NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

FRACTIONAL FACTORIAL CONTROLLED SEQUENTIAL BIFURCATION: EFFICIENT FACTOR SCREENING THROUGH DIVIDE AND DISCARD

by

Regine Pei Tze Oh

December 2007

Thesis Co-Advisors:	Susan M. Sanchez
	Thomas W. Lucas
Second Reader:	Mark E. Nissen

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE December 2007	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE Fractional Factorial Controlled Sequential Bifurcation: Efficient Factor Screening through Divide and Discard

5. FUNDING NUMBERS

6. AUTHOR(S) Regine Pei Tze Oh

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense, the U.S. Government, the Ministry of Defence of Singapore, or that of the Singapore Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

On any given day, organizations use software simulations to make better decisions. Software simulations of real world systems are often large and rich with many parameters potentially affecting outcomes. Faced with a multitude of parameters, decision makers may not know or may lose sight of the few truly critical factors. Thus, screening algorithms are essential in order to identify the factors that most impact outcome measures. This enables experimenters to better utilize their resources by focusing on truly important factors.

Fractional Factorial Controlled Sequential Bifurcation (FFCSB) is a newly proposed two-phase screening procedure for large-scale simulation experiments. This thesis evaluates the performance of FFCSB from accuracy and efficiency perspectives. FFCSB is also compared to existing algorithms, Controlled Sequential Bifurcation (CSB) and Fractional Factorial (FF), in order to understand the relative merits and weaknesses of each algorithm. FFCSB delivers consistent accuracy guarantees across more factor patterns and offers efficiency savings over CSB. FFCSB and FF are equally matched in accuracy; however, FFCSB is more robust to non-ideal settings of control parameters and scales better with increasing response model size; conversely FFCSB can be less efficient than FF. A first-case application of FFCSB on the Hierarchy organizational model yields results in agreement with prior research, as well as providing interesting hypotheses for further exploration. The Hierarchy model serves as a benchmark to compare innovative Command and Control structures for enabling more effective warfare.

14. SUBJECT TERMS Screening Algorithms, Fractional Factorial Controlled Sequential Bifurcation, FFCSB, Command and Control, C2

15. NUMBER OF PAGES
112

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

FRACTIONAL FACTORIAL CONTROLLED SEQUENTIAL BIFURCATION: EFFICIENT FACTOR SCREENING THROUGH DIVIDE AND DISCARD

Regine Pei Tze Oh
Civilian, DSO National Laboratories, Singapore
B.A.Sc. (Honors), University of Toronto, 2001

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
December 2007

Author: Regine Pei Tze Oh

Approved by: Susan M. Sanchez
Thesis Co-Advisor

Thomas W. Lucas
Thesis Co-Advisor

Mark E. Nissen
Second Reader

James N. Eagle
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

On any given day, organizations use software simulations to make better decisions. Software simulations of real world systems are often large and rich with many parameters potentially affecting outcomes. Faced with a multitude of parameters, decision makers may not know or may lose sight of the few truly critical factors. Thus, screening algorithms are essential in order to identify the factors that most impact outcome measures. This enables experimenters to better utilize their resources by focusing on truly important factors.

Fractional Factorial Controlled Sequential Bifurcation (FFCSB) is a newly proposed two-phase screening procedure for large-scale simulation experiments. This thesis evaluates the performance of FFCSB from accuracy and efficiency perspectives. FFCSB is also compared to existing algorithms, Controlled Sequential Bifurcation (CSB) and Fractional Factorial (FF), in order to understand the relative merits and weaknesses of each algorithm. FFCSB delivers consistent accuracy guarantees across more factor patterns and offers efficiency savings over CSB. FFCSB and FF are equally matched in accuracy; however, FFCSB is more robust to non-ideal settings of control parameters and scales better with increasing response model size; conversely FFCSB can be less efficient than FF. A first-case application of FFCSB on the Hierarchy organizational model yields results in agreement with prior research, as well as providing interesting hypotheses for further exploration. The Hierarchy model serves as a benchmark to compare innovative Command and Control structures for enabling more effective warfare.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF EQUATIONS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS

CEP	Center for Edge Power
CSB	Controlled sequential bifurcation
FF	Fractional factorial
FFCSB	Fractional factorial controlled sequential bifurcation
FTE	Full Time Equivalent
MOE	Measure of Effectiveness
MOP	Measure of Performance
POW-ER	Projects, Organizations and Work for Edge Research
SB	Sequential bifurcation
SEED	Simulation Experiments and Efficient Design

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

On any given day, organizations use software simulations to make better decisions. Software simulations of real world systems are often large and rich with many parameters potentially affecting outcomes. Faced with a multitude of parameters, decision makers may not know or may lose sight of the few truly critical factors. Thus, screening algorithms are essential in order to identify the factors that most impact outcome measures. This enables experimenters to better utilize their resources by focusing on truly important factors.

Fractional factorial controlled sequential bifurcation (FFCSB) is a newly proposed screening procedure for large-scale simulation experiments and offers several enhancements over conventional screening algorithms. First, FFCSB dramatically reduces the need for *a priori* knowledge on the direction of factor effects, which is often a condition for optimal performance of conventional algorithms and has proven difficult to meet. FFCSB also does not require *a priori* knowledge of the number of experiments required for factor classification. It conducts sufficient experiments to complete classification. Second, FFCSB scales well for large scale models with thousands of factors. Third, FFCSB provides accuracy guarantees in its factor classification. Fourth, FFCSB provides a savings in computation.

This thesis conducts controlled experiments to evaluate the performance of FFCSB from accuracy and efficiency perspectives. FFCSB is also compared to existing screening methods, Controlled Sequential Bifurcation (CSB) and Fractional Factorial (FF), in order to understand the relative merits and weaknesses of each algorithm. FFCSB assumes a main effects response model. This basic model is varied for experimentation via: (1) increasing factor counts from 2^3 to 2^{10}, (2) different types of response variance (homogeneous versus heterogeneous), (3) different magnitudes of response variances, and (4) different factor patterns (mix of factor effect direction). In a first-case application, FFCSB is used to support current research in Computation Organization Theory.

For experiments on models displaying homogeneous response variances:

1. FFCSB fulfills accuracy guarantees for all factor patterns. It maintains consistent performance for all factor patterns, model sizes and variance magnitudes.
2. FFCSB is more robust than CSB in handling a mix of factor effects and offers up to 25% in computation savings. The mix of factor effects causes CSB to fare poorly as factors of opposite directions in the same screening group cancel out one another's effects. FFCSB averts this undesirable phenomenon via the FF pre-sorting phase to divide the entire factor space into positive and negative groups for CSB screening.
3. FFCSB and FF are equally matched in accuracy, but FFCSB can be less efficient than FF. However, FFCSB is more robust to non-ideal settings of control parameters, which often happens when exploring response models. Also, FFCSB does not require *a priori* knowledge of the number of experiments to conduct for complete factor classification, as FF does.

For experiments on models displaying heterogeneous response variances:

1. FFCSB fulfills accuracy guarantees for three of the five factor patterns simulated. It fails when there are significant percentages of opposite factor effects that are not negligible in effect and yet not critical enough to be classified. Hence, these effects distort the factor classification accuracy. In the three favorable factor patterns, FFCSB is robust to variance magnitudes and model sizes. In the two unfavorable factor patterns, FFCSB deteriorates with variance magnitudes and model size.
2. FFCSB is more robust than CSB in handling a mix of factor effects and offers at least a 30% computation savings.
3. In the three FFCSB favorable factor patterns, FFCSB fulfills accuracy guarantees better than FF. FF accuracy scales poorly with increasing model size.

In a first-case application, FFCSB is applied to the Hierarchy organizational model. In *Joint Vision* 2010, Army General John M. Shalikashvili, Chairman of the Joint Chiefs of Staff, said "The nature of modern warfare demands that we fight as a joint team. This was important yesterday, it is essential today, and it will be even more imperative tomorrow." In light of the critical Transformation drive of the U.S. military, innovative organizational models are needed to deliver better team performance.

At the Naval Postgraduate School, the Center for Edge Power (CEP) is active in multi-disciplinary research on network-centric operations to enable more powerful warfare. In collaboration with Stanford University, CEP conducts computational experimentation on various command and control structures to understand the factors that drive team performance, which can be measured in various forms (e.g., duration, risk and cost.) The experimentation is enabled by a powerful modeling environment, POW-ER—Projects, Organizations, Work and Edge Research—developed by the Virtual Design Team Research Group at Stanford. The POW-ER environment and its models are well grounded in sound research and extensively validated. In particular, the Hierarchy model is representative of many military organizations and serves as a benchmark for comparisons of new organizational forms. FFCSB is applied to identify important factors in the Hierarchy model that drive the Measure of Performance of Project Duration.

In prior computation experimentation on organizational models, researchers typically use full factorial designs of experiments. Given the computation intensity or infeasibility posed by hundreds or thousands of factors in such complex models, the designs were restricted to block changes of factor groups instead of single factor resolution. FFCSB extends the suite of tools available to tackle the question of team performance from an alternate perspective. It offers single factor resolution, allowing researchers to probe: Which single factors are most important in influencing the MOP? Table 1 lists the factor space for exploration of the Hierarchy model.

Table 1. Factor Space for Exploration of Hierarchy Model

Mission & Environment	**Network Architecture**	**Professional Competency**
(Project) Function Exception Probability (Project) Project Exception Probability (Task) Effort (Task) Learning Days (Task) Priority (Task) Requirement Complexity (Task) Solution Complexity (Task) Uncertainty (Personnel) Full Time Equivalent (Personnel-Task) Allocation (Task-Task) Successor	(Project) Priority (Project) Length Of Work-day (Project) Length Of Work-week (Project) Centralization (Project) Matrix-strength (Project) Communication Probability (Project) Noise Probability (Project) Instance Exception Probability (Meeting) Priority (Meeting) Duration (Personnel-Meeting) Allocation (Task-Task) Rework Strength	(Project) Team Experience (Personnel) Culture (Personnel) Role (Personnel) Application Experience (Personnel) Cultural Experience (Personnel) Skill Ratings

Table 2 lists the expert opinion and FFCSB findings on important factors that impact Project Duration most. FFCSB findings were in agreement with expert opinion on two out of four factors. For the important factors of Task Effort and Personnel Skill Ratings, FFCSB further quantified that these factors are important only for missions and personnel associated with the critical path. Such findings were possible because FFSCB offers single factor resolution.

Table 2. FFCSB Findings in Partial Agreement with Expert Opinion

Expert Opinion Important Factor	FFCSB Finding
Manpower available (FTE)	Relatively not important.
Task Effort	Relatively important. Only for missions with minimum float on critical path
Application Experience	Relatively not important.
Skill Ratings	Relatively important. Only for personnel working on missions of shorter duration and lying on critical path

In addition, FFCSB provided interesting observations.

1. There were other relatively important factors that drive Project Duration in the Hierarchy model.
    a. Project Exception Probability (probability that a subtask will fail and generate rework for failure dependent tasks).
    b. Task Requirement Complexity & Task Solution Complexity—Only for missions on critical path.
    c. Team Experience (Familiarity of team working together).
2. There were no important factors in the Network Architecture subspace.
3. Counter to intuition, higher Team Experience led to longer Project Duration. This mirrors a similar finding from earlier research. Had the original intuition been used with conventional screening algorithms, this factor could have distorted factor classification.
4. The Hierarchy model has a 3-tier command chain that models the Command, Coordination and Operations layers in a Joint Task Force. There were more important factors associated with the Operations layer than the other layers.

5. There were more uncontrollable or difficult to control factors (e.g., Project Exception Probability, Task Requirement Complexity, Task Solution Complexity and Team Experience) than controllable or easy to control factors (e.g., Skill Ratings).

There are limitations to the FFCSB application to any model. FFCSB assumes a main effects model, and interactions can distort the accuracy of factor classification. The nature of the response variance (homogeneous or heterogeneous) and its magnitude are unknown. Both model characteristics could have bearings on the FFCSB findings and accuracy guarantees. Particular to the Hierarchy model, the observations of this FFCSB exploration are unique to the factor space organization and ranges of exploration. Hence, the findings are not conclusive of the Hierarchy model. The important factor classification and observations are meant to provide direction for researchers in future work and optimize their experimentation budget on truly important factors. This first-case FFCSB application on a real-world simulation model has produced results that are coherent with critical path analysis and that agree with earlier research on similar models. Hence, it is an encouraging sign that FFCSB can serve as a complementary tool to better understand complex simulation models.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Foremost, I would like to thank my advisors, Professor Lucas, Professor Sanchez and Professor Nissen. Your guidance, feedback and positive encouragement were instrumental to the completion of this work. To Mr. Steve Upton of the SEED center and Mr. Marc Ramsey of Stanford University, thank you for your invaluable support in the application. To the great folks at the SEED center, you have been a wonderfully encouraging and responsive audience. Your feedback improved my work products.

To my colleagues at DSO National Laboratories, thank you for the opportunity and for believing in me. The year has been an amazing eye opener into the world of Operations Research. Exciting possibilities lay ahead.

Family and friends have always been my pillars of strength. To Grandma, thank you for helping me keep my promise. To Wui Liang and our families, your love and support are the foundation of my achievements. To the Michelsons, your warmth and laughter enriches my every day. To Jim and Gil, you always had the right words to motivate and to comfort. Thank you.

To life, thank you. What doesn't kill us will make us stronger.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. MOTIVATION

On any given day, organizations use software simulations to make better decisions, which range from designing robust systems, optimizing production flows, planning humanitarian missions to optimizing battlefield logistics. Such systems of interest are often complex, with many factors potentially affecting outcomes. Consequently, the software simulations used to study them are large and rich with many simulation parameters. Faced with a multitude of simulation parameters, decision makers may not know or may lose sight of the few factors that truly impact the outcome measure. Thus, screening algorithms are essential in order to identify the factors that most impact outcome measures. This enables experimenters to better utilize their resources by focusing on truly important factors.

B. REAL WORLD SIGNIFICANCE OF SCREENING EXPERIMENTS

Screening algorithms are fast and efficient methods for identifying important factors, especially when only a few factors are truly important amidst the multitude of factors being simulated and considered. This latter fact, also known as the Pareto principle, has often been empirically proven true (Bettonvil and Kleijnen, 1997).

Group screening is a subset of screening algorithms that conducts batch tests in order to classify the important factors in as few rounds of testing as possible, while meeting the accuracy performance of the algorithm. Groups of factors are aggregated for testing. If there is an indication that some of the factors are significant, they are further divided into smaller groups for additional testing. If there is no indication, the group of factors is discarded and no longer considered. As such, group screening is extremely useful when experimental resources are scarce.

Screening was first developed to be used on physical experiments, such as drug testing on humans, manufacturing runs and crop-mix planning in agriculture. These physical experiments were often one-shot experiments with few opportunities for replications, and the motivation of the screening was to optimize identification of significant factors in as few experiments as possible. As software simulations evolved in importance and gained widespread application, screening algorithms have also evolved to screen factors with computer simulation experiments. The relatively low cost and ease of software experimentation have proved invaluable to decision making in many fields. As the speed of computing increased exponentially with decreasing costs, expectations of modeling complexity also rose. Computer simulations grew bigger, with more factors for realism and detail. Consequently, screening algorithms for simulation experiments have also become more important. Using "just enough" computation resources, these algorithms enable decision makers to focus on the most critical of decision factors—those which truly matter.

C. RELATED WORK LEADING TO FFCSB

Early group screening algorithms revolved mainly around physical experiments, which tend to be costly, difficult to control and possibly non-repeatable. In World War II, group screening was used to cheaply and quickly test new recruits for syphilis (Dorfman, 1943). Given this nature of the physical experiments, traditional screening methods typically prioritize the minimum number of experiments to categorize all factors over the accuracy performance of factor categorization (Shen and Wan, 2005).

In the last decade, group screening has been applied to simulation experiments. Owing to the distinct differences between both forms of experiments, screening algorithms for simulation experiments are not subject to the same constraints as those for physical experiments. It is much easier to change factor levels in simulation experiments than in physical experiments. Consequently, it is easier and more possible to control and study more factors in simulation experiments than in physical experiments (Wan, Ankenman and Nelson, 2003). Furthermore, screening algorithms for simulation experiments can capitalize on the sequential nature of simulation experiments. Unlike screening algorithms for physical experiments, screening algorithms for simulation experiments place greater emphasis on the accuracy performance of correctly classifying important factors.

Bettonvil and Kleijnen's (1997) "Searching for important factors in simulation models with many factors: Sequential bifurcation" first proposed the Sequential Bifurcation (SB) step-down procedure for deterministic computer simulations, and Cheng's (1997) "Searching for important factors: Sequential bifurcation under uncertainty" extended the algorithm to accommodate stochastic responses with homogeneous variances. The SB algorithm and its variants depart from many traditional screening algorithms with its sequential characteristic, meaning that new design points evolve with experimental results from previous design points. Thus, information is accumulated over experiments, unlike in traditional screening methods for physical experiments. Performance-wise, the SB algorithm is highly efficient in terms of experiment count when significant factors are sparse and clustered, albeit with "no accuracy guarantee in the stochastic simulation case" (Shen and Wan, 2005).

Wan, Ankenman and Nelson (2003) proposed Controlled Sequential Bifurcation (CSB), which builds upon the SB procedure from Bettonvil and Kleijnen (1997). CSB enhances SB by providing accuracy guarantees: Type I Error and power. Type I Error is defined as the probability of incorrectly declaring an unimportant factor as important, while power is defined as the probability of correctly declaring an important factor as important. CSB accuracy performance is robust for response meta-models with heterogeneous variances.

In general, SB and CSB algorithms work well within certain boundaries. They assume a main effects meta-model for the simulation response and require *a priori* knowledge of the directions of the factors to avoid cancellation between factors of opposite directions within the same group. The efficiency of these algorithms improves when significant factors are aggregated and sorted for screening. In reality, knowledge on significant factors is often non-existent or imperfect. Hence, it is difficult to avoid factor cancellation and to realize the optimal performance of CSB or SB algorithms.

Fractional factorial CSB (FFCSB) is a hybrid algorithm that eliminates the need for *a priori* knowledge of the direction of factor effects by using a nearly-saturated FF design to prescreen factors before the conventional CSB algorithm. Sanchez, Wan and Lucas (2005) provide empirical results to illustrate that accuracy performance guarantees are met as well as efficiency enhancements over the CSB algorithm.

The SEED (Simulation Experiments and Efficient Design) Center for Data Farming at the Naval Postgraduate School is interested in tools to facilitate large-scale experimental designs. FFCSB adds to this ever-expanding suite of tools.

D. DESCRIPTION OF FFCSB ALGORITHM

FFCSB is a two-phase hybrid algorithm for simulation factor screening, as illustrated below. The first phase of FFCSB uses a nearly-saturated fractional factorial (FF) design to estimate the direction of factor effects and classify factors into a positive and a negative group of factors. The second phase of FFCSB uses the CSB algorithm to screen factors within the each group of factors.

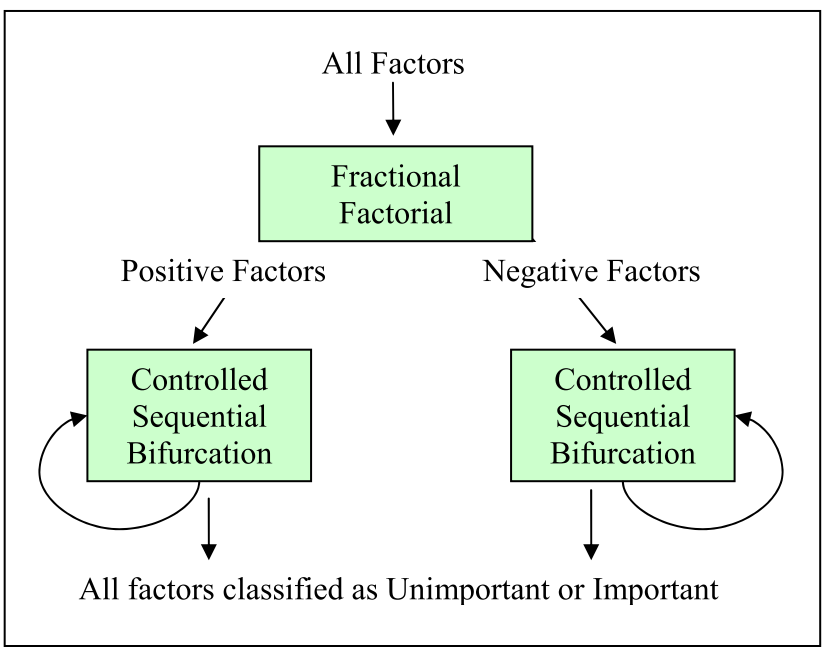

Figure 1. FFCSB: Two Phase Hybrid Algorithm for Factor Screening

Similar to SB and CSB, FFCSB assumes a main effects model over the factor space of exploration. The following response meta-model describes the observed output Y from a simulation experiment as a function of K factors of interest with factor effects β_i ($i \in 1,2,...K$) and normally distributed errors ε with variance σ^2.

$$Y = \beta_O + \sum_{i=1}^{K} \beta_i x_i + \varepsilon_i \text{ where } \varepsilon_i \sim N(0, \sigma^2)$$

Equation 1. Basic Main Effects Response Model

The main effects only assumption is not restrictive for a group screening algorithm for the purpose of quickly identifying important factors for more detailed post-FFCSB experimentation. Interactions are not identified, but they are not lost if their parent factors are identified for further study.

The FFCSB implementation for this thesis uses Resolution III FF designs and the fully sequential version of CSB as proposed in Wan, Ankenman and Nelson (2003, 2006). CSB uses two factor thresholds: (1) threshold for importance Δ_0 and (2) threshold for criticality Δ_1. CSB guarantees Type I Error no greater than α for unimportant factors (defined by effect magnitude $|\beta| < \Delta_0$) and power no less than γ for critical factors (defined by effect magnitude $|\beta| > \Delta_1$). CSB works best when all factor effects are of the same sign and factor effects are ordered. Both conditions, when met, imply that critical factors will not cancel one another's effects during the screening experiments, and unimportant factors can get eliminated quickly, thus saving on experiments.

The FF stage dramatically reduces the need for *a priori* knowledge of factor effect direction and prepares the factors for CSB. It categorizes the entire factor space by the direction of factor effect and reduces the possibility of including two critical factors of different directions in the same group for CSB classification. With the initial factor effect estimate, factors can be sorted for CSB classification. The Resolution III FF design used in this implementation is most efficient for estimating main effects, but can be potentially confounded with two-factor interactions, if any. The following table details the structure of the FFCSB algorithm.

Table 3. Structure of FFCSB

Initialization:
 Create two empty LIFO queues for groups, NEG and POS.
Phase 1:
 Conduct a saturated or nearly-saturated fractional factorial experiment and estimate $\hat{\beta}_1, \ldots, \hat{\beta}_k$. Order the estimates so that $\hat{\beta}_{[1]} \leq \ldots \leq \hat{\beta}_{[z]} < 0 < \hat{\beta}_{[z+1]} \leq \ldots \leq \hat{\beta}_{[K]}$. Add factors {[1], . . . , [z]} to the NEG queue, and factors {[z + 1], . . . , [K]} to the POS queue.
Phase 2:
For queue = POS and queue = NEG, do
 While queue is not empty, do
  Remove: Remove a group from the queue.
  Test:
   Unimportant:
    If the group effect is unimportant ($<\Delta_0$), then classify all factors in the group as *unimportant*.
   Important (size=1):
    If the group effect is important ($>\Delta_0$) and of size 1, then classify the factor as *important*.
   Important (size>1):
    If the group effect is important ($>\Delta_0$) and the size is greater than 1, then split the group into two subgroups such that all factors in the first subgroup have smaller [i]'s (ordered indices) than those in the second subgroup. Add each subgroup to the LIFO queue.
  End Test
 End While
End For

E. THESIS OBJECTIVE

FFCSB has been newly proposed and requires further experimentation in order to understand its strengths and weaknesses. Thus, a series of controlled experiments is set up to evaluate the performance of FFCSB under different response model configurations. In addition, the performance of FFCSB is compared with that of CSB and FF.

In a first-case application, the FFCSB algorithm is applied to a simulation model, the Hierarchy organizational model provided by the Center for Edge Power (CEP). The Hierarchy model and its computational environment POW-ER—Projects, Organizations, Work and Edge Research—are work products stemming from collaborative research and development between the Naval Postgraduate School and Stanford University. FFCSB extends the suite of tools that researchers can use for computation experimentation and research on organization studies.

F. THESIS ORGANIZATION

The document is organized into six chapters. Chapter I provides the motivation and purpose of the thesis research. Chapter II provides the experimental setup used to compare the algorithms. Chapters III and IV describe the performance evaluation of FFCSB by itself and in comparison with CSB and FF. Chapter V describes the Hierarchy organizational model and the results of applying FFCSB to the model. Chapter VI concludes the research with insight and recommendations for future work.

THIS PAGE INTENTIONALLY LEFT BLANK

II. EXPERIMENTAL SETUP & CONTROL

A. CHAPTER OVERVIEW

When applying algorithms to real world problems, it is useful to know the relative strengths and weaknesses of the candidate algorithms, as well as their applicability to the problem of interest. Controlled experiments provide a good way to evaluate the models and conditions under which an algorithm succeeds or fails. This section documents the experimental setup to evaluate the FFCSB algorithm against the CSB and FF algorithms. The structure of the controlled experiments is adapted from Sanchez, Wan and Lucas (2005), where the FFCSB algorithm was proposed.

B. EVALUATION CRITERIA

The three algorithms of interest are evaluated via two Measures of Effectiveness (MOEs): accuracy and efficiency. The primary MOE of accuracy is the performance guarantee offered by each algorithm, and it takes priority over the secondary MOE of efficiency. The primary MOE of accuracy measures the probability of correct classification of factors by each algorithm and is quantified by two Measures of Performance (MOPs) as illustrated in Figure 2: (1) Type I Error and (2) power. The former is the probability of incorrectly classifying an unimportant factor (defined by effect magnitude below threshold for importance Δ_0), and the latter is the probability of correctly classifying a critical factor (defined by effect magnitude above threshold for criticality Δ_1). The secondary MOE of efficiency measures the computation savings in the average number of simulations runs required by each algorithm in order to provide its accuracy guarantee. Several MOPs, e.g., percentage savings and computation equivalence, are used to quantify efficiency wherever appropriate.

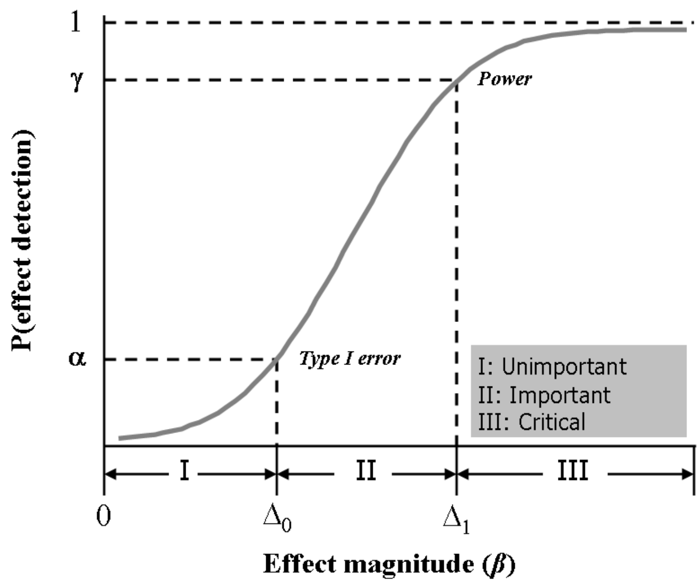

Figure 2. Generic Illustration of Desired Performance of Screening Procedures. From Wan et al., 2003.

C. EXPERIMENTAL SETUP & CONTROL

1. Response Meta-Model

The FFCSB algorithm assumes a main effects response model, as described in Equation (1). This basic response model is set up with factors such that the magnitudes of their effects are equally spaced between -5 and +5.

a. Variation: Factor Count

The controlled experiments study algorithm performance with increasing factor count. Hence, response models are generated with different number of factors $K = 2^N$, where $N = 3, 4, \ldots, 10$. FFCSB can be applied to bigger response models, even beyond 2^{10} factors.

b. Variation: Factor Patterns

The controlled experiments study algorithm performance on response models with different proportions of positive and negative factor effects. Hence, response models are generated with five different factor patterns, as listed in the following table. The direction of the factor effect indicates the subsequent positive or negative change in the response variable as a result of increasing the factor value.

Table 4. Factor Patterns Used in Controlled Experiments

Factor Pattern	Percentage of factors with positive effects	Percentage of factors with negative effects
None Negative	100%	0%
Small Negative	12.5%	87.5%
Medium Negative	25%	75%
Large Negative	37.5%	62.5%
Half Negative	50%	50%

Figure 3 depicts the proportions of positive and negative effects by factor pattern. The first factor pattern of "None Negative" has only positive factor effects. The percentage of negative factor effects increases up to 50% in the last factor pattern of "Half Negative."

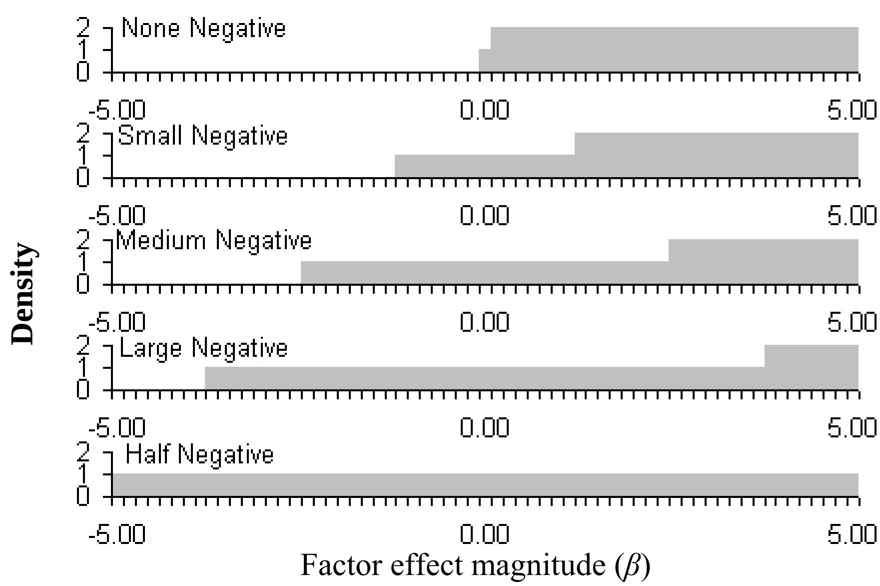

Figure 3. Distribution of Factor Effect Direction for Various Factor Patterns. From Sanchez, Wan and Lucas, 2005.

c. Variation: Homogeneous versus Heterogeneous Variance

The controlled experiments study algorithm performance on response models with two types of response variance: (1) homogeneous and (2) heterogeneous. Homogenous response variance implies that the errors in the response model are constant and independent of the response magnitude. On the other hand, heterogeneous response variance implies that the errors are dependent on the response magnitude. Heterogeneous responses are often encountered in real world systems. The controlled experiments scale the heterogeneous variance with some percentage of the simulated response. The following equations denote both types of variances in relation to the response.

$$Y = \beta_O + \sum_{i=1}^{K} \beta_i x_i + \varepsilon_i \quad \text{where} \quad \varepsilon_i \sim N(0, \sigma^2)$$

Equation 2. Response Model with Homogeneous Variance

$$Y = \beta_O + \sum_{i=1}^{K} \beta_i x_i + \varepsilon_i \quad \text{where} \quad \varepsilon_i \sim N(0, (mY)^2)$$

Equation 3. Response Model with Heterogeneous Variance:

d. Variation: Magnitudes of Response Variance

The controlled experiments study algorithm performance on response models with different magnitudes of response variance. Response models with homogeneous variances have the errors terms modeled as $\varepsilon_i \sim N(0,\sigma^2)$ where ${\sigma_\varepsilon}^2$ =1, 2, 4 or 8. Response models with heterogeneous variances have errors modeled as $\varepsilon_i \sim N(0,(mY)^2)$ where σ_ε = 0.05Y, 0.10Y, 0.15Y or 0.20Y.

2. CSB Algorithm Parameter Settings

CSB uses factor thresholds (Δ_0, Δ_1) for hypothesis testing on factor significance in order to provide the accuracy guarantees of power (γ) and Type I Error (α). The discussed parameters are configured as follows. The parameters of Δ_0 and Δ_1 are set relative to the response model specified for the controlled experiments, so that when all factors are ranked in ascending order of factor effect magnitude, the bottom 40% are unimportant, and the top 20% are critical.

Table 5. CSB Parameter Settings Used in Controlled Experiments

FFCSB parameter	Value
Power γ	0.95
Type I Error α	0.05
Threshold for unimportant factors Δ_0	2
Threshold for critical factors Δ_1	4

3. Experiment & Measurements Methodology

Each variation of the control experiments is repeated up to 1000 times using different random seeds. Experiments for larger factor counts (K=512, 1024) are repeated up to 500 times due to the longer computation time requirements. MOPs are then averaged over the total number of experiments conducted.

D. MODIFICATION TO FF TO PROVIDE ACCURACY GUARANTEE

The controlled experiments aim to compare the accuracy and efficiency performance of FFCSB against that of CSB and FF. However, FF designs by themselves do not provide accuracy guarantees. Hence, it is unfair to compare FFCSB against the plain FF designs. In the controlled experiments, the FF algorithm consists of the Resolution III FF design with replications and a statistical decision criterion to classify factors as critical or unimportant. This form of accuracy guarantee makes it viable to compare the three algorithms.

1. FF Statistical Decision Criteria for Factor Classification

The proposed FF statistical decision criteria unify results from two hypothesis tests to classify the statistical significance of each factor as unimportant or critical. It uses the same accuracy control parameters as the CSB algorithm, i.e., Δ_0, Δ_1, α and γ. There are three additional parameters required by the criteria: (1) estimate of factor effect ($\hat{\beta}$), (2) standard deviation of factor effect estimate ($\hat{\sigma}_{Factor}$) and (3) degrees of freedom (ν). $\hat{\beta}$ is the estimate of the factor effect from the FF pre-sorting stage and $\hat{\sigma}_{Factor}$ is the corresponding standard deviation of the estimate. ν is the number of degrees of freedom available for the hypothesis test.

The following two tables describe the Stage 1 and 2 hypothesis tests of the criterion. Stage 1 tests that the factor is critical; while Stage 2 tests that the factor is unimportant. Both tests use estimates of the factor effect and standard deviation to compute the test statistic.

Table 6. Stage 1: Test that Factor is Critical & Guarantees Power > γ

Null Hypothesis:	H_O: $\lvert\beta\rvert \geq \Delta_1$ (definition of critical factor)
Alternative Hypothesis:	H_A: $\lvert\beta\rvert < \Delta_1$
Test Statistic:	$T_O = \dfrac{\lvert\hat{\beta}\rvert - \Delta_1}{\hat{\sigma}_{Factor}}$
Rejection Criteria:	$T_O <$ T(1-γ, ν) from Student's T Distribution

If H_O is rejected

Factor is temporarily labeled as <u>not critical</u>.

Else

Factor is temporarily labeled as <u>critical</u>

Table 7. Stage 2: Test that Factor is Unimportant & Guarantees Type I Error < α

Null Hypothesis:	H_O: $\lvert\beta\rvert \le \Delta_0$ (definition of unimportant factor)
Alternative Hypothesis:	H_A: $\lvert\beta\rvert > \Delta_0$
Test Statistic:	$T_O = \dfrac{\lvert\hat{\beta}\rvert - \Delta_0}{\hat{\sigma}_{Factor}}$
Rejection Criteria:	$T_O > T(1-\alpha, \nu)$ from Student's T Distribution

If H_O is rejected
    Factor is temporarily labeled as <u>not unimportant</u>.
Else
    Factor is temporarily labeled as <u>unimportant</u>.

When there is strong evidence, the decision criteria is able to declare the factor unambiguously as either critical or important in Stage 1 and 2. When there is insufficient or conflicting data, the criteria does not classify a factor. Results of both hypothesis tests are unified according to the following table.

Table 8. Stage 3: Unification of Hypothesis Tests Results

		Stage 1 Results	
		Critical	Not Critical
Stage 2 Results	Unimportant	Unclassified	Unimportant
	Not Unimportant	Critical	Important

2. Illustration of FF Statistical Decision Criteria Operation

Three numerical examples are provided to illustrate the operation of the FF statistical decision criteria. Example 1 illustrates the "desired" operating state for the criteria where there is sufficient simulation data for decisive testing. Examples 2 and 3 illustrate the problem of non-classification when simulation data are insufficient, leading to ambiguous conclusions. The accuracy control parameters are: $\Delta_0=2$, $\Delta_1=4$, $\alpha=0.05$ and $\gamma=0.95$.

Example 1: $\hat{\sigma}_{Factor}$ **= 0.5,** ν **= 72**

Stage 1: $|\hat{\beta}| < (-1.6602^{1} \times \hat{\sigma}_{Factor} + \Delta_1) = 3.17$ is rejected and is Not Critical.

Stage 2: $|\hat{\beta}| > (1.6602^{2} \times \hat{\sigma}_{Factor} + \Delta_0) = 2.83$ is rejected and is Not Unimportant.

Unification: Factors with $|\hat{\beta}|<2.83$ are unambiguously classified as unimportant.

Factors with $|\hat{\beta}|>3.17$ are unambiguously classified as critical.

Factors with $2.83<|\hat{\beta}|<3.17$ are unambiguously classified as important.

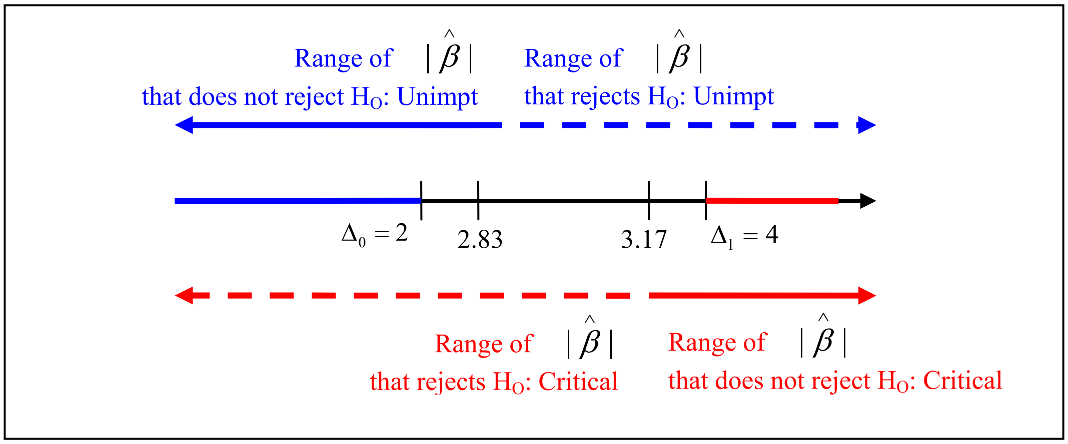

Figure 4. FF Statistical Decision Criteria: Example 1 Classification

---

[1] $T(1-\gamma,\nu) = T(1-0.95,72) = -1.6602$ where T is the Student's T Distribution.

[2] $T(1-\alpha,\nu) = T(1-0.05,72) = +1.6602$ where T is the Student's T Distribution.

Example 2: $\hat{\sigma}_{Factor}$ **= 1,** ν **= 100**

Stage 1: $|\hat{\beta}| < (-1.6602 \times \hat{\sigma}_{Factor} + \Delta_1) = 2.34$ is rejected and is Not Critical.

Stage 2: $|\hat{\beta}| > (1.6602 \times \hat{\sigma}_{Factor} + \Delta_0) = 3.66$ is rejected and is Not Unimportant.

Unification: Factors with $|\hat{\beta}|<2.34$ are unambiguously classified as unimportant

Factors with $|\hat{\beta}|>3.66$ are unambiguously classified as critical.

Factors with $2.34<|\hat{\beta}|<3.66$ cannot be classified.

With factor effects ranging from -5 to 5, this is a large range for un-classification.

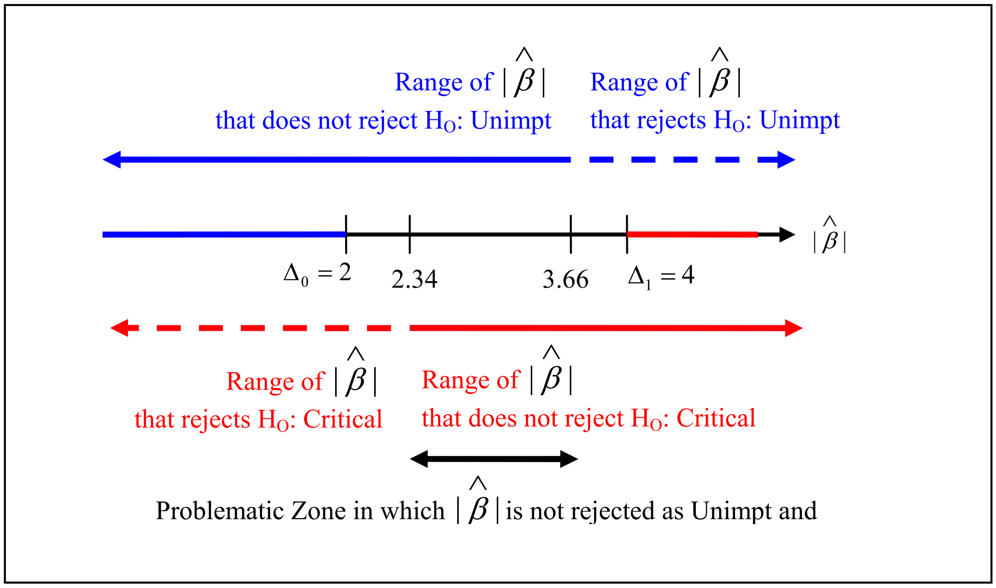

Figure 5. FF Statistical Decision Criteria: Example 2 Non-classification

Example 3: $\hat{\sigma}_{Factor}$ **= 1.5,** ν **= 16**

Stage 1: $|\hat{\beta}| < (-1.6602 \times \hat{\sigma}_{Factor} + \Delta_1) = 1.21$ is rejected and is Not Critical.

Stage 2: $|\hat{\beta}| > (1.6602 \times \hat{\sigma}_{Factor} + \Delta_0) = 4.79$ is rejected and is Not Unimportant.

Unification: Factors with $|\hat{\beta}|<1.21$ are unambiguously classified as unimportant

Factors with $|\hat{\beta}|>4.79$ are unambiguously classified as critical.

Factors with $1.21<|\hat{\beta}|<4.79$ cannot be classified.

With factor effects ranging from -5 to 5, this is an unacceptable range for un-classification.

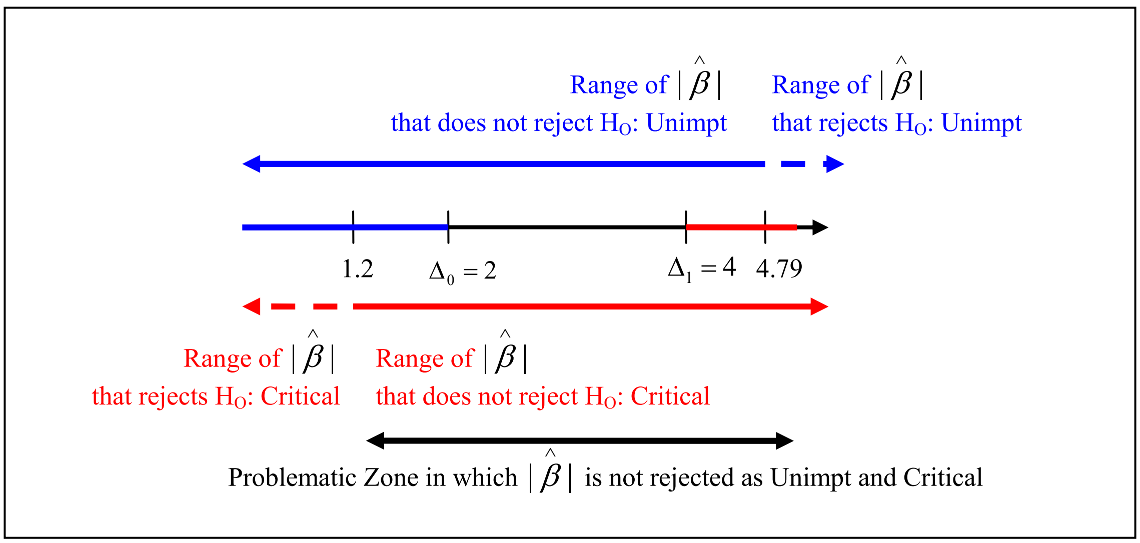

Figure 6. FF Statistical Decision Criteria: Example 3 Non-classification

III. PERFORMANCE EVALUATION OF FFCSB UNDER HOMOGENEOUS RESPONSE VARIANCES

A. CHAPTER OVERVIEW

In this section, the three algorithms, FFCSB, CSB and FF, are applied to response models with homogeneous response variances. The response models are varied with different numbers of factors, different factor patterns and different magnitudes of homogeneous variance. First, the accuracy performance of FFCSB is presented. Next, comparisons are drawn between FFCSB and the other two algorithms using the accuracy and performance MOEs. Lastly, the algorithms are evaluated for their relative strengths and weaknesses. The graphs in this chapter are best viewed in color.

B. PERFORMANCE OF FFCSB UNDER HOMOGENEOUS RESPONSE VARIANCE

1. Accuracy: FFCSB Fulfills Performance Guarantees Comfortably and is Consistent across Factor Patterns, Variance Magnitudes and Model Size

The "Power & Type I Error" figure (7) illustrates the accuracy performance of FFCSB under the "None Neg" factor pattern and for various magnitudes of homogeneous variance. Such graphs will be used consistently to represent and compare the accuracy performance of the algorithms which are of interest in this thesis. The left graph plots the power performance (Recall: probability of correctly classifying a critical factor) versus number of factors in the response model. The right graph plots the Type I Error performance (Recall: probability of incorrectly classifying an unimportant factor) versus number of factors in the response model. The different colored lines within each figure plot the respective accuracy performance for increasing magnitudes of homogenous variances, ranging from σ_{ε}^{2}=1, 2, 4 or 8. The variance magnitude is indicated in the legend.

Figure 7 shows that FFCSB easily fulfills the power and Type I Error accuracy guarantees. The left graph shows that FFCSB demonstrates power performance better (i.e., more) than γ=0.95, and the right graph shows Type I Error performance better (i.e., less) than α=0.05. The accuracy guarantees are consistent over factor count. They are robust against increasing magnitudes of variance, as the clustered accuracy lines represent similar performance for all variance magnitudes simulated. In this figure and others in this thesis, FFCSB performs better than expected at small factor counts of 8 and hence forming a misleading "dip" in performance at factor counts of 16 and 32. The spread of factor effects at small factor counts could have led to this anomaly of better than expected performance. Such fluctuations even out towards larger factor counts.

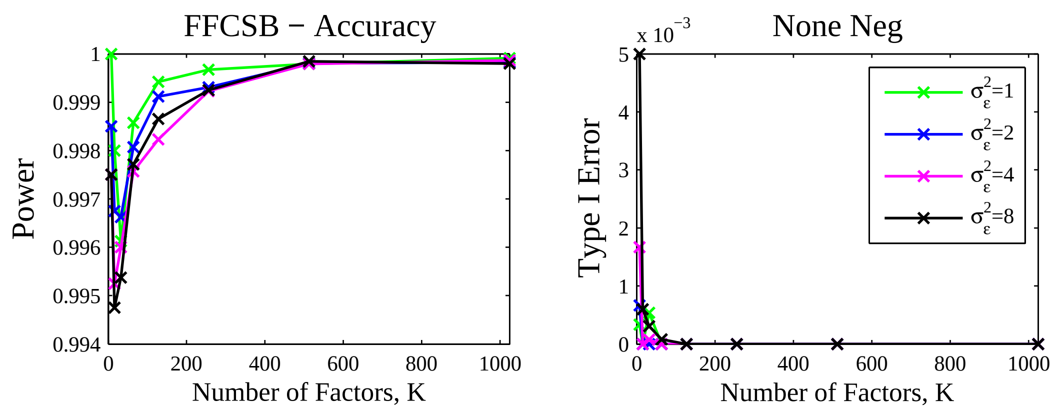

Figure 7. FFCSB Accuracy for "None Neg" & Various Homogeneous Variances

The following "Power & Type I Error" figure (8) illustrates the accuracy performance of FFCSB under constant homogeneous variance (σ_ε^2=1) and for various factor patterns. Each colored line represents FFCSB accuracy performance under a specific factor pattern, which is indicated in the legend. The figure illustrates that FFCSB power and Type I Error performances are within the guarantees and consistent across the factor patterns.

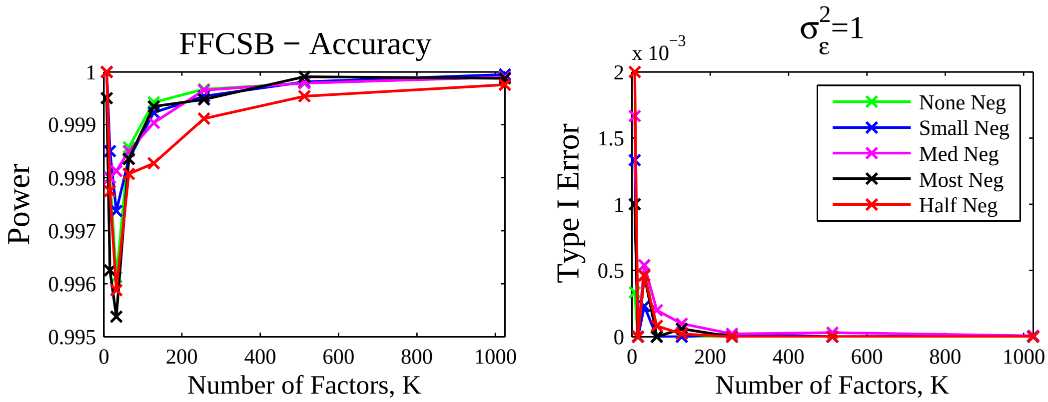

Figure 8. FFCSB Accuracy for Unit Homogeneous Error ($\sigma_{\varepsilon}^{2}=1$) & Various Factor Patterns

The following figures (9-13) present the results of FFCSB accuracy performance categorized by factor patterns. In each factor pattern, FFCSB accuracy performance is robust against magnitudes of homogeneous variances. In addition, the accuracy performances are similar across factor patterns.

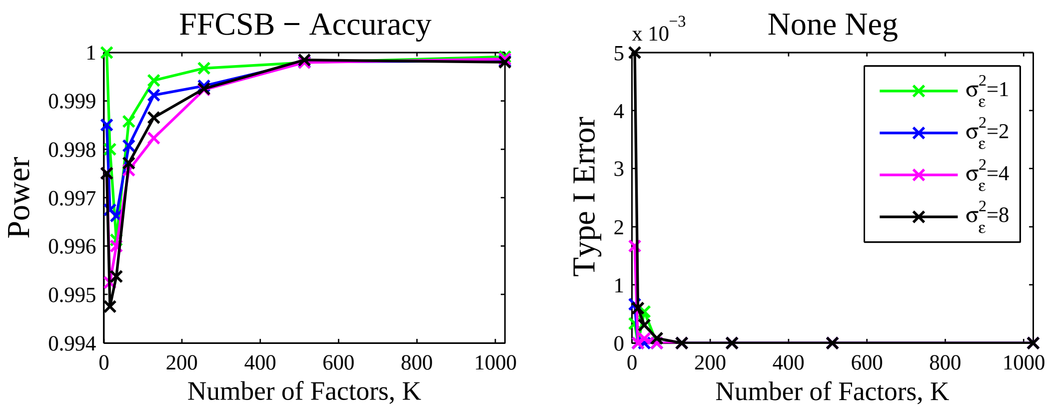

Figure 9. FFCSB Accuracy for "None Neg" & Various Homogeneous Variances

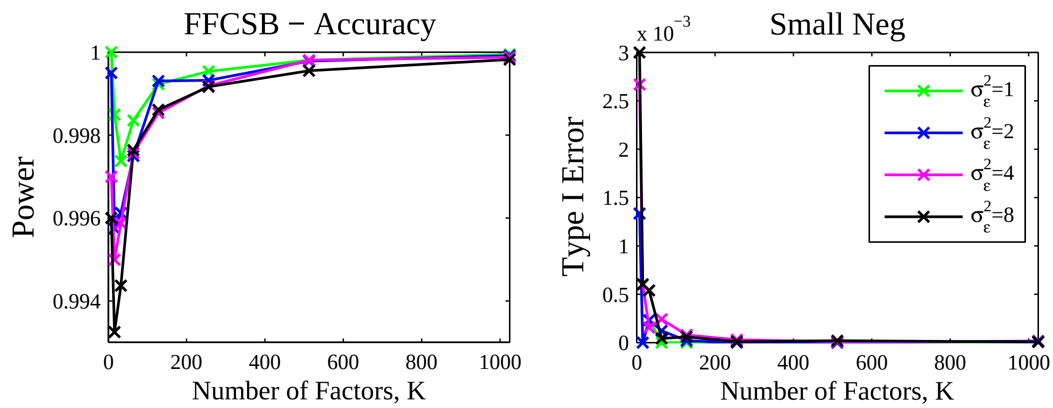

Figure 10. FFCSB Accuracy for “Small Neg” & Various Homogeneous Variances

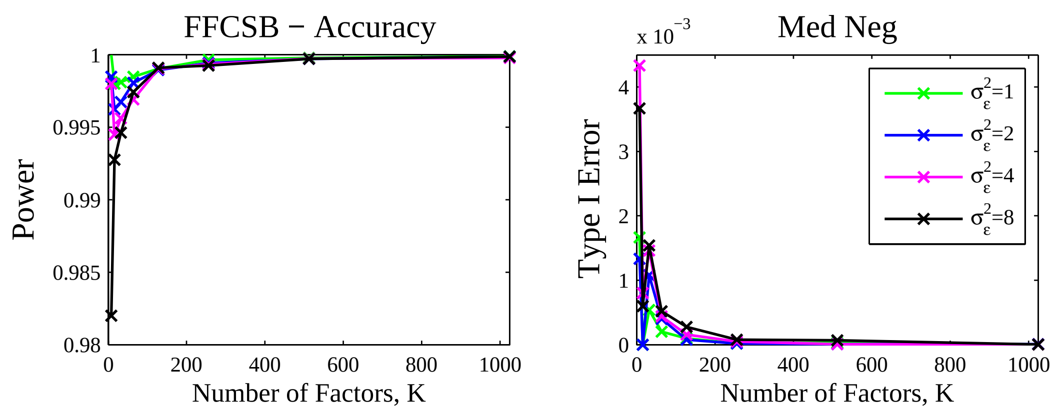

Figure 11. FFCSB Accuracy for “Med Neg” & Various Homogeneous Variances

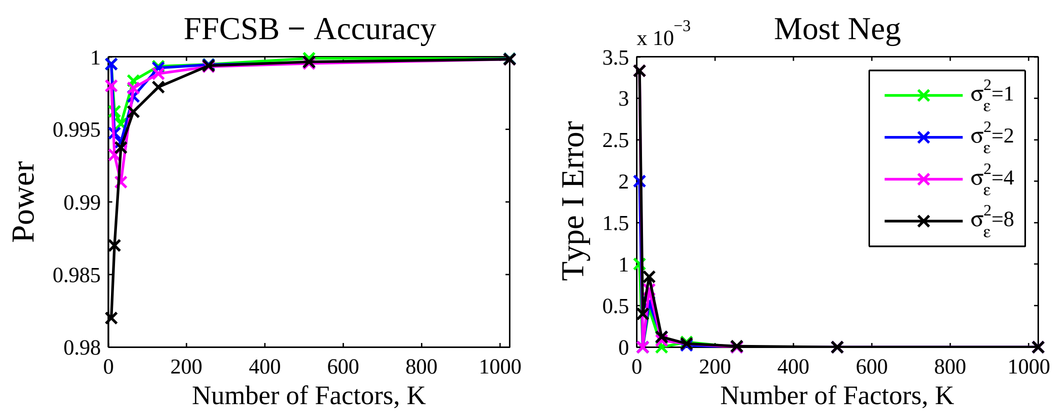

Figure 12. FFCSB Accuracy for “Most Neg” & Various Homogeneous Variances

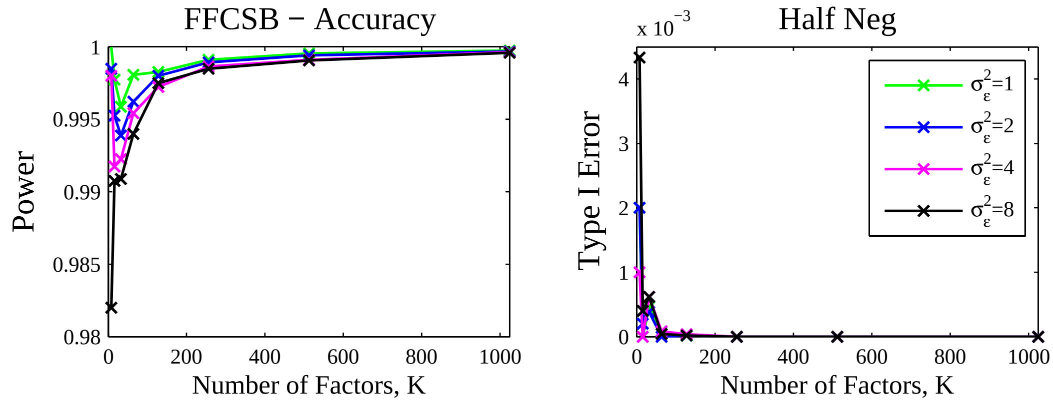

Figure 13. FFCSB Accuracy for “Half Neg” & Various Homogeneous Variances

C. COMPARISON OF FFCSB & CSB UNDER HOMOGENEOUS RESPONSE VARIANCE

1. Accuracy: FFCSB is Robust for All Factor Patterns while, as Expected, CSB Fails for Increasing Factor Negativity

Testing under the different factor patterns reveals the strength of FFCSB over CSB. FFCSB is able to realize the performance guarantees of power and Type I Error under all factor patterns. On the other hand, CSB always guarantees Type I Error, but fails to guarantee power for increasing percentages of factor negativity because factors with different directions may cancel each other out in the group screening process. These observations are drawn from Figure 14. The top graphs compare the power performance of FFCSB versus CSB. FFCSB provides consistent performance, while CSB fails by the third factor pattern of “Med Neg,” providing only 80% power. The lower graphs show that both algorithms provide Type I Error guarantee in all factor patterns.

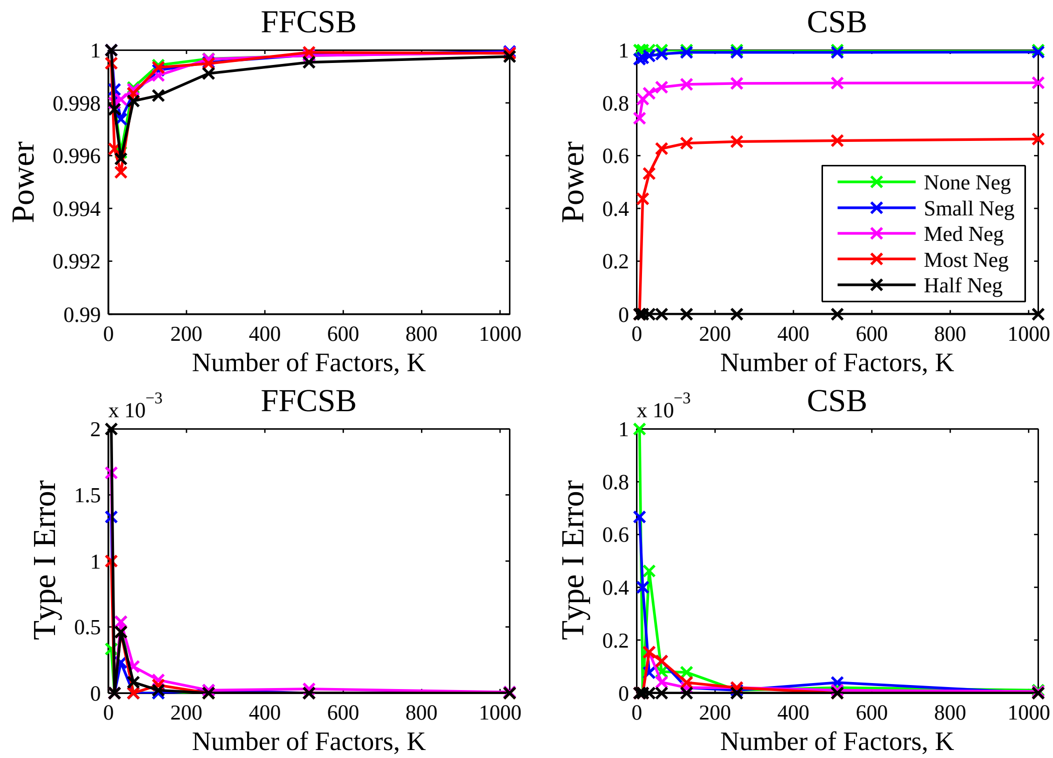

Figure 14. Accuracy Comparison of FFCSB versus CSB for Various Factor Patterns

Thus, without *a priori* knowledge of the factor pattern in a real-world problem, it is “safer” to apply FFCSB than CSB. Henceforth, further comparison of FFCSB versus CSB is meaningful only for factor patterns of “None Neg” and “Small Neg.” Under the factor patterns of “None Neg” and “Small Neg,” both FFCSB and CSB are robust to the level of homogeneous variance in the response model. This is observed from the equally good performance for all simulated variance magnitudes and the proximity of the plotted lines in the Figures 15 and 16.

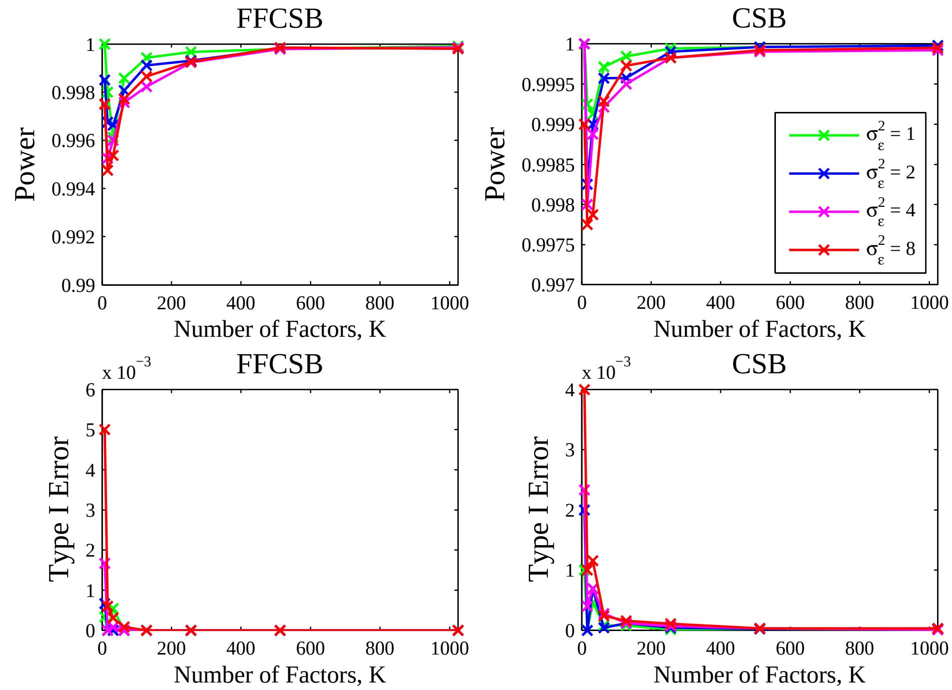

Figure 15. Accuracy Comparison of FFCSB versus CSB for "None Neg" & Various Homogeneous Variances

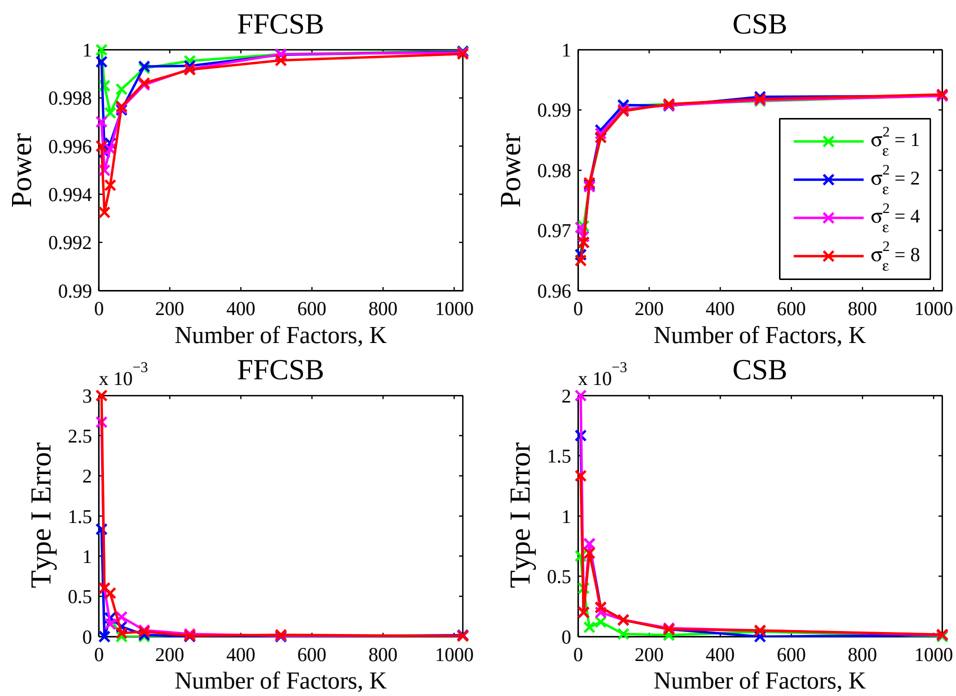

Figure 16. Accuracy Comparison of FFCSB versus CSB for "Small Neg" & Various Homogeneous Variances

2. Efficiency: FFCSB Saves Up to 25% Computation Effort on Large Response Models as Compared to CSB.

The computation resources for FFCSB and CSB are compared. Figure 17 plots the average simulation runs required by each algorithm versus the number of factors in the response model. The left and right graphs are for the factor patterns of "None Neg" and "Small Neg" respectively. The upward sloping lines suggest that both FFCSB and CSB require proportionally more computation resources with the number of factors in the response model. Within each graph, the same colored pair of lines represents the FFCSB and CSB runs for the same variance magnitude. The different colored lines represent both algorithms for different variance magnitudes, ranging from $\sigma_\varepsilon^{\ 2} = 1$ to 8. The increasing height of each pair of FFCSB and CSB efficiency lines shows that both FFCSB and CSB require more computation resources with increased response variance.

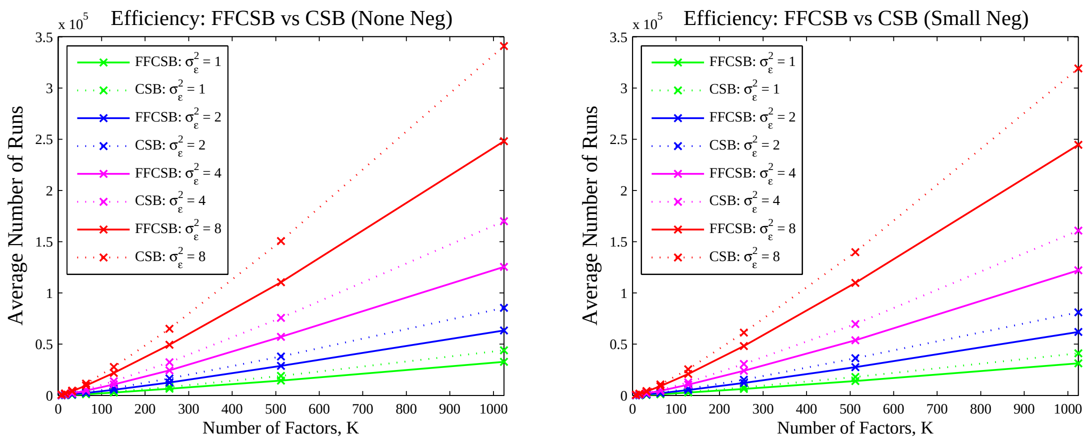

Figure 17. Efficiency of FFCSB versus CSB (Average Runs)

It is, however, more insightful to express the relative efficiency by expressing the ratio of the number of runs saved by employing FFCSB to the number of runs required by CSB. These are plotted in Figure 18 as percentage savings in average runs. FFCSB presents potential average savings of up to 25% for larger response models with 200 or more factors, relative to the computation effort required by CSB. Thus, FFCSB yields an equally good answer as CSB in three quarters of the time.

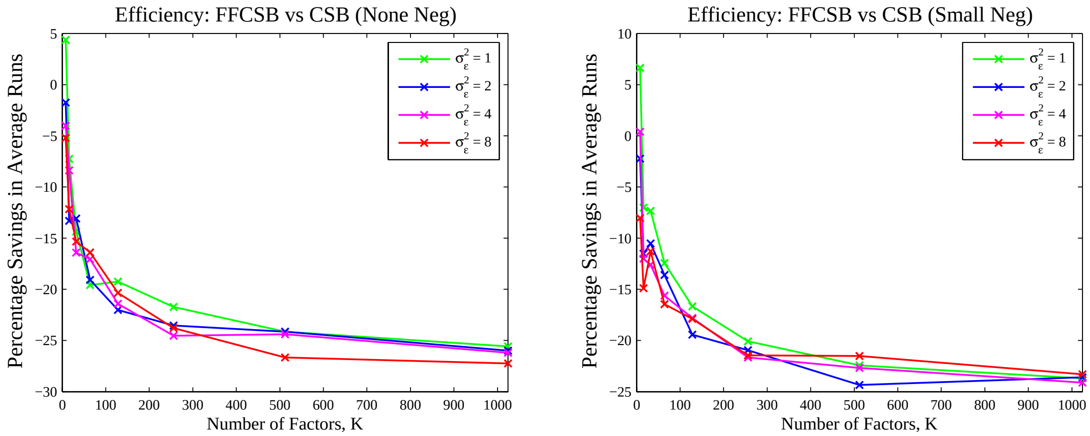

Figure 18. Efficiency of FFCSB versus CSB (Percentage Savings)

D. COMPARISON OF FFCSB & FF UNDER HOMOGENEOUS RESPONSE VARIANCE

1. Accuracy: FFCSB Performs as Well as FF

FFCSB and FF (replications supplemented with statistical decision criteria) are equally matched in performance over the factor patterns and homogeneous variances simulated. The following two "Power and Type I Error" figures (19-20) present FFCSB versus FF performance in "None Neg" and small homogeneous variance (σ_ε^2=1), as well as in "None Neg" and large homogeneous variance (σ_ε^2=8). They compare FFCSB versus FF accuracy performance on the same graph (as opposed to using two graphs in FFCSB versus CSB). The red lines present FFCSB power and Type I Error performance. The remaining colored lines present FF performance with different replications (replications of 2, 25, 50, 75, 100, 150 or 200), as indicated in the legend. Analysis of both figures suggests that FFCSB and FF deliver equally good accuracy guarantees over the range of homogeneous variances simulated and over the factor count of the response model. In Figure 20 ("None Neg" and large homogeneous variance), the FF-2 (replications) fails accuracy guarantees for smaller models, where there are fewer critical factors and each misclassification is more severe. This suggests that more FF replications are required to provide accuracy guarantee with increased response variance.

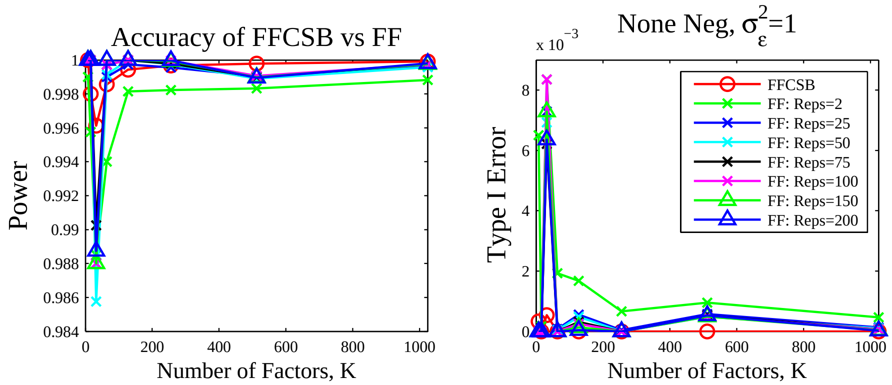

Figure 19. Accuracy Comparison of FFCSB versus FF for “None Neg” & Small Homogenous Variance (σ_ε^2 =1)

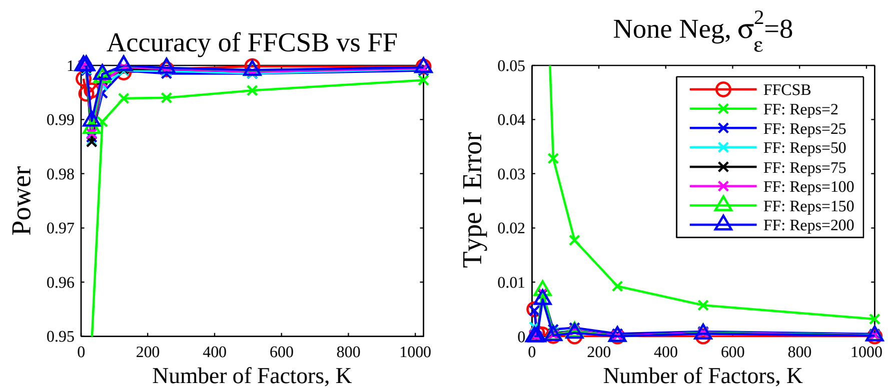

Figure 20. Accuracy Comparison of FFCSB versus FF for “None Neg” & Large Homogenous Variance (σ_ε^2 =8)

The matched performance of FFCSB and FF continue for the remaining factor patterns of “Small Neg” to “Half Neg.” Each pair of graphs in Figures 21 through 24 present the FFCSB versus FF accuracy comparison for the factor pattern specified in the title and under the largest homogeneous variance (σ_ε^2 =8) simulated. The green line of FF-2 replications is plotted to remind the reader of the lower bound of FF performance. A minimum of 2 replications is simulated in order to generate sufficient degrees of freedom for the hypothesis testing in the FF statistical decision criteria.

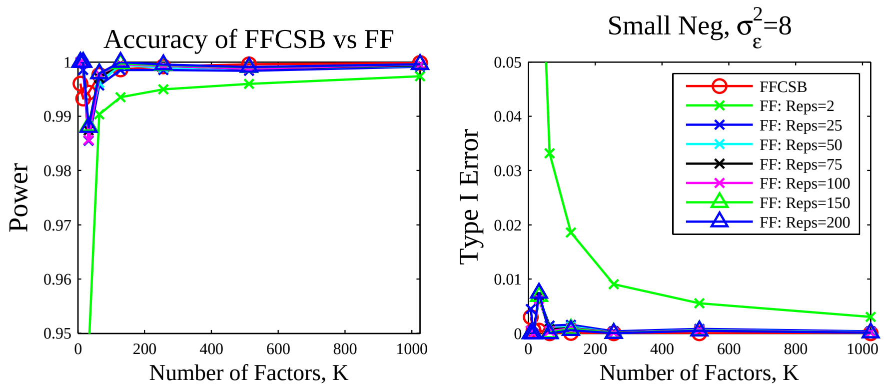

Figure 21. Accuracy Comparison of FFCSB versus FF for "Small Neg" & Large Homogenous Variance ($\sigma_\varepsilon{}^2$=8)

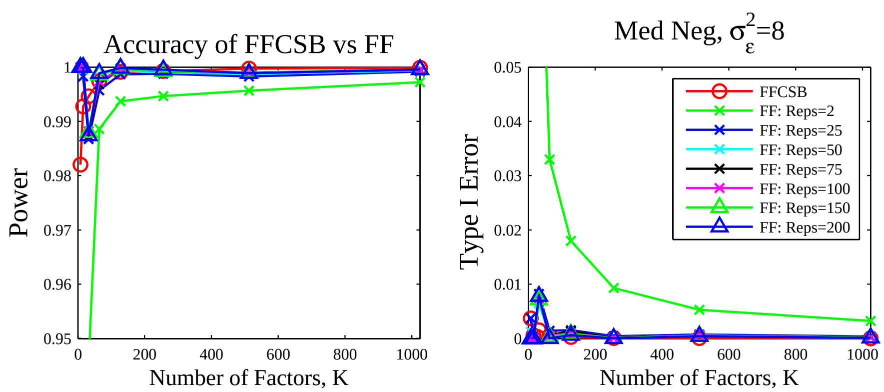

Figure 22. Accuracy Comparison of FFCSB versus FF for "Med Neg" & Large Homogenous Variance ($\sigma_\varepsilon{}^2$=8)

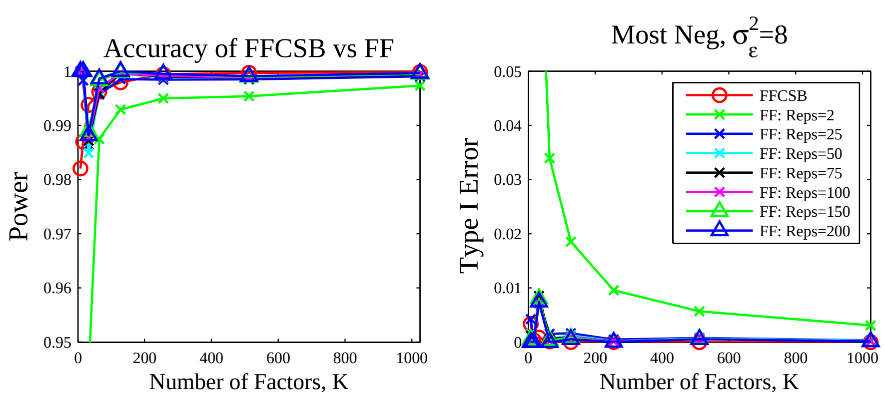

Figure 23. Accuracy Comparison of FFCSB versus FF for "Most Neg" & Large Homogenous Variance ($\sigma_\varepsilon{}^2$=8)

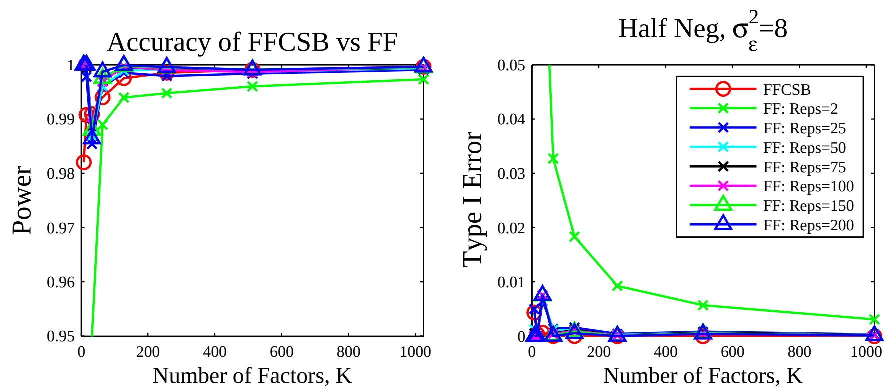

Figure 24. Accuracy Comparison of FFCSB versus FF for "None Neg" & Large Homogenous Variance ($\sigma_\varepsilon^{\ 2}$ =8)

2. Efficiency: FFCSB is Less Efficient than FF

FF has an efficiency advantage over FFCSB. Figure 25 presents the average number of simulations required by FFCSB and FF for factor classification. The red line represents the FFCSB average run count, while the colored lines represent the FF average run counts for different replications (indicated in legend).

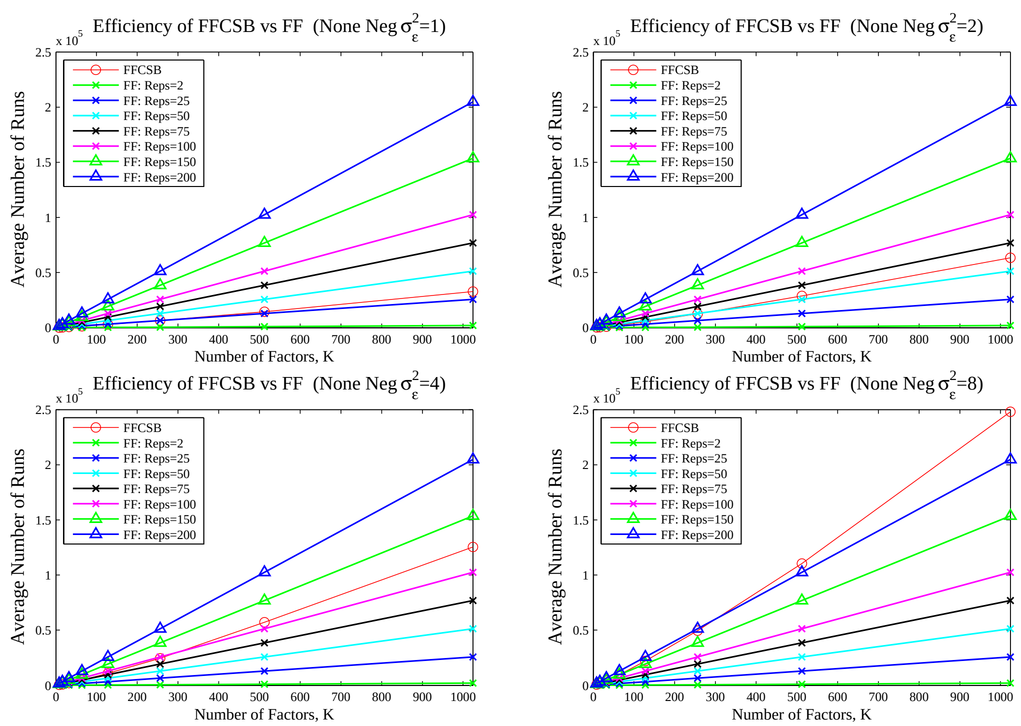

Figure 25. Computationally Equivalents: FFCSB & FF-25 for "None Neg" and $\sigma_\varepsilon^{\ 2}$ = 1. FFCSB & FF-200 for "None Neg" and $\sigma_\varepsilon^{\ 2}$ = 8.

In the uppermost left graph ("None Neg" with small variance σ_ε^2 = 1), FFCSB and FF-25 are computationally equivalent, i.e., they use the same average number of experiments to complete classification. Computational equivalence does not equate to accuracy equivalence. However, in this case, it has been presented earlier that both designs fulfill the accuracy guarantees. In the lowermost right graph ("None Neg" with large variance σ_ε^2 = 8), FFCSB and FF-150 are computationally equivalent and also fulfill accuracy guarantees. However, a smaller design, such as FF-25, could have realized the accuracy guarantees adequately. The top-right and bottom-left graphs show the increasing computation requirements by FFCSB with increasing variance. In general, FFCSB is not as efficient as FF because FF can deliver equivalent accuracy guarantees with fewer average runs. Additional analysis of the figure above also shows that FFCSB requires approximately double the computational resources for every doubling of homogeneous variances. For instance, FFCSB is (approximately) computationally equivalent to FF-25, FF-50, FF-100 and FF-200 for the homogeneous variances of $\sigma_\varepsilon^2=1$, $\sigma_\varepsilon^2=2$, $\sigma_\varepsilon^2=4$ and $\sigma_\varepsilon^2=8$ respectively. These efficiency observations hold true across factor patterns. In addition, it is observed that the factor pattern does not affect the computation requirements of FFCSB. The average runs for each level of response variance remains constant across the factor patterns simulated.

The five factor patterns simulated in the thesis are challenging for screening algorithms because they include many intermediate factors that are not critical and not unimportant. For other factor patterns, such as sparse factor effects, FFCSB is much more efficient and requires far fewer experiments (Sanchez, Wan and Lucas, 2005).

3. Tweaking Δ_1 for a Challenge: FFCSB Maintains Accuracy Guarantees while FF Fluctuates under Non-Ideal Control Settings

To distinguish further between FFCSB and FF, the response model is modified to make the screening more difficult. The default simulation settings were maximum factor magnitude of 5 and the threshold for critical factors was set at Δ_1 = 4 for CSB accuracy control. An extra set of experiments were conducted with the maximum factor magnitude changed from 5 to 4. The critical factor threshold remains at $\Delta_1 = 4$.

a. *Accuracy: Under Non-Ideal Control Settings, FFCSB is Robust against Model Size, Factor Patterns and Variance Magnitude*

Figures 26 and 27 summarize the accuracy performance of FFCSB under non-ideal control settings. From Figure 26, the clustering of the accuracy lines (different colors representing factor patterns indicated in legend) indicates that FFCSB offers consistent accuracy guarantees regardless of factor patterns. FFCSB fulfills Type I Error performance comfortably. However, the algorithm does not quite meet power performance guarantees for smaller response models with less than 10 factors. It delivers power of 92% and better, compared to the desired guarantee of 95%. The small number of factors causes FFCSB to fail to fulfill performance guarantees if there was a single misclassification. From Figure 27, the clustering of the accuracy lines (different colors representing variances indicated in legend) indicates consistent performance across different magnitudes of homogenous variances.

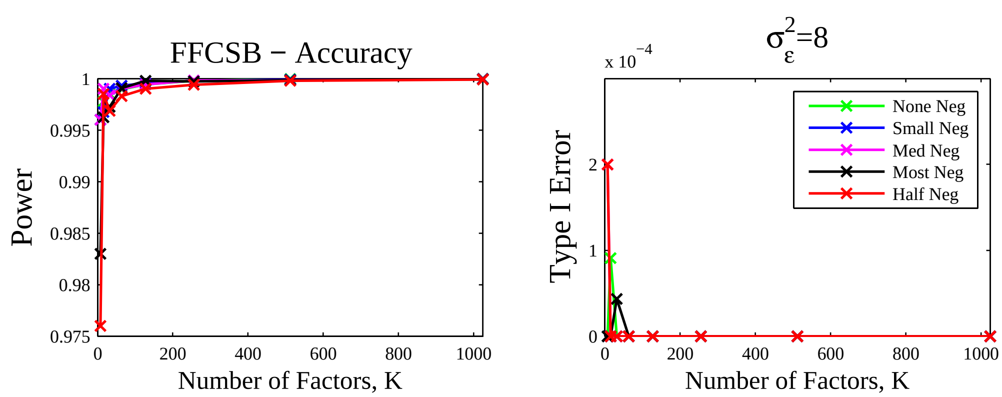

Figure 26. Consistent Accuracy of FFCSB for Various Factor Patterns & Large Homogeneous Variance ($\sigma_\varepsilon^2 = 8$)

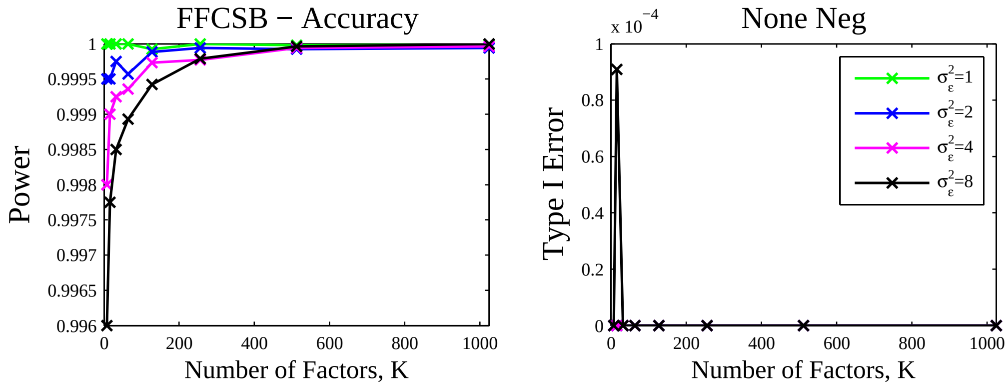

Figure 27. Consistent Accuracy of FFCSB for ”None Neg” & Various Homogeneous Variances

b. *Accuracy: FFCSB is Robust to Control Parameter Settings while FF is More Sensitive and Drops to Minimum Accuracy Guarantee*

This series of experiments reveal the sensitivity of FFCSB and FF to the selection of control parameter Δ_1. The results in Figure 28 reveal the robust accuracy performance of FFCSB when test conditions are taken to the limit. With a more challenging response model, FFCSB maintains stellar accuracy performance well beyond FF and proves to be robust to the settings of the control parameter Δ_1. Within sufficient replications (FF-25 is computationally equivalent), FF fluctuates around the power guarantee of 0.95 while meeting the Type I Error guarantee comfortably. In addition, FF power performance does not improve correspondingly with more replications, as observed from the Power graphs of Figure 28 where the FF-25 to FF-200 lines are all clustered and overlapping at power = 0.95.

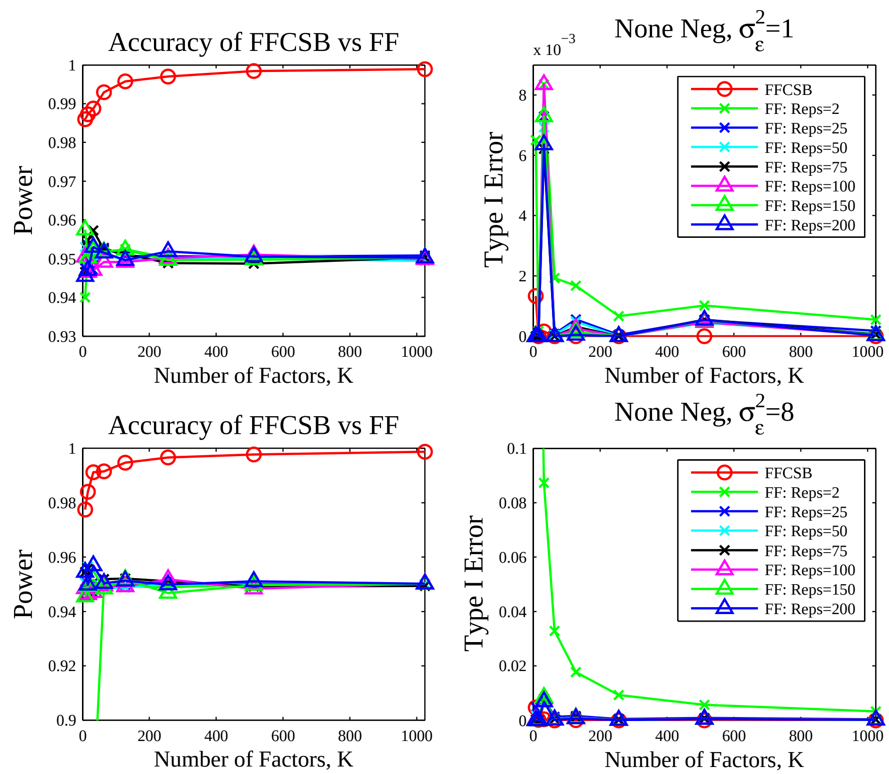

Figure 28. Accuracy of FFCSB versus FF for Δ_1 = Max Factor Magnitude = 4

c. *Efficiency: FFCSB Delivers Better Performance than its Computational Equivalent FF Design*

Figure 29 compares the average runs required by FFCSB and FF to complete factor classification. The left graph illustrates that FFCSB and FF-25 are computationally equivalent under small variance ($\sigma_\varepsilon^{\ 2}$=1). The right graphs illustrate that FFCSB and FF-200 are computationally equivalent under large variance ($\sigma_\varepsilon^{\ 2}$=8). However, it has been presented that FFCSB delivers better accuracy performance than any of the simulated FF replications design. Hence, under the challenging control settings, FFCSB has proven robust to the non-ideal settings and is more efficient than FF for the same accuracy performance. The efficiency comparisons are identical for the other factor patterns.

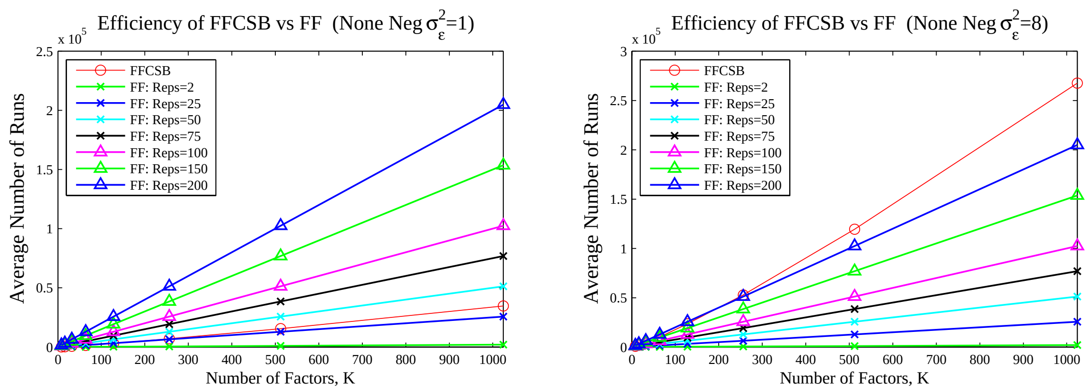

Figure 29. Efficiency of FFCSB versus FF for Δ_1 = Max Factor Magnitude = 4

E. AREAS IN WHICH ALGORITHMS EXCEL

FFCSB has proven its accuracy and efficiency in the series of experiments on response models with homogeneous variance and various factor patterns. FFCSB maintains robust accuracy performance across factor patterns, for various magnitudes of homogenous response variance and it scales well for large models, even up to 1024 factors.

Relatively, CSB falls short by failing with response models with significant degrees of mixed factor direction. CSB's vulnerability to mixed factor directions cancelling out one another is averted with the FF pre-sorting stage in FFCSB. For the factor patterns that CSB can handle, FFCSB provides up to a 25% computation savings.

On the other hand, FF proves to be matched to FFCSB in performance and more efficient for the general experiments. Both algorithms scale well with increased factor count and increased homogeneous response variance. However, FF is more sensitive to the control parameters settings and may fail accuracy guarantees if control parameters are not set "ideally" or are set at the limit. FFCSB proves to be robust to such stringent control parameters and fulfills accuracy guarantees comfortably.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. PERFORMANCE EVALUATION OF FFCSB UNDER HETEROGENEOUS RESPONSE VARIANCES

A. CHAPTER OVERVIEW

In this section, the three algorithms, FFCSB, CSB and FF, are applied to response models with heterogeneous response variances. The response models are varied with different numbers of factors, different factor patterns and different variance scaling. First, the accuracy performance of FFCSB is presented. Next, comparisons are drawn between FFCSB and the other two algorithms using the accuracy and performance MOEs. Lastly, the algorithms are evaluated for their relative strengths and weaknesses. The graphs in this chapter are best viewed in color.

B. PERFORMANCE OF FFCSB UNDER HETEROGENEOUS RESPONSE VARIANCE

1. Accuracy: FFCSB Fulfills Performance Guarantees in Three Out of Five Factor Patterns Simulated

Generally, heterogeneous variances are more realistic because they occur more frequently than homogeneous variances in real world systems. The heterogeneity poses a greater challenge to the algorithms. All three algorithms either fail to fulfill accuracy guarantees or take more computation power to maintain accuracy guarantees than in the case of homogenous variances.

The "Power and Type I Error" figure (30) illustrates the accuracy performance of FFCSB under the mildest case of heterogeneous variance (σ_ε =0.05Y). Within each graph, the different colored lines depict FFCSB accuracy performance for the various factor patterns, as indicated in the legend. In all five factor patterns, FFCSB fulfills power and Type I Error guarantees of 95% and 5% respectively across the different amounts of factor effect negativity in the different factor patterns. This is supported by the accuracy plots lying within the respective accuracy limits in the figure.

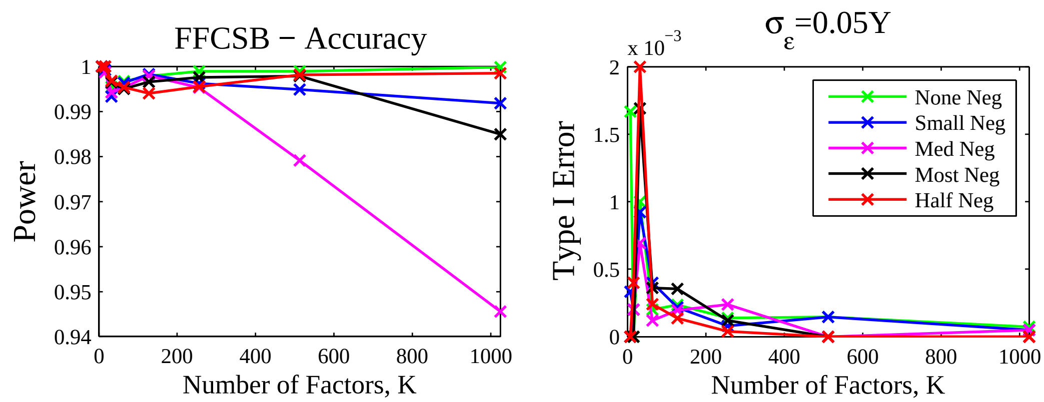

Figure 30. FFCSB Maintains Accuracy Guarantees in Various Factor Patterns for Mild Heterogeneity (σ_{ε}=0.05Y)

There is a slight violation of the power performance of FFCSB for the "Med Neg" factor pattern. In fact, FFCSB does not scale well with factor count for this factor pattern, as can be seen by deteriorating performance at 512 and 1024 factors. This factor pattern poses the greatest challenge amongst all because it has a significant percentage of negative factors that are not negligible in effect and yet not critical enough to be classified. Thus, they cannot easily be eliminated as unimportant nor classified as critical. They remain within the experiments and cause errors in the factor classification. Nevertheless, for mild heterogeneity (σ_{ε}=0.05Y), the average FFCSB power performance of 94.56% over 1000 randomization runs is not statistically different from the guarantee of 95%.

Next, Figures 31 through 33 present FFCSB accuracy performance for increasing heterogeneity. Each pair of "Power and Type I Error" graphs is plotted for increasing heterogeneity, as indicated in the title. As the magnitude of heterogeneity increases, FFCSB deteriorates in power performance, while meeting Type I Error guarantees comfortably. The power performance for factor patterns of "Med Neg" and "Most Neg" deteriorates the fastest. The missing data points represent experiments with excessive computation time.

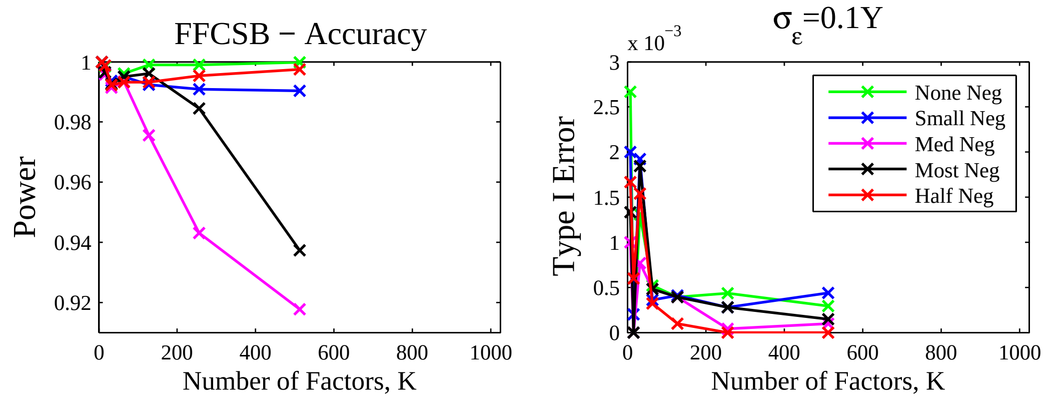

Figure 31. FFCSB Accuracy for σ_ε =0.10Y & Various Factor Patterns

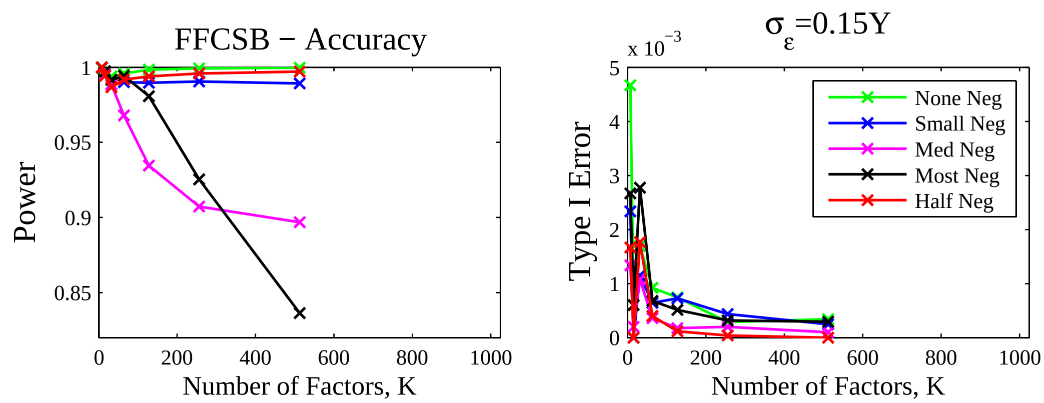

Figure 32. FFCSB Accuracy for σ_ε =0.15Y & Various Factor Patterns

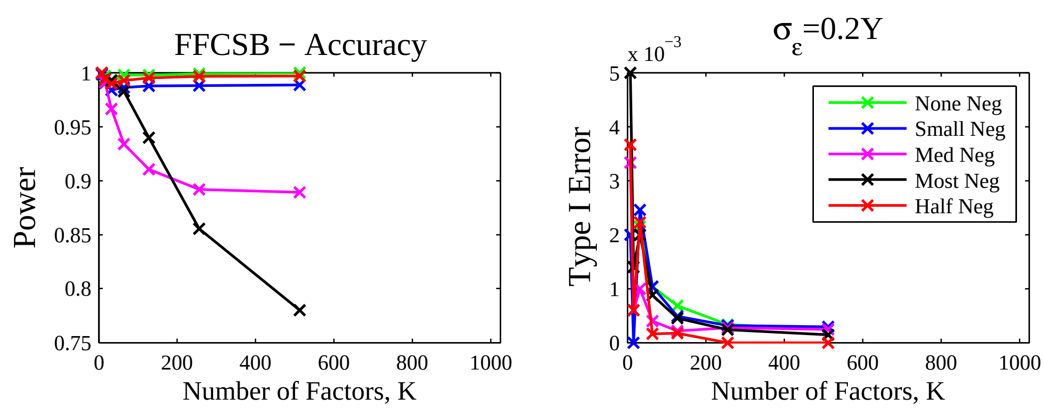

Figure 33. FFCSB Accuracy for σ_ε =0.20Y & Various Factor Patterns

FFCSB displays robustness to heterogeneity in certain factor patterns and deterioration in others. The following figures illustrate FFCSB accuracy performance under the extreme factor patterns of "None Neg" and "Half Neg," as well as under "Small Neg." Here, FFCSB is robust to heterogeneity because the FF pre-sorting stage is effective at sorting factors by effect direction based on the 1-replication estimate. Within the same CSB group for classification, factors of opposite effect direction do not exist or are negligible in effect. Thus, critical effects with the same factor effect direction in each group dominate the response and classification is accurate.

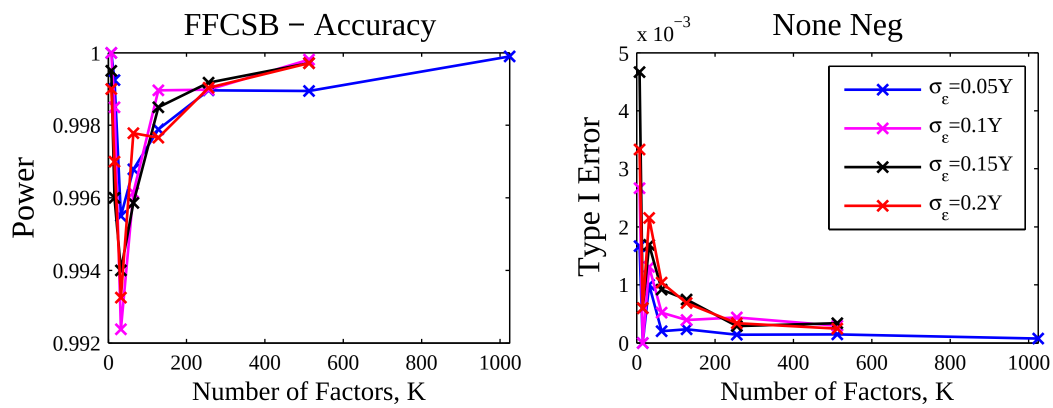

Figure 34. FFCSB Robust Against Heterogeneity for "None Neg"

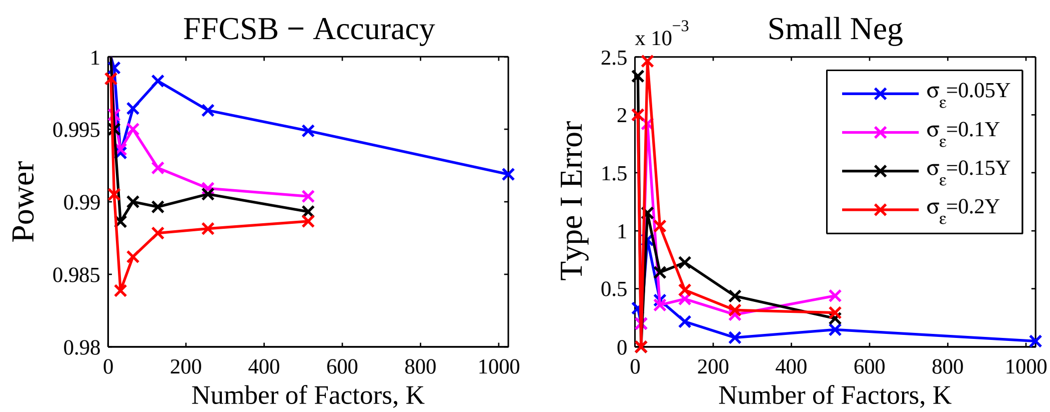

Figure 35. FFCSB Robust Against Heterogeneity for "Small Neg"

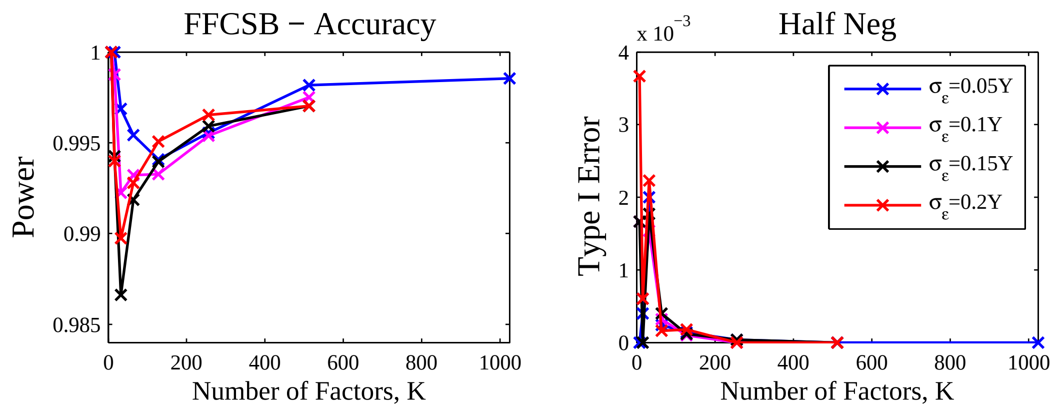

Figure 36. FFCSB Robust Against Heterogeneity for “Half Neg”

The next figures for FFCSB accuracy performance for the intermediate factor patterns of “Med Neg” and “Most Neg” paint a different picture. FFCSB power performance deteriorates with increasing heterogeneity, increasing model size and increasing percentage of negative factors. As previously explained, the increasing percentage of negative factors injects significant noise within the experiments and yet can neither be eliminated nor classified. Hence, it becomes more difficult to classify the critical factors as the level of heterogeneity increases. Interestingly, Type I Error does not suffer.

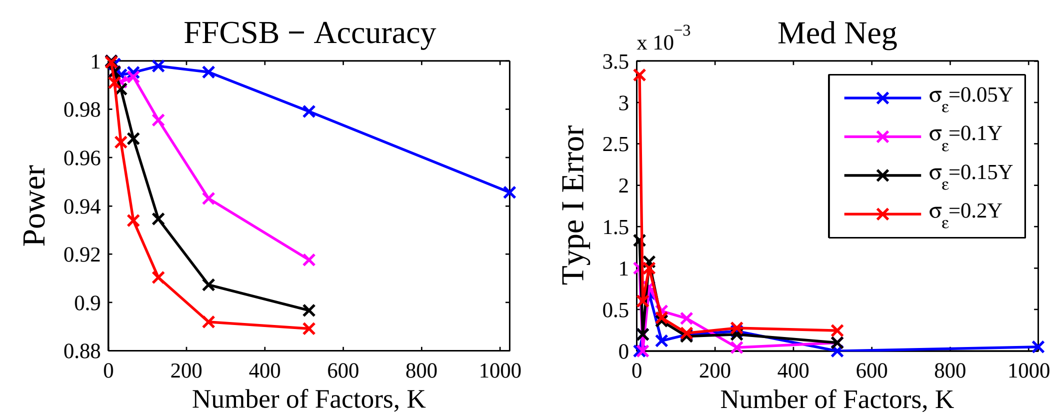

Figure 37. Failing FFCSB Accuracy for “Med Neg”

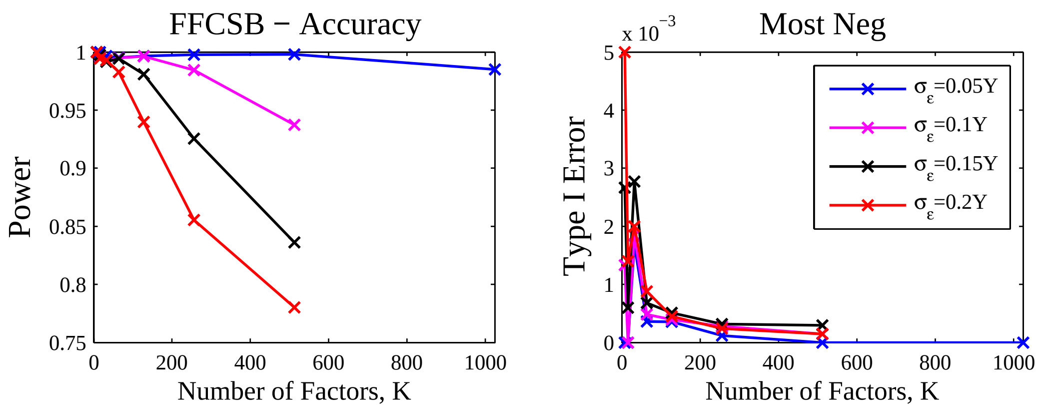

Figure 38. Failing FFCSB Accuracy for "Most Neg"

C. COMPARISON OF FFCSB WITH CSB UNDER HETEROGENEOUS RESPONSE VARIANCE

1. Accuracy: FFCSB is Robust for All Factor Patterns while CSB Fails for Increasing Factor Negativity

The heterogeneous response experiments yield comparison findings between FFCSB and CSB similar to that for the homogeneous response experiments. The "Power and Type I Error" Figure 39 illustrates the accuracy performance of both algorithms. The upper graphs suggest that CSB fails to realize power guarantees beyond the "Small Neg" factor pattern while FFCSB fulfills accuracy guarantees for all factor patterns. The lower graphs support that both algorithms fulfill the Type I Error guarantees comfortably. There are missing data points for CSB due to excessive computation requirements. Simulation times exceeded reasonable timing (days) for collection.

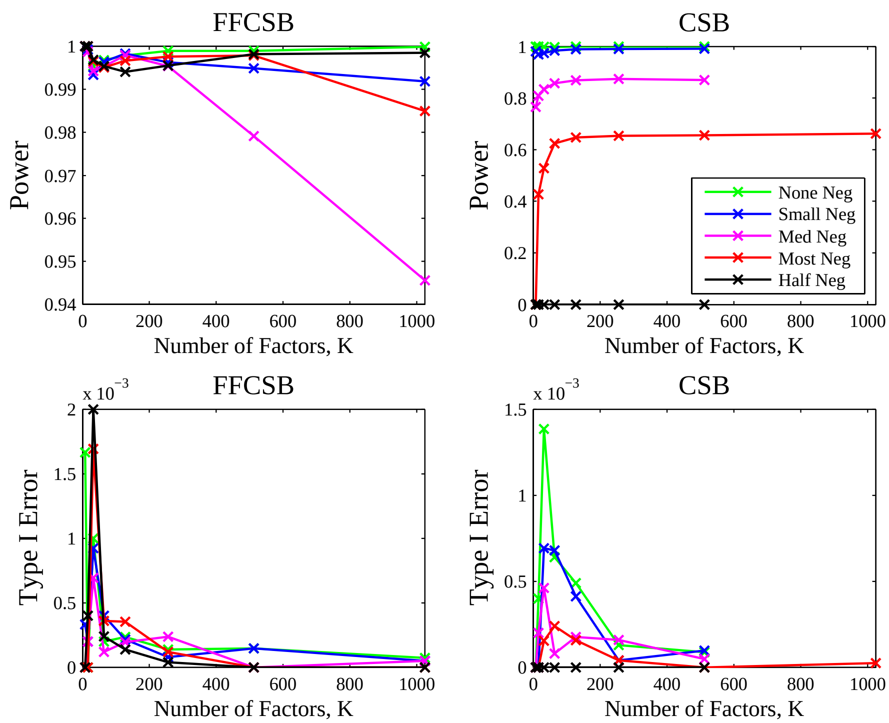

Figure 39. Accuracy Comparison of FFCSB versus CSB for Mild Heterogeneity (σ_{ε}=0.05Y) & Various Factor Patterns

Within the factor patterns of "None Neg" and "Small Neg" (Figures 40-41), both FFCSB and CSB meet accuracy guarantees and are robust to the magnitude of heterogeneity.

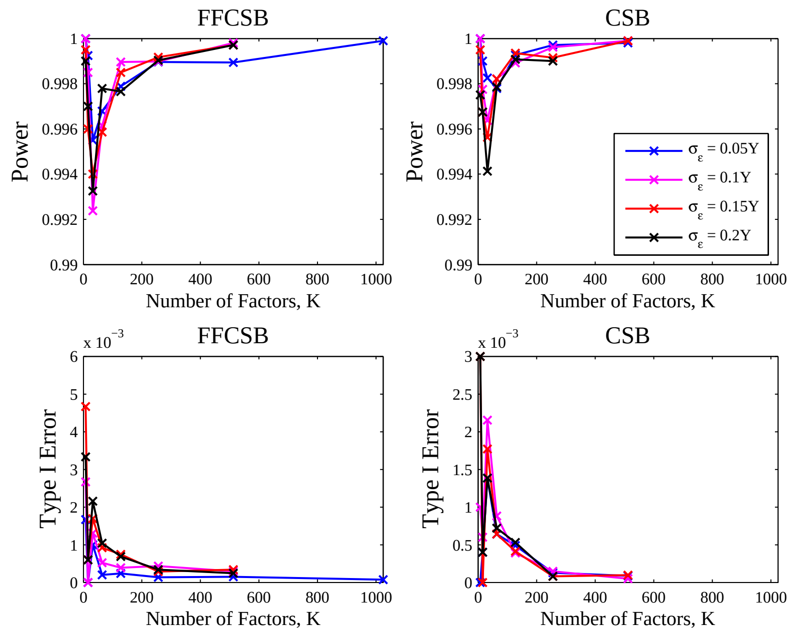

Figure 40. Accuracy Comparison of FFCSB versus CSB for "None Neg"

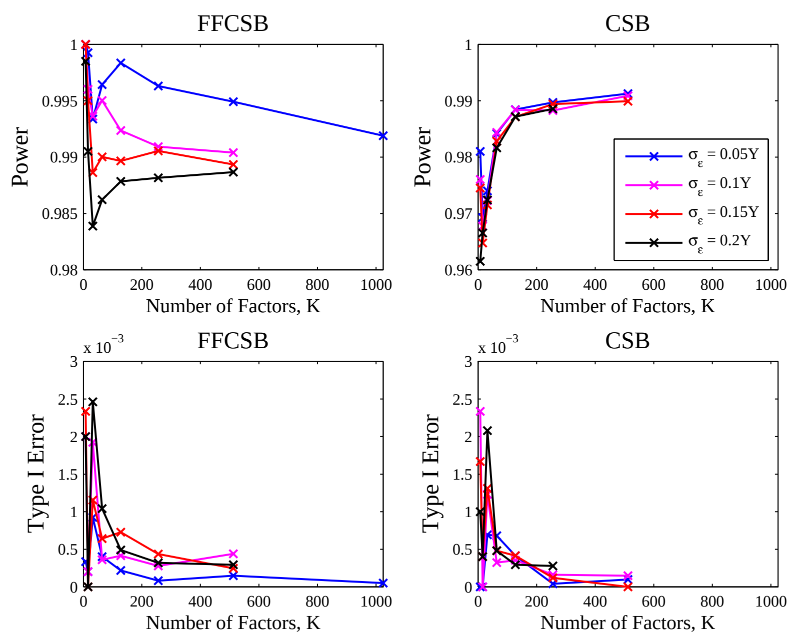

Figure 41. Accuracy Comparison of FFCSB versus CSB for "Small Neg"

2. Efficiency: FFCSB Offers at Least 30% Computation Savings Compared to CSB for Models with Higher Factor Counts or with Higher Response Heterogeneity

The computation resources for both algorithms are compared for factor patterns of "None Neg" and "Small Neg." Figure 42 presents the efficiency comparison of FFCSB versus CSB with the average runs and percentage savings presentation used in the homogeneous experiments. In both graphs, the growing gaps between each pair of same-colored lines suggests that CSB requires many more runs than FFCSB with increasing factor count. The gaps widen as the level of response heterogeneity increases. The missing data points for CSB and FFCSB are due to excessive computation requirements.

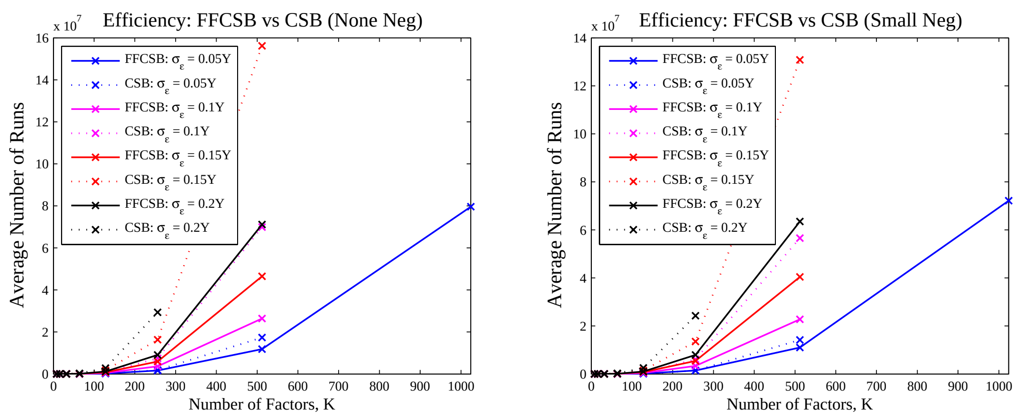

Figure 42. Efficiency Comparison of FFCSB versus CSB for Various Heterogeneous Variances. "None Neg" (Left) & "Small Neg" (Right)

Figure 43 presents the percentage savings for applying FFCSB versus CSB. FFCSB offers a significant computation savings advantage over CSB when applied on response models with more factors or with larger response heterogeneity. For larger response models with more than 200 factors and higher response heterogeneity, FFCSB offers at least a 30% computation savings as compared to CSB. On the other hand, for smaller response models with less than 100 factors and lower response heterogeneity, FFCSB consumes up to 50% more computation resources than CSB. Thus, FFCSB lends itself to application on bigger models more readily than CSB. Even when response heterogeneity is low, FFCSB computation savings set in at factors of 300 or more.

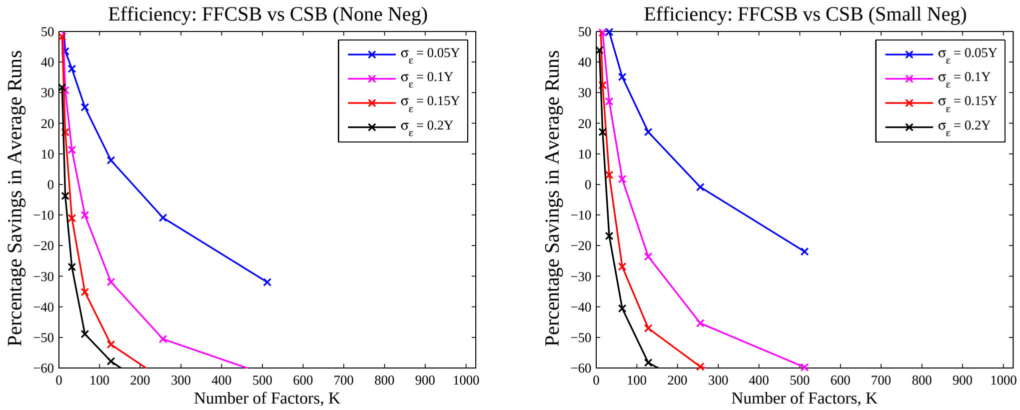

Figure 43. FFCSB Offers 30% or More Computation Savings Compared to CSB

D. COMPARISON OF FFCSB AND FF UNDER HETEROGENEOUS RESPONSE VARIANCE

1. Accuracy: FFCSB Performs Better than FF for "None Neg" and "Small Neg" and Equally Well for "Half Neg"

In the controlled experiments with heterogeneous response variances, FFCSB has proven to be robust in accuracy performance under the factor patterns of "None Neg," "Small Neg" and "Half Neg." FFCSB deteriorates in power performance under the remaining factor patterns of "Med Neg" and "Most Neg." Hence, the comparison of FFCSB versus FF under heterogeneous variances will use this division of factor patterns.

The next "Power and Type I" figure (44) compares the accuracy performance of FFCSB and FF under "None Neg" and severe heterogeneity (σ_ε =0.20Y). Within each graph, the red line represents FFCSB accuracy performance, while the remaining colored lines represent FF with a different number of replications. The replications are indicated in the legend. This figure is presented first in this section as it crystallizes two important comparison findings and prepares the reader for a latter appreciation of the FFCSB versus FF comparison. First, while FFCSB is robust to factor count, FF does not scale well with factor count under severe heterogeneity. FFCSB accuracy performance remains close to ideal despite increasing factor count. However, FF accuracy performance deteriorates rapidly, as seen from the steep drop-offs of the power lines and steep ascents of the Type I Error lines. Second, FF takes increasingly more replications in order to meet both performance guarantees at larger factor counts.

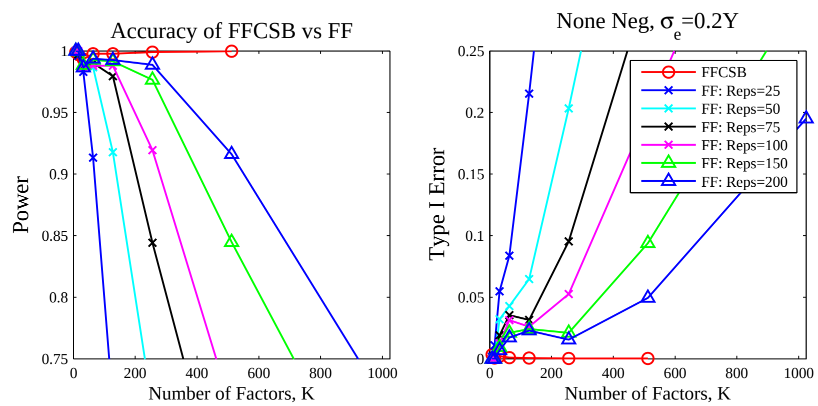

Figure 44. Accuracy Comparison of FFCSB versus FF for Severe Heterogeneity (σ_ε=0.20Y) & "None Neg"

With the understanding of FF performance characteristics, the following analysis will require multiple graphs to support the observations. Next, the first series of FFCSB versus FF graphs (Figure 45) compares the accuracy performance of FFCSB and FF under "None Neg" and various magnitudes of heterogeneity. Each pair of graphs presents the FFCSB accuracy performance for a different magnitude of heterogeneity, as indicated in the graph title.

The "None Neg" factor pattern is one of three favorable factor patterns for FFCSB, where the algorithm constantly fulfills accuracy guarantees for increasing factor counts and heterogeneity. Graphically, this is supported by the red lines in all graphs constantly at power = 1 and Type I Error = 0. On the other hand, the FF lines depict deteriorating accuracy performances with increased factor counts and increased heterogeneity. The latter is observed from the same colored-lines in different graphs (i.e., FF with the same number of replications but subjected to different magnitudes of heterogeneous variance) providing worse performance guarantees as heterogeneity increases. The same observations apply to the accuracy comparison of FFCSB and FF under the second favorable factor pattern of "Small Neg" (Figure 46, p. 49).

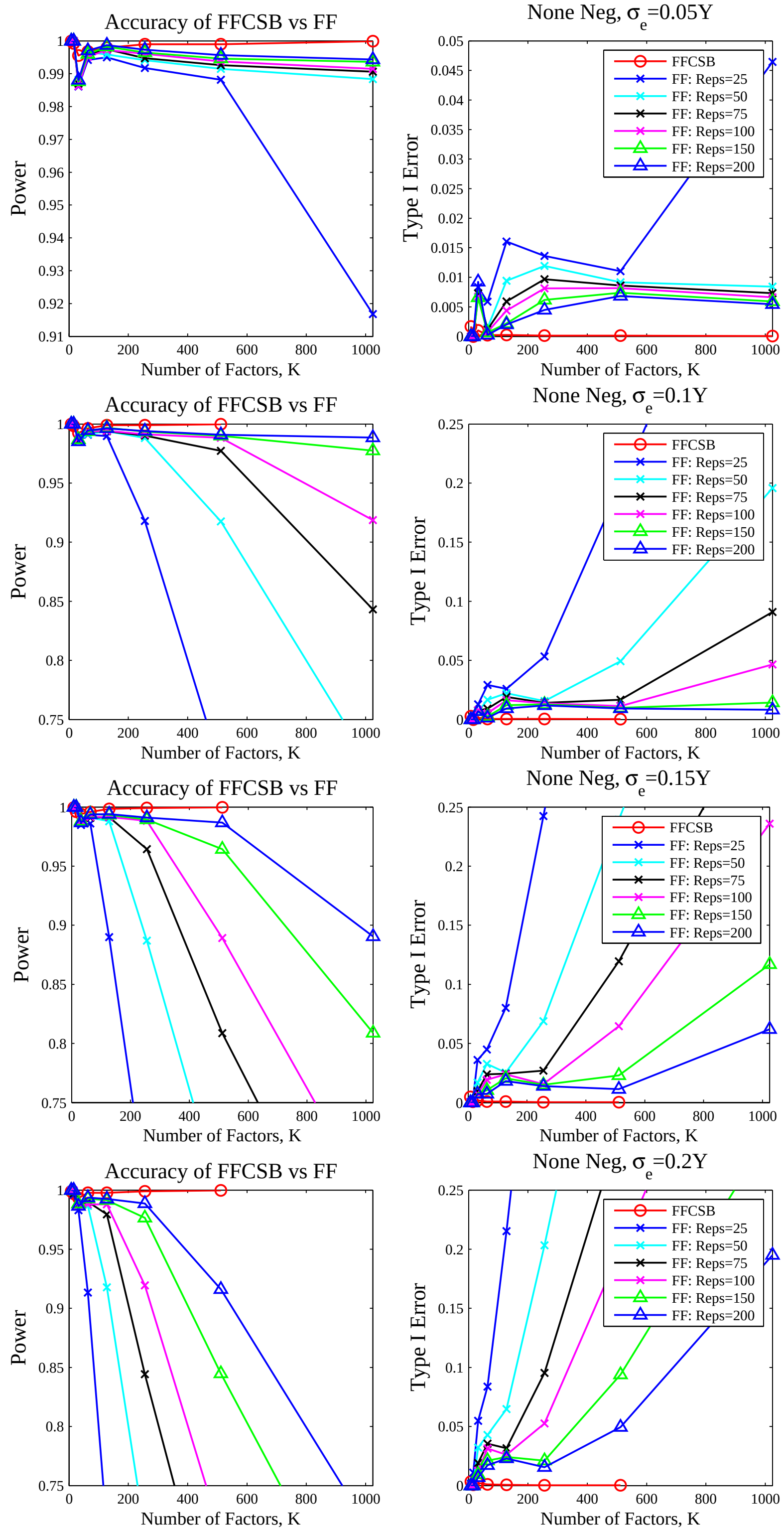

Figure 45. Accuracy Comparison of FFCSB versus FF for "None Neg" & Various Heterogeneous Variances (σ_ε=0.05Y, 0.10Y, 0.15Y or 0.20Y)

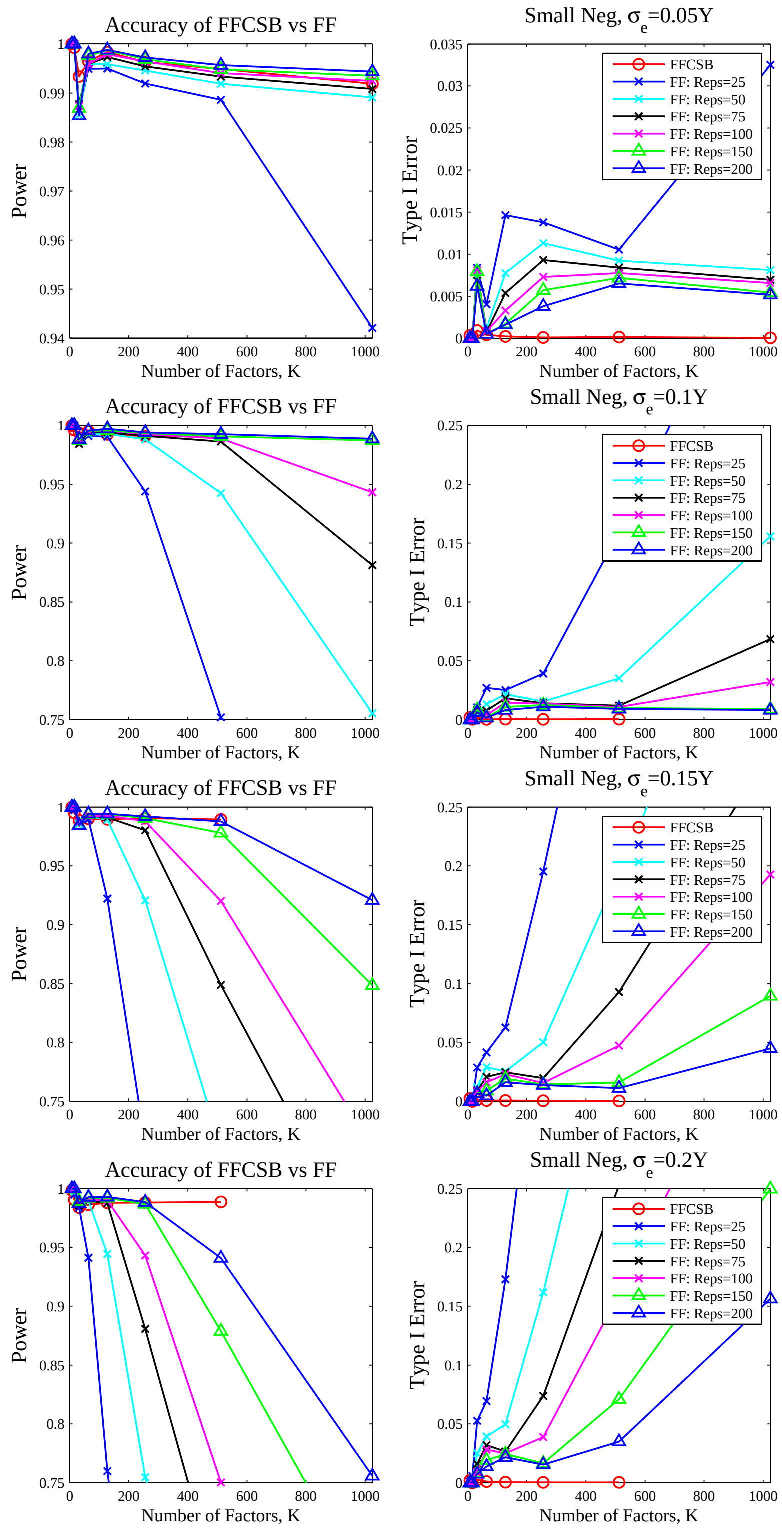

Figure 46. Accuracy Comparison of FFCSB versus FF for "Small Neg" & Various Heterogeneous Variances (σ_{ε}=0.05Y, 0.10Y, 0.15Y or 0.20Y)

The third series of graphs (Figure 47, p. 51) compares the accuracy performance of FFCSB and FF under "Half Neg" and various magnitudes of heterogeneity. This is the third favorable case for FFCSB accuracy performance. Both FFCSB and FF perform equally well. In this factor pattern, FF bucks previous trends and provides consistent performance guarantees for all simulated factor counts and heterogeneity. This suggests that FF provides optimum performance when there is a balanced mix of factor effect directions.

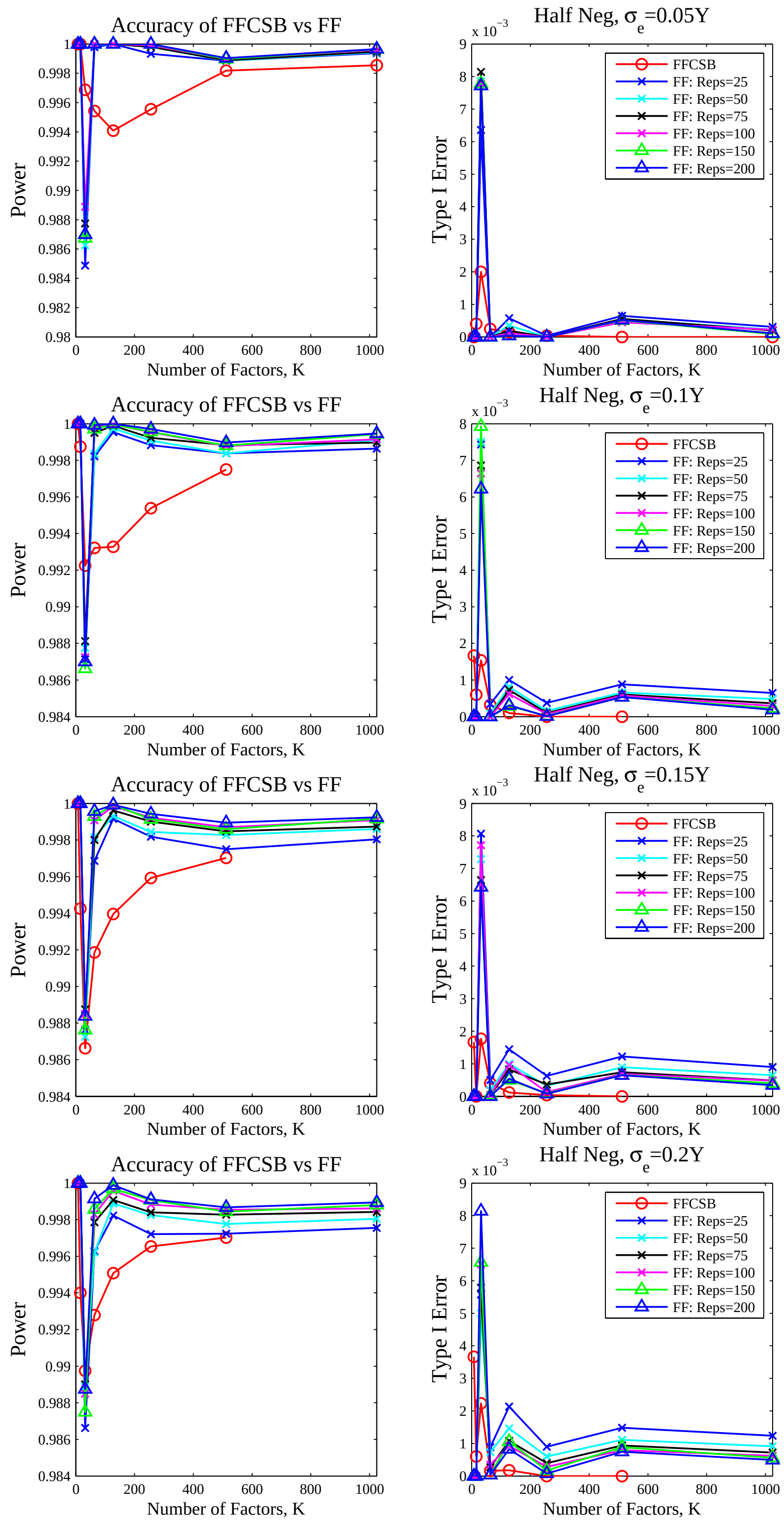

Figure 47. Accuracy Comparison of FFCSB versus FF for “Half Neg” & Various Heterogeneous Variances (σ_ε =0.05Y, 0.10Y, 0.15Y or 0.20Y)

Thus far, FFCSB has been compared with FF in the former's favorable operating cases. Next, the fourth and fifth series of figures (Figures 48-49 on pp. 53-54) present the FFCSB versus FF accuracy comparisons for "Med Neg" and "Most Neg" factor patterns respectively, with increasing magnitudes of heterogeneity. Where FFCSB has failed, FF displays its characteristic non-scalability with increased factor count or with increased heterogeneity. However, the deterioration in FF performance in "Med Neg" and "Most Neg" appears less severe than in the "None Neg" and "Small Neg."

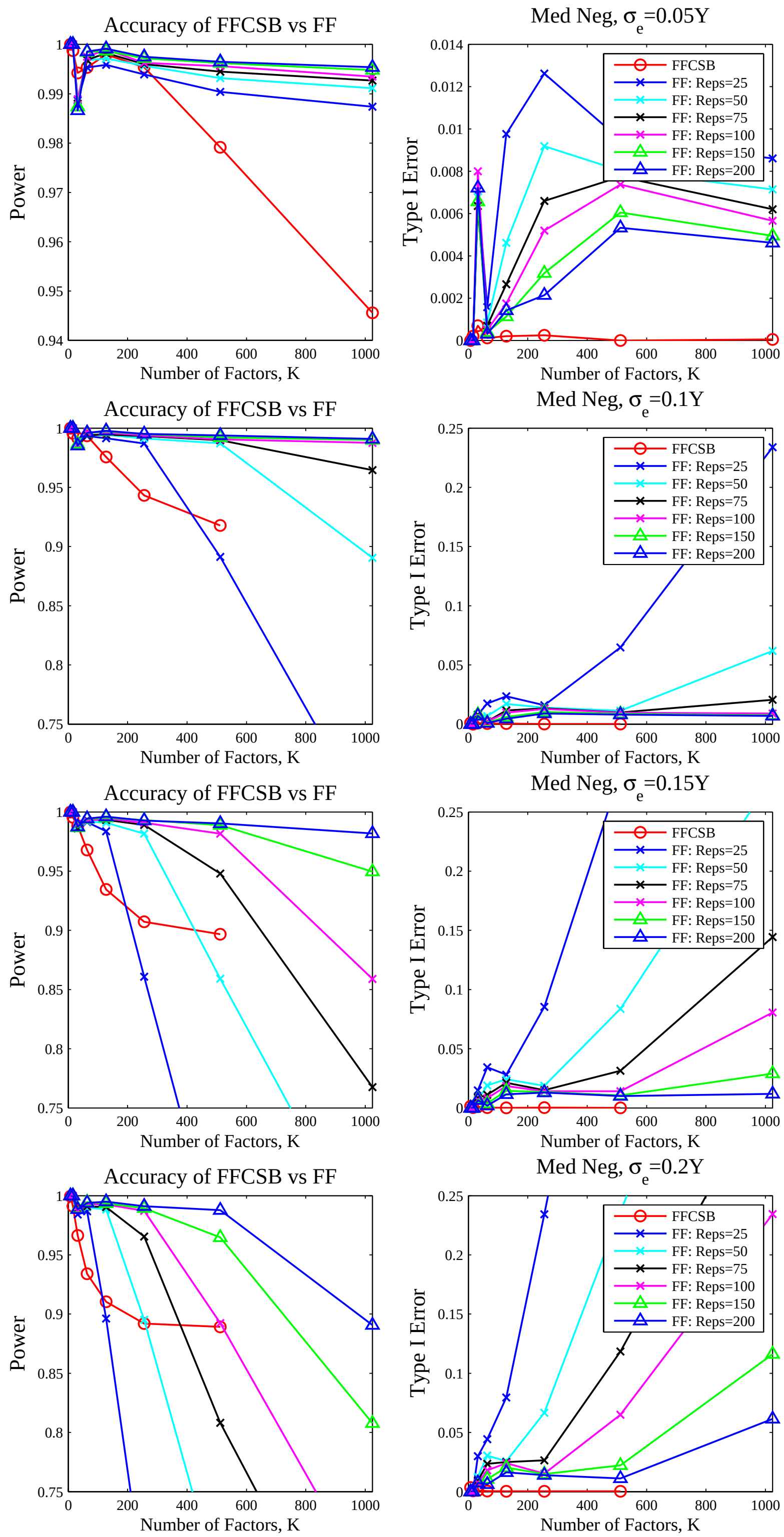

Figure 48. Accuracy Comparison of FFCSB versus FF for "Med Neg" & Various Heterogeneous Variances (σ_ε =0.05Y, 0.10Y, 0.15Y or 0.20Y)

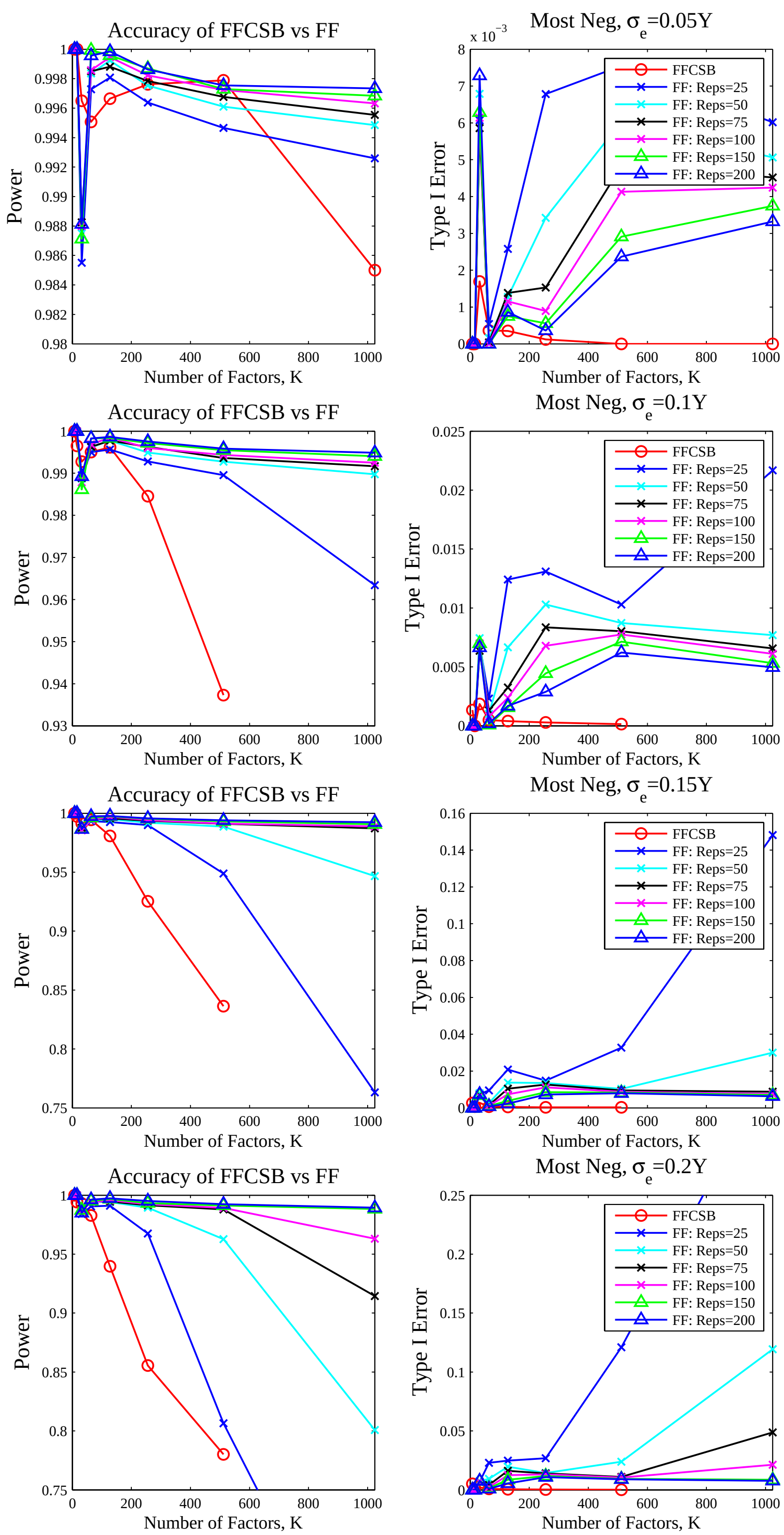

Figure 49. Accuracy Comparison of FFCSB versus FF for “Most Neg” & Various Heterogeneous Variances (σ_ε =0.05Y, 0.10Y, 0.15Y or 0.20Y)

The observations on FF prompt analysis on FFCSB versus FF from an alternative perspective. The last series of figures (50-51) presents FFCSB versus FF accuracy performance under severe heterogeneity (σ_ε=0.20Y) for various factor patterns. Indeed, FF demonstrates improved performance guarantee with an increasing percentage of negative factor effects in the response model. This characteristic can work against or for FF. FF delivers good performance with a balanced mix of factor effects (50% positive and 50% negative) in the response model, but the algorithm fails when the factor effects in the response model are lopsided in either direction. This finding may be utilized to reinforce the FFCSB algorithm. Additional replications of the FF design can be made at the pre-sorting phase of FFCSB to gain confidence in its factor sorting by effect direction, while the second phase of CSB would provide the accuracy guarantee.

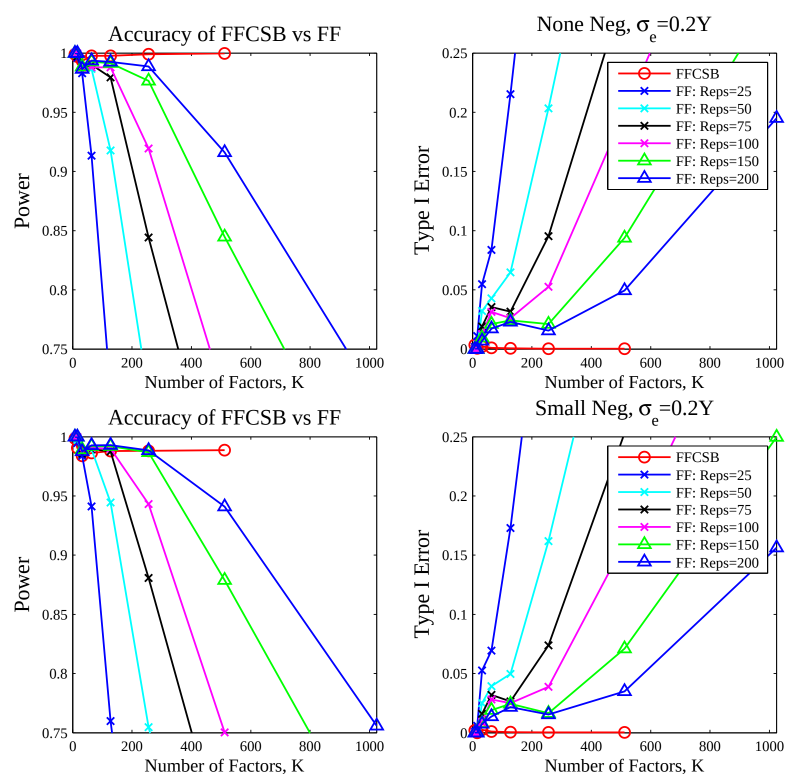

Figure 50. Accuracy Comparison of FFCSB versus FF for Severe Heterogeneity (σ_ε=0.20Y) & (Top) "None Neg" & (Bottom) "Small Neg"

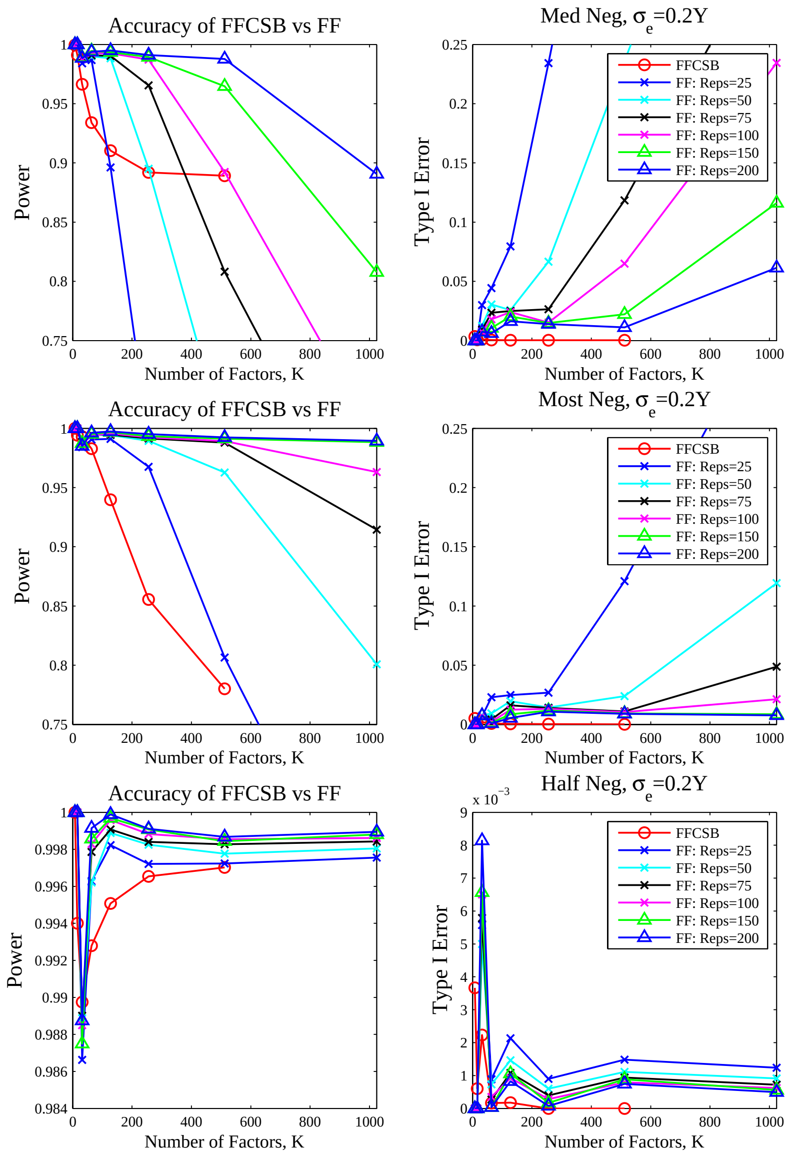

Figure 51. Accuracy Comparison of FFCSB versus FF for Severe Heterogeneity (σ_{ε}=0.20Y) & (Top) “Med Neg”, (Middle) “Most Neg” & (Bottom) “Half Neg”

2. Efficiency: FFCSB Expected to be Less Efficient than FF. However, FF Suffers from Implementation Issues for Large Factor Count

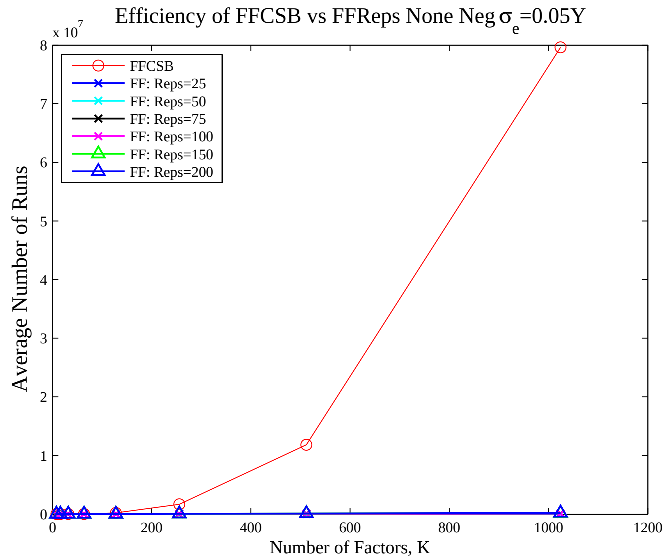

Figure 52. Characteristic Efficiency Performance of FFCSB versus FF

Figure 52 depicts the characteristic efficiency performance of FFCSB versus the FF replications simulated. Towards the left of the graph where model sizes are small, FFCSB is as efficient as FF, as illustrated by the proximity of the red line to the remaining colored lines. Towards the right of the graph where model sizes are large, FFCSB requires many more runs to meet accuracy guarantees than the FF replications simulated. FF deteriorates in performances (from earlier accuracy discussion) when working with a fixed run budget insufficient to decipher the response model. FFCSB is expected to be less efficient than FF, especially as model size increases. There is also another limitation to taking large FF replications. The Matlab implementation developed in this thesis was unable to compute large FF replications (in ranges of 1000s of replications) due to out of memory error. Thus, there is an implementation limitation to large-scale FF replications, whereas FFCSB does not have the problem as it is a sequential algorithm.

E. AREAS IN WHICH ALGORITHMS EXCEL

Generally, heterogeneous variances are more realistic because they occur more frequently than homogeneous variances in real world systems. The heterogeneity poses a greater challenge to the algorithms. All three algorithms either fail to fulfill accuracy guarantees or take more computation power to maintain accuracy guarantees than in the case of homogenous variances.

FFCSB has proved its accuracy in the series of experiments on response models with heterogeneous variance and various factor patterns. FFCSB fulfills accuracy guarantees in the three factor patterns of “None Neg,” “Small Neg” and “Half Neg.” In these factor patterns, FFCSB maintains robust accuracy performance against heterogeneity and scales well for large models, even up to 1024 factors. In the two intermediate factor patterns of “Med Neg” and “Most Neg,” FFCSB fulfills accuracy guarantees under mild heterogeneity ($\sigma_{\varepsilon}=0.05Y$), but fails for any larger heterogeneity. FFCSB fails in these cases because they have a significant percentage of negative factors that are not negligible in effect and yet not critical enough to be classified. Thus, they cannot be eliminated as unimportant nor classified as critical. They remain within the experiments and cause errors in the factor classification.

Similar to findings from the homogeneous variance experiments, CSB fulfills accuracy guarantees in “None Neg” and “Small Neg.” CSB fails the remainder factor patterns with increasingly significant percentages of mixed factor effect direction. CSB performs poorly in efficiency in these heterogeneous variance experiments. CSB data points are missing due to excessive computation requirements for a single point. In the factor patterns that CSB fulfills performance guarantees (i.e., “None Neg” and “Small Neg”), FFCSB provides at least a 30% computation savings when applied to models with larger factor counts (200 or more) or with higher response variance heterogeneity.

In general, FF does not scale well with larger models beyond 1000 factors and deteriorates with increased heterogeneity. However, it improves with an increasingly balanced mix of factor effect direction. The comparison for FFCSB and FF is divided into FFCSB favorable operating factor patterns (“None Neg,” “Small Neg” and “Half

Neg") and unfavorable operating factor patterns ("Med Neg" and "Most Neg"). In the favorable factor patterns of "None Neg" and "Small Neg," FF fails to meet performance guarantees for larger factor counts and increased heterogeneity, while FFCSB maintains near ideal accuracy guarantees. In the favorable factor pattern of "Half Neg," both algorithms fulfill performance guarantees comfortably. In the unfavorable factor patterns of "Med Neg" and "Most Neg," FF displays accuracy deterioration with factor count and heterogeneity, albeit to a less severe degree than in "None Neg" and "Small Neg." These observations suggest that more replications of FF in the FFCSB pre-sorting phase would help improve effectiveness of the pre-sorting phase, and consequently the overall effectiveness of FFCSB.

THIS PAGE INTENTIONALLY LEFT BLANK

V. APPLICATION OF FFCSB TO HIERARCHY ORGANIZATIONAL MODEL

> Transformation is "a process that shapes the changing nature of military competition and cooperation through new combinations of concepts, capabilities, people and organizations that exploit our nation's advantages and protect against our asymmetric vulnerabilities to sustain our strategic position, which helps underpin peace and stability in the world."
>
> U.S. Transformation Planning Guidance (April 2003)

A. CHAPTER OVERVIEW

In a first-case application, FFCSB is used to support current research in Computation Organization Theory. This section describes the FFCSB application on the Hierarchy organizational model and compares expert opinion with FFCSB findings.

B. MOTIVATION OF FFCSB APPLICATION

In *Joint Vision 2010*, Army General John M. Shalikashvili, Chairman of the Joint Chiefs of Staff, said that "The nature of modern warfare demands that we fight as a joint team. This was important yesterday, it is essential today, and it will be even more imperative tomorrow." In light of the critical Transformation drive of the U.S. military, innovative organizational models are needed to deliver better team performance.

At the Naval Postgraduate School, the Center for Edge Power (CEP) has keen interest in research on network-centric operations (e.g., organization, command and control (C2), management, doctrine and personnel) to enable more powerful warfare. In collaboration with Stanford University, CEP conducts computational experimentation on various C2 structures to understand the factors that drive team performance, which can be measured in various forms (e.g., duration, risk and cost.) Compared to field experiments and human-in-the-loop experiments, computation experimentation is relatively inexpensive, fast and easy to conduct. It generates knowledge of the comparative strengths and weaknesses of organizational forms. Such knowledge helps decision makers arrive at better decisions and avoid costly mistakes in warfare. FFCSB extends CEP's suite of computational tools to explore large and complex computer simulations of organizational behavior and identify important factors that drive the bottom-line of organization performance. In this application, FFCSB will identify important factors in the Hierarchy model that drive the Measure of Performance of Project Duration. The Hierarchy is representative of the prevalent structure in militaries and serves as a benchmark for comparisons of new organizational forms. The model is studied with factor ranges spanning two contrasting mission-environmental contexts: the Industrial Age and the 21st Century.

C. EXPERIMENTATION TOOLS FOR ORGANIZATION THEORY

1. POW-ER Computation Experimentation Tool

These complex computer simulations of organizational behavior are developed in POW-ER—Projects, Organizations and Work for Edge Research—a virtual environment for computational modeling of C2 organizations and processes. POW-ER builds upon collaborative research and development between NPS and Stanford University. The organizational models are formulated from well-accepted organizational theory. The computation tool has been validated extensively and thoroughly via: "1) internal validation against micro-social science research findings and against observed micro-behaviors in real-world organizations, 2) external validation against the predictions of macro-theory and against the observed macro-experience of real-world organizations, and 3) model cross-docking experiments against the predictions of other computational models with the same input data sets" (Orr and Nissen 2006, p. 8, Levitt et al, 2005). The POW-ER environment uses agent-based simulation to emulate micro-behaviors (e.g., trust, learning, skill sets compatibility, skill competency, centralization) and discrete-event-simulation to emulate processes (e.g., meetings, exception occurrences, rework, process quality). Organizational performance is measured by quantitative metrics, e.g., project duration, project risk, project cost.

2. Computational Experimentation for Organizational Studies

Using the POW-ER environment, researchers have conducted modeling, simulation and analysis of comparative performance of alternate C2 approaches, including different organization structures, work processes, technologies and personnel. Research and experimentation results have been published in a series of recent works. First, Nissen (2005) laid the fundamentals by defining the Hierarchy and Edge organization models from theory and comparing their performance in the Industrial Age and 21st Century mission contexts. Second, Orr and Nissen (2006) defined four more organization models and compared the performance of the six organizations in the Industrial Age and 21st Century mission contexts. Third, Gateau et al., (2007) articulated an organizational design space, using only three parameters of centralization, hierarchy and application experience to characterize organization models. Most recently, Mackinnon et al., (2007) calibrated and compared the impact of learning and forgetting micro-behaviors on the Hierarchy and Edge organizational models in the Industrial Age and 21st Century mission contexts

3. FFCSB: An Alternative Approach to Tackle the Same Question

The Hierarchy organization model is modeled by three sets of structural factors: (1) organization structure (2) communication structure (3) work structure (Nissen 2005, p. 11). The Industrial Age and 21st Century mission contexts are modeled by three manipulations of mission factors: (1) mission and environmental context, (2) network architecture and (3) professional competency (Nissen 2005, p. 14).

Researchers typically used full factorial experimental designs to explore organizational performance over different organizational structures and mission contexts. Nissen (2005) used a 2 organizations $\times$ 2 scenarios design, while Orr and Nissen (2006) used a larger 6 organizations $\times$ 2 scenarios $\times$ 4 manipulations design. Mackinnon et al., (2007) keeps simulation parameters constant between the Hierarchy and other organizations in order to isolate performance change due to learning and forgetting micro-behaviors only. Given that there are hundreds or thousands of factors in such complex organization models, it is computationally expensive or infeasible to conduct full factorial designs on individual factors. Instead, the design of experiments would use the six groups of factors listed above and change multiple factors within a group as one variation. Experimental results of organizational performance were analyzed over the entire organization's model changes and mission changes (Nissen 2005) or single block change (Orr and Nissen 2006, Mackinnon et al., 2007). Through analyzing the relative impact of each variation on individual organization performance, the researchers drew practical insights. For instance, Orr and Nissen inferred that: "professional competency improvements to the Hierarchy/Machine Bureaucracy can produce even more dramatic results in terms of agility as those associated with adopting the Edge organizational form. Hence, a change in professional competency can be substituted to a large degree for a change in organizational form. Unlike the substitution effects noted above for the network architecture manipulation, however, the converse does not hold for professional competency: changing organizational form does not compensate for a reversion to an efficiency-oriented organization and knowledge-flow approach" (2006, p. 16).

FFCSB offers an alternative approach to tackle the same question. It offers single factor resolution and allows researchers to probe questions such as: For an organization model, which are the topmost important single factors, either organizational or mission, driving the Measure of Performance? Without group screening algorithms, it would have required an exorbitant amount of experimentation resources to conduct full factorial experiments to identify performance enhancement (or deterioration) due to single factors. FFCSB overcomes this limit by efficient division and experimentation of the entire factor space, and gradually limiting the scope of search for important factors. Through group screening of singular factors, FFCSB can shed light on significant individual factors within each structural or mission factor block that have the most impact on the outcome of interest.

D. MODEL DESCRIPTION & SIGNIFICANCE

1. Hierarchy Organizational Model

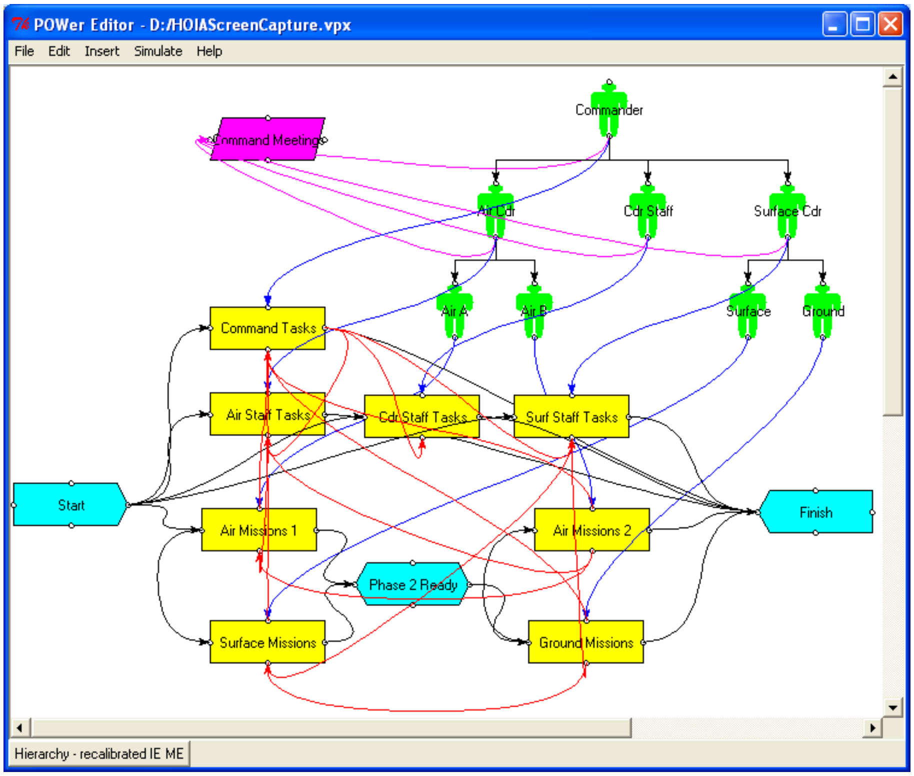

Figure 53. Hierarchy Organizational Model in POW-ER

Figure 53 is a screen-capture of the Hierarchy model in the POW-ER environment. The figure illustrates the personnel hierarchy and mission structure in the Hierarchy model. Personnel are grouped and communicate over a 3-tier command chain, which emulates the Command, Coordination and Operations levels in a Joint Task Force Hierarchy (Nissen 2005). There are four tasks executed sequentially via two phases. Tasks are linked to each other and to project milestones. Tasks can flow completed work down the chain, or flow rework (additional work to rectify earlier mistakes). Personnel are linked to work on meetings and tasks. Operations level personnel act directly on tasks, while Command and Coordination level personnel act directly upon their specialized tasks while indirectly supporting operations tasks.

2. Measure of Performance: Project Duration

Earlier quoted works compared organizational performances using seven MOPs: duration, cost, project risk, maximum backlog, work volume, rework volume and coordination volume. This FFCSB application focuses on the first MOP of interest: (Project) Duration. Duration is defined as "the predicted time to perform a project, in working days, which includes both direct and indirect (i.e., coordination, rework and decision latency) work" (Orr and Nissen, 2006).

3. Factor Exploration Space

Table 9 lists the factors identified in the Hierarchy model for the FFCSB application. FFCSB was applied to the Hierarchy model with this entire factor space in one exploration. However, this single exploration took weeks to run, without yielding results. The sequential nature of FFCSB meant that the experiments could not be parallelized. There were unusually long simulation times of the Hierarchy model, possibly due to combinations of factors that were either unreasonable or stressed the model too much.

Table 9. Factor Space for Exploration of Hierarchy Model

Mission & Environment	Network Architecture	Professional Competency
(Project) Function Exception Probability (Project) Project Exception Probability (Task) Effort (Task) Learning Days (Task) Priority (Task) Requirement Complexity (Task) Solution Complexity (Task) Uncertainty (Personnel) Full Time Equivalent (Personnel-Task) Allocation (Task-Task) Successor	(Project) Priority (Project) Length Of Work-day (Project) Length Of Work-week (Project) Centralization (Project) Matrix-strength (Project) Communication Probability (Project) Noise Probability (Project) Instance Exception Probability (Meeting) Priority (Meeting) Duration (Personnel-Meeting) Allocation (Task-Task) Rework Strength	(Project) Team Experience (Personnel) Culture (Personnel) Role (Personnel) Application Experience (Personnel) Cultural Experience (Personnel) Skill Ratings

In order to keep within the computation resources and time constraints for this section, the entire factor space was divided into three subspaces for separate FFCSB exploration. Hence, three smaller and faster explorations were conducted instead of one big exploration. The division of the factor space followed the three manipulations of mission context factors: (1) mission and environmental context, (2) network architecture and (3) professional competency. In addition, the three sets of structural factors: (1) organization structure (2) communication structure and (3) work structure were subsumed under these factor subspaces. This division of factor space was intended to mirror that in the literature as closely as possible, but was not exact. The factor ranges of exploration were derived from the default values of the Hierarchy model in the contrasting mission contexts of Industrial Age and 21st Century.

4. Expert Opinion on Significant Factors

Among the factors identified for exploration, subject matter experts (SMEs) identified the following as important:

1. Mission & Environment
    a. (Personnel) Full Time Equivalent
    b. (Task) Effort
2. Professional Competency
    a. (Personnel) Application Experience
    b. (Personnel) Skill Ratings

E. FFCSB FINDINGS ON SIGNIFICANT FACTORS

The following tables (10-11) summarize the FFCSB findings of important factors in the Hierarchy model that impact Project Duration most. There were no factors classified as important in the Network Architecture factor subspace.

Table 10. Important Factors in Mission & Environment Factor Subspace

Object	Attribute	Factor Effect on Duration
Mission	Project Exception Probability	+
Surface Missions	Effort	+
Surface Missions	Solution Complexity	+
Ground Missions	Effort	+
Ground Missions	Requirement Complexity	+
Ground Missions	Solution Complexity	+

Table 11. Important Factors in Professional Competency Factor Subspace

Object	Attribute	Factor Effect on Duration
Mission	Team Experience	+
Air A (Personnel)	Skill Ratings	-
Ground (Personnel)	Skill Ratings	-

In the first factor subspace of Mission & Environment, SMEs identified the factors of Full Time Equivalent (FTE) and Effort as important. FTE measures the equivalent of manpower resources available and Task Effort quantifies the time effort requirement of the task. Contrary to expert opinion, FFCSB did not classify any FTE factors as important over the factor range of exploration. Thus, FTE is not as important as the other factors in this subspace in impacting the Project Duration. In line with expert opinion, FFCSB classified Effort factors as important, but only for Surface Missions and Ground Missions out of all eight missions in the Hierarchy model. Critical path analysis of the Hierarchy model explains why factors associated with only these two missions showed up consistently as important.

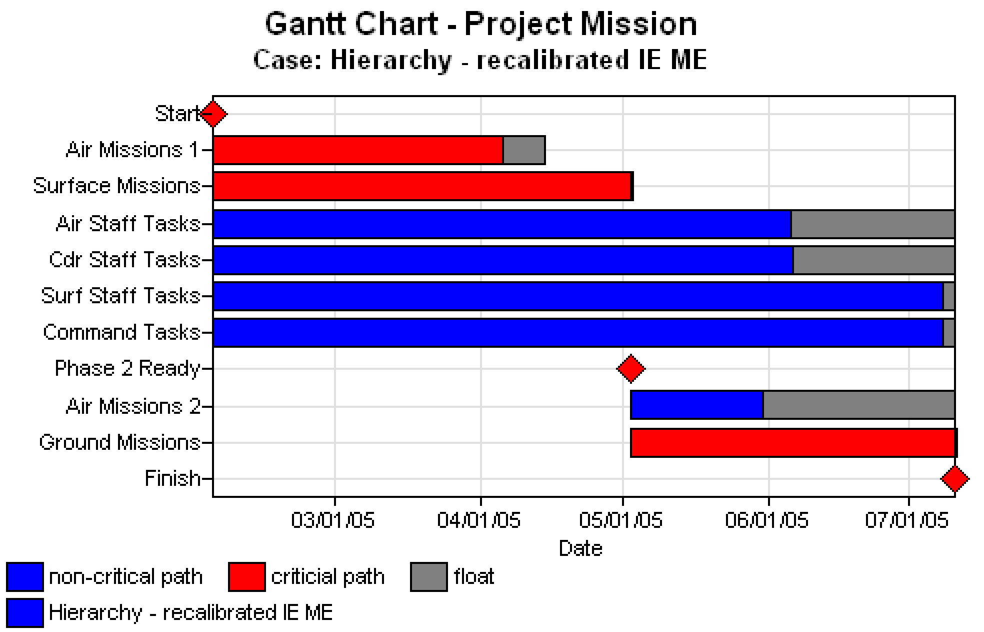

Figure 54. Critical Path Analysis of Hierarchy model shows Air Missions 1, Surface Missions and Ground Missions on Critical Path (Best viewed in color)

The red bars in Figure 54 depict the critical path of the project simulated in the Hierarchy model. Following the red bars, the Air Missions 1, Surface Missions and Ground Missions are on the critical path. Of these three missions, the Surface Missions and Ground Missions have minimum float, i.e., there is no allowance for shifting these missions in time. Hence, these two missions are crucial to the MOP of Project Duration. Besides the Task Effort factor, FFCSB also classified the Solution Complexity factors of the Surface and Ground Missions as important, as well as the Requirements Complexity of the Ground Missions. Thus, FFCSB has further quantified expert opinion by flagging those factors associated with missions on the critical path only and with specific characteristics.

In addition, FFCSB classified the global factor of Project Exception Probability (PEP) as important. PEP is the probability that a subtask will fail and generate rework for failure dependent tasks. This factor is significant for the Hierarchy model that is characterized by sequential and interdependent tasks and hence, suffers a longer Project Duration in the event of increased PEP.

In the second factor subspace of Network Architecture, there are no factors classified as important. This finding is in agreement with SMEs, who did not expect any important factors in this subspace. A set of (relatively computationally expensive) Resolution V Fractional Factorials design was used to verify the factor coefficients in this factor group. The results confirmed that the factor coefficients were relatively small in magnitude and hence, practically insignificant.

In the third factor subspace of Professional Competency, experts identified Skill Ratings and Application Experience factors as important. FFCSB classified the Skill Ratings of the Air A and Ground personnel as important, but not that of the Surface personnel. These three groups of personnel are operations personnel and directly responsible for the missions on the critical path. The contrast between the three missions is that the Surface Missions require a considerably longer effort of 21 months versus that of the Ground Missions (6.5 months) and Air Missions 1 (11 months). These findings suggest that Skill Levels may be more critical for missions that lie on the critical path and have relatively shorter Effort requirements. FFCSB did not classify Application Experience as important. However, interestingly, FFCSB classified Team Experience as important and positively related to the MOP. Team Experience quantifies the degree of familiarity that team members have in working with one another as a team. In other words, this finding suggests that more team experience leads to longer Project Duration in the Hierarchy model. This counter-intuitive finding may have been observed in earlier research and experimentation. Ramsey and Levitt (2005) summarized high level findings from Horii, Jin and Levitt's "Modeling and Analyzing Cultural Influences on Team Performance through Virtual Experiments" (2004) on the impact of cultural differences in project teams: "Japanese-style organizations were more effective, with either US or Japanese agents, at performing tasks with high interdependence when the team experience of members was low." The Hierarchy model studied in this application shares common characteristics of centralized authority, high formalization, and multiple hierarchies with the Japanese-style organization modeled in Horii, Jin and Levitt (2004 pp. 3). In addition, these experiments had used the MOPs of Project Duration and Quality Risk to quantify team performance, while this FFCSB application only used

Project Duration. Hence, there is common ground to compare the similarity of both findings. Had the original intuition on Team Experience been applied with conventional screening algorithms, this factor could have distorted screening findings.

Lastly, there were two general observations of interest. First, there were more important factors associated with the Operations layer of the JTF structure than the other layers. Recall that the Hierarchy model has a 3-tier command chain that models the Command, Coordination and Operations layers in a JTF. Second, there were more uncontrollable or difficult to control factors (e.g., Project Exception Probability, Task Requirement Complexity, Task Solution Complexity and Team Experience) than controllable or easy to control factors (e.g., Skill Ratings.)

F. WAY AHEAD

The FFCSB application of the Hierarchy model was conducted at the International Data Farming Workshop 15 in Singapore in November, 2007. A team of four international data-farming enthusiasts collaborated on the simulation and analysis of this exploration for a week. The FFCSB application produced many delightful surprises. Part of the important factor classification was in line with expert opinion and part of it ran contrary to expectations. There were new findings of important factors that were justified by critical path analysis and in agreement with earlier research and experimentation. Overall, this particular FFCSB application has confirmed expert opinion, flagged out new important factors and produced some interesting hypothesis, all for further exploration.

There are limitations to the FFCSB application to any model. FFCSB assumes a main effects model and interactions can distort the accuracy of factor classification. The nature of the response variance (homogeneous or heterogeneous) and its magnitude are unknown. Both model characteristics can have bearings on the FFCSB findings and accuracy guarantees. Particular to the Hierarchy model, the observations of this FFCSB exploration are unique to the factor space organization and ranges of exploration. Hence, the findings are not conclusive of the Hierarchy model. The important factor classification and observations are meant to provide direction for researchers in future work and optimize their experimentation budget on truly important factors. This first-case FFCSB application on a real-world simulation model has produced results that are coherent with critical path analysis and that agree with earlier research on similar models. Hence, it is an encouraging sign that FFCSB can serve as a complementary tool to better understand complex simulation models.

VI. CONCLUSIONS

A. CONCLUSIONS

FFCSB is a newly proposed screening algorithm that offers enhancements over conventional screening algorithms. In the series of controlled experiments, more has been learnt about its performance, from accuracy and efficiency perspectives.

Figure 55 summarizes the comparison of FFCSB, CSB and FF for experiments on response models exhibiting homogeneous variances. The key findings are:

1. FFCSB fulfills accuracy guarantees for all factor patterns. It maintains consistent performance for all factor patterns, model sizes and variance magnitudes.
2. FFCSB is more robust than CSB in handling mix of factor effects and offers up to a 25% computation savings. The mix of factor effects causes CSB to fare poorly as factors of opposite directions in the same screening group cancel out one another's effects. FFCSB averts this undesirable phenomenon via the FF pre-sorting phase to divide the entire factor space into positive and negative groups for CSB screening.
3. FFCSB and FF are equally matched in accuracy, but FFCSB can be less efficient than FF. However, FFCSB is more robust to non-ideal settings of control parameters, which often happens when exploring response models. Also, FFCSB does not require *a priori* knowledge of the number of experiments to conduct for complete factor classification, as FF does.

Figure 56 summarizes the comparison of FFCSB, CSB and FF for experiments on response models exhibiting heterogeneous variances. The key findings are:

1. FFCSB fulfills accuracy guarantees for three of the five factor patterns simulated. It fails when there are significant percentages of opposite factor effects that are not negligible in effect and yet not critical enough to be classified. Hence, these effects distort the factor classification accuracy. In the three favorable factor patterns, FFCSB is robust to variance magnitudes and model sizes. In the two unfavorable factor patterns, FFCSB deteriorates with increased variance and model size.
2. FFCSB is more robust than CSB in handling a mix of factor effects and offers at least a 30% computation savings.

3. In the three FFCSB favorable factor patterns, FFCSB fulfills accuracy guarantees better than FF. FF accuracy scales poorly with increasing model size.

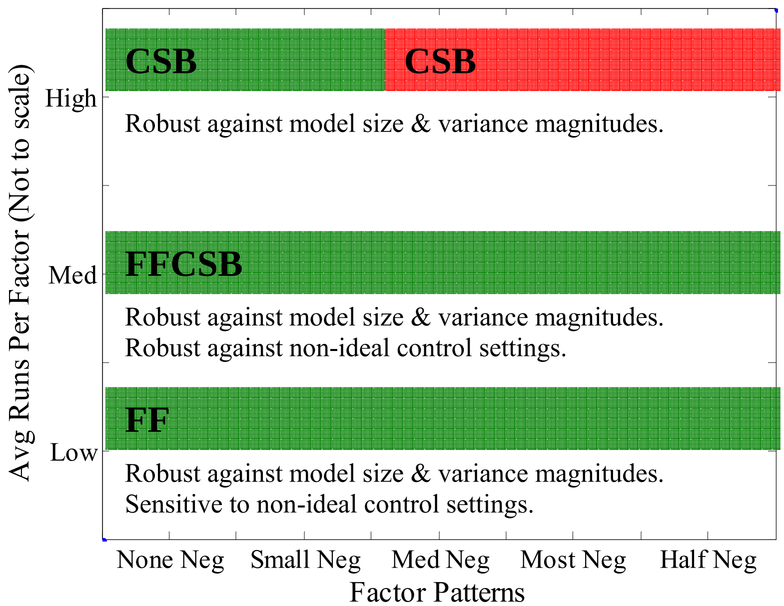

Figure 55. Comparison of FFCSB with CSB & FF for Homogeneous Variances

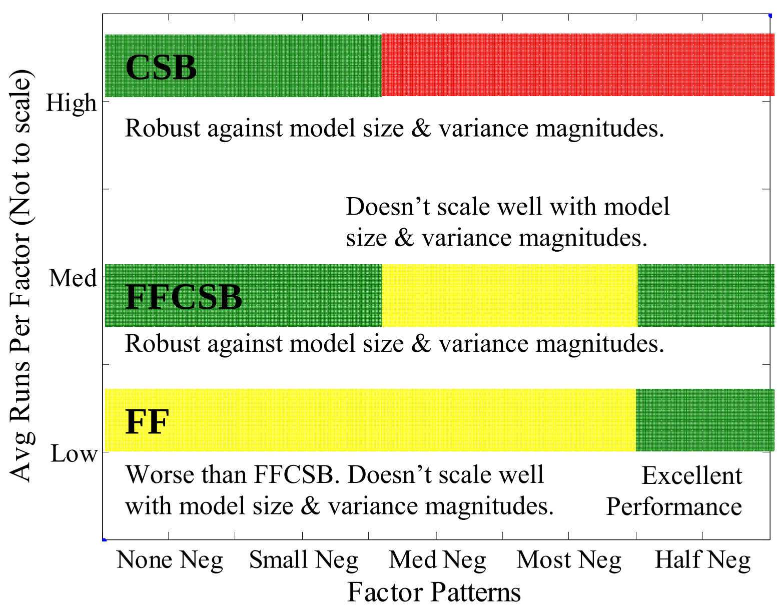

Figure 56. Comparison of FFCSB with CSB & FF for Heterogeneous Variances

In a first-case application, FFCSB is used to support current research in Computation Organization Theory. FFCSB is applied to identify important factors in the Hierarchy model that drive the Measure of Performance of Project Duration. FFCSB findings were in agreement with expert opinion on two out of four factors. In addition, FFCSB provided interesting observations:

1. There were other relatively important factors that drive Project Duration in the Hierarchy model.
    a. Project Exception Probability (probability that a subtask will fail and generate rework for failure dependent tasks)
    b. Task Requirement Complexity & Task Solution Complexity – Only for missions on critical path
    c. Team Experience (Familiarity of team working together)
2. There were no important factors in the Network Architecture subspace.
3. Counter to intuition, higher Team Experience led to longer Project Duration. This mirrors a similar finding from earlier research. Had the original intuition been used with conventional screening algorithms, this factor could have distorted factor classification.
4. The Hierarchy model has a 3-tier command chain that models the Command, Coordination and Operations layers in a Joint Task Force. There were more important factors associated with the Operations layer than the other layers.
5. There were more uncontrollable or difficult to control factors (e.g., Project Exception Probability, Task Requirement Complexity, Task Solution Complexity and Team Experience) than controllable or easy to control factors (e.g., Skill Ratings).

The important factor classification and observations are meant to provide direction for researchers in future work and optimize their experimentation budget on truly important factors. This first-case FFCSB application on a real-world simulation model has produced results that are coherent with critical path analysis and that agree with earlier research on similar models. Hence, it is an encouraging sign that FFCSB can serve as a complementary tool to better understand complex simulation models.

B. FUTURE WORK RECOMMENDATIONS

The thesis has helped shaped understanding of the performance envelope of FFCSB, as well as comparatively to other screening methods, CSB and FF. More controlled experiments can be conducted to further this understanding of the algorithm and the circumstances under which different screening algorithms can offer maximum benefits.

The application of FFCSB on a real-world simulation model has produced encouraging results. Continued exploration of the Hierarchy model with different factor space organization and factor ranges would form a good sensitivity analysis study of the FFCSB application on the model. Exploration of a competing model, the Edge organization model, would form an interesting study in itself and allow for meaningful contrasts between both competing organizational forms.

APPENDIX

Table 12 lists the factors identified in the Hierarchy model for FFCSB application and their range of exploration.

Table 12. Factors & Ranges in Hierarchy Model for FFCSB Application

Object	Factor	Hierarchy Industrial Age	Hierarchy 21st Century	FFCSB Exploration	
				Low	High
Project	priority	Medium	Medium	Low	High
	work-day	480	480	360	600
	work-week	2400	2400	1440	3600
	team-experience	Low	Low	Low	Medium
	centralization	High	High	Medium	High
	formalization	High	High	Medium	High
	matrix-strength	Low	Low	Low	Medium
	communication-prob	0.1	0.1	0.05	0.2
	noise-prob	0.3	0.3	0.01	0.6
	func-except-prob	0.1	0.2	0.05	0.4
	proj-except-prob	0.1	0.2	0.05	0.4
	inst-except-prob	0	0	0.01	0.4
Meeting	priority	High	High	Medium	High
	duration	2 hours	2 hours	0.5 hours	4 hours
Personnel	Culture	Generic	Generic	American	Japanese
	Role	(Various)	(Same)	PM	ST
	App Experience	med	Low	Low	Medium
	Cclt Experience	Medium	Medium	Low	High
	FTE	(Various)	(Same)	0.5 * Default	2 * Default
	Skill Ratings	Medium	Medium	Low	High
Task	Effort	(Various)	(Same)	0.5 * Default	2 * Default
	Learning Days	0	0	0	90
	Priority	Medium	Medium	Low	High
	Requirement Complexity	Medium	High	Medium	High
	Solution Complexity	Medium	High	Medium	High
	Uncertainty	Medium	High	Medium	High
Meeting Assignment		0.1-1.0		0.1	1.0
Task Assignment	Allocation	0.9-1.0	1.0	0.7	1.0

Object	Factor	Hierarchy Industrial Age	Hierarchy 21st Century	FFCSB Exploration	
Successor	TimeLag	0	0	0.0 pct-complete	0.5 pct-complete
Rework	Strength	(Various) 0.15,0.3,1.0	0.1	0.15	0.3

LIST OF REFERENCES

Bettonvil, B., and Kleijnen, J.P.C. (1997). "Searching for important factors in simulation models with many factors: Sequential bifurcation," *European Journal of Operational Research, 96*(1), pp. 180-194.

Center for Edge Power, Naval Postgraduate School. http://www.nps.navy.mil/GSOIS/cep/mission.htm [Last Accessed October 10, 2007].

Cheng, R.C.H. (1997). "Searching for important factors: Sequential bifurcation under uncertainty," *Proceedings of the 1997 Winter Simulation Conference*, pp. 275-280. Available at http://www.informs-sim.org/wsc97papers/0275.PDF [Last Accessed December 5, 2007]

Dorfman, R. (1943). "The detection of defective numbers of large populations," The Annals of Mathematical Statistics 14: pp. 436-440.

Gateau, J.B., Leweling, T.A., Looney, J.P. and Nissen, M.E. (2007). "Hypothesis Testing of Edge Organizations: Modeling the C2 Organization Design Space," *Proceedings International Command & Control Research & Technology Symposium*, Newport, RI (June 2007) Available at http://www.nps.navy.mil/GSOIS/cep/docs/2007/ICCRTS07_Gateau.pdf [Last Accessed October 10, 2007].

Horii, T., Jin, Y. and Levitt, R.E. (2004). "Modeling and Analyzing Cultural Influences on Team Performance through Virtual Experiments," (2004). Available at http://crgp.stanford.edu/publications/conference_papers/TamakiCMOT.pdf [Last Accessed October 23, 2007]

Joint Vision 2010 (1996). Office of the Joint Chiefs of Staff. Available at http://www.dtic.mil/jv2010/jv2010.pdf [Last Accessed November 27, 2007].

Levitt, R.E., Orr, R.J. and Nissen, M.E. (2005). "Validating the Virtual Design Team (VDT) Computational Modeling Environment," *The Collaboratory for Research on Global Projects*, Working Paper Series #25, pp. 1-15. (September 2005) Available at: http://crgp.stanford.edu/publications/working_papers/WP25.pdf [Last Accessed October 10, 2007].

MacKinnon, D.J., Ramsey, M., Levitt, R.E. and Nissen, M.E. (2007). "Hypothesis Testing of Edge Organizations: Empirically Calibrating an Organizational Model for Experimentation," *Proceedings International Command & Control Research & Technology Symposium*, Newport, RI (June 2007). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2007/ICCRTS07_MacKinnon.pdf [Last Accessed October 10, 2007].

Nissen, M.E. (2005). "Hypothesis Testing of Edge Organizations: Specifying Computational C2 Models for Experimentation," *Proceedings International Command & Control Research Symposium,* McLean, VA (June 2005). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2005/Nissen64-ICCRTS2005.pdf [Last Accessed October 10, 2007].

Orr, R.J. and Nissen, M.E. (2006). "Computational Experimentation on C2 Models," *Proceedings International Command and Control Research and Technology Symposium*, Cambridge, UK (September 2006). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2006/Orr-Nissen057-ICCRTS2006.pdf [Last Accessed October 10, 2007].

Ramsey, M. and Levitt, R.E. (2005). "A Computational Framework for Experimentation with Edge Organizations," *Proceedings International Command & Control Research & Technology Symposium*, McLean, VA (June 2005). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2005/Ramsey-Levitt-ICCRTS2005.pdf [Last Accessed October 10, 2007].

Sanchez, S.M., Wan, H. and Lucas, T.W. (2005). "A two-phase screening procedure for simulation experiments," *Proceedings of the 2005 Winter Simulation Conference*, pp. 223-230. Available at http://www.informs-sim.org/wsc05papers/023.pdf [Last Accessed March 23, 2007].

SEED (Simulation Experiments and Efficient Design) Center for Data Farming, Naval Postgraduate School. http://harvest.nps.edu/ [Last Accessed December 5, 2007].

Shen, H., and Wan, H. (2005). "Controlled sequential factorial design for simulation factor screening," *Proceedings of the 2005 Winter Simulation Conference,* pp. 467-474. Available at http://www.informs-cs.org/wsc05papers/052.pdf [Last accessed October 23, 2007]

Wan, H., Ankenman, B. and Nelson, B.L. (2003). "Controlled sequential bifurcation: A new factor-screening method for discrete-event simulation," *Proceedings of the 2003 Winter Simulation Conference,* pp. 565-573. Available at http://www.informs-sim.org/wsc03papers/070.pdf [Last Accessed October 23, 2007]

INITIAL DISTRIBUTION LIST

1. Dudley Knox Library
Naval Postgraduate School
Monterey, California

2. Professor Thomas W. Lucas
Naval Postgraduate School
Monterey, California

3. Professor Susan M. Sanchez
Naval Postgraduate School
Monterey, California

4. Professor Mark E. Nissen
Naval Postgraduate School
Monterey, California

5. Professor Hong Wan
Purdue University
West Lafayette, Indiana

6. Professor Jack Kleijnen
Tilburg University
Tilburg, The Netherlands

7. Col. Eileen A. Bjorkman, USAF
Joint Test and Evaluation Center
Suffolk, Virginia

8. Mr. Timothy D. Beach
Joint Test and Evaluation Center
Suffolk, Virginia

9. Mr. Joseph Ting
DSO National Laboratories
Singapore

10. Mr. Kheng Kang Chan
DSO National Laboratories
Singapore

11. Mr. Kah Chee Yee
    DSO National Laboratories
    Singapore

12. Mr. Chwee Seng Choo
    DSO National Laboratories
    Singapore

13. Regine Pei Tze Oh
    Naval Postgraduate School
    Monterey, California