Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

CONTROLLED SEQUENTIAL FACTORIAL DESIGN FOR SIMULATION FACTOR SCREENING

Hua Shen
Hong Wan

School of Industrial Engineering
Purdue University
West Lafayette, IN 47907-2023, U.S.A.

ABSTRACT

Screening experiments are performed to eliminate unimportant factors so that the remaining important factors can be more thoroughly studied in later experiments. Sequential screening methods are specifically fit for simulation experiments. They are usually more efficient than one-step procedures. The challenge is to prove the "correctness" of the result.

This paper proposes Controlled Sequential Factorial Design (CSFD) for discrete-event simulations. It combines sequential hypothesis-testing procedures with the traditional factorial design to control the Type I Error and power for each factor under heterogeneous variances conditions. CSFD requires minimum assumptions and demonstrates robust performance with different system conditions.

1 INTRODUCTION

Many screening strategies have been proposed for screening purpose (Trocine and Malone 2000, 2001 and Campolongo, et al. 2000). However, most research has concentrated on designs for physical experiments. Because of the high cost of conducting physical experiments, the traditional screening methods usually emphasize using the fewest number of runs to estimate as many effects as possible, the correctness of the results is considered secondary. In addition, simulation experiments have many unique properties that are not stressed in traditional screening design. One of the most important ones is that sequential methods are favored in simulation due to the comparative ease of switching between factor settings.

Controlled Sequential Bifurcation (CSB) (Wan, Ankenman, and Nelson, 2003, 2005a) is a new factor-screening method specifically designed for simulation experiments. CSB incorporates a statistical hypothesis testing procedure into sequential bifurcation and controls the probability of Type I error for each factor and power at each bifurcation step. Wan, et al. (2005b) later proposed an improved version of CSB, called CSB-X, using a fold-over design. CSB-X has the same error control as CSB for main effects even in the presence of even-order interaction effects. Both methods allow heterogeneous variances but require all main effects are positive (or negative) to avoid effect cancellation.

Although these CSB methods are attractive for many simulation applications, there are also some limitations on them. First, both CSB and CSB-X have restriction on their response models, (CSB is based on a main-effect-only regression model and CSB-X allows main effects and even-order interaction effects), and are designed to classify only main effects. When higher-order interactions exist, the results of the classification can be misleading. Second, in CSB methods the signs of the main effects have to be known beforehand to avoid effect cancellation within groups. Moreover, CSB methods prefer important effects being clustered and sorted to obtain higher efficiency. But neither of these is applicable in some cases. Third, CSB methods do not fully utilize the information available before and during the experiment. Most simulation replications generated in previous design points will not be useful in the later screening procedure and new simulation replications are usually needed after each bifurcation step.

The limitations of CSB inspire us to propose a new approach which can overcome these limitations while still keeping the error rate control. It is well known that factorial experiment design can estimate all main and interaction effects and needs no prior information of the effects. However, factorial design has typically been used in physical experiments and the major concern is the limitation on the number of runs available. So in most cases, factorial design is a one-stage procedure and assumes equal variance for different factor settings. Our purpose here is to propose a sequential factorial design which can take full advantage of previously generated information and provide desired error control for simulation experiments.

Sequential factorial design was first mentioned by Davies and Hay (1950) in their guideline for the use of factorial designs in industrial research. They briefly dis-

cussed how the features of industrial experiments make it necessary to have sequential factorial designs and the possible benefits of using it. Hunter (1964) used a sequential factorial approach to estimate the coefficients of a linear regression model. With an initial set of orthogonal estimates, least square estimates of all the coefficients are updated by a "predictor-corrector" equation at the conclusion of each run. This method has the flexibility to stop the experiment early and to increase the model and block size when further investigation is necessary. However, no discussion of the stopping rules of the sequential procedure was provided. Zacks (1968) tried to obtain the optimal sequential designs of fractional factorial experiments using Bayesian analysis. Assuming that the prior distribution of the parameters is normal and the loss function is quadratic, the Bayesian analysis concludes that the best sequential procedure is a fixed sample procedure. Gilmour and Mead (1995) presented Bayesian-approached stopping rules for sequential fractional factorial designs in which the purpose is to find the optimal combination of factor levels. By computing the Bayesian estimation of the regression coefficients, the method obtains the posterior distribution of the difference between the optimal response and the response from the predicted optimum. The sequential procedure stops when some given criterion is satisfied by the posterior distribution. Sensitivity analysis showed that the procedure performance depends on the variance information of the prior distribution.

In this paper, we propose a controlled sequential factorial design (CSFD) for stochastic simulations. In addition to the simultaneous Type I error and power control, CSFD has the flexibility to classify any desired main effect or interaction effect. The resolution (which effects can be independently screened) only depends on the factorial design initially selected. No prior information of the effects is required. Unlike CSB which usually needs to generate new replications at each bifurcation, CSFD uses all previously generated replications in the later classification. In most cases, after first several effects are classified, there is enough data to classify all the other effects. Moreover, with the option of a fractional factorial design, CSFD can be more efficient than CSB in many cases.

The paper is organized as follows: The underlying response model and the objective of screening are discussed in Section 2. Section 3 describes CSFD methods and two hypothesis testing procedures. Section 4 presents empirical evaluations of CSFD comparing to CSB (and CSB-X). Future research is discussed in Section 5.

2 MODEL DESCRIPTION

Suppose there are in total of L factors, a general metamodel includes all main effects and interactions is given as follows:

$$\begin{aligned} Y &= \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_L z_L \\ &\quad + \beta_{12} z_1 z_2 + \cdots + \beta_{L-1,L} z_{L-1} z_L \\ &\quad + \cdots + \beta_{12\ldots L} z_1 z_2 \ldots z_L + \varepsilon \end{aligned} \tag{1}$$

Here $\beta = \{\beta_1, \beta_2, \ldots, \beta_{12\ldots L}\}$ are the effect coefficients of corresponding effects. The level settings, $\mathbf{z} = (z_1, z_2, \ldots, z_L)$, are assumed to be deterministic. The error term, ε, on the other hand, is a random variable; in this paper we assume it is a Nor$(0, \sigma^2(\mathbf{z}))$ random variable where $\sigma^2(\mathbf{z})$ is unknown and may depend on $\mathbf{z}$. In practice, the model may include any subset of the effects. Usually if an interaction exists, the main effects and low-order interactions of all factors involved should also be included.

The objective of our screening procedure is to classify desired effects into two groups: important ones and unimportant ones. For those effects with effect coefficients $\leq \Delta_0$, CSFD should control the Type I Error of declaring them important to be $\leq \alpha$; and for those effects with effect coefficients $\geq \Delta_1$, the power of identifying them as important should be $\geq \gamma$. Those factors whose effect coefficients fall between Δ_0 and Δ_1 are considered important and we want CSFD to have reasonable, though not guaranteed, power to identify them. Δ_0 and Δ_1 are the thresholds of importance and critical respectively; α and γ are user-specified confidence parameters. In practice, the determination of both thresholds and the factor settings are usually associated with the cost to achieve the system changes. Wan, et al. (2005a) proposed a cost model using the cost of changing factor to determine the thresholds and factor levels. The model guarantees that the comparison of effects are based on the same cost. For more details, the readers is referred that paper. The selections of levels and thresholds, on the other hand, will not influence the performance of CSFD. After determining the levels of factors, CSFD will then code them from -1 to +1 (Montgomery 2001).

3 CONTROLLED SEQUENTIAL FACTORIAL DESIGN (CSFD)

The first step of CSFD is to select a factorial design which will be sequentially implemented. The design will determines which effects can be screened and which cannot. Factorial design is the most widely used experiment design in practice. We will only give a short introduction here for 2-level factorial design, which is the most common one. For detailed description of factorial design and analysis, please refer to Montgomery (2001).

For L factors with two levels (low and high) each, there are in total 2^L possible treatment combinations. If a design includes all 2^L treatment combinations, it is a full factorial design; if only a fraction of those are included, it is a fractional factorial design. The key to factorial design is the design matrix X. For the 3-factor full factorial design given in Table 1, the design matrix is as below.

Table 1: 3-Factor Full Factorial Design

Run	Factorial Effect							Response
	A	B	C	AB	AC	BC	ABC	
1	-	-	-	+	+	+	-	Y_1
2	+	-	-	-	-	+	+	Y_2
3	-	+	-	-	+	-	+	Y_3
4	+	+	-	+	-	-	-	Y_4
5	-	-	+	+	-	-	+	Y_5
6	+	-	+	-	+	-	-	Y_6
7	-	+	+	-	-	+	-	Y_7
8	+	+	+	+	+	+	+	Y_8

$$X = \begin{pmatrix} +1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & +1 & -1 & +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & +1 & -1 & +1 & -1 & -1 \\ +1 & -1 & +1 & +1 & -1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \end{pmatrix}$$

In general, the design matrix has the form of $X = (X'_{R_1}, X'_{R_2}, \ldots, X'_{R_I})' = (X_{C_1}, X_{C_2}, \ldots, X_{C_J})$, where X_{R_i} and X_{C_j} stand for the ith row and the jth column of the design matrix respectively. Each row of the design matrix represents a treatment combination, i.e, a design point. Low level are coded as -1s and high levels are coded as +1s. For example, the first row represents a run with all three factors at low level. It is easy to conclude that $E[Y_i] = X_{R_i}\beta'$. Each column is associated with the level settings of a specific effect (main or interaction) at all combinations. For example, the fourth column of the design matrix is associated with the interaction effect AB. The first column is associated with β_0 and are always all +1s. Notice that the columns associated with main effects, in our case the second to the fourth column, determine the experiments. The other columns are simply the multiplication of the columns of the involved factors. For example, the AB column is the product of columns A and B. In the first combination, the AB effect is at high level, and in the second combination, the AB effect is at low level.

For a linear regression model $Y = X'\beta$, where $Y = \{Y_1, Y_2, \ldots, Y_I\}'$, the estimated effect coefficient $\hat{\beta}_j = \frac{1}{2^L} X'_{Cj} Y$. In other words, the estimated coefficient of jth effect is the half of the effect value, which is the average of all response in which the effect is at high level minus the average of all responses in which the effect is at low level. Notice that the estimation uses observations from all design points. The matrix expression of the estimation is $\hat{\beta} = (X'X)^{-1}X'Y$. For a full factorial design, all main and interaction effects can be estimated independently; for a fractional factorial design, some of the effects are confounded with others. The resolution of a fractional factorial design characterizes the degree of confounding. For example, a resolution III design can independently estimate all main effects, but some main effects are confounded with two-factor or higher-order interactions. A resolution V design has all main effects and two-factor interaction estimators independent to each others. On the other hand, the higher the resolution, the more design points are required. For more details of effect estimation in fractional factorial design, please also refer to Montgomery (2001).

Many books (e.g. Wu and Hamada, 2000) and software packages provide recommended designs for various resolutions and various values of L, the number of factors to be screened. If no recommended fractional factorial design is easily found for a particular value of L, there are many ways to construct it. Below we will explain how resolution III fractional factorial designs can be constructed for a large number factor screening experiments where only main effects exist. (This design will be used in the empirical evaluation.) For more details of methods to construct fractional factorial designs, please refer to Wu and Hamada (2000).

For a L-factor main-effects-only model, a resolution III fractional factorial design only needs 2^m design points, where m is an integer satisfying $2^{m-1} \leq L < 2^m$, to provide an estimate of each of the L main effect coefficients. We will use a 6-factor example given in Table 2 to explain the idea of constructing this design.

Table 2: Construction of Resolution III Fractional Factorial Design

Run	A	B	C	D = AB	E = AC	F = BC
1	-	-	-	+	+	+
2	+	-	-	-	-	+
3	-	+	-	-	+	-
4	+	+	-	+	-	-
5	-	-	+	+	-	-
6	+	-	+	-	+	-
7	-	+	+	-	-	+
8	+	+	+	+	+	+

For $L = 6$, $m = 3$. As shown in Table 2, The full 2^3 factorial design is first based on the first 3 factors, A, B and C. The rest of the main effects (D, E and F) are assigned to the last three columns of AB, AC and BC respectively so their main effects are confounded with corresponding interaction effects. The level settings of main effects confounded with interaction effects at each design

point are set to be the same as the level settings of the corresponding interaction effects. Thus, we have a design matrix and all main effects can be estimated.

3.1 CSFD Procedure

A generic structure of CSFD methods is given in Figure 1. Like other sequential factorial designs, CSFD first generates a small number of random samples at the beginning of the screening. Hypothesis testings are then performed sequentially to classify desired effects. Each hypothesis testing is based on all then available samples and new observations will be generated whenever larger sample size is needed to guarantee the specified Type I error and power control.

Initialization

(1) Form a queue of all desired effects.

(2) Generate n_0 replications.

While queue is not empty, do

(1) Remove an effect from the queue.

(2) Use factorial design to obtain the sample estimates and sample variance of the effect coefficient.

(3) Classify the effect using a testing procedure with error control. Generate new replications if necessary.

End While

Figure 1: Structure of CSFD

In CSFD, the observations are generated in batches with one observation from each design point. Each batch is called one *replication* and each replication can provide one estimate for every desired effects. For example, for a full factorial experiment with L factors, one replication consists of 2^L observations, one at each of the 2^L design points. Since CSFD classifies the desired effects one at a time, no assumption of the signs of the effects is required. For each effect, CSFD first computes an estimate of the effect coefficient from each available replication. An overall estimate of the effect coefficient and its sample variance are then computed based on these estimates. Finally, a hypothesis testing procedure will be used to classify this effect and determine if more replications are necessary. The notation used to define CSFD are given below.

- There are in total L indexed factors and there are K effects of interest, $1 \leq K < 2^L$.
- β_k: kth effect coefficient, $k = 1, \ldots, K$
- n_0: Number of initial replications generated at the beginning of the screening procedure
- n_k: Number of available replications at the beginning of the classification of kth effect. $n_1 = n_0$.
- j: Index of available replications, $j = 1, \ldots, n_k$
- $B_j(k)$: The j^{th} estimate of kth effect coefficient
- $\bar{B}_k = \sum_{j=1}^{n_k} B_j(k)/n_k$
- $S^2(k) = \sum_{j=1}^{n_k} \left(B_j(k) - \bar{B}_k\right)^2 /(n_k - 1)$: Sample variance $S^2(k)$ is computed based on the first n_k replications. $S^2(k)$ will not be updated when more replications are generated.

3.2 Performance of CSFD

The process to classify an effect is actually to sequentially test the following hypotheses to determine if each effect is important.

$$H_0 : \beta_k \leq \Delta_0 \text{ vs. } H_1 : \beta_k > \Delta_0.$$

We say a testing procedure is qualified if it guarantees that for any $k = 1, \ldots, K$,

- $\Pr\{\text{Declare } k\text{th effect important} \mid \beta_k \leq \Delta_0\} \leq \alpha$, and
- $\Pr\{\text{Declare } k\text{th effect important} \mid \beta_k \geq \Delta_1\} \geq \gamma$.

Then it is easy to prove the following result.

Theorem 1 *Given a qualified testing procedure, CSFD guarantees that*

$$\Pr\{\textit{Declare effect } k \textit{ important } \mid \beta_k \leq \Delta_0\} \leq \alpha$$

and

$$\Pr\{\textit{Declare effect } k \textit{ important } \mid \beta_k \geq \Delta_1\} \geq \gamma$$

for any $k = 1, \ldots, K$.

The selection of the testing procedure determined the efficiency of CSFD. Two qualified tests, 2-stage test and fully sequential (FSQ) test, have been proposed before and are introduced below. These testing procedures are given in Figures 2 and 3 respectively. For more details of 2-stage test, please refer to Wan, et al. (2005a); for FSQ test, please see Wan, et al. (2005b) and Kim and Nelson (2001).

(1) If $\bar{B}_I(k) \leq U_I(k)$ and $n_k \geq N(k)$, then classify the kth effect as unimportant,

(2) else if $\bar{B}_I(k) \leq L(k)$, then classify the effect as unimportant,

(3) else if $\bar{B}_I(k) > U_I(k)$, then classify the effect as important,

(4) else generate $(N(k) - n_k)^+$ replications,

(a) If $\bar{B}_{II}(k) < U_{II}(k)$, then classify the effect as unimportant,

(b) else classify the effect as important.

Figure 2: 2-stage Testing Procedure

(1) Set $r(k) = n_k$.
(2) If $r(k) > M(k)$,
  (a) If $\bar{B}(k) < (\Delta_1 + \Delta_0)/2$, then classify the effect as unimportant,
  (b) else classify the effect as important.
(3) If $r(k) \leq M(k)$,
  (a) If $r(k)(\bar{B}(k) - (\Delta_1 + \Delta_0)/2) \leq \lambda r(k) - a(k)$, then classify the effect as unimportant,
  (b) else if $r(k)(\bar{B}(k) - (\Delta_1 + \Delta_0)/2) \geq a(k) - \lambda r(k)$, then classify the effect as important,
  (c) else generate one more replication; update $\bar{B}(k)$; set $r(k) = r(k) + 1$; go to Step (3).

Figure 3: Fully Sequential Testing Procedure

Below notations are used exclusively for CSFD methods with 2-stage testing procedure.

- h: A constant satisfying $\Pr(T \leq t_{1-\sqrt{1-\alpha},n_k-1} - h) = (1-\gamma)/2$, where T is a t-distributed random variable with $n_k - 1$ degrees of freedom
- $N(k) = \lceil h^2 S^2(k)/(\Delta_1 - \Delta_0)^2 \rceil$: The minimum number of replications required to achieve the specified type I error and power
- $\bar{B}_I(k) = \sum_{j=1}^{n_k} \frac{B_j(k)}{n_k}$ and $\bar{B}_{II}(k) = \sum_{j=1}^{N(k)} \frac{B_j(k)}{N(k)}$: The overall estimates of kth effect coefficient in first and second stages
- $U_I(k) = \Delta_0 + t_{1-\sqrt{1-\alpha},n_k-1} S(k)/\sqrt{n_k}$ and $U_{II}(k) = \Delta_0 + t_{1-\sqrt{1-\alpha},n_k-1} S(k)/\sqrt{N(k)}$
- $L(k) = \Delta_0 - t_{\frac{1-\gamma}{2},n_k-1} S(k)/\sqrt{n_k}$

The following notations are used exclusively for CSFD methods with FSQ testing procedure.

- $\lambda = (\Delta_1 - \Delta_0)/4$
- $\eta = (\exp(\varphi) - 1)/2$, where $\varphi = 2\ln(2\alpha)/(1-n_k)$.
- $a(k) = 2\eta(n_k - 1)S^2(k)/(\Delta_1 - \Delta_0)$
- $M(k) = \lfloor a(k)/\lambda \rfloor$
- $\bar{B}(k) = \sum_{j=1}^{r(k)} B_j(k)/r(k)$, where $r(k)$ is the number of current available replications

4 EMPIRICAL EVALUATION

In this section, we present some numerical results of artificial examples to compare CSFD and CSB (CSB-X). The results demonstrated below are based on fully sequential tests, which usually requires less simulation observations than two-stage tests. For comparison of the two tests see Wan, Ankenman and Nelson (2005b). If two-stage tests are used, similar conclusion can be made. Normal errors are assumed with mean 0 and standard deviation, $\sigma = m * (1 +$ size of the group effect). That is, the standard deviation is proportional to the size of the effect being screened. Common random numbers were not employed. For each case considered, both CSFD and CSB procedures are applied 1000 times and the percentage of times the kth effect is declared important is recorded; this is an unbiased estimate of Pr{Declare effect k important}. The number of runs reported for each method is the multiplication of number of replications used and number of observations for each replication, i.e., the number of design points. It is the total computational effort required for screening.

4.1 CSFD vs. CSB in Presence of Interactions

In this section, we compare the effectiveness and efficiency of CSFD and CSB (or CSB-X) methods on two sets of factors ($L = 10$) with interactions. The simulation parameters are listed in Table 3. In the implementation of CSB methods, the number of initial simulation runs at each bifurcation step is set to be 25, which is the same as in Wan, et al. (2005a). The numbers of initial simulation replications for CSFD ranges from 2 to 5. The influence of initial sample size on the efficiency of the method will be discussed in Section 4.3.

Table 3: Parameters for Small Scale Cases

Parameter	Value
Δ_0	2
Δ_1	4
α	0.05
γ	0.95
σ	$m*$(1 + size of the group effect)
m	0.1, 1

4.1.1 Second-Order-Interaction Cases

We first apply CSFD, CSB and CSB-X on cases where only main effects and second-order interaction effects exist. For those cases, a 2_V^{10-3} fractional factorial design is sufficient for CSFD.

Case 1. Main effects: $(\beta_1, \beta_2, \ldots, \beta_{10})$ =(2, 2, 2, 2.44, 2.88, 3.32, 3.76, 4.2, 4.64, 5). Interaction effects: $(\beta_{12}, \beta_{46}, \beta_{58})$ = (1.75, -2.5, 3.9). All other second or higher order interactions are zero. For an effective screening procedure, the probability that $\beta_1, \ldots, \beta_3$ is declared important should be smaller than $\alpha = 0.05$, but for $\beta_8, \ldots, \beta_{10}$ it should be greater than $\gamma = 0.95$.

Tables 4 and 5 presents P(DI) (probability of declaring important) and average number of simulation runs of case 1, small variance and large variance cases respectively. We can see that in the presence of second-order interaction, both CSB-X and CSFD give desired screening results, CSB on the other hand, gives misleading results for some fac-

Table 4: Screening Results in Case 1 ($m = 0.1$)

Effect	CSB	CSB-X	CSFD
$\beta_1 = 2$	0.000	0.000	0.000
$\beta_2 = 2$	1.000	0.000	0.000
$\beta_3 = 2$	0.000	0.000	0.000
$\beta_4 = 2.44$	0.031	0.000	0.000
$\beta_5 = 2.88$	0.382	0.302	0.073
$\beta_6 = 3.22$	0.000	0.872	1.000
$\beta_7 = 3.76$	0.975	0.985	1.000
$\beta_8 = 4.2$	1.000	1.000	1.000
$\beta_9 = 4.64$	1.000	1.000	1.000
$\beta_{10} = 5$	1.000	1.000	1.000
$\beta_{1,2} = 1.75$	N/A	N/A	0.000
$\beta_{4,6} = -2.5$	N/A	N/A	0.000
$\beta_{5,8} = 3.9$	N/A	N/A	1.000
# of runs	343	279	256

Table 5: Screening Results in Case 1 ($m = 1.0$)

Effect	CSB	CSB-X	CSFD
$\beta_1 = 2$	0.013	0.008	0.000
$\beta_2 = 2$	0.958	0.009	0.004
$\beta_3 = 2$	0.007	0.007	0.002
$\beta_4 = 2.44$	0.098	0.074	0.051
$\beta_5 = 2.88$	0.373	0.363	0.357
$\beta_6 = 3.22$	0.000	0.785	0.865
$\beta_7 = 3.76$	0.973	0.969	0.991
$\beta_8 = 4.2$	1.000	0.995	0.999
$\beta_9 = 4.64$	1.000	0.999	1.000
$\beta_{10} = 5$	1.000	0.999	1.000
$\beta_{1,2} = 1.75$	N/A	N/A	0.000
$\beta_{4,6} = -2.5$	N/A	N/A	0.050
$\beta_{5,8} = 3.9$	N/A	N/A	0.992
# of runs	22247	10734	1569

tors. Even unimportant interaction effects may affect the classification of the main effects in CSB. For example, in both small variance and large variance cases, the existence of unimportant interaction $\beta_{1,2}$ affects the effectiveness of CSB in classifying effect β_2. The average numbers of simulation runs show that CSFD is as efficient as CSB in small variance case and is more efficient than CSB in large variance case.

4.1.2 3rd-and-Higher-Order-Interaction Case

We then apply CSFD and CSB-X on a case where third and higher order interaction effects also exist. A 2^{10} full factorial design is used in CSFD for the general case.

Case 2. Parameter settings and effect coefficients of case 2 are the same as those of case 1 except for the addition of two non-zero 3rd-order interaction effects, $(\beta_{123}, \beta_{789}) =$ (1.9, -4.5). All other 3rd-order and higher order interaction effects are zero.

Table 6 shows that CSFD is effective in classifying both main effects and interaction effects, but CSB-X is not effective when odd-order interaction effects exist. For example, CSB-X misclassifies β_3 and β_9 in all 1000 trials.

Table 6: Screening Results in Case 2

	$m = 0.1$		$m = 1.0$	
Effect	CSB-X	CSFD	CSB-X	CSFD
$\beta_1 = 2$	0.000	0.000	0.006	0.000
$\beta_2 = 2$	0.000	0.000	0.009	0.000
$\beta_3 = 2$	1.000	0.000	0.987	0.000
$\beta_4 = 2.44$	0.004	0.000	0.083	0.008
$\beta_5 = 2.88$	0.324	0.000	0.373	0.273
$\beta_6 = 3.22$	0.855	1.000	0.789	0.918
$\beta_7 = 3.76$	0.985	1.000	0.971	1.000
$\beta_8 = 4.2$	0.998	1.000	0.993	1.000
$\beta_9 = 4.64$	0.000	1.000	0.000	1.000
$\beta_{10} = 5$	1.000	1.000	1.000	1.000
$\beta_{1,2} = 1.75$	N/A	0.000	N/A	0.000
$\beta_{4,6} = -2.5$	N/A	0.000	N/A	0.005
$\beta_{5,8} = 3.9$	N/A	1.000	N/A	1.000
$\beta_{1,2,3} = 1.9$	N/A	0.000	N/A	0.000
$\beta_{7,8.9} = -4.5$	N/A	1.000	N/A	1.000
# of runs	279	2048	11632	3515

4.2 CSFD vs. CSB on Large-Scale Cases

In this section we compare the efficiency of CSFD and CSB methods on screening problems with large scale number of factors. We will consider two cases with 200 factors and 500 factors respectively and in all cases main-effects-only model is assumed.

The simulation parameters are same as those of previous cases given in Table 3 except that the variance factor m takes values of 0.01, 0.1 and 0.3. For both 200 factors and 500 factors cases, the initial number of simulation runs at each bifurcation step of CSB is 25 and the initial number of replications of CSFD is 2. In both cases, there are 4% of the factors which are important. The important factors have effect coefficient equal to 5 and the unimportant factors have effect coefficient equal to 0.

For each case, there are three scenarios. The first scenario has all important factors clustered together with the smallest indices so that the number of important groups is as small as possible at each step of CSB. The second scenario has the important factors evenly spread so there are the maximum number of important groups remaining at each step. The third scenario has the important factors randomly spread.

Because there is no interaction exists, CSFD can use resolution III fractional factorial designs discussed in Section 3. For these 200 factors and 500 factors cases, CSFD needs 256 and 512 simulation runs in each replication respectively. Both methods provide the desired error control for main

effects. The average numbers of total simulation runs of CSB and CSFD methods on large scale cases are listed in Table 7. We can see that the efficiency of CSFD is more robust to different system configurations. When variance is relatively large, CSFD is more efficient than CSB.

Table 7: Number of runs in Large Scale Cases

	$m = 0.01$		$m = 0.1$		$m = 0.3$	
Scenarios	CSB	CSFD	CSB	CSFD	CSB	CSFD
200, 1st	80	512	811	512	6440	515
200, 2nd	245	512	1828	512	12550	512
200, 3rd	220	512	1633	512	10652	512
500, 1st	145	1024	5676	1024	51903	1070
500, 2nd	598	1024	19036	1024	160916	1060
500, 3rd	521	1024	15183	1024	133799	1060

4.3 Further Discussions

1. The relationship of the number of total simulation runs with the initial replication number n_0, obtained by applying CSFD with full factorial design and FSQ testing procedure on case 2 when $m = 1$, is demonstrated in Figure 4. This graph is consistent with Figure 18.1 in Kim and Nelson (2006), which presents the typical form of expected total simulation runs as a function of n_0. These figures show that for each case there exists an "optimal" initial replication number; a too small n_0 usually leads to huge penalty and a too large n_0 is usually unnecessary. However this optimal number is usually unknown because of the unknown variance.

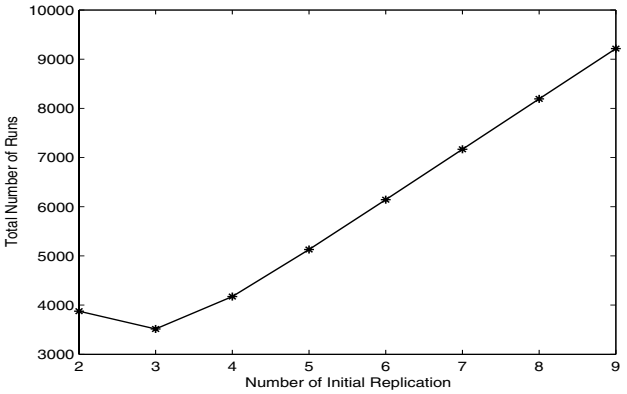

Figure 4: Total Simulation Runs vs. n_0

The reason of this "convexity" is that when n_0 is too small, it produces large $S^2(k)$ and $N(k)$ which then leads to a large total replication number. Also notice that the right side of the curve is actually a straight line. This implies that when n_0 is too large, all of the sequential testing procedures reach conclusion in the first stage. This "convexity" property is also true for the CSB methods.

2. With other system parameters fixed, the larger the variance, usually the larger the "optimal" number of initial replications and the larger the total number of simulation runs.
3. We observed that the number of total runs sometimes decreases as the number of factors increases. This is because when the number of factors increases, the number of simulation runs needed to form one replication in factorial design increases exponentially. The significantly larger number of simulation runs in one replication produces more accurate estimates of the effect coefficients and smaller sample variance $S^2(k)$, which then leads to smaller minimum required replication number $N(k)$. This property depends on the variance structure of different levels and will be explored further in future.

5 CONCLUSION

CSFD offers a factor-screening approach which not only provides simultaneous error and power control but also has the flexibility to be applied to different situations. When interaction effects exist, CSFD requires no prior effect information and is able to classify any desired effect; when there is no interaction or no high-order interaction, CSFD with fractional design can perform the classification with high efficiency. CSFD has also shown robust performance for system with unknown variance. Future research will concentrate on developing hybrid methods of grouping screening and factorial design for better performance.

ACKNOWLEDGMENT

This research was partially supported by Purdue Research Foundation (PRF). Additional thanks go to Professor Bruce Ankenman from Northwestern University.

REFERENCES

Campolongo, F., J. P. C. Kleijnen, and T. Andres. 2000. Screening Methods, In *Sensitivity Analysis*. ed. A. Saltelli, K. Chan, E. M. Scott. New York: John Wiley & Sons Inc.

Gilmour, S. G. and R. Mead. 1995. Stopping rules for sequences of factorial designs. *Applied Statistics* 44(3): 343-355.

Hunter, J. S. 1964. Sequential factorial estimation. *Technometrics* 6(1): 41-55.

Kim, S.-H. and B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transaction on Modeling and Computer Simulation* 11(3): 251-273.

Kim, S.-H. and B. L. Nelson. 2006. Chapter 18 in Elsevier Handbooks in Operations Research and Management Science: Simulation. ed. S. G. Henderson and B. L. Nelson. Elsevier. Available online via <http://users.iems.northwestern.edu/~nelsonb/Publications/Chapter18Revised012504.pdf> [accessed May 10, 2005].

Montgomery, D. C. 2001. *Design and analysis of experiments, 5th Edition*. New York: John Wiley & Sons Inc.

Myers, R. H. and D. C. Montgomery. 1995. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. New York: John Wiley & Sons Inc.

Trocine, L. and L. Malone. 2000. Finding important independent variables through screening designs: A comparison of methods. In *Proceeding of the 2000 Winter Simulation Conference*, ed. J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 749–754. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers. Available online via <http://www.informs-cs.org/wsc00papers/098.PDF> [accessed July 6, 2003].

Trocine, L. and L. Malone. 2001. An overview of newer, advanced screening methods for the initial phase in an experimental design. In *Proceeding of the 2001 Winter Simulation Conference*, ed. B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, 169–178. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers. Available online via <http://www.informs-cs.org/wsc01papers/020.PDF> [accessed July 6, 2003].

Wan, H., B. Ankenman, and B. L. Nelson. 2003. Controlled sequential bifurcation: A new factor-screening method for discrete-event simulation. In *Proceeding of the 2003 Winter Simulation Conference*, ed. S. Chick, P. J. Sánchez, D. Derrin, and D. J. Morrice, 565-573. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Wan, H. 2004. Simulation factor screening with controlled sequential bifurcation. Ph.D thesis. Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208-3119, USA.

Wan, H., B. Ankenman, and B. L. Nelson. 2005a. Controlled sequential bifurcation: A new factor-screening method for discrete-event simulation. Operations Research, to appear.

Wan, H., B. Ankenman, and B. L. Nelson. 2005b. Simulation factor screening with controlled sequential bifurcation in the presence of interactions. Working Paper.

Wu, J. C. F., and M. Hamada. 2000. *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York: John Wiley & Sons, Inc.

Zacks, S. 1968. Bayes sequential design of fractional factorial experiment for the estimation of a subgroup of pre-assigned parameters. *The Annals of Mathematical Statistics* 39(3): 973-982.

AUTHOR BIOGRAPHIES

HONG WAN is an Assistant Professor in the School of Industrial Engineering at Purdue University. Her current research is in factor-screening methods for discrete-event simulation. The other projects in which she has participated include industrial product design and quality control; and software for strategic management of transportation systems. Her research interests include experimental design and quality control, simulation of manufacturing and financial systems, and simulation input analysis. She has taught a variety of courses and is a member of INFORMS and ASA. Her e-mail and web addresses are <hwan@purdue.edu> and <http://web.ics.purdue.edu/~hwan/index.htm>

HUA SHEN is doctoral student in the School of Industrial Engineering at Purdue University. His dissertation topic is factor-screening methods for discrete-event simulation. The other projects in which he has participated include supply chain management and statistical consulting service. His research interests include experimental design and quality control, simulation of manufacturing systems, and simulation input analysis. His email address is <shenh@purdue.edu>.