Statistics and Applications {ISSN 2454-7395(online)}
Volume 19, No. 1, 2021 (New Series), pp 307–317

On Iterative Analysis of Orthogonal Saturated Factorial Designs

Daniel T. Voss
Department of Mathematics and Statistics
Wright State University, Dayton, Ohio, USA

Received: 25 November 2020; Revised: 11 February 2021; Accepted: 18 February 2021

Abstract

Orthogonal saturated factorial designs are useful for screening a few important factors from many. Independent effect estimates can be normalized to have common variance but, with no independent estimate of variability, tests are based on the comparison of larger estimates to smaller ones under an assumption of effect sparsity. Early methods of analysis were proposed by Daniel (1959) and Birnbaum (1959), and subsequent work by Zahn (1975ab) rekindled interest in the problem. They each suggested methods to be applied iteratively, but justifications are generally empirical.

Analytical results establishing control of error rates remain limited. Voss (1988, 1999), Holm, Mark, and Adolfsson (2005), and Voss and Wang (2006a) provided a class of closed step-down tests shown to be of family-wise size-α, utilizing the order statistics of the normalized estimates or the corresponding sums of squares. These are non-iterative tests, utilizing the effect estimates as k order statistics, comparing the ith largest estimate to a critical value based on the distribution of the ith largest of k estimates. However, the step-down tests would be more powerful if conducted iteratively—namely, testing the effect with the ith largest estimate using a critical value based on the largest of i estimates, rather than the ith largest of k estimates. Iterative tests also require the tabulation of fewer critical values.

In this paper, simulations are used to support the conjecture that certain iterative step-down tests for analysis of orthogonal saturated designs do strongly control the family-wise error rate. Some insight is also garnered to guide efforts for an analytical proof.

Key words: Closed test; Family-wise error rate; Iterative testing; Screening experiment.

AMS Subject Classifications: Primary 62K15, 62F07, 62F03; secondary 62F35, 62L10.

1. Introduction

Orthogonal saturated factorial designs are useful for screening a few important factors from many. Such designs yield independent effect estimates that can be normalized to have common variance. Such designs provide no independent estimate of variability, but larger

Correponding Author: Daniel T. Voss
Email: dan.voss@wright.edu

estimates can be compared to smaller estimates, so a standard premise for the analyses is an assumption of *effect sparsity*—namely, that only a few of the effects under study are substantial.

Daniel (1959) introduced the use of half-normal probability plots for the graphical analysis of the normalized effect estimates, and he proposed a corresponding testing procedure. Birnbaum (1959) also proposed tests for the analysis of such designs. They each suggested that their proposed methods be applied iteratively—namely, given k effects and corresponding estimates, if the effect with largest normalized estimate is asserted nonzero, the remaining $k-1$ effect estimates are then analyzed as if said largest estimate was never part of the data, and this process is iterated till the largest remaining estimate is not significantly nonzero. Zahn (1975ab) proposed and evaluated several variations on the iterative analyses of Daniel (1959), including revised test statistics and critical values, comparing methods empirically. His work renewed interest in the analysis of orthogonal saturated factorial designs, but establishing control of error rates remained an open problem.

Many methods of analysis of orthogonal saturated factorial designs have been proposed over the years, but most authors have relied on simulation studies to justify the methods. Progress on analytic justification has been slow, despite long interest in the problem. Voss (1988) proposed a family-wise size-α step-down test of the effects based on the order statistics of the normalized estimates. Subsequently, Voss (1999) provided a rigorous proof that said test strongly controls the family-wise size of the test, characterizing the procedure as a closed, step-down test (see Marcus, Peritz, and Gabriel, 1976), the proof utilizing an obscure stochastic ordering lemma of Alam and Rizvi (1966) and Mahamunulu (1967). Holm, Mark, and Adolfsson (2005) and Voss and Wang (2006a) also provided step-down tests strongly controlling error rates. For reviews of methods of analysis of orthogonal factorial designs, see Hamada and Balakrishnan (1998) and Voss and Wang (2006b).

To make our discussion concrete, consider here the step-down test procedure and statistics utilized by Voss (1988). The test statistics are

$$ss_{(i)}/qmse, \quad i = 1, \ldots, k,$$

for independent, normalized effect estimators $\hat{\theta}_i \sim N(\theta_i, \sigma)$, where $ss_{(i)} = \hat{\theta}^2_{(i)}$ are the corresponding order statistics of the sums of squares, and where $qmse = \sum_{i=1}^{\nu} ss_{(i)}/\nu$ is the *quasi mean squared error* obtained as the average of the ν smallest sums of squares, for specified ν.

For the *step-down test* proposed by Voss (1988), one asserts $\theta_h \neq 0$ if $\hat{\theta}_h = \hat{\theta}_{(j)}$ and $ss_{(i)}/qmse > c(\alpha, i, k)$ for all $i = j, ..., k$. In other words, if $ss_{(15)}/qmse > c(\alpha, 15, 15)$, then the effect corresponding to the largest order statistic $\hat{\theta}_{(15)}$ is asserted to be nonzero and one continues; else one stops. If $ss_{(15)}/qmse > c(\alpha, 15, 15)$ and $ss_{(14)}/qmse > c(\alpha, 14, 15)$, then one also asserts the effect corresponding to the second-largest order statistic to be nonzero and one continues; else one stops. The test procedure continues stepping down in this manner, testing each order statistic's effect in turn starting with the largest and stepping down, continuing as long as an assertion is made.

Voss (1999) showed that the above test procedure is a *closed step-down test* and strongly controls the family-wise error rate to be α if one uses critical value $c(\alpha, i, k)$ such that

$$P(SS_{(i)}/QMSE) > c(\alpha, i, k)) = \alpha, \tag{1}$$

for $SS_{(i)}/QMSE$ (i=1,..., k) the order statistics of the test statistics under the complete null distribution—namely, assuming all effects are zero. A test *strongly controls* the family-wise error rate to be at most α if the probability of any false assertions is at most α for all parameter configurations.

However, many authors have proposed that such step-down tests be conducted iteratively in the following sense. As the step-down test is conducted, if an effect is asserted to be nonzero, then the test proceeds as if that effect where never considered. For example, suppose $k = 15$ effects are initially analyzed. If the effect with largest absolute estimate (*i.e.* corresponding to $ss_{(15)}$) is asserted nonzero, then one proceeds to test the remaining 14 effects as if there never was a 15th effect, and one iteratively steps down in this way until one fails to make an assertion. Such is the case if the ith critical value $c(\alpha, i, k)$ is the upper α quantile of the distribution of $SS_{(i)}/QMSE$, the ith largest of i order statistics (rather than the ith largest of k), under the complete null distribution. Call this variation the *iterative step-down test*, rather than the closed step-down test. Such an iterative approach conjures up a sense of statistical magic since, when testing the effect of any estimate smaller than the largest, one would in essence and reality be ignoring the fact that effects with larger estimates have already been inferred to be nonzero.

There are two advantages to the iterative step-down test. First, as observed by Voss (1988), it is more powerful than the closed step-down test, having smaller critical values after the first. This is not obvious, since the test statistic numerator $SS_{(i)}$ and denominator $QMSE$ are each stochastically larger as a function of i estimators than as a function of k. However, this seems to be born out in practice. For example, Table 1 contains the critical values for the closed and iterative step-down tests for $k = 15$ effects, $\nu = 8$, and $\alpha = 0.10$, 0.05, 0.01, based on 999,999 simulated null samples. The critical values are by definition the same for the largest estimate, but the critical values for the iterative test are substantially smaller for testing all remaining estimates. The second advantage of the iterative test is that fewer critical values need be tabulated, since the iterative test only uses critical values corresponding to the largest order statistics, whereas the closed test requires critical values for each order statistic in a sample of size k.

The focus of this paper is the following conjecture, where *strong control of the error rate* means control under all effect parameter configurations, assuming the standard assumptions of normality and homogeneity of error variances.

Iterative step-down test conjecture. The iterative step-down test, using critical values satisfying equation (1), strongly controls family-wise error rate at the specified level.

A colleague and I have for years sought a rigorous proof of this conjecture, to establish its statistical magic as a happy reality. Alas, each time we 'found' a proof, we subsequently found a hole in it. While we hope someone will succeed where we have thus far failed, it is good to believe that what one hopes to prove is or may well be true. The simulation

Table 1: Critical values for the closed and iterative step-down tests for $k = 15$, $\nu = 8$, and $\alpha = 0.01$, 0.05, 0.10

α	Test Type	c15	c14	c13	c12	c11	c10	c9	c8
0.01	closed	151.2	123.4	98.13	76.31	56.78	39.79	24.66	6.360
	iterative	151.2	84.28	55.25	37.79	25.87	17.35	10.67	5.002
0.05	closed	81.75	67.22	53.93	42.03	31.60	22.42	14.11	5.434
	iterative	81.75	46.80	31.30	21.88	15.40	10.60	6.936	4.054
0.10	closed	60.23	49.51	39.93	31.30	23.65	16.88	10.78	4.907
	iterative	60.23	35.11	23.76	16.83	12.01	8.433	5.712	3.631

results presented in this paper strongly support our belief that the iterative step-down test conjecture is true. A secondary goal is to provide insight that may facilitate analytical proof of the conjecture.

In each of the simulations presented here, the step-down test statistics of Voss (1988) were utilized, but with the sharper critical values $c(\alpha, i, i)$ of the iterative test. Without loss of generality, fix $\sigma = 1$. Thus, pseudo random estimates were generated as $\hat{\theta}_i = \theta_i + \epsilon_i$, for ϵ_i pseudo random $N(0, 1)$, with each effect θ_i as specified, whether zero, a nonzero constant, or $N(0, \sigma = 5)$. In each case, the step-down test was stopped if the statistic $ss_{(\nu)}/qmse$ was significantly large, even though the closed step-down test if continued would still strongly control the error rate. All computations were done using the SAS software.

In the simulations, the following events were of interest. Let IA denote the event of an *incorrect assertion*—namely, that any effect with mean zero is asserted to be nonzero. Let MN denote the event that $ss_{(m)}/qmse > c(\alpha, m, k)$ for $ss_{(m)}$ the sum of squares corresponding to the *maximum nonactive* effect estimate. This condition is necessary but not sufficient for an incorrect assertion, since step-down testing may stop sooner, so $P(IA) \leq P(MN)$. Hence, in search of an analytic proof, showing $P(MN) \leq \alpha$ would establish the conjecture concerning iterative testing. Finally, let $P(A)$ denote the probability of an assertion (correct or not), and $P(CA)$ the probability of a *correct assertion*. Better understanding of the behavior of these probabilities may help someone prove the iterative step-down conjecture.

In Section 2, consider the common case of 15 estimates, corresponding for example to analysis of a 2^{15-11}_{III} fraction. For the setting, we ran an extensive simulation involving 100 distinct randomly chosen effect configurations for each number of active (nonzero) effects from one to seven, plus 100 replications of the null scenario, using family-wise significance level $\alpha = 0.01$. Section 3 contains the results of a similar simulation but using $\alpha = 0.10$. Subsequent sections present simulations with systematically chosen non-null parameter configurations. In Section 4, we consider a small simulation with only five effects and only one active effect. In Section 5, we revisit the common case of 15 estimates, with from zero to eight active effects, but systematically varying the values of the active effects over the values 0,3,6,9. The simulation in Section 6 likewise involves 15 estimates, but with from zero to three active effects with values varying over the values 0.0001, 2, 4, 6, 8. Conclusions are summarized in Section 7.

2. Simulation for 15 Effects, with Zero to Seven Random $N(0, 5)$ Active Effects, with $\alpha = 0.01$

In this section, we present the results of three iterations of an extensive three replicate simulation, including 700 different randomly selected non-null parameter configurations, providing substantial evidence of the conjecture that the iterative step-down test under considerations does strongly control family-wise error rates.

In particular, consider again the case of 15 estimates, here with from zero to seven active effects. For each number of active effects 1,2,. . . ,7, we generated 100 random parameter configurations, where the active effects were independent $N(0, \sigma = 5)$, yielding 700 distinct non-null parameter configurations. Also included were 100 replications of the null parameter configuration, giving 800 cases in total. For each of these 800 cases, 10,000 sets of 15 estimates were generated by adding a $N(0, 1)$ error to each active or null effect, and the iterative step-down test was applied to each of the 10,000 sets of estimates set using $\alpha = 0.01$. In each of the 800 cases, the 10,000 tests were used to estimate the probability of an incorrect assertion, $P(IA)$. This same process was replicated three times, using the same 800 cases or parameter configurations but using different $N(0, 1)$ errors in each replicate, giving three estimates of $P(IA)$ for each of the 800 cases. The results are as follows.

For each of the 800 cases or parameter configurations, given an estimated value of $P(IA)$ from each of the three replicates, the minimum of the three values was saved. Only 11 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.01$, compared to 20 false positives in the 100 null cases. Furthermore, of the 11 non-null cases so flagged, the largest estimate of $\min P(IA)$ was only 0.0110. This simulation strongly supports the truth of the conjecture that iterative application of the step-down test strongly controls the family-wise error rate.

While the above results seem convincing, one might want further evidence that the 11 (of 700) non-null parameter configurations yielding values of $\min P(IA)$ between 0.01 and 0.011 were indeed false positives. To this end, we repeated the above process two more times, using the same 700 non-null parameter configurations each time, but generating different random errors. In the second iteration of the simulation, only 13 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.01$, compared to 17 false positives in the 100 null cases. In the third iteration of the simulation, only 13 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.01$, compared to 15 false positives in the 100 null cases. More importantly, while the three iterations respectively flagged 11, 13 and 13 of 700 non-null parameter configurations as having $\min P(IA) > 0.01$, comparing the results of the three iterations of the three-replicate simulations, none of the 700 non-null parameter configurations where flagged in all three iterations. In other words, of the 700 randomly chosen non-null parameter configurations, there was no non-null parameter configuration for which $\min P(IA)$ exceeded $\alpha = 0.01$ for all three iterations of the simulation.

This simulation concerned the fairly common case of $k = 15$ effects, with $\nu = 8$ effects to form the denominator and a family-wise test size of $\alpha = 0.01$. In this case, the simulation strongly supports the truth of the conjecture that iterative application of the step-down test strongly controls the family-wise error rate.

3. Simulation for 15 Effects, with from Zero to Seven Random $N(0,5)$ Active Effects, with $\alpha = 0.10$

The prior section described three iterations of a simulation for $k = 15$ effects, using $\nu = 8$ effects to form the denominator, and with $\alpha = 0.01$. One may well prefer to use a larger value of α for a screening experiment and family-wise control of test size. In this section, we present the results of one iteration of the simulation presented in the prior section, but with $\alpha = 0.10$ rather than 0.01. As in the prior section, the simulation includes 700 different randomly selected non-null parameter configurations, plus 100 iterations for the null parameter configuration. The results for $\alpha = 0.10$ are as follows.

For each of the 800 cases or parameter configurations, the minimum estimated value of $P(IA)$ was computed over the three replications. Only 5 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.10$, compared to 11 false positives in the 100 null cases. Furthermore, of the 5 non-null cases so flagged, the largest estimate of $\min P(IA)$ was only 0.1027, with the other estimates ranging from 0.1002 to 0.1005.

With more false positives (11 out of 100 null cases) than possible true positives (5 out of 700 non-null cases), and given the small estimates of $\min P(IA)$ in the prospective non-null cases, the results are again encouraging for $\alpha = 0.10$. In short, this simulation also strongly supports the truth of the conjecture that iterative application of the step-down test strongly controls the family-wise error rate.

4. A Small Simulation with Five Effects with One Active

In this Section, we consider a small simulation with only five effects ($k = 5$) and at most one active effect θ, forming *qmse* from the $\nu = 3$ smallest sums of squares. The values considered for the 'active' effects are $\theta = 0, 1, \ldots, 10$, with $\theta = 0$ treated as active but negligible for estimating probabilities. The simulation included 10,000 runs for each value of θ, using $\alpha = 0.10$. Simulation results are provided in Table 2.

Table 2: Simulation for $k = 5$, $\nu = 3$, $\alpha = 0.1$, and one active effect θ

θ	$P(IA)$	$P(CA)$	$P(A)$	$P(MN)$
0	0.089	0.028	0.100	0.093
1	0.074	0.060	0.104	0.081
2	0.068	0.156	0.168	0.078
3	0.077	0.308	0.310	0.087
4	0.094	0.501	0.501	0.099
5	0.096	0.671	0.671	0.097
6	0.097	0.809	0.809	0.098
7	0.099	0.901	0.901	0.099
8	0.100	0.953	0.953	0.100
9	0.097	0.981	0.981	0.097
10	0.100	0.993	0.993	0.100

As noted above, the case with $\theta = 0$ is treated in the simulation as a non-zero but negligibly small effect. Since this is in essence the null case, any assertion is really an incorrect assertion, so it is not surprising that $P(A)$ essentially equals $\alpha = 0.10$. Not surprisingly, the results suggest that going from the null case to having one negligible active effect (approximated by and corresponding to $\theta = 0$) causes $P(IA)$ to drop discretely from 0.10 to 0.08909—the value listed for $P(IA)$ in Table 2. One might then anticipate $P(IA)$ being monotone increasing in θ for $\theta > 0$. Interestingly though, such is not the case. Instead, $P(IA)$ actually decreases as θ goes from 0 (*i.e.* negligible) to 1 to 2, then increases for $\theta > 2$; the slight exception when $\theta = 9$ is probably just simulation error. It is surprising that $P(IA)$ is not monotone in nonzero θ. This lack of monotonicity may be useful in proving the conjecture, if one can prove concavity, for example. It is perhaps also noteworthy that the greatest disparity between $P(IA)$ and $P(MN)$ is when θ is small, *i.e.* about 2 or 3.

Note that the probability of an incorrect assertion, $P(IA)$, is at most 0.10 for all nonzero values of θ, suggesting error rate control, so this simulation supports the iterative testing conjecture.

5. Simulation for 15 Effects, with Seven or Fewer Active Effects with Values 3, 6 or 9, with $\alpha = 0.01$

In this Section, we consider the common case of 15 estimates (*e.g.* corresponding to a 2^{15-11}_{III} fraction), with from zero to seven active effects, systematically varying the values of the last seven effects $\theta_9, \ldots, \theta_{15}$ to have nondecreasing values 0, 3, 6 or 9, yielding 120 distinct parameter configurations. Any effect with value zero is treated as inactive in estimating probabilities.

Simulation results are provided in Table 3 (after references). Only a few of the parameter configurations involving an effect of size three are shown, since $P(IA)$ is well below α in most such cases, including all cases not displayed. A few general observations are in order. The probability of making any assertions, $P(A)$, and the probability of making any correct assertions, $P(CA)$, are largest when there are a few large effects. The probability of an incorrect assertion, $P(IA)$, and the probability of the necessary condition for an incorrect assertion, $P(MN)$, are nearly equal when all active effects are very large, *i.e.* when all active effects are highly likely to be asserted to be nonzero. Most importantly for our purposes, note that the probability of an incorrect assertion, $P(IA)$, never exceeds $\alpha = 0.01$ by more than a negligible amount attributable to simulation error, supporting the conjecture.

6. Simulation for 15 Effects, with Three or Fewer Active Effects with Values 0.0001, 2, 4, 6, or 8, with $\alpha = 0.01$

In the prior section, it was seen that $P(IA)$ and $P(MN)$ are nearly equal when all active effects are very large, and that the test had more power when there were a few large effects. In this section, we examine whether the inclusion of a few small active effects among only a small number of active effects sheds any light on the relative behavior of $P(IA)$ and $P(MN)$, in case this helps in the quest for an analytic proof of the conjecture. In particular, we revisit the common case of 15 estimates, but with from zero to three active effects, systematically varying the values of the last three effects $\theta_{13}, \theta_{14}, \theta_{15}$ to have nondecreasing values 0.0001, 2, 4, 6, or 8. This yields 56 distinct parameter configurations. Any effect with value zero is treated as inactive in estimating probabilities.

Simulation results are provided in Table 4 (after references). To save table space, the results are not shown for any parameter configurations with $\theta_{13} = 2$; these cases all yield $0.003 < P(IA) < 0.007$ so are not interesting.

A few general observations are in order. The probability of making any assertions, $P(A)$, and the probability of making any correct assertions, $P(CA)$, are largest when there are a few large effects. Conversely, consider the three parameter configurations when all active effects are small—namely, when the active effects consist of one, two or three effects of size 0.0001. These are close to the complete null case, when all 15 effects are zero, so it is not surprising that $P(A)$ is approximately α in each case. That said, incorrect assertions are more likely than correct assertions, since there are simply more null effects.

In examining the behavior of $P(IA)$ and $P(MN)$, perhaps the most interesting observations is that these can lack monotonicity in the effects. For example, in the cases with 14 null effects, these probabilities decrease as the lone active effect increases from 0.0001 to 2, but then they increase. It seems obvious that the conjecture should be true if all active effects are either very large or very small. In particular, any very large effects will almost surely correspond to the largest estimates and almost surely be asserted to be nonzero, after which the step-down test will proceed as if they were never in the picture; if so, then large effects shouldn't cause the size of the iterative step-down test to exceed α. Also, the estimates of any very small effects will behave like null effect estimates, except asserting any of them to be active would be correct assertions, so the existence of very small effects may cause $P(IA)$ to be less than α. In view of this, perhaps an analytic proof of the conjecture would follow if one could establish that either $P(IA)$ or $P(MN)$ is concave in each nonzero effect.

7. Concluding Remarks

In the analysis of orthogonal saturated designs, certain closed step-down tests are known to provide strong family-wise control of error rates. Many authors have proposed applying such step-down tests iteratively, but it remains an open problem to establish analytically the conjecture that iterative step-down tests strongly control family-wise error rates. The various simulations presented in this paper strongly support this conjecture. Then may also provide some insight that will be helpful in the search for an analytic proof of the conjecture. In particular, while the probability of making any incorrect assertions is apparently largest in the null case and when all active effects are very large, interestingly, this probability is apparently not monotone in the value of active effects. Simulations suggest that the behavior of the probability of incorrect assertions may be concave, but clearly it is not monotone.

Acknowledgements

It is my pleasure to contribute this paper for inclusion in the special volume to honor the late Professor Aloke Dey. It was my pleasure to have known him—a gentleman and a scholar.

References

Alam, K. and Rizvi, M. H. (1966). Selection from multivariate normal populations. *Annals of the Institute of Statistical Mathematics*, **18**, 307–318.

Birnbaum, A. (1959). On the analysis of factorial experiments without replication. *Technometrics*, **1**, 343–357.

Daniel, C. (1959). Use of half-normal plots in interpreting factorial two-level experiments. *Technometrics*, **1**, 311–341.

Hamada, M. and Balakrishnan, N. (1998). Analyzing unreplicated factorial experiments: a review with some new proposals. *Statistica Sinica*, **8**, 1-–41.

Holm, S., Mark, S. and Adolfsson, T. (2005). A step-down test for effects in unreplicated factorial trials. *Communication in Statistics–Theory and Methods,* **34(2)**, 405–416.

Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. *Annals of Mathematical Statistics*, **38**, 1079–1091.

Marcus, R., Peritz, E. and Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, **63**, 655–660.

Voss, D. T. (1988). Generalized modulus-ratio tests for analysis of factorial designs with zero degrees of freedom for error. *Communication in Statistics–Theory and Methods* **17**, 3345–3359.

Voss, D. T. (1999). Analysis of orthogonal saturated designs. *Journal of Statistical Planning and Inference*, 111–130.

Voss, D. T. and Wang, W. (2006a). On adaptive testing in orthogonal saturated designs. *Statistica Sinica*, **16**, 227–234.

Voss, D. T. and Wang, W. (2006b) Analysis of orthogonal saturated designs. In: Dean A., Lewis S. (eds) *Screening.* Springer, New York, NY.

Zahn, D. A. (1975a). Modifications of and revised critical values for the half-normal plots. *Ph.D. thesis*, Harvard Univ.

Zahn, D. A. (1975b). An empirical study of the half-normal plot. *Technometrics*, **17**, 201–211.

Table 3: Simulation for $k = 15$, $\nu = 8$, $\alpha = 0.01$, and up to seven active effects of size 3, 6 or 9

No. Null	θ_9–θ_{15}	$P(A)$	$P(CA)$	$P(IA)$	$P(MN)$
15	0 0 0 0 0 0 0	0.0102	0.0000	0.0102	0.0102
14	0 0 0 0 0 0 3	0.0791	0.0780	0.0081	0.0083
14	0 0 0 0 0 0 6	0.5852	0.5852	0.0099	0.0099
14	0 0 0 0 0 0 9	0.9477	0.9477	0.0101	0.0101
13	0 0 0 0 0 3 3	0.0891	0.0889	0.0071	0.0076
13	0 0 0 0 0 3 6	0.4899	0.4899	0.0078	0.0081
13	0 0 0 0 0 3 9	0.9102	0.9102	0.0082	0.0084
13	0 0 0 0 0 6 6	0.5986	0.5986	0.0104	0.0104
13	0 0 0 0 0 6 9	0.9098	0.9098	0.0104	0.0104
13	0 0 0 0 0 9 9	0.9478	0.9478	0.0102	0.0102
12	0 0 0 0 3 3 3	0.0726	0.0725	0.0049	0.0055
:	: : : : : : :	:	:	:	:
12	0 0 0 0 3 9 9	0.9043	0.9043	0.0078	0.0082
12	0 0 0 0 6 6 6	0.5386	0.5386	0.0094	0.0095
12	0 0 0 0 6 6 9	0.8515	0.8515	0.0101	0.0101
12	0 0 0 0 6 9 9	0.9031	0.9031	0.0097	0.0097
12	0 0 0 0 9 9 9	0.9258	0.9258	0.0100	0.0100
:	: : : : : : :	:	:	:	:
111	0 0 0 3 9 9 9	0.8569	0.8569	0.0075	0.0080
11	0 0 0 6 6 6 6	0.4408	0.4408	0.0100	0.0102
11	0 0 0 6 6 6 9	0.7567	0.7567	0.0099	0.0099
11	0 0 0 6 6 9 9	0.8243	0.8243	0.0099	0.0099
11	0 0 0 6 9 9 9	0.8531	0.8531	0.0100	0.0100
11	0 0 0 9 9 9 9	0.8730	0.8730	0.0104	0.0104
:	: : : : : : :	:	:	:	:
10	0 0 3 9 9 9 9	0.7623	0.7623	0.0071	0.0078
10	0 0 6 6 6 6 6	0.3137	0.3137	0.0094	0.0096
10	0 0 6 6 6 6 9	0.6143	0.6143	0.0101	0.0104
10	0 0 6 6 6 9 9	0.6960	0.6960	0.0096	0.0097
10	0 0 6 6 9 9 9	0.7321	0.7321	0.0098	0.0099
10	0 0 6 9 9 9 9	0.7557	0.7557	0.0095	0.0096
10	0 0 9 9 9 9 9	0.7721	0.7721	0.0096	0.0096
:	: : : : : : :	:	:	:	:
9	0 3 9 9 9 9 9	0.5901	0.5901	0.0062	0.0075
9	0 6 6 6 6 6 6	0.1789	0.1789	0.0082	0.0097
9	0 6 6 6 6 6 9	0.4178	0.4178	0.0084	0.0095
9	0 6 6 6 6 9 9	0.4958	0.4958	0.0099	0.0104
9	0 6 6 6 9 9 9	0.5361	0.5361	0.0099	0.0102
9	0 6 6 9 9 9 9	0.5658	0.5658	0.0096	0.0098
9	0 6 9 9 9 9 9	0.5814	0.5814	0.0094	0.0094
9	0 9 9 9 9 9 9	0.5989	0.5989	0.0100	0.0101
:	: : : : : : :	:	:	:	:
8	3 9 9 9 9 9 9	0.3261	0.3261	0.0012	0.0070
8	6 6 6 6 6 6 6	0.0614	0.0614	0.0004	0.0094
8	6 6 6 6 6 6 9	0.1839	0.1839	0.0010	0.0104
8	6 6 6 6 6 9 9	0.2340	0.2340	0.0014	0.0101
8	6 6 6 6 9 9 9	0.2611	0.2611	0.0018	0.0100
8	6 6 6 9 9 9 9	0.2839	0.2839	0.0023	0.0100
8	6 6 9 9 9 9 9	0.2995	0.2995	0.0026	0.0096
8	6 9 9 9 9 9 9	0.3051	0.3051	0.0030	0.0101
8	9 9 9 9 9 9 9	0.3173	0.3173	0.0029	0.0098

Table 4: Simulation for $k = 15$, $\nu = 8$, $\alpha = 0.01$, and up to three active effects of size 0.0001, 2, 4, 6, or 8

No. Null	θ_{13}	θ_{14}	θ_{15}	$P(A)$	$P(CA)$	$P(IA)$	$P(MN)$
15	0	0	0	0.0102	0.0000	0.0102	0.0102
14	0	0	0.0001	0.0099	0.0009	0.0094	0.0094
14	0	0	2	0.0257	0.0221	0.0072	0.0074
14	0	0	4	0.1998	0.1995	0.0097	0.0098
14	0	0	6	0.5826	0.5826	0.0100	0.0100
14	0	0	8	0.8767	0.8767	0.0102	0.0102
13	0	0.0001	0.0001	0.0102	0.0020	0.0094	0.0095
13	0	0.0001	2	0.0251	0.0221	0.0066	0.0069
13	0	0.0001	4	0.2007	0.2005	0.0083	0.0085
13	0	0.0001	6	0.5811	0.5811	0.0094	0.0094
13	0	0.0001	8	0.8774	0.8774	0.0094	0.0095
13	0	2	2	0.0293	0.0280	0.0048	0.0052
13	0	2	4	0.1561	0.1560	0.0065	0.0069
13	0	2	6	0.5067	0.5067	0.0070	0.0073
13	0	2	8	0.8309	0.8309	0.0073	0.0075
13	0	4	4	0.2146	0.2146	0.0087	0.0090
13	0	4	6	0.4970	0.4970	0.0096	0.0099
13	0	4	8	0.8180	0.8180	0.0098	0.0100
13	0	6	6	0.5988	0.5988	0.0100	0.0100
13	0	6	8	0.8252	0.8252	0.0101	0.0101
13	0	8	8	0.8831	0.8831	0.0102	0.0102
12	0.0001	0.0001	0.0001	0.0100	0.0030	0.0085	0.0086
12	0.0001	0.0001	2	0.0264	0.0235	0.0069	0.0071
12	0.0001	0.0001	4	0.2014	0.2012	0.0080	0.0082
12	0.0001	0.0001	6	0.5810	0.5810	0.0090	0.0090
12	0.0001	0.0001	8	0.8782	0.8782	0.0089	0.0090
12	0.0001	2	2	0.0280	0.0269	0.0049	0.0052
12	0.0001	2	4	0.1556	0.1555	0.0061	0.0064
12	0.0001	2	6	0.5055	0.5055	0.0063	0.0066
12	0.0001	2	8	0.8288	0.8288	0.0068	0.0071
12	0.0001	4	4	0.2182	0.2182	0.0079	0.0082
12	0.0001	4	6	0.4966	0.4966	0.0091	0.0092
12	0.0001	4	8	0.8151	0.8151	0.0093	0.0096
12	0.0001	6	6	0.5976	0.5976	0.0096	0.0096
12	0.0001	6	8	0.8237	0.8237	0.0095	0.0095
12	0.0001	8	8	0.8840	0.8840	0.0098	0.0098
:	:	:	:	:	:	:	:
12	4	4	4	0.1815	0.1815	0.0078	0.0085
12	4	4	6	0.3991	0.3991	0.0089	0.0092
12	4	4	8	0.7293	0.7293	0.0084	0.0089
12	4	6	6	0.4879	0.4879	0.0087	0.0089
12	4	6	8	0.7386	0.7386	0.0091	0.0093
12	4	8	8	0.8095	0.8095	0.0090	0.0093
12	6	6	6	0.5384	0.5384	0.0100	0.0101
12	6	6	8	0.7428	0.7428	0.0101	0.0101
12	6	8	8	0.8109	0.8109	0.0100	0.0101
12	8	8	8	0.8442	0.8442	0.0102	0.0102