

A Step-Down Lenth Method for Analyzing Unreplicated Factorial Designs

KENNY Q. YE

State University of New York at Stony Brook, Stony Brook, NY 11794-3600

MICHAEL HAMADA

Los Alamos National Lab, Los Alamos, NM 87545

C. F. J. WU

University of Michigan, Ann Arbor, MI 48109-1027

Unreplicated factorial designs are frequently used in industrial experiments. A commonly used method to identify active effects from such experiments is the half-normal plot. Many formal testing methods have been developed to overcome the subjectivity of using this graphical method. Among them, the Lenth (1989) method is simple, yet powerful, as shown by Hamada and Balakrishnan (1998). In this paper, we propose a step-down version of the Lenth method.

It is compared via simulation with the original Lenth method and with stepwise methods proposed by Venter and Steel (1998). It is shown that the step-down Lenth method is better than the original Lenth method and the Venter and Steel step-down method. The Venter and Steel step-up method controlled by the same experimentwise error rate has more power, but it also has a higher individual error rate. Critical values used in the step-down Lenth method are provided.

Introduction

UNREPLICATED factorial designs are very popular in industrial experiments because of their cost efficiency. In most cases, the estimated factorial effects (also called contrasts) $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_I$ are independently distributed with the same standard error if there are no dispersion effects. For information on dispersion effects with unreplicated factorial designs, see McGrath and Lin (2001). With no dispersion effects, and appealing to the central limit theorem, the distributions of the contrasts are assumed to be nor-

mal. Consequently, a half-normal plot of the absolute contrasts is often used to identify the active factorial effects (Daniel (1959)). With half-normal quantiles on the horizontal axis and absolute value of the contrasts on the vertical axis, the active effects are identified as those corresponding to the absolute contrasts falling above the line through the small absolute contrasts thought to be associated with inactive effects. To overcome the subjectivity of using the half-normal plot, many formal testing methods have been proposed. For an extensive review of these methods see Hamada and Balakrishnan (1998). Their simulation study shows that the Lenth method is one of the most powerful and also one of the most simple.

Dr. Ye is an Assistant Professor in the Department of Applied Mathematics and Statistics. His email address is kye@ams.sunysb.edu.

Dr. Hamada is a Technical Staff Member in Statistical Sciences.

Dr. Wu is H. C. Carver Professor in the Department of Statistics and Professor in the Department of Industrial and Operations Engineering.

In evaluating a method, two kinds of error rates are of concern. The first one is the *individual error rate* (IER), i.e., the proportion of inactive individual effects declared active. While the IER is adequate for screening experiments, a more conservative approach is to control the *experimentwise error rate* (EER),

which is the error rate of at least one inactive effect being declared active.

Zahn (1975) first proposed using a step-down method to control the EER. Once the largest absolute factorial contrast is declared active, this method carries out a subsequent test on the remaining contrasts. The procedure stops when the current largest absolute contrast is not declared active. Though not proved mathematically, a simulation study supports the conjecture that this step-down method controls the EER at the specified significance level. Voss (1988) proposes a step-down method which is mathematically proved to control the EER at the specified significance level. His method, however, is too conservative and will not be further discussed in this paper.

Venter and Steel (1998) propose a step-up method which starts with a group of the smallest absolute contrasts. If the effect corresponding to the largest absolute contrast in this group is not declared active, then the next largest absolute contrast is included in the group for the next step. The procedure stops when the largest absolute contrast in the group is declared active. All factorial effects corresponding to the remaining larger absolute contrasts are also declared active. Through a careful and extensive simulation study, they show that step-up methods tend to have higher power than step-down methods. However, more computational power is required to obtain the critical values used in the step-up methods.

A Step-Down Lenth Method

Let $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_I$ denote I mutually orthogonal estimated factorial effects (i.e., the contrasts of a factorial design). Assuming that there are only a few active effects, Lenth (1989) uses a pseudo standard error (PSE) to estimate the standard deviation of $\hat{\theta}_i$:

$$PSE = 1.5 \cdot \text{median}_{|\hat{\theta}_i| < 2.5s_0} |\hat{\theta}_i|, \quad (1)$$

where

$$s_0 = 1.5 \cdot \text{median} |\hat{\theta}_i|.$$

Lenth (1989) then calculates t -like statistics by dividing the $\hat{\theta}_i$ by PSE , which we will refer to as Lenth statistics:

$$t_{\text{Lenth},i} = \frac{\hat{\theta}_i}{PSE}. \quad (2)$$

He suggests using a t -distribution with $I/3$ degrees of freedom as an approximation to the $t_{\text{Lenth},i}$ reference distribution for controlling the IER. This approximation is based on empirical observation rather

than asymptotic theory. More accurate IER critical values are obtained via simulation; see Loughin (1998) and Ye and Hamada (2000).

The EER critical value at significance level α with I contrasts is the $(1 - \alpha) \times 100$ -percentile of the $\max |\hat{\theta}_i|/PSE$ distribution under the null hypothesis $H_0 : \theta_1 = \theta_2 = \dots = \theta_I = 0$. Once the effect corresponding to the largest absolute contrast is declared active, say θ_I , it is natural to test the largest absolute contrast of the remaining $I - 1$ contrasts. The corresponding critical values, however, should be calculated under the null hypothesis $H_0 : \theta_1, \theta_2 = \dots = \theta_{I-1} = 0$. This method can be repeated until the effect corresponding to the largest absolute contrast of the remaining contrasts is not declared active.

Therefore, we propose a step-down version of the Lenth method for controlling EER. Let $|\hat{\theta}|_{(1)} \leq |\hat{\theta}|_{(2)} \leq \dots \leq |\hat{\theta}|_{(I)}$ be the order statistics of I absolute contrasts. Obtain the test statistics

$$t_i = \frac{|\hat{\theta}|_{(i)}}{PSE_i},$$

where PSE_i is the pseudo standard error of $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \dots, \hat{\theta}_{(i)}$, the signed contrasts corresponding to the absolute contrasts $|\hat{\theta}|_{(1)}, |\hat{\theta}|_{(2)}, \dots, |\hat{\theta}|_{(i)}$. Let C_α^i denote the EER critical value at significance level α of the original Lenth method with i contrasts. If $t_i > C_\alpha^i$ for all $i > I - k$, then the largest k factorial effects are declared active. Values of C_α^i for i between 4 and 35 are obtained by simulation and listed in Table 1. The description of the simulation can be found in Ye and Hamada (2000).

The critical values given in Table 1 apply to a number of designs, some of which are listed in Table 2. For example, the 2^k full factorial designs are listed, but the results can be used for 2^{l-p} fractional factorial designs for which $2^{l-p} = 2^k$; e.g., a 2^{4-1} design (a half fraction of a 2^4 design) can estimate seven effects ($I = 7$). Similarly, the results can be used for 3^{l-p} fractional factorial designs. Table 2 also indicates that the results can be used for various Plackett-Burman (PB) designs found in Plackett and Burman (1946). Finally, two mixed-level orthogonal arrays are listed in Table 2, the OA18 and the OA36. The OA18 accommodates one two-level factor and seven three-level factors. Besides 15 main effects, the other two degrees of freedom can be assigned to the interaction between the two-level factor and one of the three-level factors. The OA36 accommodates 11 two-level factors and 12 three-level factors. Both

TABLE 1. Critical Value of Lenth Method for EER α and Number of Contrasts I

α	I										
	4	5	6	7	8	9	10	11	12	13	14
0.001	42.47	40.59	24.28	24.51	18.93	18.25	15.39	14.79	12.99	12.75	11.54
0.002	29.86	28.00	19.06	18.88	14.94	14.51	12.48	12.14	10.78	10.62	9.77
0.003	24.58	22.79	16.36	16.07	13.02	12.64	11.06	10.82	9.64	9.53	8.82
0.004	21.01	19.45	14.66	14.32	11.84	11.47	10.10	9.91	8.92	8.77	8.22
0.005	18.83	17.28	13.50	13.10	11.03	10.61	9.46	9.26	8.43	8.24	7.79
0.006	17.15	15.67	12.60	12.13	10.36	9.95	8.96	8.73	8.03	7.85	7.44
0.007	15.73	14.34	11.90	11.40	9.80	9.45	8.54	8.32	7.70	7.54	7.16
0.008	14.70	13.33	11.30	10.79	9.38	9.01	8.19	7.98	7.44	7.26	6.92
0.009	13.78	12.51	10.78	10.23	9.00	8.63	7.90	7.70	7.21	7.04	6.73
0.010	13.03	11.83	10.33	9.75	8.68	8.32	7.65	7.45	7.00	6.83	6.55
0.020	8.89	8.18	7.77	7.18	6.80	6.47	6.19	5.97	5.78	5.63	5.51
0.030	7.07	6.66	6.52	6.03	5.89	5.57	5.45	5.24	5.16	5.02	4.96
0.040	5.95	5.81	5.74	5.33	5.30	5.01	4.97	4.78	4.75	4.62	4.60
0.050	5.14	5.24	5.17	4.87	4.87	4.62	4.62	4.45	4.45	4.33	4.33
0.060	4.52	4.83	4.73	4.53	4.53	4.33	4.35	4.20	4.22	4.10	4.12
0.070	4.02	4.51	4.37	4.26	4.26	4.10	4.12	4.00	4.03	3.92	3.95
0.080	3.59	4.23	4.07	4.04	4.03	3.92	3.93	3.83	3.86	3.77	3.80
0.090	3.17	4.00	3.82	3.85	3.83	3.75	3.77	3.69	3.72	3.64	3.68
0.100	2.67	3.79	3.59	3.69	3.65	3.62	3.63	3.56	3.60	3.53	3.56
0.200	1.91	2.28	2.23	2.42	2.39	2.59	2.48	2.74	2.72	2.80	2.81
0.300	1.61	1.88	1.90	2.06	2.07	2.17	2.17	2.25	2.25	2.31	2.32
0.400	1.41	1.63	1.68	1.80	1.84	1.93	1.95	2.02	2.04	2.09	2.10

α	I										
	15	16	17	18	19	20	21	22	23	24	25
0.001	11.09	10.39	10.07	9.54	9.35	8.93	8.82	8.44	8.18	8.06	7.97
0.002	9.44	8.93	8.73	8.33	8.17	7.82	7.75	7.50	7.32	7.17	7.14
0.003	8.57	8.14	8.01	7.65	7.53	7.22	7.18	6.95	6.85	6.70	6.65
0.004	8.00	7.63	7.52	7.21	7.10	6.83	6.79	6.59	6.49	6.37	6.33
0.005	7.59	7.25	7.15	6.88	6.77	6.55	6.50	6.31	6.23	6.13	6.08
0.006	7.25	6.96	6.87	6.62	6.52	6.32	6.28	6.10	6.03	5.93	5.89
0.007	6.99	6.73	6.63	6.40	6.31	6.13	6.09	5.93	5.87	5.78	5.73
0.008	6.77	6.52	6.43	6.22	6.14	5.97	5.93	5.79	5.73	5.64	5.59
0.009	6.57	6.35	6.27	6.06	5.99	5.83	5.80	5.66	5.61	5.52	5.47
0.010	6.40	6.20	6.11	5.93	5.86	5.71	5.68	5.55	5.50	5.42	5.37
0.020	5.38	5.28	5.19	5.11	5.05	4.97	4.93	4.87	4.83	4.78	4.74
0.030	4.84	4.79	4.72	4.67	4.61	4.57	4.53	4.49	4.46	4.43	4.40
0.040	4.50	4.47	4.39	4.37	4.32	4.30	4.26	4.24	4.21	4.19	4.16
0.050	4.24	4.23	4.16	4.15	4.11	4.10	4.06	4.05	4.02	4.01	3.98
0.060	4.03	4.04	3.97	3.98	3.94	3.94	3.90	3.90	3.87	3.87	3.84
0.070	3.87	3.89	3.82	3.84	3.79	3.81	3.77	3.77	3.75	3.75	3.73
0.080	3.73	3.75	3.69	3.72	3.67	3.69	3.65	3.66	3.64	3.65	3.63
0.090	3.61	3.64	3.58	3.61	3.57	3.59	3.56	3.57	3.55	3.56	3.54
0.100	3.51	3.54	3.49	3.52	3.48	3.50	3.47	3.48	3.46	3.48	3.46
0.200	2.84	2.85	2.87	2.89	2.89	2.91	2.91	2.93	2.93	2.95	2.94
0.300	2.36	2.37	2.41	2.41	2.44	2.45	2.48	2.48	2.55	2.56	2.61
0.400	2.14	2.16	2.20	2.21	2.24	2.25	2.28	2.29	2.31	2.32	2.34

TABLE 1. Continued

α	I									
	26	27	28	29	30	31	32	33	34	35
0.001	7.75	7.56	7.38	7.35	7.17	7.16	7.03	6.97	6.80	6.72
0.002	6.93	6.84	6.66	6.64	6.47	6.47	6.37	6.31	6.21	6.15
0.003	6.49	6.41	6.25	6.23	6.09	6.10	6.02	5.96	5.89	5.84
0.004	6.19	6.13	6.00	5.97	5.84	5.86	5.77	5.73	5.65	5.62
0.005	5.96	5.91	5.80	5.77	5.66	5.67	5.59	5.54	5.48	5.46
0.006	5.78	5.73	5.64	5.60	5.51	5.52	5.44	5.40	5.34	5.32
0.007	5.63	5.58	5.50	5.47	5.38	5.38	5.32	5.28	5.23	5.21
0.008	5.50	5.46	5.39	5.36	5.27	5.28	5.22	5.18	5.13	5.11
0.009	5.40	5.36	5.29	5.26	5.18	5.18	5.13	5.10	5.05	5.03
0.010	5.31	5.27	5.20	5.18	5.10	5.10	5.05	5.02	4.98	4.95
0.020	4.70	4.68	4.63	4.62	4.58	4.58	4.54	4.52	4.50	4.48
0.030	4.38	4.36	4.32	4.32	4.29	4.29	4.26	4.25	4.23	4.22
0.040	4.15	4.13	4.11	4.11	4.09	4.08	4.06	4.05	4.04	4.03
0.050	3.98	3.97	3.95	3.94	3.93	3.93	3.91	3.90	3.90	3.89
0.060	3.84	3.83	3.82	3.81	3.81	3.80	3.79	3.78	3.78	3.77
0.070	3.73	3.72	3.71	3.70	3.70	3.69	3.69	3.68	3.68	3.67
0.080	3.63	3.62	3.62	3.61	3.61	3.60	3.60	3.60	3.60	3.59
0.090	3.54	3.53	3.54	3.53	3.53	3.52	3.52	3.52	3.52	3.51
0.100	3.47	3.45	3.46	3.45	3.46	3.45	3.46	3.45	3.45	3.45
0.200	2.96	2.96	2.97	2.97	2.99	2.98	3.00	2.99	3.01	3.00
0.300	2.62	2.65	2.65	2.67	2.68	2.70	2.71	2.71	2.73	2.73
0.400	2.35	2.37	2.38	2.40	2.40	2.42	2.43	2.44	2.45	2.46

designs can be found in Wu and Hamada (2000). With the three-level and mixed-level designs, care needs to be taken so that all contrasts have the same standard error and are mutually orthogonal.

An Example

Montgomery (1991) presents an unreplicated experiment using a 2^4 design to study the filtration rate of a pressure vessel. The four factors are temperature (A), pressure (B), concentration of formaldehyde (C), and the stirring rate (D). The design and filtration rate response data are shown in Table 3.

The contrasts for the fifteen factorial effects are listed in descending order (in absolute value) in the second column of Table 4. Using Equation (1), the PSE is 2.6325. The third column of Table 4 lists the absolute Lenth statistics for all fifteen effects using Equation (2). From Table 1, the critical value for the EER at $\alpha = 0.1$ and $I = 15$ is $C_{0.1}^{15} = 3.51$. Using the original Lenth method, the effects A, AC, AD, D, and C are declared active since their Lenth statistics exceed 3.51. For the step-down Lenth method, the

test statistic for A is 8.217, so we declare A as active. In the next step, we recalculate the PSE using the remaining fourteen contrasts. For this example, it is

TABLE 2. Partial List of Designs to Which Step-Down Lenth Method is Applicable

Design	Number of Contrasts (I)
2^3 factorial	7
3^2 factorial	8
12-run PB	11
2^4 factorial	15
OA18	17
20-run PB	19
24-run PB	23
3^3 factorial	26
28-run PB	27
2^5 factorial	31
OA36	35
36-run PB	35

TABLE 3. Design and Response Data,
Filtration Rate Experiment

Run	Factor				Filtration Rate
	A	B	C	D	
1	—	—	—	—	45
2	+	—	—	—	71
3	—	+	—	—	48
4	+	+	—	—	65
5	—	—	+	—	68
6	+	—	+	—	60
7	—	+	+	—	80
8	+	+	+	—	65
1	—	—	—	+	43
2	+	—	—	+	100
3	—	+	—	+	45
4	+	+	—	+	104
5	—	—	+	+	75
6	+	—	+	+	86
7	—	+	+	+	70
8	+	+	+	+	96

the same as the PSE with 15 contrasts. The test statistic for AC is 6.887, which exceeds the critical values $C_{0.1}^{14} = 3.565$. Note that in the original Lenth method the significance of AC is based on the critical value for $I = 15$ instead of for $I = 14$. Continuing with the step-down method, the test statistics AD, D, and C all exceed their corresponding critical values $C_{0.1}^{13}$, $C_{0.1}^{12}$, and $C_{0.1}^{11}$. However, the test statistic

for ABD is less than $C_{0.4}^{10} = 1.952$, so it is not declared active. Therefore, using the step-down Lenth method with the EER less than 0.1, we declare effects A, AC, AD, D, and C as active. In this example, the original Lenth statistics and step-down Lenth statistics happened to be the same for all the effects calculated, because the PSE remained the same in each step. To illustrate that they can be different, we calculate the test statistic for B in the step-down method. The PSE is now 2.4453 which yields a step-down Lenth statistic of 1.280, which is larger than 1.189, the original Lenth statistic. This example also illustrates the robustness of the PSE; it is the same with or without all the active effects, and it is one reason for the success of the original Lenth method.

A Synthetic Example

This example was generated to illustrate the power of the step-down Lenth method. The fifteen contrasts used in the example are randomly generated from fifteen independent normal distributions $N(\theta_i, 1)$, $\hat{\theta}_i, i = 1, 2, \dots, 15$, where $\theta_1 = \theta_2 = \dots = \theta_8 = 0$ are inactive effects and $\theta_9 = 1$, $\theta_{10} = 1.5$, $\theta_{11} = 2$, $\theta_{12} = 3$, $\theta_{13} = 3.5$, $\theta_{14} = 4$, and $\theta_{15} = 5$ are active effects.

Table 5 shows the true effects, the chosen set of contrasts and the corresponding Lenth statistics. Using the original Lenth method, the largest two Lenth statistics are $t_{Lenth,15} = 3.63$ and $t_{Lenth,14} = 2.60$.

TABLE 4. Contrasts and Lenth Statistics, Filtration Rate Experiment

Effect	Contrast	Statistic	
		Original Lenth	Step-Down Lenth
A	21.63	8.217	8.217
AC	-18.13	6.887	6.887
AD	16.63	6.317	6.317
D	14.63	5.557	5.557
C	9.88	3.753	3.753
ABD	4.13	1.569	1.569
B	3.13	1.189	1.280
BCD	-2.63	0.999	N/A
BC	2.38	0.904	N/A
ABC	1.88	0.714	N/A
ACD	-1.63	0.619	N/A
ABCD	1.38	0.524	N/A
CD	-1.13	0.429	N/A
BD	-0.38	0.144	N/A
AB	0.13	0.049	N/A

TABLE 5. Contrasts and Lenth Statistics, Synthetic Experiment

Effect	Contrast	Rank	Statistic	
			Original Lenth	Step-Down Lenth
$\theta_{15} = 5$	5.52	15	3.63	3.63
$\theta_{14} = 4$	3.95	14	2.60	3.96
$\theta_{13} = 3.5$	2.72	12	1.79	3.89
$\theta_{12} = 3$	3.08	13	2.02	4.40
$\theta_{11} = 2$	0.28	4	0.19	NA
$\theta_{10} = 1.5$	2.60	11	1.71	3.98
$\theta_9 = 1$	1.36	8	0.90	NA
$\theta_8 = 0$	-1.47	9	0.97	NA
$\theta_7 = 0$	0.67	7	0.44	NA
$\theta_6 = 0$	-0.25	3	0.16	NA
$\theta_5 = 0$	-2.09	10	1.38	3.21
$\theta_4 = 0$	0.50	6	0.33	NA
$\theta_3 = 0$	-0.02	1	0.01	NA
$\theta_2 = 0$	0.43	5	0.29	NA
$\theta_1 = 0$	0.22	2	0.14	NA

In Table 1, we observe that $C_{0.1}^{15} = 3.51$. Therefore θ_{15} is the only contrast to be declared active at significance level 0.1. Using the step-down Lenth method, first we obtain $t_{15} = 3.63 > C_{0.1}^{15} = 3.51$; effect θ_{15} is declared active. Next, excluding the largest contrast which was just declared active, we calculate the Lenth statistic using the remaining 14 contrasts and obtain $t_{14} = 3.96 > C_{0.1}^{14} = 3.56$. Consequently, effect θ_{14} is declared active. We continue this procedure and declare three additional effects active, θ_{12} , θ_{13} , and θ_{10} , whose contrasts are the next three largest in the order given. The procedure stops after we see that $t_{10} = 3.21 < C_{0.1}^{10} = 3.63$. Overall, five out of the seven active effects are declared active while none of the inactive effects are declared active. The half-normal plot for this example is shown in Figure 1. Since there is no large difference between the absolute contrasts, using the half-normal plot might not detect any effects.

This example shows that when the magnitudes of the active effects vary from small to large, both the Lenth method and half-normal plots may fail to detect them. The simulation study presented in the next section confirms this observation.

A Simulation Study

To evaluate the step-down Lenth method, we performed a simulation study to compare four methods, which are the step-down Lenth method, the

original Lenth method, the fixed RMS scaling step-down method, and the sequential RMS scaling step-up method. The last two methods were proposed and studied by Venter and Steel (1998).

The test statistics used by Venter and Steel (1998) are similar to our step-down Lenth statistics, except that the root mean square (RMS) $([1/i] \sum_{k=1}^i \theta_{(k)}^2)^{1/2}$ replaces PSE. The fixed RMS scaling method uses the same set of small absolute contrasts to calculate the RMS at each step. The number of contrasts used to calculate the RMS is l , a lower bound on the number of inactive effects, which is determined by the user. Sequential RMS scaling uses a different set of contrasts to calculate RMS at each step. In step-up methods, once a contrast is declared inactive it is used in computing RMS at the next step. However, the user has to choose the number of contrasts used in the first step. The sequential RMS scaling step-down method and fixed RMS scaling step-up method were also studied by Venter and Steel (1998). Since they were shown to be inferior, we do not include them in our study.

Our simulation study considers the case of 15 contrasts under the following six configurations:

- C1: $\theta_1 = \dots = \theta_{14} = 0, \theta_{15} = \Delta$
- C2: $\theta_1 = \dots = \theta_{12} = 0, \theta_{13} = \theta_{14} = \theta_{15} = \Delta$
- C3: $\theta_1 = \dots = \theta_{10} = 0, \theta_{11} = \dots = \theta_{15} = \Delta$
- C4: $\theta_1 = \dots = \theta_8 = 0, \theta_9 = \dots = \theta_{15} = \Delta$

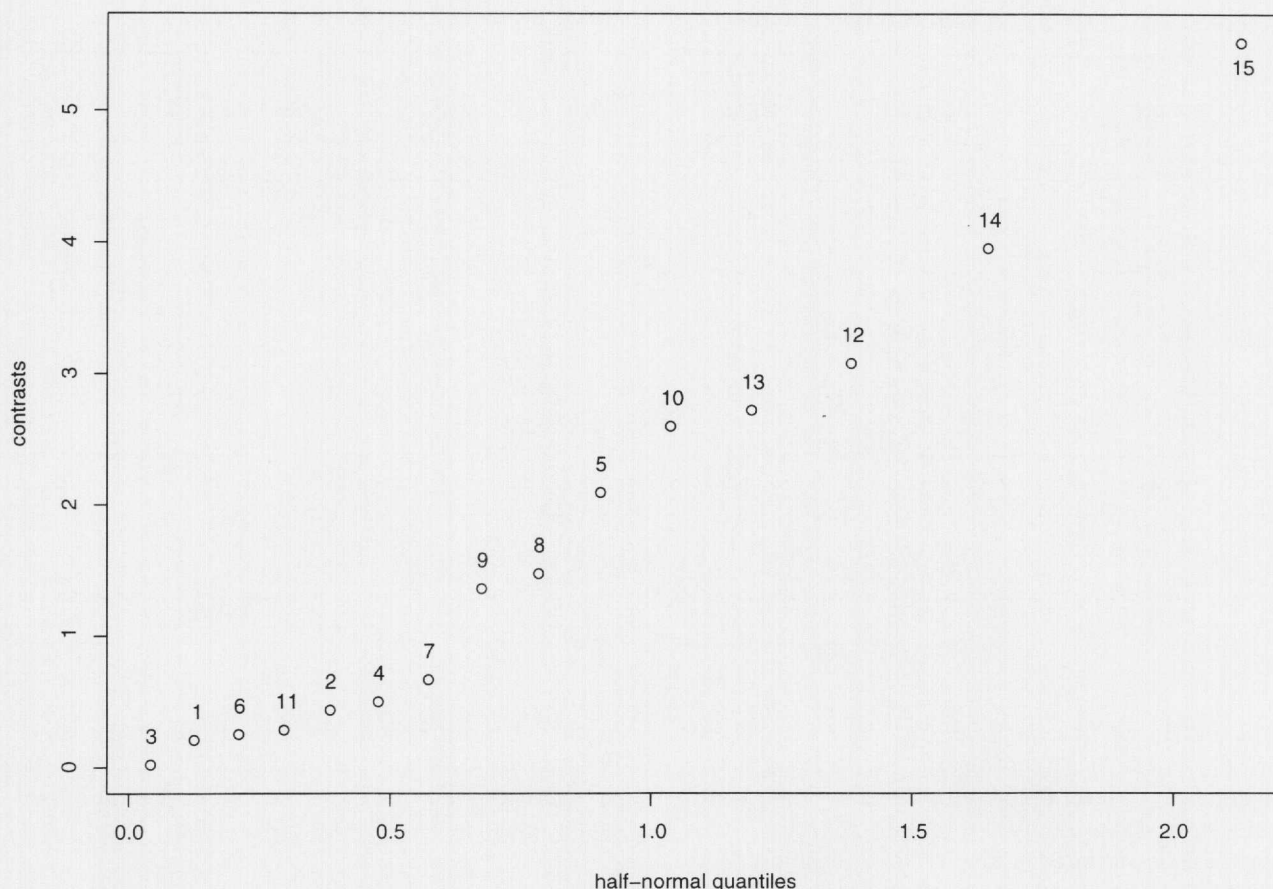


FIGURE 1. Half-Normal Plot, Synthetic Example.

C5: $\theta_1 = \dots = \theta_{12} = 0$, $\theta_{13} = \Delta$, $\theta_{14} = 2\Delta$, $\theta_{15} = 3\Delta$

C6: $\theta_1 = \dots = \theta_{10} = 0$, $\theta_{11} = \Delta$, $\theta_{12} = 2\Delta$, $\theta_{13} = 3\Delta$, $\theta_{14} = 4\Delta$, $\theta_{15} = 5\Delta$.

Here, $\Delta > 0$ is referred to as the "spacing". Three figures of merit are calculated to compare the four methods:

1. EER, the fraction of experiments in which at least one inactive effect is declared active.
2. Power, the expected fraction of active effects that are declared active.
3. IER, the expected fraction of inactive effects that are declared active.

For each configuration, the value of Δ is varied according to $0(0.5)8$. The EER, power, and IER at significance level $\alpha = 0.1$ for each of the four methods are determined based on 10,000 sets of randomly

generated $\hat{\theta}_i$'s; the behavior of these four methods is about the same at other significance levels. The critical values for RMS scaling stepwise methods were obtained from an ftp site (Vender and Steel (1999)). For the RMS scaling stepwise methods, we first used $l = 7$ as the lower bound on the number of inactive effects.

Figure 2 shows the EER of the four methods. It can be seen that the EER of the three stepwise methods is closer to the nominal level 0.1 than the original Lenth method. The three stepwise methods have comparable performance except for C4, where the step-down Lenth method is more conservative. It is to be expected that the two RMS methods for C4 have EER's closer to the nominal EER, because the number of contrasts used to calculate RMS is $l = 7$, which is close to the actual number of inactive effects (i.e., eight). While there is nothing wrong with being conservative, a likely consequence is the loss of power.

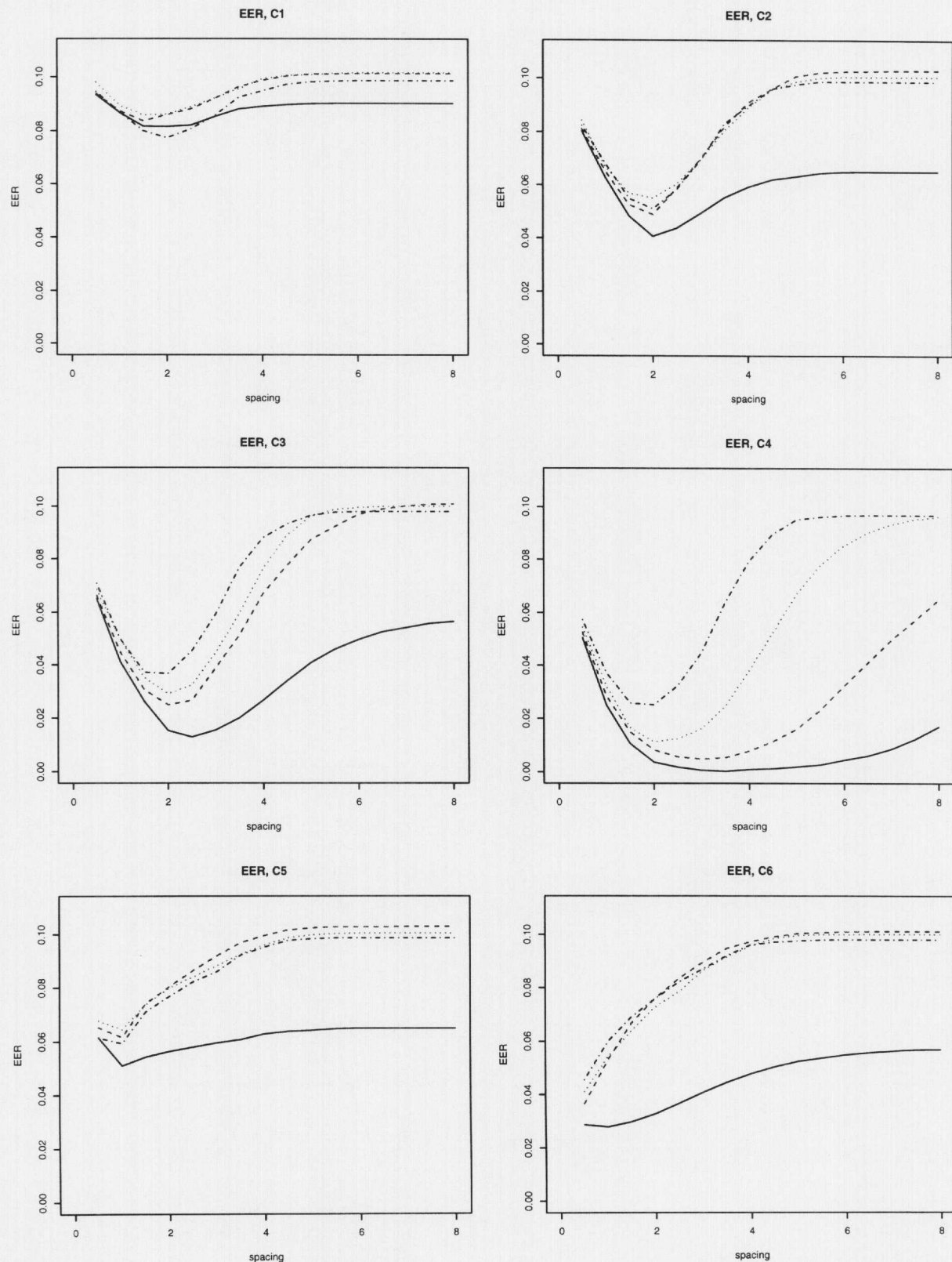


FIGURE 2. EER of Lenth, —; Step-Down Lenth, - - -; RMS Step-Down ($l = 7$),; RMS Step-Up ($l = 7$), - . - .

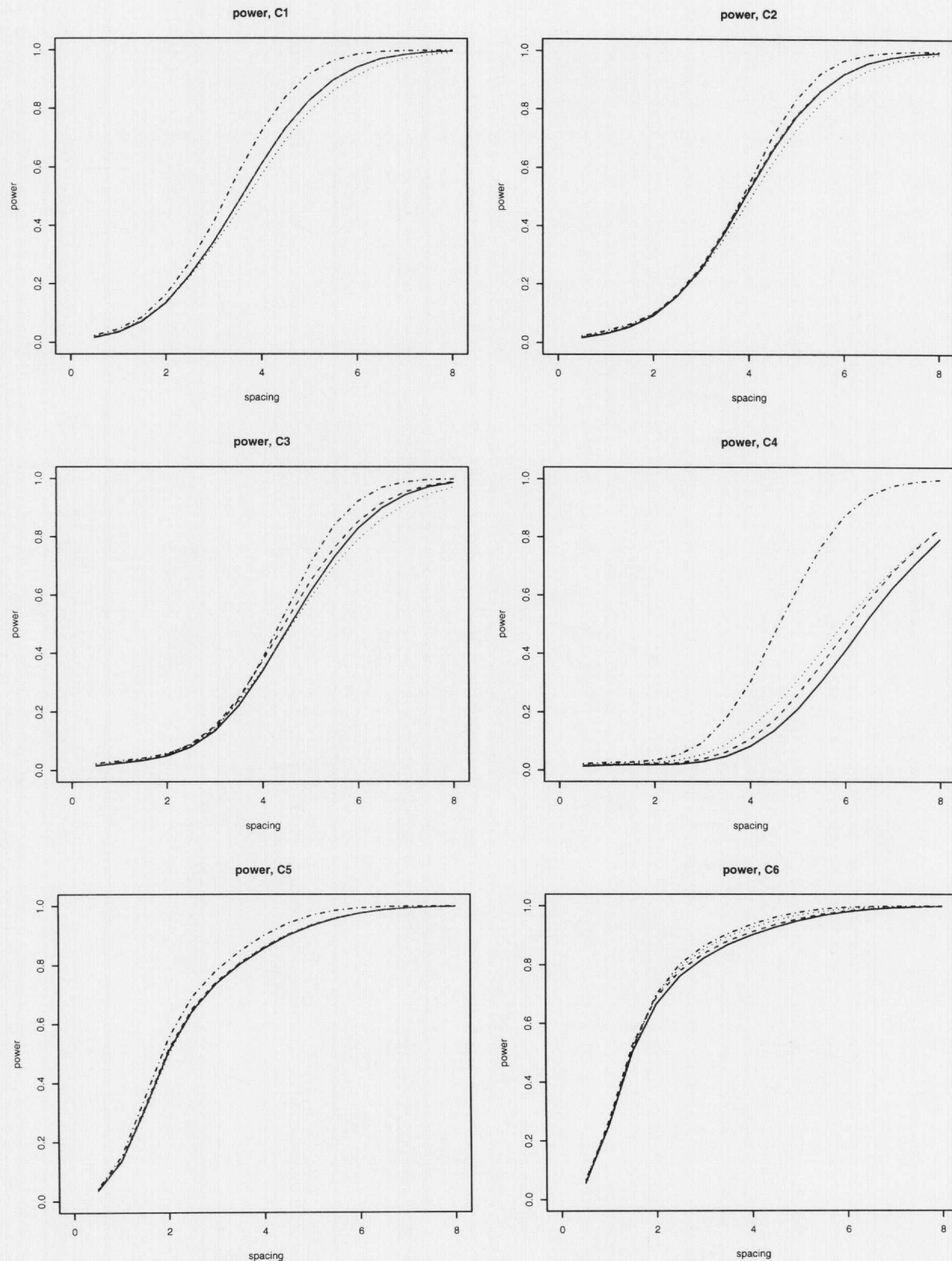


FIGURE 3. Power of Lenth, —; Step-Down Lenth, - - -; RMS Step-Down ($l = 7$),; RMS Step-Up ($l = 7$), - . - . .

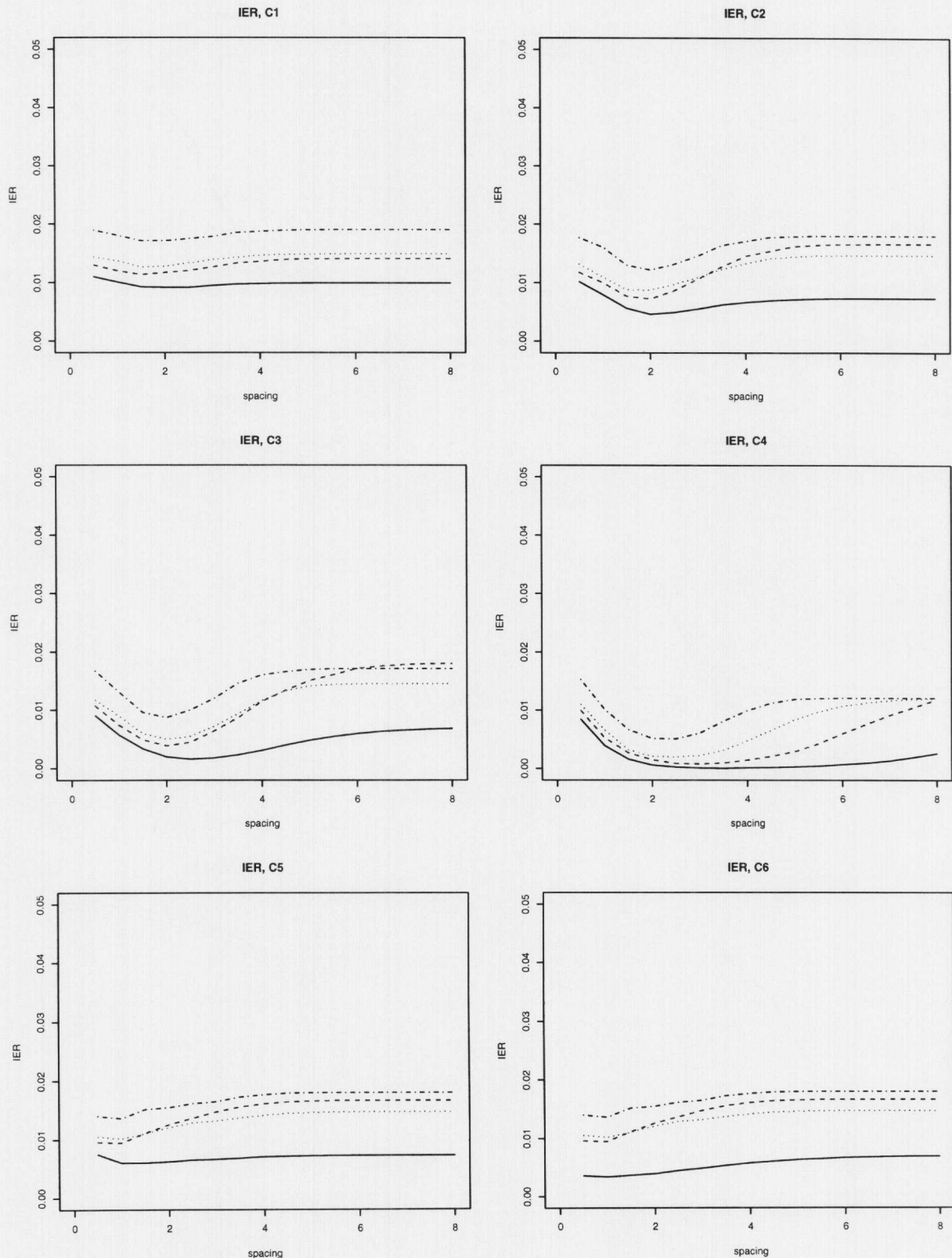


FIGURE 4. IER of Lenth, —; Step-Down Lenth, — —; RMS Step-Down ($l = 7$),; RMS Step-Up ($l = 7$), — · — ·.

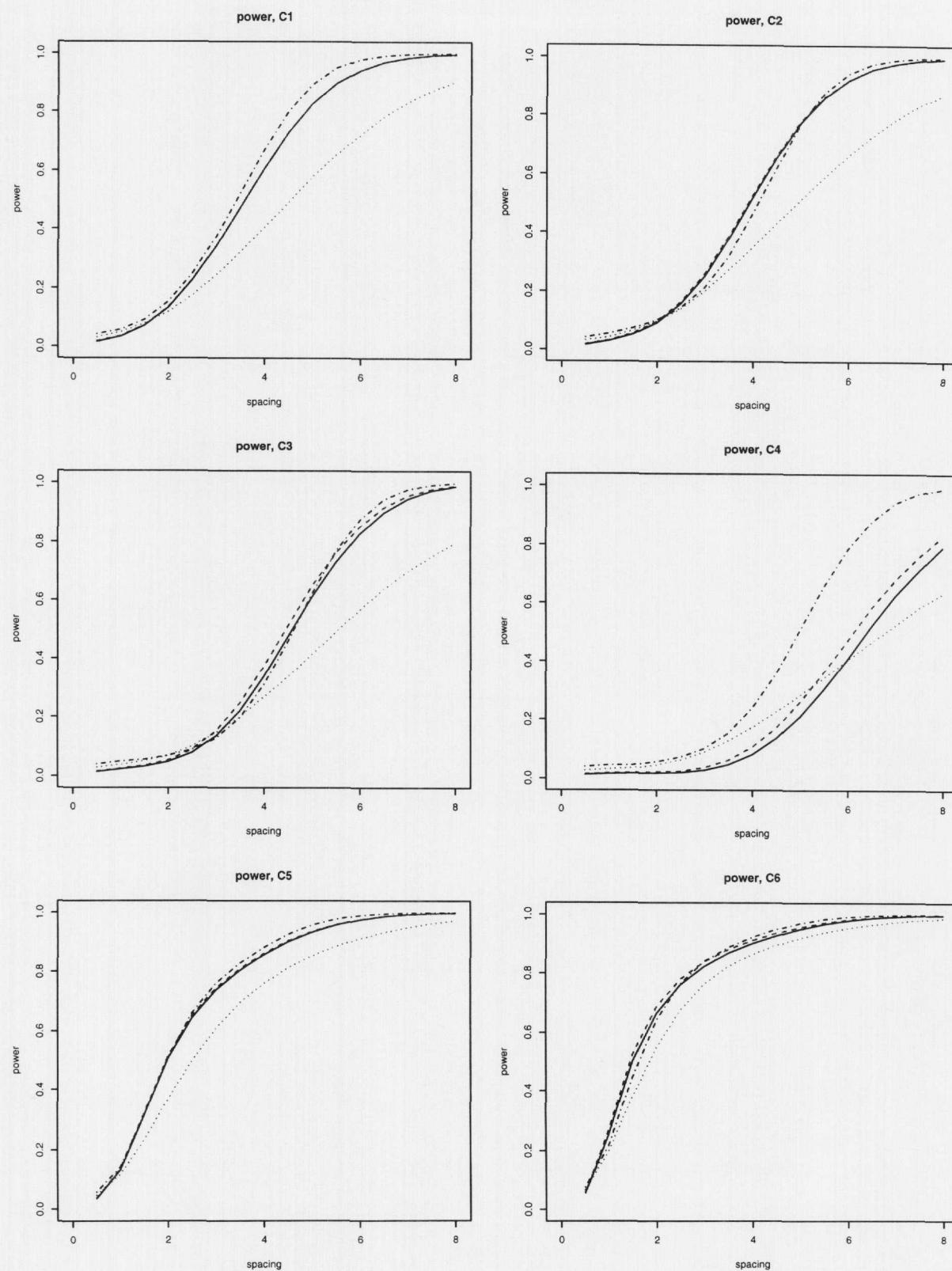


FIGURE 5. Power of Lenth, —; Step-Down Lenth, - - -; RMS Step-Down ($l = 4$),; RMS Step-Up($l = 4$), - . - .

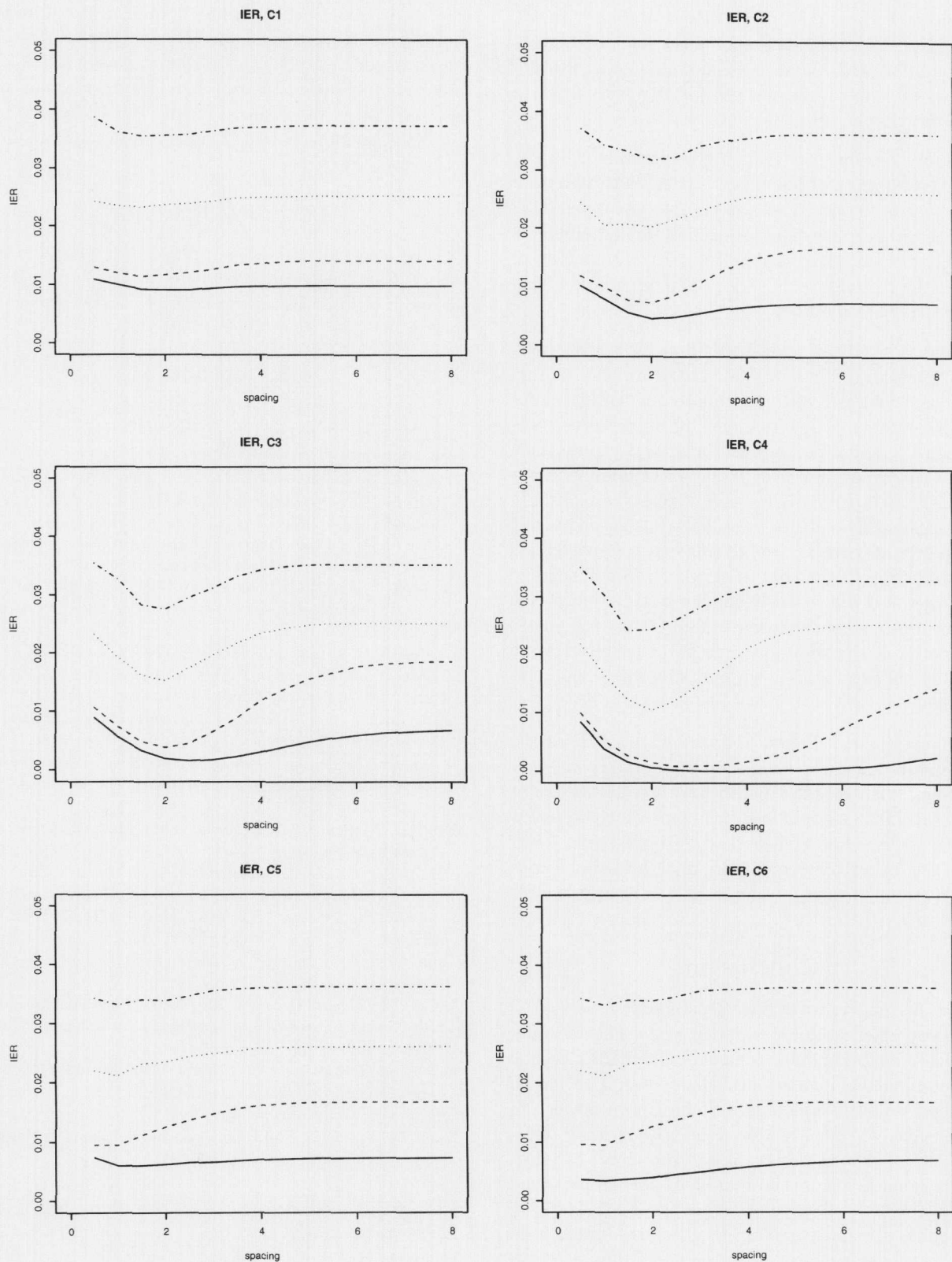


FIGURE 6. IER of Lenth, —; Step-Down Lenth, - - -; RMS Step-Down ($l = 4$),; RMS Step-Up ($l = 4$), - . - .

Figure 3 shows the power of the different methods. Except for C4, the power of the four methods is very similar, with the step-up methods being slightly better. In every case, the step-down Lenth method outperforms the original Lenth method.

Figure 4 shows the IER of the methods. Clearly, the step-up method has much higher IER rates than the other methods. A reason for this phenomenon is that the step-up method declares all the larger effects as active if an effect is declared active at one step. Thus, the method tends to declare more active effects than the other methods.

A major advantage of the proposed step-down Lenth method over the RMS stepwise methods is that users do not have to choose the lower bound on the number of inactive effects. In practice, it is often very hard for an experimenter to guess how many effects are inactive. If the lower bound l is set too large to include some active effects, then not only will those active effects be declared as inactive, but the power of correctly identifying other active effects will be reduced; the contaminated RMS will tend to be larger and will result in smaller test statistics. On the other hand, if the lower bound l is set too small, the power of the step-down method is reduced while the IER of both step-down and step-up versions increases.

Figures 5 and 6 show the power and IER of the RMS step-down and RMS step-up methods with $l = 4$, along with original Lenth and step-down Lenth methods. It is evident from the plot that the power of the RMS step-down method is significantly lower than the other three methods, and the IER of both RMS stepwise methods for $l = 4$ is nearly double that for $l = 7$.

Conclusions

In this paper, we propose a step-down version of the Lenth method for identifying active effects in unreplicated experiments. It controls the EER closer to the nominal significance level than the original Lenth method does. Consequently, it has more power. In particular, when there are a moderate number of active effects whose values range from small to large, the original Lenth method and half-normal plot tend to miss them. Compared with the RMS stepwise methods, the proposed method spares users from

having to choose l , the lower bound on the number of inactive effects, a difficult decision even for experienced experimenters. The performances of the RMS stepwise methods and step-down Lenth method are comparable in most cases. While the RMS step-up method has more power, it suffers from high individual error rates.

Acknowledgments

The authors thank two referees for helpful comments that led to improvements in the paper. The research of Mike Hamada and C. F. J. Wu was supported by NSF grant DMS 97-04649.

References

- DANIEL, C. (1959) "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments". *Technometrics* 1, pp. 311–340.
- HAMADA, M. and BALAKRISHNAN, N. (1998). "Analyzing Unreplicated Factorial Experiments: A Review With Some New Proposals (with discussion)". *Statistica Sinica* 8, pp. 1–41.
- LENTH, R. V. (1989). "Quick and Easy Analysis of Unreplicated Factorials". *Technometrics* 31, pp. 469–473.
- LOUGHIN, T. M. (1998) "Calibration of the Lenth Test for Unreplicated Factorial Designs". *Journal of Quality Technology* 30, pp. 171–175.
- MCGRATH, R. N. AND LIN, D. K. J. (2001). "Confounding of Location and Dispersion Effects in Unreplicated Fractional Factorials". *Journal of Quality Technology* 33, pp. 129–139.
- MONTGOMERY, D. C. (1991). *Design and Analysis of Experiments*, 3rd ed. John Wiley and Sons, Inc., New York, NY.
- PLACKETT, R. L. and BURMAN, J. P. (1946). "The Design of Optimum Multifactorial Experiments". *Biometrika* 33, pp. 305–325.
- VENTER, J. H. and STEEL, S. J. (1998). "Identifying Active Contrasts by Stepwise Testing". *Technometrics* 40, pp. 304–313.
- VENTER, J. H. and STEEL, S. J. ftp.sun.ac.za/local/statistics.
- VOSS, D. T. (1988). "Generalized Modulus-Ratio Tests for Analysis of Factorial Designs With Zero Degrees of Freedom for Error". *Communications in Statistics—Theory and Methods* 17, pp. 3345–3359.
- WU, C. F. J. and HAMADA, M. (2000). *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley and Sons, Inc., New York, NY.
- YE, K. and HAMADA, M. (2000). "Critical Values of the Lenth Method for Unreplicated Factorial Designs". *Journal of Quality Technology* 32, pp. 57–66.
- ZAHN, D. A. (1975). "An Empirical Study of the Half-Normal Plot". *Technometrics* 17, pp. 201–211.

Key Words: *Experimentwise Error Rate, Individual Error Rate, Power.*