OFAT. Across a population of different responses to which adaptive OFAT will be applied, the coefficients are sampled from the distributions as noted in (8) and (9). However, for any single application of adaptive OFAT to a response y, the coefficients are unknown to the experimenter but fixed from one observation to the next, consistent with the conventional assumption for regression analysis. Because factor effects typically exhibit a range of different sizes and are negative about as often as they are positive, the coefficients are modeled as random variables with mean of zero [(8) and (9)].

Inherent in the model is an assumption that the coefficients of main effects β_i and interactions β_{ij} are independent. Hierarchical probability models have been proposed for use in variable selection (Chipman, Hamada, and Wu 1997). In these models the probability that an interaction is active depends on whether the participating main effects are active or inactive. Empirical investigations have verified that interactions are significantly more likely to be active when at least one main effect involved in the interaction is active (Li, Sudarsanam, and Frey 2006). The model proposed here neglects this dependence (sometimes called "effect heredity" or "inheritance") to keep the model simple, but the main conclusions will be checked for sensitivity to this assumption.

Pure experimental error is represented by iid random variables ε_k, where the subscript k is different for each independent observation of the response y. In this model different variances are assigned to the pure experimental error, the main effects, and the two-factor interactions. Therefore we have a simple representation of the repeatability of the experimental procedure (as well as hierarchy of the system), enabling assessment of their impact on the performance of adaptive OFAT.

The variable $y_{\max}$ is the largest value of the response within the discrete factor space and is independent of experimental error, as indicated in (10). This value is useful as a basis of comparison in evaluating the performance of adaptive OFAT.

3.3 Defining Expected Improvement

We define the term "expected improvement" as it applies to experimentation on a response as given in Section 3.2. First, note that if a set of factor levels is selected within the discrete parameter space without benefit of any experimental observations, then the expected value of y is 0 because all the coefficients β have mean 0. As observations are made of the response, the experimenter can use that information in selecting subsequent points at which to make additional observations. We denote the response at the final chosen point in the parameter space as $y(x_1^*, x_2^*, \ldots, x_n^*)$. If the preference of the experimenter is for larger values of y and the decision process used is reasonable, then $E[y(x_1^*, x_2^*, \ldots, x_n^*)]$ will be positive. This rise in the expected value of y can therefore be considered an improvement rewarding the experimenter for the resources expended to make the parameter changes and observations required. In adaptive OFAT, it is also possible to consider the improvement achieved at intermediate steps such as $E[y(x_1^*, \ldots, x_{i-1}^*, \tilde{x}_i, \tilde{x}_{i+1}, \ldots, \tilde{x}_n)]$. We call these values "expected improvements." Note that these expected improvements denote the average performance of the system on repeated observations at the selected factor levels, not the value of an observation during an experiment.

It is useful to normalize the expected value of improvement so that 100% is the largest possible value. There are at least two different ways of using $y_{\max}$ to normalize the results—$E[y(x_1^*, \ldots, x_{i-1}^*, \tilde{x}_i, \tilde{x}_{i+1}, \ldots, \tilde{x}_n)/y_{\max}]$ and $E(y(x_1^*, \ldots, x_{i-1}^*, \tilde{x}_i, \tilde{x}_{i+1}, \ldots, \tilde{x}_n))/E(y_{\max})$. For all of the results reported in this article, the two quantities were nearly identical in value. In this article we report only the second one, because it is more readily applicable to the theorems presented here.

3.4 Defining Exploitation of Factor Effects

We define the term "exploit" as it applies to factor effects. We say that a factor effect has been exploited when the effect contributes positively to the value of the response y after some set of procedures has been carried out. When applied to the model (6)–(9), we say that the main effect, β_i, has been exploited whenever a procedure leads the experimenter to set the value x_i^* that meets the condition $\beta_i x_i^* > 0$. Similarly, we say that the two-factor interaction, β_{ij}, has been exploited whenever a procedure leads to values of the factors x_i^* and x_j^* satisfying the condition $\beta_{ij} x_i^* x_j^* > 0$. This definition assumes that larger values of the response y are preferred. The definition can easily be changed to accommodate scenarios in which smaller values of y are preferred.

Note that whereas main effects can in general be simultaneously exploited, two-factor interactions usually cannot. In a system with n factors, we have $\binom{n}{2}$ two-factor interactions. In most cases the degrees of freedom available to exploit interactions are far fewer than the number of interactions that we wish to exploit. In the next section we analyze the expected improvement provided by adaptive experimentation and explore how this is related to the probability of exploiting interactions.

4. THEORY

This section presents the main theoretical results of this article concerning the performance of adaptive OFAT. This section proceeds sequentially, first evaluating the early steps of the experimentation process and finally evaluating the overall performance of the completed process, including comparison with fractional factorial experiments.

4.1 The First Step in Adaptive OFAT

Theorem 1. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)], then the expected improvement after setting the first factor to x_1^* is

$$E(y(x_1^*, \tilde{x}_2, \ldots, \tilde{x}_n)) = E[\beta_1 x_1^*] + (n-1)E[\beta_{1j} x_1^* \tilde{x}_j], \tag{11}$$

where

$$E[\beta_1 x_1^*] = \sqrt{\frac{2}{\pi}} \frac{\sigma_{\mathrm{ME}}^2}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-1)\sigma_{\mathrm{INT}}^2 + (1/2)\sigma_\varepsilon^2}} \tag{12}$$

and

$$E[\beta_{1j} x_1^* \tilde{x}_j] = \sqrt{\frac{2}{\pi}} \frac{\sigma_{\mathrm{INT}}^2}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-1)\sigma_{\mathrm{INT}}^2 + (1/2)\sigma_\varepsilon^2}}. \tag{13}$$

The proofs of Theorem 1 and all other theorems are given in the Appendix.

As Theorem 1 indicates, the improvements from the first step in adaptive OFAT arise due to the main effect of the first factor and all of the $n-1$ two-factor interactions in which the first factor participates. This combination of a main effect and related interactions is often called the "conditional main effect" because it represents the effect of the factor conditioned on the current baseline settings.

An interesting property of adaptive OFAT is suggested by Theorem 1. If σ_{ME} is large compared with σ_{INT}, then most of the expected improvement is due to the main effect of the first factor. Conversely, if σ_{INT} is large compared with σ_{ME}, then most of the improvement comes from interactions.

To evaluate Theorem 1 and explore its implications, we applied the adaptive OFAT process to 10,000 responses sampled from the model (6)–(9). The value of $y_{\max}$ was determined for each of the 10,000 responses. The expected value of the response after one step of OFAT was estimated by making two observations and selecting a value for x_1^* as defined in (3). This process was repeated for many different values of $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}}$. The simulations corroborated the results of Theorem 1. The simulations also allow us to consider how the results from Theorem 1 relate to the optimal result previously defined in (10). The simulations revealed that if experimental error and interactions are moderate, then the adaptive OFAT process provides about $1/n$ of the maximum possible improvements when the first variable is set. For example, if $n = 7$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$ and $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$, then the first step of adaptive OFAT provides normalized expected improvement, $E(y(x_1^*, \tilde{x}_2, \ldots, \tilde{x}_n))/E(y_{\max})$, of slightly more than $1/7$. Experimental error has a mild negative effect, dropping the improvement by about 2% as $\sigma_\varepsilon/\sigma_{\mathrm{ME}}$ rises to unity. Interaction strength has a mild positive effect, raising the improvement by about 2% as $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ rises to unity. The reason for this positive effect is that interaction strength increases the contribution due to $n-1$ two-factor interactions. This is offset by a drop in the contribution due to the main effect and a rise in $E(y_{\max})$. The overall result is that the normalized expected improvement rises. Apparently, when using adaptive OFAT, the additional opportunities for improvement afforded by two-factor interactions outweigh the attendant disadvantages.

Although an improvement is realized due to interactions after the first step, note that none of these interactions has been exploited as defined in Section 3.3. All of the factors except for x_1 are toggled in subsequent steps of adaptive OFAT. Depending on what is observed, their final state may be different from the state after the first step. In that case the contributions due to interactions β_{1j} potentially may be reversed as the process continues.

In contrast, the first main effect β_1 has been exploited, and its contribution to the expected value is permanent. No subsequent steps in the adaptive OFAT process will affect the contribution of β_1 to the expected improvement (12). Thus the probability of β_1 being exploited can be fully determined by analyzing the behavior of the first step. This probability is given in the following theorem.

Theorem 2. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)], then the first main effect β_1 will be exploited with probability

$$\Pr(\beta_1 x_1^* > 0) = \frac{1}{2} + \frac{1}{\pi} \sin^{-1} \frac{\sigma_{\mathrm{ME}}}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-1)\sigma_{\mathrm{INT}}^2 + (1/2)\sigma_\varepsilon^2}}. \tag{14}$$

Theorem 2 shows that the probability of exploiting the main effect approaches 100%, because the main effect strength is much larger than interactions and experimental error. It also shows that the probability will drop to 50% as the experimental error increases. In a scenario with $n = 7$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$ and $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$, the probability of exploiting the first main effect is about 78%.

4.2 The Second Step in Adaptive OFAT

Theorem 3. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)], then the expected improvement after setting the second factor to x_2^* is

$$E(y(x_1^*, x_2^*, \tilde{x}_3, \ldots, \tilde{x}_n)) = 2E[\beta_1 x_1^*] + 2(n-2)E[\beta_{1j} x_1^* \tilde{x}_j] + E[\beta_{12} x_1^* x_2^*], \quad (15)$$

where

$$E[\beta_{12} x_1^* x_2^*] = \sqrt{\frac{2}{\pi}} \left[\frac{\sigma_{\mathrm{INT}}^2}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-1)\sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/2}} \right]. \quad (16)$$

Theorem 3 reveals that after the second step in the process, the response has an additional contribution due to main effects as well as added contributions due to interactions. The improvements at this stage arise due to three different contributors: two main effects that might be exploited, a two-factor interaction β_{12} that may have been exploited, and a set of $2(n-2)$ interactions that involve exactly one of the two main effects that have been toggled so far.

Following a procedure described in the preceding section, the adaptive OFAT process was applied to 10,000 responses sampled from the model (6)–(9). This simulation corroborates Theorem 2 and also reveals that if experimental error and interactions are moderate, then the adaptive OFAT process provides about $2/n$ of the possible improvements achievable when the second variable is set (appearing to be on a pace to attain all of the potential improvements). For example, if $n = 7$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$ and $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$, then the second step of adaptive OFAT provides $E(y(x_1^*, x_2^*, \tilde{x}_3, \ldots, \tilde{x}_n))/E(y_{\max})$ of very nearly 2/7. Experimental error has a mild negative effect, dropping the expected improvement by about 3% as $\sigma_\varepsilon/\sigma_{\mathrm{ME}}$ increases to unity. Interaction strength has a mild positive effect raising the ratio $E(y(x_1^*, x_2^*, \tilde{x}_3, \ldots, \tilde{x}_n))/E(y_{\max})$ by about 2% as $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ increases to unity. Note that when $n = 7$, more than half of the interactions potentially contribute to expected improvement by the time the second factor has been set. Thus, if interactions are moderately strong, then the adaptive OFAT process has many potentially large effects contributing to expected improvement.

Once the second variable is set by adaptive OFAT, the interaction β_{12} will not be affected in any way by subsequent experiments. Thus the probability of exploiting the interaction β_{12} can be determined by analyzing the process at the second step of adaptive OFAT. The following theorem arises from such an analysis.

Theorem 4. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)], then the probability that interaction β_{12} will be exploited is

$$\Pr(\beta_{12} x_1^* x_2^* > 0) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \frac{\sigma_{\mathrm{INT}}}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + (1/2)\sigma_\varepsilon^2}}. \quad (17)$$

Note that the probability given in Theorem 4 is >50% for all systems with nonzero interactions. This presents a paradox. The adaptive OFAT process confounds the two-factor interaction β_{12} with the main effects β_1 and β_2; nevertheless, the probability of exploiting the interaction is better than that provided by random chance, and the experimenter gains some improvement from the interaction on average. Thus the experimenter benefits from an effect that he or she cannot resolve.

To assess the practical implications of Theorem 4, it is useful to consider some particular numerical results. If $n = 7$ and interactions and experimental error are moderate compared with main effects, such as when $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$, then Theorem 4 indicates that the probability of exploiting the interaction β_{12} is about 58%. This probability is not very high. However, in improving a system, it is not essential (or generally possible) to exploit all of the interactions. It is, however, usually important to exploit the largest interactions present in the system. Theorem 5 gives the probability of exploiting the largest interaction assuming that the largest interaction happens to be β_{12}.

Theorem 5. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)] and the interaction β_{12} is the largest two-factor interactions in the system, then the probability that interaction β_{12} will be exploited obeys the following inequality:

$$\begin{aligned}
&\Pr(\beta_{12} x_1^* x_2^* > 0 \,|\, |\beta_{12}| > |\beta_{ij}|) \\
&\quad \ge \frac{1}{\pi}\binom{n}{2} \int_0^\infty \int_{-b}^\infty \left| \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{b}{\sigma_{\mathrm{INT}}} \right) \right|^{\binom{n}{2}-1} \\
&\qquad \times \exp\left\{ \frac{-b^2}{2\sigma_{\mathrm{INT}}^2} + \frac{-a^2}{2(\sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + (1/2)\sigma_\varepsilon^2)} \right\} \\
&\qquad \times \left(\sigma_{\mathrm{INT}} \sqrt{\sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + \frac{1}{2}\sigma_\varepsilon^2} \right)^{-1} da\, db.
\end{aligned} \quad (18)$$

Theorem 5 reveals that if $n = 7$ and interactions and experimental error are moderate compared with main effects, such as when $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$, then the probability of exploiting the interaction β_{12} is >75% if interaction β_{12} happens to be the largest two-factor interaction in the system. The results of simulations suggest that the inequality in (18) is nearly an equality for typical values of the model parameters. In addition, the inequality (18) becomes an equality in the limit as n increases. In addition, as the number of variables n rises above 4, Theorem 5 indicates that the probability of exploiting the largest interaction is sustained. For systems with 20 variables, the probability that adaptive OFAT exploits the largest two-factor interaction is about 75% for moderate values of interaction strength and experimental error. This probability is remarkably high because adaptive OFAT requires only 21 experiments and there are 190 possible two-factor interactions.

4.3 Subsequent Steps in Adaptive OFAT

As the adaptive process proceeds through subsequent steps, the mathematical results become increasingly complex. Exact closed-form solutions to the probabilities and expected values become cumbersome. However, simple bounds and approximations can still be derived that provide useful insights.

Theorem 6. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)], then the probability of exploiting a two-factor interaction, β_{ij}, is greater than or equal to the probability of exploiting the two-factor interaction, β_{12},

$$\Pr(\beta_{ij}x_i^* x_j^* > 0) \geq \Pr(\beta_{12}x_1^* x_2^* > 0). \tag{19}$$

Theorem 6 suggests that as adaptive OFAT proceeds, improvements will continue to accrue. But this theorem does not indicate that the rate of improvement will accelerate. The state of the adaptive OFAT process after k steps is graphically depicted in Figure 2. In this figure the striped squares are accumulating and the more recently added squares are darker in color, denoting a higher probability of exploiting later interactions. But also note that as the process continues, gray squares are being added at a decreasing rate.

Based on the observations arising from Figure 2 and patterns emerging from Theorems 1 and 3, we propose a closed-form expression approximating the expected improvement provided by adaptive OFAT. If adaptive OFAT [eqs. (1)–(5)] is applied to a response [eqs. (6)–(9)], then the expected improvement after setting the kth factor to x_k^* is approximately

$$\begin{aligned}
&E(y(x_1^*, x_2^*, \ldots, x_k^*, \tilde{x}_{k+1}, \ldots, \tilde{x}_n)) \\
&\quad \approx kE[\beta_1 x_1^*] + k(n-k)E[\beta_{1j}x_1^*] \\
&\qquad + \sqrt{\frac{2}{\pi}} \sum_{s=2}^{k} \Bigg[(s-1)\sigma_{\mathrm{INT}}^2 \\
&\qquad \times \left(\sqrt{\sigma_{\mathrm{ME}}^2 + \left(n - s^{\Pr(\beta_{12}x_1^* x_2^* > 0)}\right)\sigma_{\mathrm{INT}}^2 + \frac{\sigma_\varepsilon^2}{2}} \right)^{-1} \Bigg].
\end{aligned} \tag{20}$$

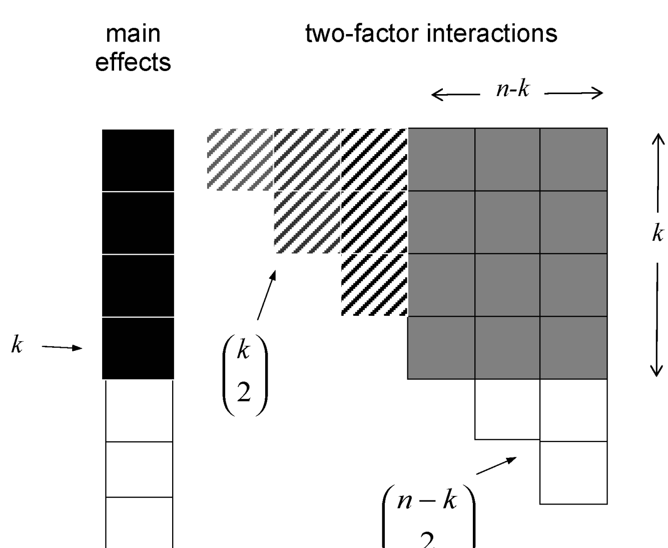

Figure 2. A Graphical Representation of the State of a Response With Seven Factors After the Fourth Step in Adaptive OFAT.

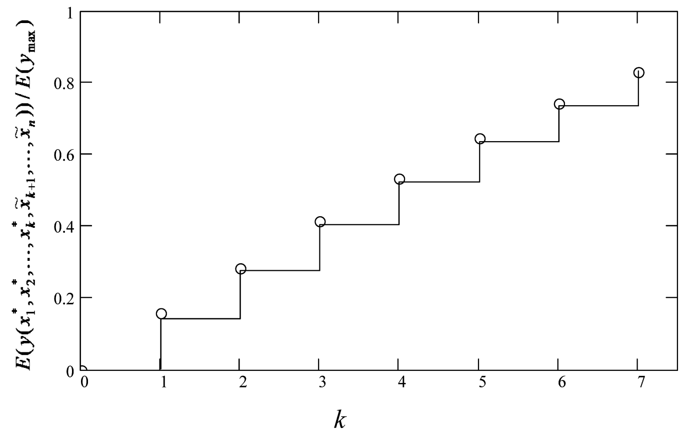

Figure 3. Evolution of the Expected Improvement (normalized by y_{max}) as the Adaptive OFAT Process Proceeds (—— eq. (20), $\sigma_\varepsilon/\sigma_{ME} = 1/4$; ○ simulation, $\sigma_{INT}/\sigma_{ME} = 1/3$).

The approximation in (20) is an extension of the results from (15) and (16) with an empirical correction factor to fit the result to computational simulations. The correction is in the denominator of the sum and takes the form of an exponent on the index s equal to the probability from (17).

Equation (20) is useful in that it captures the net result of the dynamic process depicted in Figure 2. As an indication of its accuracy, Figure 3 plots results for each of the k steps of adaptive OFAT given moderate values of interactions and experimental error compared with main effects. Simulations were used to check the results using 10,000 responses from the model (6)–(9). As the adaptive OFAT process continues (as k increases to n), the expected value rises almost linearly, dropping just slightly below its initial pace. By the time the last factor is set to x_n^*, adaptive OFAT provides 83% of the maximum possible improvement under this particular set of parameter values.

4.4 Final Results and Comparison of Theory With Simulations

The previous section provides an expression for the expected improvement at any step in the process of adaptive OFAT. Extrapolating to the last step, we can evaluate the final results of the process after $n + 1$ experiments and explore the effect of model parameters σ_{INT} and σ_ε. Results of such an analysis, using both the closed-form expression (20) and the results of simulations, are plotted in Figure 4.

If both interactions and experimental error are very low, then adaptive OFAT provides nearly all of the possible improvements. This is to be expected because, as the theory presented here indicates, adaptive OFAT can exploit all of the main effects under these conditions. For moderate values of interactions and experimental error ($\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$), adaptive OFAT yields an 82% normalized expected value of improvement. As $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ rises, the improvement drops slowly and stabilizes at about 75%.

Although not depicted in Figure 4, some results of both the theory and simulations are worth noting because they indicate the mechanisms at work. Theorem 6 shows that the probability of exploiting interactions cannot be lower than the

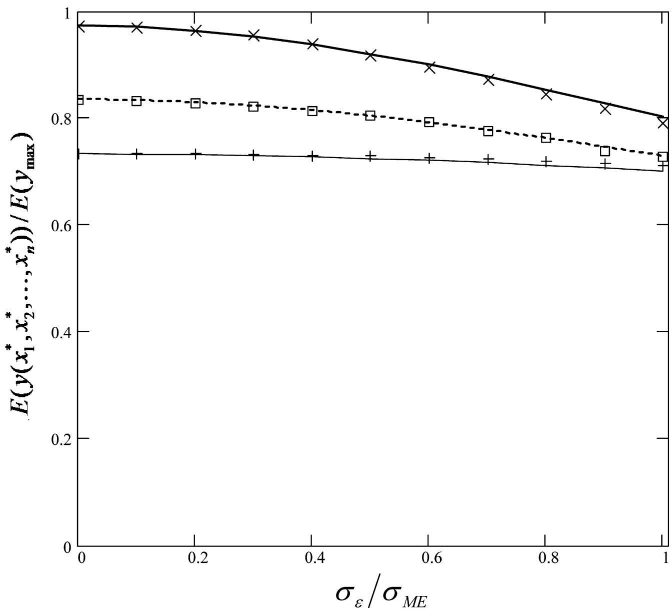

Figure 4. Expected Improvement After the Complete Adaptive OFAT Process Given a System Given $n = 7$ (—— eq. (20), × simulation, $\sigma_{INT}/\sigma_{ME} = 1/10$; ······· eq. (20), □ simulation, $\sigma_{INT}/\sigma_{ME} = 1/3$; —— eq. (20), + simulation, $\sigma_{INT}/\sigma_{ME} = 1$).

probability of exploiting the interaction β_{12}, and Theorem 5 suggests that the greatest interaction is exploited with substantially higher probability. Simulations corroborate these predictions. For moderate values of interactions and experimental error ($\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$), the interaction β_{12} is exploited in 59% of the simulations, the greatest interaction is exploited in 74% of the simulations, and interactions in general are exploited in 60% of all opportunities (across all responses simulated and all interactions within each response). These probabilities are based on the model (6)–(9), which may not be realistic. As a check, we ran a study using data from 113 published full-factorial experiments. These data enabled us to simulate adaptive OFAT by selecting tabulated values for each set of factor levels and adding pseudorandom error to simulate observation of the response in the presence of experimental error. For moderate values of experimental error (so that the contribution to the sum of squares is one-quarter that of main effects), the greatest interaction is exploited for about 80% of the responses, and interactions in general are exploited about 65% of the time across all interactions and all responses. Therefore, the results of the theory are reasonably well corroborated, and the results from published datasets are somewhat more favorable for adaptive OFAT than for the model (6)–(9). This is important because our model uses simplifying assumptions such as the independence of main effects and interactions. These simplifying assumptions do not seem to invalidate the main conclusions of the analysis.

4.5 Comparison With Fractional Factorial Experiments

The previous section enables estimates of the expected improvements provided by adaptive OFAT and reveals some of the mechanisms providing those improvements. To understand whether those improvements are large or small, it is worthwhile to have a basis of comparison. Saturated resolution III fractional factorial experiments are useful in this regard. Resolution III designs are frequently used for screening and also sometimes used in improvement processes (e.g., as "inner arrays" for robust parameter design). A saturated resolution III design (if it exists for a given n) can be carried out using the same number of experiments as adaptive OFAT (i.e., $n + 1$ experiments). Thus this section compares adaptive OFAT with an alternative with similar resource demands.

Theorem 7. If a saturated resolution III two-level fractional factorial experiment is applied to a response (6)–(9) and the factor levels are chosen among their discrete levels based on the sign of the main-effect estimates, then the expected improvement is

$$E(y_{\mathrm{III}}) = \sigma_{\mathrm{ME}}^2 \sqrt{\frac{2}{\pi}} \Big/ \sqrt{\sigma_{\mathrm{ME}}^2 + \frac{\binom{n}{2}}{n} \sigma_{\mathrm{INT}}^2 + \frac{\sigma_\varepsilon^2}{n+1}}. \tag{21}$$

Figure 5 plots results from (21) along with results from simulations. If interactions and experimental error are small, then the resolution III experiment provides nearly all of the possible improvements. In addition, the amount of improvement provided is not very sensitive to experimental error. However, the result is extremely sensitive to interactions. If interactions and experimental error are moderate compared with main effects, such as when $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}} = 1/3$ and $\sigma_\varepsilon/\sigma_{\mathrm{ME}} = 1/4$, then the expected improvement is <70% of the maximum. As $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ rise to 1, the expected improvements drop to about 20%. This seems reasonable, because when resolution III designs are used as prescribed in Theorem 7, the probability of exploiting a two-factor interaction is 50% (the same as if the factor levels were chosen at random), leading to zero net contribution due to interactions compared with random selection of factor levels.

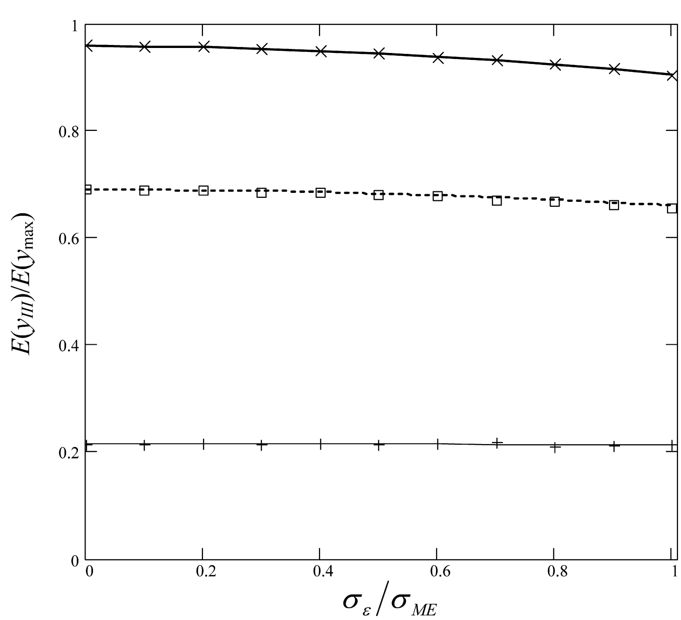

Figure 5. Expected Improvement (normalized by y_{max}) After a Saturated Resolution III Factorial Experiment Given a System Given $n = 7$ (—— Thm. 7, × simulation, $\sigma_{INT}/\sigma_{ME} = 1/10$; ······· Thm. 7, □ simulation, $\sigma_{INT}/\sigma_{ME} = 1/3$; —— Thm. 7, + simulation, $\sigma_{INT}/\sigma_{ME} = 1$).

It is interesting to consider what levels of interaction size and experimental error might motivate an experimenter to use adaptive OFAT in preference to resolution III factorial experimentation. If $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ is about $1/3$ and $\sigma_{\varepsilon}/\sigma_{\mathrm{ME}}$ is about 1.5, then the expected improvements of the two processes are roughly equal. Therefore, if an experimenter expects that interactions are moderately large compared with main effects, then it is sensible to choose adaptive OFAT if experimental error is comparable to (and perhaps slightly larger than) the main-effect coefficients, β_i, or comparable to (and perhaps slightly smaller than) the associated factor effects, $2\beta_i$. Note that this assessment is more favorable to OFAT than that of Daniel (1973), who suggested that OFAT should be used only when factor effects are three or four times the experimental error. This difference is probably due to the adaptive nature of the OFAT process studied here.

Table 2. The Largest 12 Effects on Flight Time Levels in the Electric-Powered Aircraft Experiment

Term	Coefficient
C	9.71
G	5.10
E	3.58
F	−3.24
$D \times G$	1.91
$A \times C$	1.43
$C \times F \times G$	−1.13
$E \times G$	.90
$B \times C$	.83
$D \times E \times G$	.83
$C \times D \times E \times F$	.79
B	−.79
$B \times G$	.38
$A \times F$	−.35

5. CASE STUDY

This case study concerns improvement in flight times of electric-powered model aircraft (which are currently popular among radio control hobbyists, including one of the authors). A computer-aided analysis package developed for these hobbyists, Electricalc Version 1.0E, is quite accurate when used properly. This computer program was used to explore a design space with seven two-level factors as described in Table 1. These levels are typical of the range of designs explored in student design/build/fly projects at the Massachusetts Institute of Technology. Aircraft weight and wing loading are affected by the levels of the factors. The following formulas capture those relationships:

$$\begin{aligned}\text{aircraft weight}&\\ &= 5\text{oz} + .01\text{oz} \times \text{wing area}\\ &\quad + .5\text{oz} \times \text{cells in battery}\\ &\quad + \text{number of motors}\\ &\quad \times \left[\begin{pmatrix}2.3\text{oz if SP400}\\ 3.5\text{oz if SP480}\end{pmatrix} + \begin{pmatrix}0\text{oz if } 1:1\\ .5\text{oz if } 1:1.85\end{pmatrix}\right] \qquad (22)\end{aligned}$$

and

$$\text{wing loading} = \frac{\text{aircraft weight}}{\text{wing area}}. \qquad (23)$$

A full-factorial 2^7 experiment was conducted using Electricalc based on the factors and levels in Table 1 and using (22) and (23) to set the appropriate weight and wing loading for the factors. The performance measure of interest was maximum flight duration, defined here as the battery life at the lowest throttle setting capable of level flight or a slight positive rate of climb. This performance measure might be appropriate for a customer who seeks long time on station but does not require high speed in transit or during loitering. This response value for each treatment combination was provided by Electricalc once the values of the factors and the associated weight and wing loading were entered manually. The complete results of the full-factorial experiment are available at *http://www-me.mit.edu/People/Research/danfrey.htm*.

Table 1. The Factors and Levels in the Electric-Powered Aircraft Experiment

Factor	Description	Level −	Level +
A	Propeller diameter	7 in.	8 in.
B	Propeller pitch	4 in.	5 in.
C	Gear ratio	1:1	1:1.85
D	Wing area	450 in.^{2}	600 in.^{2}
E	Cells in battery	7	8
F	Motor type	SP400 7.2V	SP480 7.2V
G	Number of motors	1	2

Based on the data from the full-factorial 2^7 experiment, the main effects and interactions were computed. The 12 largest effects are listed in Table 2. These are the only effects deemed "active" according to the Lenth method with the simultaneous margin of error criterion (Lenth 1989).

It should be noted that although a computer simulation was used to create the data for this case study, we do not suggest that adaptive OFAT is intended primarily for computer experiments. Instead, imagine a scenario in which each design evaluation involves building and evaluating an electric airplane. Under these conditions, a design method that requires only eight design variants (e.g., OFAT with $n = 7$) seems feasible, and any method that requires much more prototyping might be ruled out depending on the budget and schedule.

Table 2 suggests that this system is dominated by main effects (especially factor C, the gear ratio), but also has a number of interactions of substantial practical significance. For example, the $D \times G$ interaction represents the influence of wing area on the benefit of adding an additional motor to the aircraft. This interaction accounts for about 2 minutes of flight time, which is practically significant. It appears that this is a design problem in which interactions influence the outcomes in an important way. This influence must be quantified to relate this case study to the theory presented here.

In the model (6)–(9) the relative strength of main effects and interactions is represented by the ratio $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$. For this dataset, the ratio of standard deviation of all two-factor interactions and standard deviation of all main effects is $1/5$. But our simple model (6)–(9) includes only two-factor interactions, whereas the electric plane case study includes active three- and four-factor interactions. Therefore, a $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ of $1/5$ fails to adequately represent the strength of interactions in this system. For equivalence of our model with this case study, we find that the $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ should be around 1/3. For example, (6)–(9) with $n = 7$ and a $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ of 1/3 will typically include around eight active interactions, as observed in the electric aircraft experiment.

The data from the full-factorial experiment were used to simulate adaptive OFAT. For each trial, a starting point design and an order in which to toggle the factors were selected at random. Then the adaptive OFAT process was simulated by looking up the response in the tabulated data. Experimental error was simulated by adding a normally distributed pseudorandom number to the tabulated value to create a simulated observation. After the adaptive OFAT process selected the seven factor levels, the response at that set of levels without simulated error was stored as the outcome of the trial. This was repeated 10,000 times for each of 11 different amounts of simulated experimental error.

The results of the simulations are presented in Figure 6. The maximum flight time within this space of 128 possible discrete factor settings (50.5 minutes) and the mean (36.3 minutes) are indicated by heavy lines. The circles represent the average flight time achieved by adaptive OFAT over different starting point designs and orderings of the factors. The bars indicate the range exhibited from the 10th percentile to the 90th percentile. When experimental error was low, adaptive OFAT provided an average flight time of 48.1 minutes, representing 83% of the potential improvement (from starting design to final design) within this discrete space of factor settings. This is consistent with the prediction in (20) for low experimental error and for systems with a $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ of about 1/3.

To understand why the performance of adaptive OFAT was relatively good, it is useful to analyze the probability that effects were exploited when experimental error was low. Here the main effects were exploited with probability 75%, the two-factor interactions were exploited with probability 57%, and the largest two-factor interaction $D \times G$ was exploited with probability 66%. These values are all consistent with the prediction of Theorems 2, 4, and 5 given low experimental error and a $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ of ~1/3.

When experimental error was introduced, the performance of adaptive OFAT declined. With experimental error with a standard deviation of 3 minutes (a very large error for such an engineering experiment), adaptive OFAT provided an average flight time of 45.6, minutes, or 65% of the potential improvement. This is consistent with the prediction of (20) for moderate experimental error and moderate interactions. To visually compare the effect of experimental error in the case study with that predicted from theory, the plot in Figure 6 can be compared with the middle (dashed) line in Figure 4.

To provide a basis of comparison, the data from the full-factorial experiment were used to simulate factorial experimentation. Because adaptive OFAT required eight experiments, a 2^{7-4}_{III} design was used to maintain an equivalence of resource requirements. For each trial, 1 of 16 possible fractions was selected at random. Then the experiment was simulated by looking up the responses in the tabulated data and adding random variables to simulate experimental error as before. After data collection was complete, the main-effect estimates were used to select factor levels. The response at that set of levels without simulated error was stored as the outcome of the trial. This was repeated 10,000 times for each of 11 different amounts of simulated experimental error.

The results of the simulations are presented in Figure 7. The squares connected by dashed lines represent the average flight time achieved by factorial experimentation, and the bars indicate the range exhibited from the 10th percentile to the 90th percentile. When experimental error was low, the process provided an average flight time of 46.2 minutes, representing 74% of the potential improvement. This is substantially less than the 86% provided by adaptive OFAT under the same conditions. This is consistent with the prediction of (21) for low experimental error and for systems with $\sigma_{\mathrm{INT}}/\sigma_{\mathrm{ME}}$ of ~1/3. Also note that the range of outcomes for the factorial design, although narrower, includes worse outcomes on the low end than that of adaptive OFAT.

When experimental error was introduced, the performance of fractional factorial experimentation was relatively consistent.

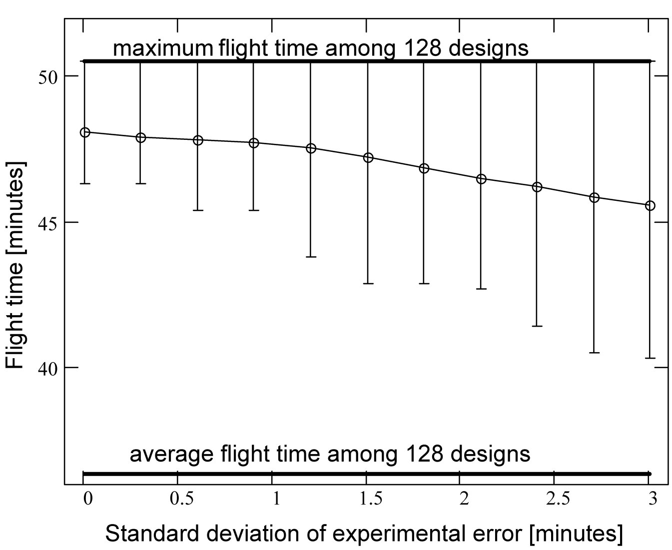

Figure 6. Flight Time After Adaptive OFAT Was Used in the Electric Aircraft Case Study (expetced value with OFAT —o——o—; range from the 10th to the 90th percentiles ⌶).

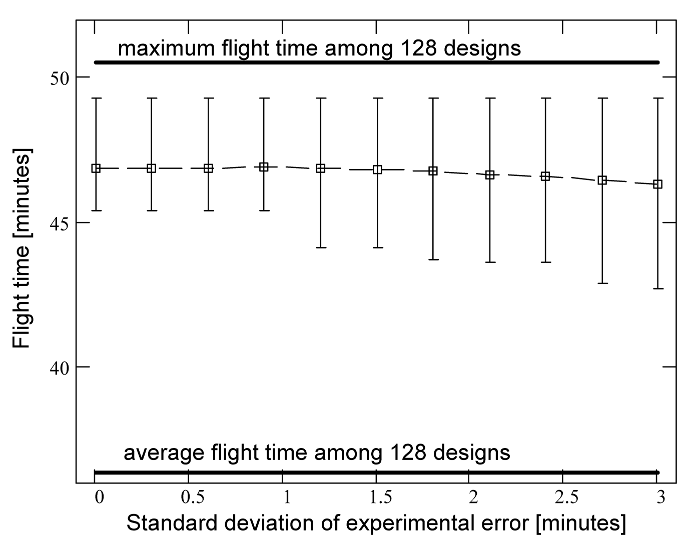

Figure 7. Flight Time After a 2^{7-4} Resolution III Design Was Used in the Electric Aircraft Case Study (expected value with 2^{7-4} --□--□--; range from the 10th to the 90th percentiles ⌶).

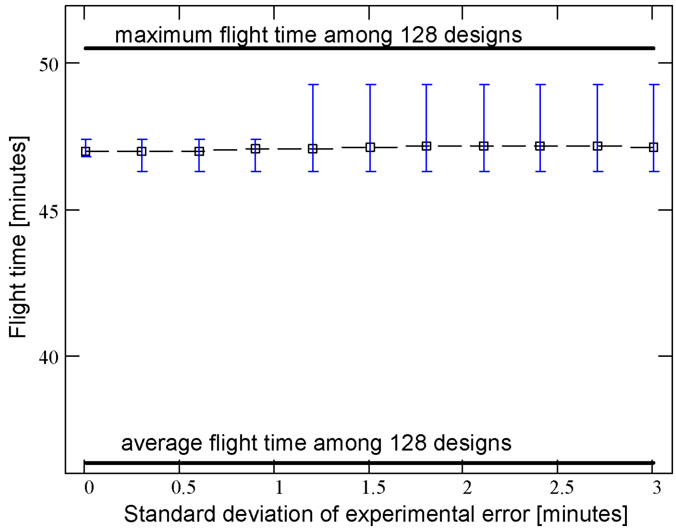

Figure 8. Flight Time After a 2^{7-3} Resolution IV Design Was Used in the Electric-Aircraft Case Study (expected value with 2^{7-3} - - □ - - □ - - ; range from the 10th to the 90th percentile).

With error having a standard deviation of 3 minutes, the process still provided an average flight time of 46.2 minutes, or 69% of the potential improvement. This was superior to the performance of adaptive OFAT under these conditions.

It is worth considering how much experimental error is present when factorial design provides more improvement than adaptive OFAT. In this case study the curves for average flight time in Figures 6 and 7 cross at an experimental error of around 2 minutes. The theory suggested that the "break-even point" is where the degree of error is about the same size as a typical main effect coefficient β_i. In this case study, the median main-effect coefficient is about 3 minutes, and the crossing point is about 2 minutes. The crossing points vary widely from case to case, and so this seems sufficiently close to be consistent with the theory.

It is also interesting to consider how a resolution IV experiment would fare in this case study. A 2^{7-3}_{IV} design was used to simulate an experimentation process using the same procedure as that used on the 2^{7-4}_{III} design. Figure 8 depicts the results. In this case study, raising the resolution of the experiment from III to IV resulted in a very small increase in expected value of flight time but a substantially reduced range of flight times. In this case study adaptive OFAT provided better results on average than resolution IV experiments as long as the standard deviation of experimental error was less than about 1.5 minutes. At very low values of experimental error, adaptive OFAT only rarely gives results as poor as those provided by resolution IV designs, which consistently fail to exploit the largest interactions.

6. CONCLUSION

This article has presented new mathematical models and associated theorems regarding adaptive experimentation. A model was proposed for adaptive OFAT that includes parameters related to the typical size of main effects, two-factor interactions, and experimental error. Six theorems were derived related to the expected improvements provided by adaptive OFAT and the probabilities that factor effects are exploited by adaptive OFAT. A closed-form expression was given for the expected improvement as a function of the number of steps completed, and this expression was validated by simulations. In addition, a theorem was given regarding expected improvements from saturated resolution III fractional factorial experimentation.

This article has shown that adaptive OFAT performs well under a broad range of conditions. Theorems 1 and 3 indicate that adaptive OFAT provides large gains in the early steps of the experimentation process, which can be an important advantage if an experiment must be terminated before completion due to budget or schedule changes. Equations (20) and (21) show that the expected improvement from adaptive OFAT is higher than that of saturated resolution III fractional factorial experiments if experimental error is low or if interactions are large. The closed-form equations presented here enable an experimenter to judge which approach is preferred under any combination of experimental error and interaction strength. The quantitative results of the theory serve as an additional check on a previously reported study using data from 66 responses to compare adaptive OFAT with fractional factorial experiments (Frey et al. 2003). The results are also consistent with a case study presented herein on design of an electric-powered aircraft.

The theorems proven here reveal some of the mechanisms causing the observed performance of adaptive OFAT. Theorem 2 shows that if main effects are large and experimental error is not too great, then adaptive OFAT exploits main effects with high probability. Theorem 4 shows that if interactions are not small, then adaptive OFAT exploits interactions with probability significantly better than random chance. Theorem 5 shows that adaptive OFAT exploits the largest interactions with even higher probability (roughly 75%). Theorem 6 suggests that these probabilities of exploitation are sustained as adaptive OFAT proceeds. A previous study based on details of five case studies also suggested mechanisms by which adaptive OFAT would exploit interactions with high probability (Frey and Jugulum 2006); therefore, the mechanisms suggested by the theory presented here have been corroborated using a different sort of analysis.

It is essential to emphasize that this article provides theorems for only two approaches that are suitable under tight resource constraints. These theorems provide a guide for choosing an approach if experimental budgets or schedules allow for only $n+1$ experiments (therefore, only resolution III designs, adaptive OFAT, and relatively few other alternatives are available). The case study suggests that raising the resolution from III to IV does relatively little to alter the conclusions discussed earlier. If budgets or schedules are expanded, a wider range of options exist including RSM. It is a worthwhile topic for future research to determine how much expected improvement other adaptive experimentation methods provide and whether these methods can be made scalable to large numbers of factors.

It is important to recognize the limitations of adaptive OFAT. If the duration of a single experiment is long, then adaptive OFAT may be ruled out because it takes at least $n+1$ times as long to run adaptive OFAT as it does to conduct a single trial. Therefore, it may be said that adaptive OFAT requires a high degree of immediacy. Further, there are many reasons that parallel experimentation may be much more economical than sequential experimentation, such as when a set of articles needed for an experiment is made most conveniently in batches. Thus adaptive OFAT is most applicable when experiments are sequential by nature, such as the industry practice of using a "mule" (a single instance of the previous model of the design) for exploring improvements. Box and Liu (1999) noted that most engineering experiments are characterized by immediacy and sequentiality, but they acknowledged that the exceptions may also be important. Another consideration is that adaptive OFAT essentially precludes randomization. When time trends exert a large influence, randomization may be required to prevent serious error.

With the cautions previously indicated, there are some interesting implications of the results shown here. In particular, viewing the results broadly, this article supports two key points:

1. The role of adaptation in experimentation is of substantial practical importance. Box and Liu (1999) have argued that iterative adaptation is essential in the planning and analysis of experiments. This idea was a motivation for the development of these theorems. It is hoped that this article provides a contribution to analysis of experimentation as a dynamic, adaptive process.
2. The effect of the designer's goals and preferences may strongly affect the planning of experiments. Hazelrigg (1998) has emphasized the use of quantitative, bottom-line metrics, such as profitability, to guide engineering design. This article begins to explore the implications of these ideas for design of experiments. If economic return is approximately linear in the engineering performance measure and the designer is approximately risk-neutral, then an experimentation method might reasonably be chosen on the basis of the expected value of improvement. Under such circumstances, Theorems 3, 4, and 6 and (20) may be used to directly evaluate adaptive OFAT, and Theorem 7 may be used for evaluating resolution III factorial design. However, there is an important trade-off inherent in this choice: Methods optimized for parameter estimation and/or model building will provide better models and perhaps more physical insight as well. These issues should be kept in mind when selecting a method of experimentation.

This work is intended as an early step in a broader research program regarding experimentation and adaptation and their role in engineering design. For example, we are interested in potential applications of adaptive OFAT in robust parameter design, especially those in which adaptive OFAT is used to structure an inner array of control factors crossed with a resolution III outer array of noise factors. In addition, the adaptive mechanisms explored here are only the simplest kind, requiring no physical knowledge. A richer theory would include consideration of the experimenter's mental models, the ways in which such models influence the planning of experiments (such as the order in which factors are considered), and the ways in which the experimental data alter one's mental models. It is also possible that adaptive OFAT will prove useful in computer experiments in which some model errors may be present, because physically reasonable predictions are made more easily when only one factor is changed. These topics are all interesting possibilities for future research.

ACKNOWLEDGMENTS

Financial support from the National Science Foundation (grant 0448972) and the support of the Ford/MIT Alliance are gratefully acknowledged. Xiang Li was very helpful in providing additional simulations that served as checks of these results. The suggestions of the referees and editors contributed to improving the article and are greatly appreciated.

APPENDIX: PROOFS OF THEOREMS

Proof of Theorem 1

Because the model of the response in (6)–(9) is symmetric about the origin, we can assume that the starting levels of the factors are $\tilde{x}_i = +1$ for all $i = 1, \dots, n$ without affecting the expected value of the outcome. Using the process in (1)–(5), the experimenter sets the first factor to $x_1^* = \mathrm{Sgn}(\beta_1 + \sum_{j=2}^n \beta_{1j} + (\varepsilon_0 - \varepsilon_1)/2)$, which leads to an expected improvement of

$$E(y(x_1^*, \tilde{x}_2, \dots, \tilde{x}_n))$$
$$= E[\beta_1 x_1^*] + E\left[\left(\sum_{j=2}^n \beta_{1j}\right) x_1^*\right] + E\left[\sum_{i=2}^n \beta_i + \sum_{i=2}^{n-1} \sum_{j=i+1}^n \beta_{ij}\right].$$

The last term reduces to 0 because the model (6)–(9) is symmetric about the origin. To simplify the first two terms, we observe that if A and B are independent and $A \sim \mathrm{N}(0, \sigma_A^2)$ and $B \sim \mathrm{N}(0, \sigma_B^2)$, then the expected value $E[A \cdot \mathrm{Sgn}(A + B)]$ can be computed by integrating over the region where $A + B > 0$ and the region where $A + B \le 0$, giving

$$E[A \cdot \mathrm{Sgn}(A+B)]$$
$$= \int_{-\infty}^{\infty} \left[\int_{-b}^{\infty} a \cdot \phi(a, \sigma_A^2) \phi(b, \sigma_B^2)\, da + \int_{-\infty}^{-b} -a \cdot \phi(a, \sigma_A^2) \phi(b, \sigma_B^2)\, da\right] db$$
$$= \sqrt{\frac{2}{\pi}} \frac{\sigma_A^2}{\sqrt{\sigma_A^2 + \sigma_B^2}},$$

where $\phi(x, \sigma^2)$ is the normal probability distribution with mean 0 and variance σ^2. Applying this identity to $E[\beta_1 x_1^*]$, where $\sigma_A^2 = \sigma_{\mathrm{ME}}^2$ and $\sigma_B^2 = (n-1)\sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/2$, gives the result in (12), and applying this identity to $E[(\sum_{j=2}^n \beta_{1j}) x_1^*]$, where $\sigma_A^2 = (n-1)\sigma_{\mathrm{INT}}^2$ and $\sigma_B^2 = \sigma_{\mathrm{ME}}^2 + \sigma_\varepsilon^2/2$, gives the result in (13).

Proof of Theorem 2

Borrowing the notation and results from the proof of Theorem 1, we find that

$$\Pr(\beta_1 x_1^* > 0)$$
$$= \Pr(\beta_1 > 0, x_1^* = +1) + \Pr(\beta_1 < 0, x_1^* = -1)$$
$$= \Pr\left(\beta_1 > 0, \beta_1 + \sum_{j=2}^n \beta_{1j} + \frac{\varepsilon_0 - \varepsilon_1}{2} > 0\right)$$

$$+\Pr\left(\beta_1 < 0, \beta_1 + \sum_{j=2}^{n}\beta_{1j} + \frac{\varepsilon_0 - \varepsilon_1}{2} < 0\right)$$

$$= 2\Pr\left(\beta_1 > 0, \beta_1 + \sum_{j=2}^{n}\beta_{1j} + \frac{\varepsilon_0 - \varepsilon_1}{2} > 0\right),$$

where the last equality holds because the model (6)–(9) is symmetric about the origin. To simplify this expression, we observe that if $A \sim \mathrm{N}(0, \sigma_A^2)$ and $B \sim \mathrm{N}(0, \sigma_B^2)$, then

$$\Pr(A > 0, A + B > 0) = \int_0^\infty \int_{-a}^\infty \phi(a, \sigma_A^2)\phi(b, \sigma_B^2)\, db\, da$$
$$= \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\left(\frac{\sigma_A}{\sqrt{\sigma_A^2 + \sigma_B^2}}\right).$$

Applying this identity where $\sigma_A^2 = \sigma_{\mathrm{ME}}^2$ and $\sigma_B^2 = (n-1) \times \sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/2$ gives the result in Theorem 2.

Proof of Theorem 3

Using the process in (1)–(5) the second factor is chosen,

$$x_2^* = \mathrm{Sgn}\left(2\beta_2 + 2\beta_{12}x_1^* + 2\sum_{j=3}^{n}\beta_{2j} + \varepsilon_0\left(\frac{1 + x_1^*}{2}\right) + \varepsilon_1\left(\frac{1 - x_1^*}{2}\right) - \varepsilon_2\right).$$

The expected improvement at this stage is

$$E(y(x_1^*, x_2^*, \tilde{x}_3, \dots, \tilde{x}_n))$$
$$= E\left[\left(\beta_1 + \sum_{j=3}^{n}\beta_{1j}\right)x_1^*\right] + E\left[\left(\beta_2 + \sum_{j=3}^{n}\beta_{2j}\right)x_2^*\right]$$
$$+ E[\beta_{12}x_1^*x_2^*] + E\left[\sum_{i=3}^{n}\beta_i + \sum_{i=3}^{n-1}\sum_{j=i+1}^{n}\beta_{ij}\right].$$

The last term reduces to 0. The first and second terms can be simplified using the identity $E[A \cdot \mathrm{Sgn}(A + B)] = \sqrt{2/\pi}\sigma_A^2 \times (\sqrt{\sigma_A^2 + \sigma_B^2})^{-1}$ with both terms contributing $E[\beta_1 x_1^*] + (n - 2)E[\beta_{1j}x_1^*\tilde{x}_j]$. Now consider the third term. Note that

$$x_1^*x_2^* = \mathrm{Sgn}\left(\beta_{12} + \frac{\varepsilon_0}{4} - \frac{\varepsilon_1}{4} + \left(\beta_2 + \sum_{j=3}^{n}\beta_{2j} + \frac{\varepsilon_0}{4} + \frac{\varepsilon_1}{4} - \frac{\varepsilon_2}{2}\right)x_1^*\right)$$

and, therefore, $E[\beta_{12}x_1^*x_2^*]$ can be simplified using the observation that if $A \sim \mathrm{N}(0, \sigma_A^2)$, $B1, B2 \sim \mathrm{N}(0, \sigma_B^2)$, and $C \sim \mathrm{N}(0, \sigma_C^2)$ are all mutually independent and $x^* \in \{-1, +1\}$, then

$$E\big[A \cdot \mathrm{Sgn}\big(A + B1 - B2 + (C + B1 + B2)x^*\big)\big]$$
$$= \sqrt{\frac{2}{\pi}}\left[\frac{\sigma_A^2}{\sqrt{\sigma_A^2 + 4\sigma_B^2 + \sigma_C^2}}\right].$$

Applying this identity where $\sigma_A^2 = \sigma_{\mathrm{INT}}^2$ and $\sigma_B^2 = \sigma_\varepsilon^2/16$ and $\sigma_C^2 = \sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/4$, we have

$$E[\beta_{12}x_1^*x_2^*] = \sqrt{\frac{2}{\pi}}\frac{\sigma_{\mathrm{INT}}^2}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-1)\sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/2}}.$$

Sum the parts derived earlier, and the result in Theorem 3 follows.

Proof of Theorem 4

The probability to be determined is

$$\Pr(\beta_{12}x_1^*x_2^* > 0) = \Pr(\beta_{12} > 0, x_1^* = 1, x_2^* = 1)$$
$$+ \Pr(\beta_{12} > 0, x_1^* = -1, x_2^* = -1)$$
$$+ \Pr(\beta_{12} < 0, x_1^* = 1, x_2^* = -1)$$
$$+ \Pr(\beta_{12} < 0, x_1^* = -1, x_2^* = 1).$$

Consider the first term of the foregoing expression. Substituting in the values of the final settings of the variables x_1^* and x_2^* and simplifying, we have

$$\Pr\left(\beta_{12} > 0, \beta_{12} + \beta_1 + \sum_{j=3}^{n}\beta_{1j} + \frac{\varepsilon_0}{2} - \frac{\varepsilon_1}{2} > 0, \beta_{12} + \beta_2 + \sum_{j=3}^{n}\beta_{2j} + \frac{\varepsilon_0}{2} - \frac{\varepsilon_2}{2} > 0\right).$$

Observe that if $\mathbf{z} \sim \mathrm{N}(\mathbf{0}, \mathbf{K})$, then

$$\Pr(\mathbf{z}_1 > 0, \mathbf{z}_2 > 0, \mathbf{z}_3 > 0)$$
$$= \frac{1}{8} + \frac{1}{4\pi}\left[\sin^{-1}\left(\frac{K_{12}}{\sqrt{K_{11}K_{22}}}\right) + \sin^{-1}\left(\frac{K_{13}}{\sqrt{K_{11}K_{33}}}\right) + \sin^{-1}\left(\frac{K_{23}}{\sqrt{K_{22}K_{33}}}\right)\right].$$

We may apply this identity to the first term with $\mathbf{K}_{11} = \mathbf{K}_{12} = \mathbf{K}_{13} = \sigma_{\mathrm{INT}}^2$, $\mathbf{K}_{22} = \mathbf{K}_{33} = \sigma_{\mathrm{ME}}^2 + (n-1)\sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/2$, and $\mathbf{K}_{23} = \sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/4$, giving the result

$$\frac{1}{8} + \frac{1}{2\pi}\tan^{-1}\left(\frac{\sigma_{\mathrm{INT}}}{\sqrt{\sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + \sigma_\varepsilon^2/2}}\right)$$
$$+ \frac{1}{4\pi}\tan^{-1}\left(\left(\sigma_{\mathrm{INT}}^2 + \frac{\sigma_\varepsilon^2}{4}\right)\left(\left[\sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + \frac{\sigma_\varepsilon^2}{2}\right] \times \sqrt{\sigma_{\mathrm{ME}}^2 + n\sigma_{\mathrm{INT}}^2 + \frac{\sigma_\varepsilon^2}{2}}\right)^{-1}\right).$$

Applying the same procedure to the other three terms gives expressions in a similar form, with only the signs of the terms changing. Summing up the four expressions, we get the expression in Theorem 4.

Proof of Theorem 5

From Theorem 4, we have an expression for $\Pr(\beta_{12}x_1^*x_2^* > 0)$ as the sum of probabilities of four possible cases. Adding the condition $|\beta_{12}| > |\beta_{ij}|$, the sum of the first two cases can be expressed by defining random variables A, B, and C so that

$$\begin{aligned}
&\Pr(\beta_{12} > 0, \beta_{12} + A > 0, \beta_{12} + B > 0 | |\beta_{12}| > |\beta_{ij}| \cdot) \\
&\qquad + \Pr(\beta_{12} > 0, \beta_{12} + A < 0, -\beta_{12} - C < 0 | |\beta_{12}| > |\beta_{ij}| \cdot) \\
&\quad \geq \Pr(\beta_{12} > 0, \beta_{12} + A > 0 | |\beta_{12}| > |\beta_{ij}|),
\end{aligned}$$

where the inequality holds due to the following properties of the random variables: β_{12}, A, B, and C are symmetric about the origin; B and C are identically distributed; and B and C are positively correlated with A. The same reasoning applies to the sum of the third and fourth cases, leading to the conclusion that $\Pr(\beta_{12} x_1^* x_2^* > 0 | |\beta_{12}| > |\beta_{ij}|) \geq 2 \Pr(\beta_{12} > 0,\ \beta_{12} + A > 0 | |\beta_{12}| > |\beta_{ij}|)$. Applying Bayes's rule, we find that $\Pr(\beta_{12} > 0, \beta_{12} + A > 0 | |\beta_{12}| > |\beta_{ij}|) = \int_0^\infty \Pr(A \geq -b | |\beta_{12}| > |\beta_{ij}|, b = \beta_{12} \cdot) \Pr(b = \beta_{12} | |\beta_{12}| > |\beta_{ij}| \cdot)\, db$.

Removing the condition $|\beta_{12}| > |\beta_{ij}|$ from $\Pr(A \geq -b | |\beta_{12}| > |\beta_{ij}|, b = \beta_{12} \cdot)$ makes the event $\beta_{12} > 0, \beta_{12} + A > 0$ more likely; therefore,

$$\begin{aligned}
&\int_0^\infty \Pr(A \geq -b | |\beta_{12}| > |\beta_{ij}|, b = \beta_{12} \cdot) \\
&\qquad \times \Pr(b = \beta_{12} | |\beta_{12}| > |\beta_{ij}| \cdot)\, db \\
&\quad \geq \int_0^\infty \int_{-b}^\infty \Pr\left(a = \beta_1 + \sum_{j=3}^{n} \beta_{1j} + \frac{\varepsilon_0 - \varepsilon_1}{2}\right) \\
&\qquad\qquad \times \Pr(b = \beta_{12} | |\beta_{12}| > |\beta_{ij}| \cdot)\, da\, db.
\end{aligned}$$

Given the assumptions of the model (6)–(9), we have

$$\begin{aligned}
&\Pr\left(a = \beta_1 + \sum_{j=3}^{n} \beta_{1j} + \frac{\varepsilon_0 - \varepsilon_1}{2}\right) \\
&\qquad = \phi\left(a, \sigma_{\mathrm{ME}}^2 + (n-2)\sigma_{\mathrm{INT}}^2 + \frac{1}{2}\sigma_\varepsilon^2\right),
\end{aligned}$$

where $\phi(x, \sigma^2)$ represents a normal distribution with mean 0 and variance σ^2. To find an expression for $\Pr(b = \beta_{12} | |\beta_{12}| > |\beta_{ij}| \cdot)$, note that

$$f(x) = m \left| \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{|x|}{\sigma}\right)\right|^{m-1} \phi(x, \sigma^2)$$

is the probability distribution function for the largest absolute value among a set of m normally distributed random variables of mean 0 and equal variance σ^2 (Gumbel 1958). Because the model in (6)–(9) has $\binom{n}{2}$ two-factor interactions that are iid by $\mathrm{N}(0, \sigma_{\mathrm{INT}}^2)$, we find that

$$\begin{aligned}
&\Pr(b = \beta_{12} | |\beta_{12}| > |\beta_{ij}|) \\
&\qquad = \binom{n}{2} \left| \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{|b|}{\sigma_{\mathrm{INT}}}\right)\right|^{\binom{n}{2}-1} \phi(b, \sigma_{\mathrm{INT}}^2).
\end{aligned}$$

The inequality in Theorem 5 follows.

Proof of Theorem 6

First, consider the probability exploiting interaction β_{13}. The process in (1)–(5) leads the experimenter to set

$$\begin{aligned}
x_3^* = \operatorname{sign}\Bigg(& 2\beta_3 + 2\sum_{i=1}^{2} \beta_{i3} x_i^* + 2\sum_{j=4}^{n} \beta_{3j} \\
& + \left[\frac{1 + x_1^*}{2}\right]\left[\frac{1 + x_2^*}{2}\right]\varepsilon_0 + \left[\frac{1 - x_1^*}{2}\right]\left[\frac{1 + x_2^*}{2}\right]\varepsilon_1 \\
& + \left[\frac{1 - x_2^*}{2}\right]\varepsilon_2 - \varepsilon_3 \Bigg).
\end{aligned}$$

The probability $\Pr(\beta_{13} x_1^* x_3^* > 0)$ contains four terms following the form of $\Pr(\beta_{12} x_1^* x_2^* > 0)$ in the proof of Theorem 4. The sum of the first two terms is

$$\begin{aligned}
\geq 2 \Pr\Bigg(& \beta_{13} > 0, \beta_1 - \beta_{12} - \beta_{13} + \sum_{j=2}^{n} \beta_{1j} + \frac{\varepsilon_0 - \varepsilon_1}{2} > 0, \\
& \beta_2 + \beta_{12} - \beta_{23} - \sum_{i=1}^{n} \beta_{i3} + \sum_{j=4}^{n} \beta_{3j} + \frac{\varepsilon_3 - \varepsilon_2}{2} > 0 \Bigg),
\end{aligned}$$

because substituting $\beta_{13} > 0$ for $\beta_{13} < 0$ in the second term makes the joint event less likely. This expression is $\geq \Pr(\beta_{12} \times x_1^* x_2^* > 0)$, because the substitution of β_{13} for β_{12} in the expression in Theorem 4 makes the joint event less likely. A similar process holds for every case in which the interaction β_{13} is exploited. Repeating the reasoning applied to β_{13} for every β_{ij} proves Theorem 6.

Proof of Theorem 7

Given $n+1$ observations of y as defined in (6)–(9) are made using a saturated resolution III fractional factorial experiment with a design matrix $\mathbf{X}$. It follows that the estimate for each main-effect coefficient β_j will be

$$\beta_j + \frac{1}{n+1} \sum_{i=1}^{n+1} \varepsilon_i \mathbf{X}_{ij} + \frac{1}{n+1} \sum_{i=1}^{n+1} \sum_{\substack{j=1 \\ j \neq i}}^{n} \beta_{ij} \mathbf{X}_{ij}.$$

The choice of factor levels is set according to the sign of the main-effect estimates. Therefore, the expected value is

$$\begin{aligned}
E\Bigg[\sum_{j=1}^{n} \beta_j \operatorname{Sgn}\Bigg(& \beta_j + \frac{1}{n+1} \sum_{i=1}^{n+1} \varepsilon_i \mathbf{X}_{ij} \\
& + \frac{1}{n+1} \sum_{i=1}^{n+1} \sum_{j=1, j \neq i}^{n} \beta_{ij} \mathbf{X}_{ij} \Bigg) \Bigg].
\end{aligned}$$

Given the result from the proof of Theorem 1 regarding $E[A \cdot \operatorname{Sgn}(A+B)]$ and the properties of the design matrix $\mathbf{X}$, the result of Theorem 7 follows.

[Received May 2005. Revised November 2005.]

REFERENCES

Box, G. E. P. (1999), "Statistics as a Catalyst to Learning by Scientific Method. Part 2: A Discussion," *Journal of Quality Technology*, 31, 16–29.

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, New York: Wiley.

Box, G. E. P., and Liu, P. T. Y. (1999), "Statistics as a Catalyst to Learning by Scientific Method. Part 1: An Example," *Journal of Quality Technology*, 31, 1–15.

Box, G. E. P., and Wilson, K. B. (1951), "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society*, Ser. B, 13, 1–38.

Chipman, H. M., Hamada, M., and Wu, C. F. J. (1997), "Bayesian Variable-Selection Approach for Analyzing Designed Experiments With Complex Aliasing," *Technometrics*, 39, 372–381.

Czitrom, V. (1999), "One-Factor-at-a-Time versus Designed Experiments," *The American Statistician*, 53, 126–131.

Daniel, C. (1973), "One-at-a-Time Plans," *Journal of the American Statistical Association*, 68, 353–360.

Frey, D. D., Engelhardt, F., and Greitzer, E. M. (2003), "A Role for One Factor at a Time Experimentation in Parameter Design," *Research in Engineering Design*, 14, 65–74.

Frey, D. D., and Jugulum, R. (2006), "The Mechanisms by Which Adaptive One-Factor-at-a-Time Experimentation Leads to Improvement," *ASME Journal of Mechanical Design*, to appear.

Friedman, M., and Savage, L. J. (1947), "Planning Experiments Seeking Maxima," in *Techniques of Statistical Analysis*, eds. C. Eisenhart, M. W. Hastay, and W. A. Wallis, New York: McGraw-Hill, pp. 365–372.

Gumbel, E. J. (1958), *Statistics of Extremes*, New York: Columbia University Press.

Hazelrigg, G. A. (1998), "A Framework for Decision-Based Engineering Design," *Journal of Mechanical Design*, 120, 653–658.

Hill, W. J., and Hunter, W. G. (1966), "A Review of Response Surface Methodology: A Literature Survey," *Technometrics*, 8, 571–590.

Koita, R. (1994), *Strategies for Sequential Design of Experiments*, unpublished thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science.

Korn, G. A., and Korn, T. M. (1961), *Mathematical Handbook for Scientists and Engineers*, Mineola, NY: Dover Publications.

Lenth, R. V. (1989), "Quick and Easy Analysis of Unreplicated Factorials," *Technometrics* 31, 469–473.

Li, X., Sudarsanam, N., and Frey, D. D. (2006), "Regularities in Data From Factorial Experiments," *Complexity*, 11 (5), 32–45.

Logothetis, N., and Wynn, H. P. (1994), *Quality Through Design*, Oxford, U.K.: Clarendon Press.

McDaniel, W. R., and Ankenman, B. E. (2000), "Comparing Experimental Design Strategies for Quality Improvement With Minimal Changes to Factor Levels," *Quality and Reliability Engineering International*, 16, 355–362.

Myers, R. H., Khuri, A. I., and Carter, W. (1989), "Response Surface Methodology: 1966–1988," *Technometrics*, 31, 137–157.

Qu, X., and Wu, C. F. J. (2005), "One-Factor-at-a-Time Designs of Resolution V," *Journal of Statistical Planning and Inference*, 131, 407–416.

Wu, C. F. J., and Hamada, M. (2000), *Experiments: Planning, Analysis, and Parameter Design Optimization*, New York: Wiley.