

## Research Article

# A Confidence Region for Zero-Gradient Solutions for Robust Parameter Design Experiments

Aili Cheng,<sup>1</sup> John Peterson,<sup>2</sup> and Pallavi Chitturi<sup>1</sup>

<sup>1</sup>Department of Statistics, Temple University, 1810 North 13th Street, Philadelphia, PA 19122, USA

<sup>2</sup>Quantitative Sciences, GlaxoSmithKline Pharmaceuticals, 1250 South Collegeville Road, Collegeville, PA 19426, USA

Correspondence should be addressed to John Peterson, john.peterson@gsk.com

Received 11 March 2011; Accepted 28 June 2011

Academic Editor: Myong (MK) Jeong

Copyright © 2011 Aili Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the key issues in robust parameter design is to configure the controllable factors to minimize the variance due to noise variables. However, it can sometimes happen that the number of control variables is greater than the number of noise variables. When this occurs, two important situations arise. One is that the variance due to noise variables can be brought down to zero. The second is that multiple optimal control variable settings become available to the experimenter. A simultaneous confidence region for such a locus of points not only provides a region of uncertainty about such a solution, but also provides a statistical test of whether or not such points lie within the region of experimentation or a feasible region of operation. However, this situation requires a confidence region for the multiple-solution factor levels that provides proper simultaneous coverage. This requirement has not been previously recognized in the literature. In the case where the number of control variables is greater than the number of noise variables, we show how to construct critical values needed to maintain the simultaneous coverage rate. Two examples are provided as a demonstration of the practical need to adjust the critical values for simultaneous coverage.

## 1. Introduction

Robust Parameter Design (RPD) is also called Robust Design or Parameter Design in the literature [1, 2]. The concept of RPD was introduced in the United States by Genichi Taguchi in the early 1980s. It is a methodology that takes both the mean and variance into consideration for product or process optimization. Taguchi [3] divided the predictor variables into two categories: control variables and noise variables. Control variables are easy to control while noise variables are either difficult to control or uncontrollable at a large scale. In practice, we would like to find a range of control variables such that (1) the variance caused by the change of noise variables is minimized and (2) the mean response is close to target. Multiple optimization design and analysis methods have been developed to achieve these two goals simultaneously, ranging from the traditional Taguchi methods to the more sophisticated response surface alternatives. (See [2, 4, 5] for detailed reviews on these methods.)

Under certain conditions, Myers et al. [6] proposed a way to construct a confidence region of the control variables where the variability transmission by the noise variables is minimized to zero. Although they only focused on the variance part, there are many such situations in which focus is placed entirely on the process variance (see p506 in [5]). For instance, if the process mean can be optimized using certain control factors that do not interact with noise variables or impact the noise variance (i.e., “tuning factors”), then one can seek other control factors which can drive the noise variance to zero. Even if the process mean cannot be driven to the target by tuning factors alone, it is nonetheless illuminating to consider both the confidence region for the minimum process variance and the confidence region for the optimal process mean [5, 6]. If the number of noise variables is not greater than the number of control factors (for examples, see [7–9] and see pages 491–492, 499 in [5]), then the conditions for a minimum process variance and a zero-gradient solution will coincide.

Furthermore, a zero-gradient solution is quite useful in that, even for very large noise variation, the transmission of that noise to the output of the process can be made negligible (or very small) by utilization of zero-gradient (or near-zero-gradient) operating conditions. Therefore, it is desirable to have a way to statistically test for the existence of such a solution within the experimental region or a region of feasible operation (see p506 in [5]). A simultaneous confidence region for the locus of points forming a zero-gradient solution forms such a test and also provides a graphical measure of uncertainty about the zero-gradient solution.

The existence of a zero-gradient solution is especially interesting for mixture experiments with control factor ingredients and noisy process variables. If a zero-gradient solution for the ingredient mixture can be found, that is, within the experimental region, and the confidence region for the zero-gradient solution intersects one or more of the mixture simplex boundaries, then this implies that it may be possible to remove one or more mixture ingredients and still maintain very low noise variance. Removal of one or more mixture ingredients may help to reduce production cost [10].

Myers et al. [6] proposed a confidence region in control variables based upon the standard response surface model as shown in (1) for incorporating noise variables (see, for examples, [4, 11, 12])

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\boldsymbol{\gamma} + \mathbf{x}'\boldsymbol{\Delta}\mathbf{z} + \varepsilon, \quad (1)$$

where  $\mathbf{x}$  is a vector of  $k$  control variables,  $\mathbf{z}$  is a vector of  $h$  noise factors,  $\beta_0$  is the intercept,  $\boldsymbol{\beta}$  is a  $k \times 1$  vector of coefficients for the main effects of control variables,  $\boldsymbol{\gamma}$  is a  $h \times 1$  vector of coefficients for the main effects of noise variables,  $\mathbf{B}$  is a  $k \times k$  matrix whose diagonals are the coefficients for the quadratic terms of control variables and whose off-diagonals are one-half of the control variable interaction effects, and  $\boldsymbol{\Delta}$  is a  $k \times h$  matrix of control by noise variable interaction effects.  $\varepsilon$  is the random error term. It is assumed here that  $\varepsilon \sim N(0, \sigma^2)$ .

Assuming that noise variables have mean zero and variance-covariance matrix  $\mathbf{V}_z$ , the variance of the response in Model (1) is

$$\sigma^2[y(\mathbf{x}, \mathbf{z})] = (\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x})'\mathbf{V}_z(\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}) + \sigma^2. \quad (2)$$

Here the variance of the response is divided into two parts: the variance transmitted by the noise variables (represented by the first term) and the constant variance ( $\sigma^2$ ) due to modeling error and other factors not considered in the model. In other words, it is the noise variables that lead to the variance heterogeneity in the response. Since the changes of noise variables are inevitable in practice, Myers et al. [6] proposed that the “minimum process variance” can be reached by setting the slope of the noise variables,  $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}$ , equal to zero and, therefore, eliminating the noise variance part from the response variance. A confidence region for such control variable values can be constructed by inverting a hypothesis test of the form:

$$H_0 : \boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x} = \mathbf{0}, \quad (3)$$

for each  $\mathbf{x}$ -point. To simplify the notation, let  $\boldsymbol{\psi}$  be a  $p \times 1$  vector that contains all the elements of the noise variable's main effect vector  $\boldsymbol{\gamma}$  and the interaction matrix  $\boldsymbol{\Delta}$ , that is,  $\boldsymbol{\psi} = (\gamma_1, \delta_{11}, \delta_{12}, \dots, \delta_{1k}, \gamma_2, \delta_{21}, \delta_{22}, \dots, \delta_{2k}, \dots, \gamma_h, \delta_{h1}, \delta_{h2}, \dots, \delta_{hk})'$ , where  $\gamma_i, i = 1, 2, 3, \dots, h$ , is the  $i$ th element of vector  $\boldsymbol{\gamma}$  and  $\delta_{ij}, i = 1, 2, 3, \dots, h, j = 1, 2, 3, \dots, k$ , is the element of matrix  $\boldsymbol{\Delta}'$  in the  $i$ th row and  $j$ th column. Let  $\mathbf{M}(\mathbf{x}) = [\mathbf{I}_h \otimes (1, \mathbf{x}')]'$ , where  $\mathbf{I}_h$  is an identity matrix of dimension  $h$ ,  $\mathbf{x}$  is a  $k \times 1$  vector of the corresponding control variables that interact with noise variables, and  $\otimes$  denotes the Kronecker product. The null hypothesis in (3) can then be written as

$$H_0 : \mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}, \quad (4)$$

where  $\mathbf{M}(\mathbf{x})$  is a  $h \times p$  matrix and  $\mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}$ . Now the  $1 - \alpha$  confidence region in control variables for zero variance due to noise variables can be defined as

$$\{\mathbf{x} : \mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0} \text{ is not rejected at level } \alpha\}. \quad (5)$$

Furthermore, let  $Q(\mathbf{x}; \hat{\boldsymbol{\psi}})$  denote the test statistic for  $H_0$  which is

$$(\mathbf{M}(\mathbf{x})\hat{\boldsymbol{\psi}})' [\mathbf{M}(\mathbf{x})\hat{\mathbf{V}}_{\hat{\boldsymbol{\psi}}}\mathbf{M}(\mathbf{x})']^{-1} (\mathbf{M}(\mathbf{x})\hat{\boldsymbol{\psi}}), \quad (6)$$

where  $\hat{\boldsymbol{\psi}}$  is the estimate of  $\boldsymbol{\psi}$ ;  $\hat{\mathbf{V}}_{\hat{\boldsymbol{\psi}}}$  is the usual unbiased estimate of the variance of  $\hat{\boldsymbol{\psi}}$ ,  $\mathbf{V}_{\hat{\boldsymbol{\psi}}}$ .

Let  $\hat{\boldsymbol{\Psi}} \sim N(\boldsymbol{\psi}, \mathbf{V}_{\hat{\boldsymbol{\psi}}})$ , Myers et al. [6] have shown that, for each fixed  $\mathbf{x}$ ,  $Q(\mathbf{x}; \hat{\boldsymbol{\Psi}}) \sim hF(h, \nu)$ , where  $\hat{\boldsymbol{\Psi}}$  is a vector of random variables (i.e.,  $\hat{\boldsymbol{\Psi}}$  is an estimator of  $\boldsymbol{\psi}$ , whereas  $\hat{\boldsymbol{\psi}}$  is an estimate from the actual data),  $\nu$  is the residual degrees of freedom (df), and  $F(h, \nu)$  is  $F$  distribution with numerator df equal to  $h$  and denominator df equal to  $\nu$ . They then conclude that the  $100(1 - \alpha)$  percent confidence region in (5) is

$$C_{\mathbf{x}}(c_{\alpha}) = \{\mathbf{x} : Q(\mathbf{x}; \hat{\boldsymbol{\psi}}) \leq c_{\alpha}\}, \quad (7)$$

where the critical value  $c_{\alpha} = hF(1 - \alpha, h, \nu)$ .

Note that the confidence region in (7) (called the “MKG confidence region” from now on) and the critical value  $c_{\alpha} = hF(1 - \alpha, h, \nu)$  were derived based on two critical assumptions: (1) the minimum of the variance due to noise variables is zero; (2) the solution to the zero-gradient equation in (4) is unique. There are situations where the first assumption cannot be met due to the fact that the solution to (4) is either outside the experimental region or does not exist. (However, the approach proposed in this article may provide the determination of such existence in a statistically significant way.) Assuming that the first assumption is met (like the two examples in Section 3), the second assumption is only true when  $h \geq k$ . Notice that (4) represents a series of  $h$  equations with  $k$  unknown control variables. As recognized by Myers et al. (see page 506 in [5]), the equation will result in a single point solution when  $h = k$ , a line or hyperplane when  $h < k$ , and a single point solution or no solution when  $h > k$ . In other words, the MKG confidence region provides the correct critical value for the zero-gradient solution (if it exists) only when there are at least as many noise variables as control variables.

However, when the number of noise variables is less than the number of control variables, multiple solutions can exist to the zero-gradient equation in (4). In such situations, use of the MKG region will provide below nominal simultaneous coverage. As such, a confidence region which covers all the solutions simultaneously needs to be developed. In practice, statistical inference for the multiple-solution problems is important as this gives the experimenter more options with regard to finding the zero-gradient factor settings. The objective of this paper is to generalize the MKG confidence region such that it will provide the adequate coverage for both the single-solution case (where  $h \geq k$ ) and multiple-solution case (where  $h < k$ ).

The rest of this paper will be organized as follows. Section 2 will focus on the derivation of such a generalized confidence region and the corresponding critical value required for inverting the associated null hypothesis. In Section 3, we give two examples to demonstrate the difference in simultaneous coverage between the MKG confidence region and our proposed confidence region when the number of control factors exceeds the number of noise variables. Section 4 provides a summary of the results.

## 2. A Generalized Confidence Region Approach

**2.1. The Multiple Zero-Gradient Solution Problem.** To address the multiple solution situation, the hypothesis in (4) is generalized to  $H_0 : \mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}$ , for all  $\mathbf{x} \in \mathcal{L}$ , where  $\mathcal{L}$  is the linear subspace representing either a unique single solution (i.e., a point) or multiple solutions (i.e., a line, or a hyperplane). In other words, the confidence region could be a collection of either points or linear subspaces (of dimension  $\geq 1$ ) depending on whether the solution to the equation  $\mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}$  is unique or not. Therefore, we propose to generalize the MKG confidence region to

$$C_{\mathcal{L}}(c_{\alpha}) = \{\mathcal{L} : Q(\mathbf{x}; \hat{\boldsymbol{\psi}}) \leq c_{\alpha} \quad \forall \mathbf{x} \in \mathcal{L}\}, \quad (8)$$

where  $\mathcal{L}$  represents the linear subspace of the space defined by the elements of  $\mathbf{x}$ , which are solutions to  $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x} = \mathbf{0}$ . Here  $\mathbf{x}$  is a  $k \times 1$  vector.  $\mathcal{L}$  has dimension  $d$ , where  $d$  is defined as  $k - h$ , if  $h < k$ , 0 otherwise. Therefore, when the solution to  $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x} = \mathbf{0}$  is a single point,  $d = 0$ ; otherwise,  $d > 0$ . In this section, we derive values for  $c_{\alpha}$  in (8). When  $d > 0$  and  $\mathcal{L}$  is, therefore, not a point, computation of the confidence region in (8) may appear difficult due to the replacement of  $x$ -points by  $\mathcal{L}$  subspaces. However, in this section, we will also show that the confidence region in (8) is equivalent to one based on pointwise gridding (of the type done for the MKG confidence region computations).

We call the confidence region given in (8) the generalized zero-gradient (GZG) confidence region. As indicated by the definition, the MKG confidence region is a special case of the GZG confidence region where  $\mathcal{L}$  is a point and  $d = 0$ . The MKG confidence region is correct, and the critical value is  $c_{\alpha} = hF(1 - \alpha, h, \nu)$  for  $h \geq k$ , if a solution exists. (If  $h \geq k$  and the MKG region is the null set, then there is statistically significant evidence that a solution does not exist.) The next question is what value should  $c_{\alpha}$  take when  $d > 0$ ? For  $d >$

0, note that a  $100(1 - \alpha)\%$  GZG confidence region should contain the zero-gradient solution set  $\mathcal{L}$ , with probability  $1 - \alpha$  before the experiment is performed. It is worth pointing out that when  $h < k$ , the GZG confidence region in (8) is a simultaneous confidence region problem in that a line or hyperplane will be included in the confidence region only if all the points on the line or hyperplane satisfy the criterion in (8). Therefore, the GZG confidence region in (8) can also be expressed as

$$\{\mathcal{L} : Q_{\mathcal{L}} \leq c_{\alpha}\}, \quad (9)$$

where  $Q_{\mathcal{L}} = \max_{\mathbf{x} \in \mathcal{L}} Q(\mathbf{x}; \hat{\boldsymbol{\psi}})$  and  $\mathcal{L} = \{\mathbf{x} : \mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}\}$ . To find the critical value  $c_{\alpha}$  for  $h < k$ , we need to first investigate the distribution of the test statistic  $Q_{\mathcal{L}}$  when  $H_0 : \mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}$  is true.

When  $h = 1$ ,  $\mathbf{M}(\mathbf{x}) = (1, \mathbf{x}')$ , which is a  $1 \times p$  vector. Based on Miller's theorem (see p65 and p113 in [13]), the critical value should be  $c_{\alpha} = (d+1)F(1 - \alpha, (d+1), \nu)$ , where  $d$  is the dimension of the solution set,  $\mathcal{L}$ , for the equation  $\mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}$ . Here,  $\nu$  is the degrees of freedom of the residuals. Note that when  $h = k = 1$ ,  $d = 0$  means that the solution is a point, which is a linear space of dimension zero. The critical value  $(d+1)F(1 - \alpha, (d+1), \nu)$  then becomes the MKG critical value  $F(1 - \alpha, 1, \nu)$  because  $d = 0$ .

For the  $k > h > 1$  case, the distribution of  $Q_{\mathcal{L}}$  is more complex. Section 2.2 addresses the full model in (1) where the experimental design is completely orthogonal or partially orthogonal so that  $\mathbf{V}_{\hat{\boldsymbol{\psi}}} = c\sigma^2\mathbf{I}$  for some positive constant,  $c$ , residual variance  $\sigma^2$ , and an identity matrix  $\mathbf{I}$ , of dimension of  $p$ . Here, an exact simultaneous confidence region is derived. For the general case, Section 2.3 proposes a simulation method based upon the multivariate  $t$ -distribution to find approximate critical values with which to construct the confidence region.

**2.2. Full Model-Orthogonal Case.** Here, we assume that the data are generated from an orthogonal design or partially orthogonal design such that  $\mathbf{V}_{\hat{\boldsymbol{\psi}}} = c\sigma^2\mathbf{I}$ . Furthermore, it is assumed that we have a full-noise-control variable interaction model, meaning that each noise variable interacts with the same set of control variables, that is, each element of interaction matrix  $\boldsymbol{\Delta}$  is nonzero. If  $\mathbf{V}_{\hat{\boldsymbol{\psi}}} = c\sigma^2\mathbf{I}$ ,  $c > 0$ , and all the elements of the interaction matrix  $\boldsymbol{\Delta}$  are nonzero, then, for  $k > h > 1$ , the distribution of the test statistic  $Q_{\mathcal{L}}$  has the same distribution as a function of a chi-square random-variable and a random Wishart's matrix (which are stochastically independent) as shown below:

$$Q_{\mathcal{L}} \sim \frac{\lambda_{\max}(\mathbf{A})}{U/\nu}, \quad (10)$$

where  $U \sim \chi^2(\nu)$ ,  $\mathbf{A} \sim \text{Wishart}(\mathbf{I}_{d+1}, h)$ , and  $\nu$  is the residual degrees of freedom. Here,  $\mathbf{I}_{d+1}$  is a  $(d+1) \times (d+1)$  identity matrix, where  $d$  is the same as defined in Section 2.1. The degrees of freedom of the Wishart distribution is  $h$ , and  $\lambda_{\max}(\mathbf{A})$  is the maximum eigenvalue of the matrix  $\mathbf{A}$ . The proof of this result is provided in Appendix A.

Using (10), the critical value,  $c_{\alpha}$ , can then be computed as the  $100(1 - \alpha)\text{th}$  percentile of the distribution of  $Q_{\mathcal{L}}$ , which

can be obtained by the simple Monte Carlo simulation from  $\chi^2$  and Wishart's distributions. Some limited tables of critical values are given in Appendix B based on (10), although computation of the critical value for any specific case using the random variable in (10) is easily accomplished. Note that the critical value determined by (10) becomes the MKG critical value,  $hF(1 - \alpha, h, \nu)$ , when  $d = 0$ , that is,  $h = k$ . This is because when  $d = 0$ ,  $\mathbf{A}$  has a Wishart  $(1, h)$  distribution which is  $\chi^2(h)$ . In other words,  $\mathbf{A}$  is not a matrix anymore but a  $\chi^2(h)$  random variable. Hence,  $\lambda_{\max}(\mathbf{A}) = \mathbf{A}$ . So  $Q_{\mathcal{L}}$  can then be written as

$$Q_{\mathcal{L}} = h \frac{\mathbf{A}/h}{U/\nu} \sim hF(h, \nu). \quad (11)$$

Furthermore, the same result in (10) holds when  $[\boldsymbol{\gamma}, \boldsymbol{\Delta}]$  has an  $h \times (k + 1)$  matrix normal distribution. (See p90-91 in [14].) (For details, see [15].)

**2.3. The General Case.** In some cases, the experimental design may be such that  $\text{Var}(\hat{\boldsymbol{\psi}})$  does not have the orthogonal  $c\sigma^2\mathbf{I}$  form, or we may wish to use a model with some "control  $\times$  noise variable" interaction terms deleted, that is,  $\boldsymbol{\Delta}$  has some zero elements. In such situations, when  $k > h > 1$ , the distribution of  $Q_{\mathcal{L}}$  does not have a simple form and may depend upon  $\boldsymbol{\psi}$  even under  $H_0$ . Nonetheless, in such situations, it is still possible to obtain approximately conservative simultaneous confidence regions for control variables associated with zero-gradient solutions. We provide such a construction as follows.

Recall that  $Q_{\mathcal{L}}$  in (9) is a function of  $\boldsymbol{\psi}$ , and consider

$$Q_{\max} = \max_{\boldsymbol{\psi} \in C_{\boldsymbol{\psi}}(b_{\alpha})} Q_{\mathcal{L}} = \max_{\boldsymbol{\psi} \in C_{\boldsymbol{\psi}}(b_{\alpha})} \max_{\mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}} Q(\mathbf{x}; \hat{\boldsymbol{\Psi}}), \quad (12)$$

where  $C_{\boldsymbol{\psi}}(b_{\alpha}) = \{\boldsymbol{\psi} : (\boldsymbol{\psi} - \hat{\boldsymbol{\psi}})' \mathbf{V}_{\hat{\boldsymbol{\psi}}}^{-1} (\boldsymbol{\psi} - \hat{\boldsymbol{\psi}}) \leq b_{\alpha}\}$  and  $b_{\alpha} = F(1 - \alpha, 1, n - p)$ . Note that by using  $b_{\alpha} = F(1 - \alpha, 1, n - p)$ ,  $Q_{\max}$  is an approximate upper confidence bound for the scalar-valued quantity,  $Q_{\mathcal{L}}$ . (See Clarke, [16], for a discussion of confidence bounds on nonlinear functions of model parameters constructed from confidence regions.) Let  $c_{\alpha}^*$  denote the 100(1- $\alpha$ )th percentile of the distribution of  $Q_{\max}$  under  $H_0$ . Consider the confidence region defined by  $C_{\mathbf{x}}^{\max} = \{\mathbf{x} : Q(\mathbf{x}; \hat{\boldsymbol{\Psi}}) \leq c_{\alpha}^*\}$ . This confidence region should provide (at least approximately) a conservative simultaneous confidence region for the zero-gradient solutions. However, computation of  $c_{\alpha}^*$  (using (12)) and the associated confidence region is numerically difficult due to the complex constraints associated with the definition of  $Q_{\max}$ . Fortunately, it can be shown that

$$Q_{\max} = \max_{\mathbf{x} \in C_{\mathbf{x}}(b_{\alpha})} Q(\mathbf{x}; \hat{\boldsymbol{\Psi}}), \quad (13)$$

where  $C_{\mathbf{x}}(b_{\alpha}) = \{\mathbf{x} : Q(\mathbf{x}; \hat{\boldsymbol{\Psi}}) \leq b_{\alpha}\}$ . A proof is given in Appendix C. The expression for  $Q_{\max}$  in (13) allows for much easier computation of the  $c_{\alpha}^*$  critical value. The actual construction of the GZG confidence region from the relevant critical value will be outlined in Section 2.4.

### 2.3.1. The Critical Value Computation for the General Case.

Note that under  $H_0 : \mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}$  we can express  $Q(\mathbf{x}; \hat{\boldsymbol{\Psi}})$  as

$$(\mathbf{M}(\mathbf{x})\mathbf{t})' (\mathbf{M}(\mathbf{x})\boldsymbol{\Omega}\mathbf{M}(\mathbf{x})')^{-1} (\mathbf{M}(\mathbf{x})\mathbf{t}), \quad (14)$$

where  $\hat{\mathbf{V}}_{\hat{\boldsymbol{\psi}}} = s^2\boldsymbol{\Omega}$ ,  $s^2$  is the mean squared error,  $\boldsymbol{\Omega}$  is a known matrix computed from the design matrix, and  $\mathbf{t} = (\hat{\boldsymbol{\Psi}} - \boldsymbol{\psi})/s$ . Here,  $\mathbf{t}$  follows the multivariate  $t$  distribution with location parameter equal to zero, scale matrix  $\boldsymbol{\Omega}$ , and degrees of freedom  $\nu$ . Using (14) we can then compute the critical value,  $c_{\alpha}^*$ , using the Monte Carlo simulations as follows.

*Step 1.* Compute  $C_{\mathbf{x}}(b_{\alpha}) = \{\mathbf{x} : Q(\mathbf{x}; \hat{\boldsymbol{\Psi}}) \leq b_{\alpha}\}$ , where  $b_{\alpha} = F(1 - \alpha, 1, n - p)$ .

*Step 2.* Simulate a multivariate  $\mathbf{t}$  random vector (rv) with scale matrix  $\boldsymbol{\Omega}$  and  $\nu$  df. (This can be done by simulation of a multivariate normal rv with mean vector  $\mathbf{0}$  and variance-covariance matrix  $\boldsymbol{\Omega}$  and a chi-square random variable with  $\nu$  df. See [17] for details.)

*Step 3.* Compute  $Q_{\max}$  using the expressions in (13) and (14). (For practical reasons, computation of  $Q_{\max}$  can be done by maximization of  $Q(\mathbf{x}; \hat{\boldsymbol{\Psi}})$  over  $C_{\mathbf{x}}(b_{\alpha}) \cap R$  instead, where  $R$  is a prespecified, bounded region. This will calibrate the coverage to be simultaneous only over  $\mathcal{L}_0 \cap R$ , where  $\mathcal{L}_0$  is the true linear subspace such that  $\mathbf{M}(\mathbf{x})\boldsymbol{\psi} = \mathbf{0}$ .)

*Step 4.* Do Steps 2-3 a large number of times to estimate the 100(1- $\alpha$ )th percentile of the Monte Carlo distribution of  $Q_{\max}$ . This 100(1- $\alpha$ )th percentile is then a Monte Carlo estimate of  $c_{\alpha}^*$ .

### 2.3.2. The Coverage Rate of the Critical Value.

In order to check the accuracy of  $c_{\alpha}^*$  as a critical value, we have done some Monte Carlo simulations of the above four-step procedure using three different noise variable models in conjunction with both orthogonal and nonorthogonal designs. The statistical models used are summarized in Table 1. These models are constructed so that the zero-gradient solution exists in the experimental region. Three partially orthogonal, face-centered central composite experimental designs were assessed, with associated statistical models 1, 2, and 3, respectively. These designs employed a coded factor space with factor levels equal to  $\pm 1$  (except for the center points). The axial points in noise variables are deleted to maintain partial orthogonality. The factorial part of the designs is either full factorial (e.g., model 1 and model 2) or half factorial (e.g., model 3). The nonorthogonal designs are constructed by changing the (one) factorial point (comprised of all -1s) from  $(-1, -1, \dots, -1, -1)$  to  $(-1, -1, \dots, -1, 0)$ . The resultant sample size ( $n$ ) and the residual df ( $\nu$ ) of each composite design are both listed in Table 2. For demonstration purposes, we simply chose model parameters to be either 1 or -1, with residual error variance equal to 1.

The results of these coverage rate simulations are summarized in Table 2. The coverage rates were computed as follows. The models in Table 1 were used to compute the

TABLE 1: Noise-variable-reduced models for simulations.

Model no.	$k$	$h$	Model form
1	3	2	$y = 1 + x_1 + x_2 + x_3 + x_1^2 + x_2^2 + x_1 \times x_2 + z_1 + x_1 \times z_1 + x_2 \times z_1 + z_2 - x_3 \times z_2 + e$
2	4	2	$y = 1 + x_1 + x_2 + x_3 + x_4 + x_1^2 + x_2^2 + x_2 \times x_3 + z_1 + x_1 \times z_1 + x_2 \times z_1 + z_2 - x_3 \times z_2 + x_4 \times z_2 + e$
3	4	3	$y = 1 + x_1 + x_2 + x_3 + x_4 + x_1^2 + x_2^2 + x_2 \times x_3 + z_1 + x_1 \times z_1 + x_2 \times z_1 + z_2 + x_3 \times z_2 + z_3 + x_4 \times z_3 + e$

TABLE 2: The coverage rates for the GZG confidence region using the approximate and conservative critical values,  $\tilde{c}_\alpha^*$  and  $c_\alpha^*$ , respectively. The nominal coverage rate here is 95%. (Here  $k$  = no. of control factors,  $h$  = no. of noise variables,  $n$  = sample size, and  $\nu$  = residual df).

Model no.	$k$	$h$	$n$	$\nu$	Approximate coverage rates		Conservative coverage rates	
					Orthogonal design	Nonorthogonal design	Orthogonal design	Nonorthogonal design
1	3	2	40	28	96.1%	96.1%	96.9%	96.9%
2	4	2	74	60	96.3%	96.3%	98.0%	98.0%
3	4	3	72	57	96.1%	96.1%	97.2%	97.2%

true  $\mathcal{L}$  space,  $\mathcal{L}_0$ . For each of the three models, the region,  $R$ , was a hypercube constructed from the Cartesian product of intervals of the form  $[-10, 10]$ . For each simulation, a dataset was generated based on the model and the corresponding central composite design, a critical value  $c_\alpha^*$  was computed based on the simulated dataset, then a check was done to see if the event

$$\max_{\mathbf{x} \in \mathcal{L}_0 \cap E} Q(\mathbf{x}; \hat{\Psi}) \leq c_\alpha^* \tag{15}$$

occurred, where  $E$  is the convex hull formed by the  $\pm 1$  factor levels. For each simulated dataset, the critical value of  $c_\alpha^*$  was computed using 1000 Monte Carlo simulations. 5000 simulations were done to assess the simultaneous coverage of the GZG confidence region for the set  $\mathcal{L}_0 \cap E$ . In an attempt to reduce the conservatism of the above approach for computing  $c_\alpha^*$ , we also considered the approximate approach obtained by maximizing  $Q(\mathbf{x}; \hat{\Psi})$  over  $\mathcal{L}_0(\hat{\Psi}) \cap R$ , where  $\mathcal{L}_0(\hat{\Psi}) = \{\mathbf{x} : \mathbf{M}(\mathbf{x})\hat{\Psi} = \mathbf{0}\} = C_x(0)$ . We denote this approximate critical value by  $\tilde{c}_\alpha^*$  and use it in place of  $c_\alpha^*$  to reduce conservatism.

*Remark 1.* Because  $C_x(b_\alpha)$  is a function of the data, the relatively large region,  $R$ , was chosen for these simulations so that  $C_x(b_\alpha) \cap R$  would be extremely unlikely to be empty for any simulated dataset. In addition, we did not want to rule out situations where the confidence region was outside the experimental region. While, in practice, such extrapolated inferences must be treated with caution, nonetheless it may be desired to compute such a confidence region. Such a confidence region outside the experimental region suggests that it may not be possible to obtain a “zero-gradient” solution for noise transmission, at least within the current experimental region. However, such a confidence region just outside the experimental region may offer hope that resetting process control conditions may allow for a more robust process. Of course, additional experiments outside the current experimental region would be needed to confirm this.

*Remark 2.* Maximization of  $Q(\mathbf{x}; \hat{\Psi})$  over  $C_x(b_\alpha) \cap R$ , to compute the Monte Carlo critical value,  $c_\alpha^*$ , was accomplished by using the SAS/IML Nelder-Mead simplex algorithm,

nlpnms. This was done to make the Monte Carlo simulations of this Monte Carlo procedure tractable. Some limited simulations were also done whereby the maximization of  $Q(\mathbf{x}; \hat{\Psi})$  over  $C_x(b_\alpha) \cap R$  was computed by gridding instead. This was done to make sure that the Nelder-Mead algorithm did not stop its maximization prematurely. In all cases, each critical value,  $c_\alpha^*$ , computed using nlpnms, was slightly larger than that obtained using gridding. (Random number seeds were aligned to avoid the Monte Carlo differences in the comparisons between gridding and the use of the Nelder-Mead simplex algorithm.) For the approximate approach, maximization over  $\mathcal{L}_0(\hat{\Psi}) \cap R$  was done by gridding as this was easier to accomplish with finer gridding.

Table 2 below displays the percent of times the event in (15) occurred for each of the three models with and without an orthogonal design. If the event in (15) occurs, then that portion of the true linear subspace,  $\mathcal{L}_0$  (within  $E$ ), is entirely covered by the GZG confidence region; otherwise it is not.

Table 2 indicates that the simultaneous coverage rate of the GZG confidence region using the conservative critical value,  $c_\alpha^*$ , produces reasonably conservative results, while the approximate approach (that maximizes over  $\mathcal{L}_0(\hat{\Psi}) \cap R$ , instead) achieves closer to nominal (yet slightly conservative) coverage rates. It is interesting to note that for each approach the coverage rate appears to be insensitive to the minor departure from orthogonality that was induced by changing the (one) factorial point (comprised of all  $-1$ s) from  $(-1, -1, \dots, -1, -1)$  to  $(-1, -1, \dots, -1, 0)$ . Such a departure from design orthogonality could happen due to a design execution error or a process restriction.

*2.3.3. The Full Model Nonorthogonal Case.* Because computation of  $c_\alpha^*$  and  $\tilde{c}_\alpha^*$  requires maximization within a Monte Carlo calculation, it would be useful to assess if this can be eliminated when a full model is employed. We, therefore, conduct another simulation study to see if the critical value based upon the random variable in (10) can be used as an approximate critical value for mild departures from orthogonality. We use the same nonorthogonal designs as used in Table 2. The corresponding full-interaction models

TABLE 3: Noise variable full models for simulations.

Model no.	$k$	$h$	Model form
4	3	2	$y = 1 + x_1 + x_2 + x_3 + x_1^2 + x_2^2 + x_3^2 + x_2 \times x_3 + z_1 + x_1 \times z_1 + x_2 \times z_1 + x_3 \times z_1 + z_2 - x_1 \times z_2 + x_2 \times z_2 - x_3 \times z_2$
5	4	2	$y = 1 + x_1 + x_2 + x_3 + x_4 + x_1^2 + x_2^2 + x_2 \times x_3 + z_1 + x_1 \times z_1 + x_2 \times z_1 + x_3 \times z_1 + x_4 \times z_1 + z_2 - x_1 \times z_2 + x_2 \times z_2 - x_3 \times z_2 + x_4 \times z_2$
6	4	3	$y = 1 + x_1 + x_2 + x_3 + x_4 + x_1^2 + x_2^2 + x_2 \times x_3 + z_1 + x_1 \times z_1 + x_2 \times z_1 + x_3 \times z_1 + x_4 \times z_1 + z_2 - x_1 \times z_2 + x_2 \times z_2 - x_3 \times z_2 + x_4 \times z_2 + z_3 + x_1 \times z_3 + x_2 \times z_3 + x_3 \times z_3 - x_4 \times z_3$

TABLE 4: The simultaneous coverage rate of the critical value for the full model nonorthogonal designs (no. of simulations = 100,000, and nominal coverage rate = 95%.  $k$ : no. of control factors,  $h$ : no. of noise variables,  $n$ : sample size, and  $\nu$ : residual df).

Model no.	$k$	$h$	$n$	$\nu$	Coverage rate
4	3	2	40	24	95.2%
5	4	2	74	56	95.0%
6	4	3	72	49	95.1%

are listed in Table 3. This time the  $c_\alpha$  critical value was used with these nonorthogonal designs to assess the simultaneous coverage rate. The results are shown in Table 4. In order to assess the coverage rate gridding had to be done over a subset of  $\mathcal{L}_0$ . As a more fair comparison with the theoretical  $c_\alpha$  critical value, gridding was done over  $\mathcal{L}_0 \cap R$ , (whereas before  $R$  is a hypercube region composed of the Cartesian product of the intervals  $[-10, 10]$ , instead of  $[-1, 1]$ ). This is because the  $c_\alpha$  critical value associated with the random variable in (10) is computed by maximization over the whole linear subspace,  $\mathcal{L}_0$ .

Table 4 indicates that this minor departure from orthogonality has virtually no effect on the coverage rate of the GZG confidence region when the more convenient  $c_\alpha$  critical value is used. For more radical departures from orthogonality, it may possibly be safer to use the conservative  $c_\alpha^*$  critical value. But further robustness studies are needed to ascertain how well the more convenient  $c_\alpha$  critical value works under departures from its assumptions.

**2.4. Computation of Simultaneous Confidence Region.** For the  $k > h$  case, once we have computed the critical value, the confidence region can be computed by searching linear subspaces,  $\mathcal{L}$ , that satisfy the condition as defined in (9). However, searching over various lines or hyperplanes that span an experimental region is more computationally difficult than searching the same experimental region in a pointwise fashion. Fortunately, it can be shown that, for any given critical value, the GZG confidence region can be computed by pointwise gridding. This is because for  $C_x(c_\alpha)$  in (7) and  $C_{\mathcal{L}}(c_\alpha)$  in (8), with the same critical value,  $C_x(c_\alpha) = C_{\mathcal{L}}(c_\alpha)$ . A proof is provided in Appendix D. This equivalency shows that one can construct the GZG confidence region by simply gridding over the experimental region in a pointwise fashion.

### 3. Examples

**3.1. One Noise Variable.** This example is from Myers et al. [6]. It was originally taken from Montgomery [18] (2009, page 231). The data was generated from a  $2^4$  factorial

experiment with a total of 16 observations from a pilot plant to explore the factors that could affect the filtration rate of a chemical bonding substance. The goal is to maximize the filtration rate,  $y$ .

As in Myers et al. [6], one of the four factors, temperature, is assumed difficult to control at large scale and, therefore, treated as a noise variable  $z$ . The rest of the factors are control variables:  $x_1$ : pressure,  $x_2$ : concentration,  $x_3$ : stirring rate. The fitted model is

$$\hat{y} = 70.06 + 10.81z + 4.94x_2 + 7.31x_3 - 9.06x_2z + 8.31x_3z - 0.56x_2x_3, \quad (16)$$

with mean square error equal to 21.12 and residual df equal to 9.

The estimated slope of noise variable is

$$\hat{y} + \widehat{\Delta} \mathbf{x} = 10.81 - 9.06x_2 + 8.31x_3. \quad (17)$$

Therefore,  $k = 2$  and  $h = 1$ . The solution to the null hypothesis  $H_0 : \gamma + \Delta' \mathbf{x} = \mathbf{0}$  is a line. Then the general critical value  $2F(1-\alpha, 2, \nu)$  (based on Miller's Theorem (1981) [13]) should be used to calculate the GZG confidence region (as shown in Figure 1). The GZG confidence region in Figure 1 is clearly wider than the MKG confidence region in Myers et al. (1997, Figure 2 in [6]) where  $F(1-\alpha, 1, \nu)$  is used as the critical value. It is clear from Figure 1 that we are at least 95% confident that the zero-gradient locus of points passes through the experimental region.

Next, we do some simulations to compare the coverage rates of the GZG and MKG confidence regions. Since the true optima is not known in practice, we calculate the coverage rate for the solution to  $\hat{y} + \widehat{\Delta} \mathbf{x} = 10.81 - 9.06x_2 + 8.31x_3 = 0$ , using a simulation model equal to the fitted model in (16) with  $\sigma^2 = 21.12$ . Note that the true solution in this example is a line with infinite length. But the simulation is done only for the line within the experimental region, that is,  $[-1, 1]$  of the control variables.

Using 100,000 Monte Carlo simulations, the simultaneous coverage rate of the GZG confidence region for all of the zero-gradient solutions in the experimental region is 97% while the MKG confidence region only has 92% coverage

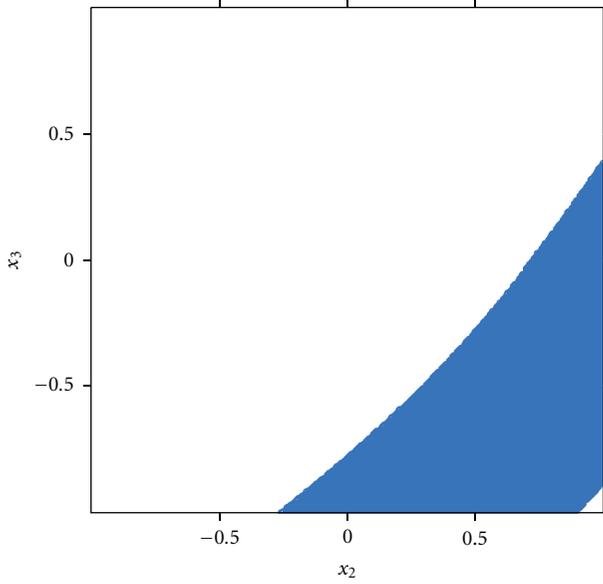


FIGURE 1: The 95% GZG confidence region for one-noise-variable and two-control-variable case.

rate. The MKG confidence region has a lower coverage rate because it was designed to contain the true optima only when the optimum is a point. Although the GZG confidence region is designed to contain all the true solutions (which could be a point, a line, or a hyperplane), the simulated coverage rate tends to exceed the nominal coverage rate because the simulation is done within a finite range of the control variables while the line or hyperplane has an infinite range in theory.

**3.2. Two Noise Variables.** This example comes from a face-centered central composite design with the factorial part being a half-fractional factorial design (see details in [19]). The objective of this study is to find the optimized condition that maximizes the yield of diacylglycerol oil, which is a natural component of various edible oils and has shown some beneficial effects as compared to the traditional triacylglycerol oil.

Five factors were studied in this experiment: reaction time (RTIME), enzyme load (ENZL), reaction temperature (RTEMP), water content (WATC), and substrate molar ratio (SUBR). Water content (WATC) is difficult to control at large scale [19] and, therefore, treated as a noise variable. For illustration purposes, substrate molar ratio is also treated as a noise variable, and the axial points corresponding to the noise variables are excluded from the analysis to obtain partial design orthogonality with respect to the noise variables (i.e., to ensure  $\mathbf{V}_{\hat{\psi}} = c\sigma^2\mathbf{I}$ ). The final model in coded factor value is as follows:

$$\begin{aligned} \hat{y} = & 57.58 + 9.12x_1 + 4.78x_2 + 11.01x_3 \\ & - 4.69x_1^2 - 9.47x_2^2 - 7.37x_3^2 - 1.61x_{23} \\ & - 2.05z_1 + 4.83z_2 - 2.92z_1x_1 - 2.07z_1x_2 \\ & - 2.12z_1x_3 - 3.17z_2x_1 + 4.90z_2x_2 - 2.41z_2x_3, \end{aligned} \tag{18}$$

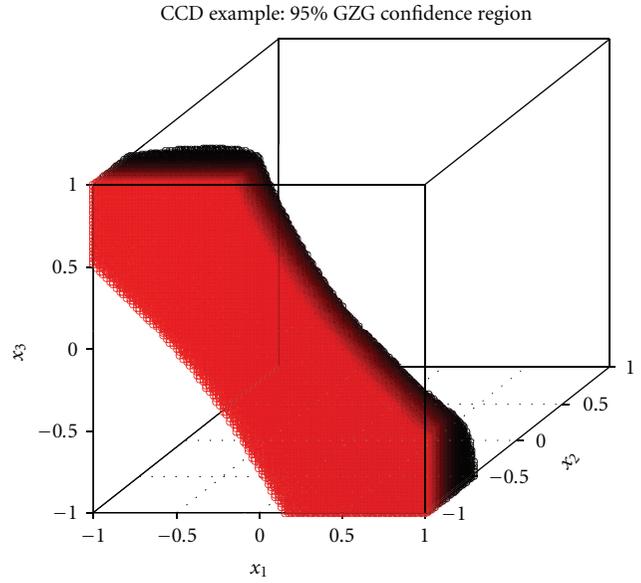


FIGURE 2: The 95% GZG confidence region for two-noise-variable and three-control-variable case.

where  $x_1$ : RTIME,  $x_2$ : ENZL,  $x_3$ : RTEMP,  $z_1$ : WATC,  $z_2$ : SUBR. Here, the residual mean squared error is equal to 2.56 with 25 observations and residual df equal to 9.

Since  $k = 3, h = 2$ , the solution to the null hypothesis is a line in a 3-dimensional space determined by control variables  $x_1, x_2$ , and  $x_3$ . Therefore, the confidence region for this line is a tube in this 3-dimensional space. A 95% GZG confidence region is shown in Figure 2. Based on (10), the GZG critical value is obtained via  $\chi^2$  and Wishart's distribution. From Figure 2, we can see that while this confidence region does not provide statistically significant evidence that the zero-gradient locus of points passes through the experimental region, it does appear that a good portion of the confidence region is within the experimental region, and hence, attainment of near-zero-gradient conditions should be feasible for this process.

As with previous example, we compare the coverage rates for the GZG and MKG confidence regions using the fitted models as true population models. Using 100,000 Monte Carlo simulations (based upon the fitted model in (18) with  $\sigma^2 = 2.56$ ), the simultaneous coverage rate is 96% while it is only 90% for the nominal 95% MKG confidence region. (Here, gridding was done over the cube formed by the Cartesian product of  $[-1, 1]$  associated with each  $x_i$ .)

#### 4. Summary

This paper shows that when the number of control variables does not exceed the number of noise variables, the MKG approach provides a confidence region for control variables associated with a zero-gradient for noise transmission. Otherwise, the MKG approach results in a confidence region that is too small for simultaneous coverage of the linear subspace of zero-gradient solutions. It is important to know that the true optimal condition represented by control variables is

either a line or a hyperplane instead of a single point when  $h < k$ . In this situation, constructing a simultaneous confidence region about the linear subspace solution is desirable in that a subspace of solutions provides the investigator with many options for setting the zero-gradient control level. Of course a confidence region also provides the experimenter with a measure of uncertainty for the optimal solution. If the confidence region is too large, further experimental runs may be needed to make more accurate inferences. If the current manufacturing set point is outside of the confidence region, this provides statistically significant evidence that reconfiguration of the set point may help improve process variability by lowering the transmission of noise through the system. The GZG confidence region for the zero-gradient conditions is proposed and is shown to provide nominal or reasonably conservative coverage rates for many noise variable experiments that occur in practice.

In the situation where there are many noise variables, it may be either costly or difficult to study all the noise effects. One way to deal with this problem is to combine the multiple noise factors into one compound noise factor with two extreme conditions as its two levels (provided certain assumptions can be satisfied). See [1, 20, 21] for discussion. If the noise factors can be combined into one compound noise factor, then we could have a situation where the number of control variables is greater than the number of noise variables. The GZG approach is directly applicable for this situation. In some cases, however, one may desire to create a compound noise factor, with more than two levels [22], or two or more compound noise factors. In either case, as long as the predictive model is in the form of (1), the GZG approach is applicable.

The GZG confidence region provides inferences about the optimal control point or points that yield a zero-gradient for the transmission of variability from the noise variables. However, there are situations where the control points corresponding to zero noise variance are either outside the experimental region or simply do not exist. In this case, it would still be useful to find the constrained optimal control point for minimum noise variance over the experimental region. A method to further generalize the confidence region for the constrained optimal point for minimum noise variance is needed and is currently under development.

## Appendices

### A. Proof of (10)

We will prove the result for the no-intercept case where  $\boldsymbol{\gamma} = \mathbf{0}$  in the hypothesis  $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x} = \mathbf{0}$ . It then follows that the result can be generalized to the intercept case where  $\boldsymbol{\gamma} \neq \mathbf{0}$ .

*Part 1.* For the no-intercept case:

When  $k > h$ , the solution,  $\mathbf{x}$ , to  $\boldsymbol{\Delta}'\mathbf{x} = \mathbf{0}$  can be expressed as the linear combination of the basis vectors for the linear subspace  $\mathcal{L}$ , that is,  $\mathbf{x}' = \mathbf{w}'\mathbf{L}$ , where  $\mathbf{w}$  is a  $d \times 1$  coefficient vector, and  $\mathbf{L}$  is a  $d \times k$  matrix where the rows consist of the basis vectors of linear subspace  $\mathcal{L}$ . (Here,  $\mathbf{w} \in \mathbb{R}^d$  and only  $\mathbf{L}$  is a function of  $\boldsymbol{\Delta}$ .) Then  $\mathbf{M}(\mathbf{x}) = \mathbf{I}_h \otimes \mathbf{w}'\mathbf{L}$ . Let  $\mathbf{z} = (\hat{\boldsymbol{\psi}} - \boldsymbol{\psi})/(\sigma\sqrt{c})$ , then  $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I}_p)$  and  $\mathbf{M}(\mathbf{x})\mathbf{z} = \mathbf{G}(\mathbf{z})\mathbf{x}$ , where  $\mathbf{G}(\mathbf{z})$  is defined as

$$\mathbf{G}(\mathbf{z}) = \begin{pmatrix} z_{11} & \cdots & z_{1k} \\ \vdots & \ddots & \vdots \\ z_{h1} & \cdots & z_{hk} \end{pmatrix}, \quad (\text{A.1})$$

where  $z_{ij} = (\hat{\delta}_{ij} - \delta_{ij})/(\sigma\sqrt{c})$ ,  $i = 1, 2, 3, \dots, h$ ,  $j = 1, 2, 3, \dots, k$ , and  $\delta_{ij}$  is the  $j$ th element on the  $i$ th row of the matrix  $\boldsymbol{\Delta}'$ . Since the vector  $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I}_p)$ ,  $\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z}) \sim \text{Wishart}(\mathbf{I}_k, h)$  by the definition of Wishart distribution (see p92 in [23]). Let  $s^2$  be the sample estimate of the residual variance  $\sigma^2$ , then the test statistic  $Q_{\mathcal{L}}$  becomes

$$\begin{aligned} Q_{\mathcal{L}} &= \max_{\mathbf{x} \in \mathcal{L}} (\mathbf{M}(\mathbf{x})(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}))' \\ &\quad \times (cs^2\mathbf{M}(\mathbf{x})\mathbf{M}(\mathbf{x})')^{-1} (\mathbf{M}(\mathbf{x})(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi})) \\ &= \frac{\sigma^2}{s^2} \max_{\mathbf{w}} \left( (\mathbf{I}_h \otimes \mathbf{w}'\mathbf{L}) \left( \frac{\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}}{\sigma\sqrt{c}} \right) \right)' \\ &\quad \times \left( (\mathbf{I}_h \otimes \mathbf{w}'\mathbf{L})(\mathbf{I}_h \otimes \mathbf{w}'\mathbf{L})' \right)^{-1} \left( (\mathbf{I}_h \otimes \mathbf{w}'\mathbf{L}) \left( \frac{\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}}{\sigma\sqrt{c}} \right) \right). \end{aligned} \quad (\text{A.2})$$

By replacing  $(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi})/(\sigma\sqrt{c})$  by  $\mathbf{z}$  in (A.2),  $Q_{\mathcal{L}}$  becomes

$$Q_{\mathcal{L}} = \frac{\sigma^2}{s^2} \max_{\mathbf{w}} \frac{\mathbf{w}'\mathbf{L}\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z})\mathbf{L}'\mathbf{w}}{\mathbf{w}'\mathbf{L}\mathbf{L}'\mathbf{w}} = \frac{\sigma^2}{s^2} \lambda_{\max}(\mathbf{D}), \quad (\text{A.3})$$

where  $\mathbf{D} = (\mathbf{L}\mathbf{L}')^{-1}\mathbf{L}\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z})\mathbf{L}'$  and  $\mathbf{w}'\mathbf{L}\mathbf{L}'\mathbf{w}$  is a scalar. (For a proof of (A.3), see the result in Problem 22.1 in Rao 1973 [24, p74].)

By the definition of an eigenvalue, it follows that

$$\left| (\mathbf{L}\mathbf{L}')^{-1}\mathbf{L}\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z})\mathbf{L}' - \lambda_{\max}\mathbf{I}_d \right| = 0. \quad (\text{A.4})$$

Note  $\mathbf{L}\mathbf{L}'$  is positive definite and symmetric, and, therefore,  $(\mathbf{L}\mathbf{L}')^{-1}$  is positive definite and symmetric as well. Therefore, both  $(\mathbf{L}\mathbf{L}')^{1/2}$  and  $(\mathbf{L}\mathbf{L}')^{-1/2}$  exist.

Now multiply both sides of (A.4) by  $(\mathbf{L}\mathbf{L}')^{1/2}$  from the left, then multiply both sides by  $(\mathbf{L}\mathbf{L}')^{-1/2}$  from the right, we get

$$\left| (\mathbf{L}\mathbf{L}')^{-1/2}\mathbf{L}\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z})\mathbf{L}'(\mathbf{L}\mathbf{L}')^{-1/2} - \lambda_{\max}\mathbf{I}_d \right| = 0. \quad (\text{A.5})$$

Let  $\mathbf{A} = (\mathbf{L}\mathbf{L}')^{-1/2}\mathbf{L}\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z})\mathbf{L}'(\mathbf{L}\mathbf{L}')^{-1/2}$ , then  $\lambda_{\max}(\mathbf{D}) = \lambda_{\max}(\mathbf{A})$ . Therefore,

$$Q_{\mathcal{L}} = \frac{\sigma^2}{s^2} \lambda_{\max}(\mathbf{A}) \sim \frac{\lambda_{\max}(\mathbf{A})}{U/v}, \quad (\text{A.6})$$

where  $U = vs^2/\sigma^2$  follows  $\chi^2(v)$  distribution.

Next, we show that  $\mathbf{A} \sim \text{Wishart}(\mathbf{I}_d, h)$ . Because  $\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z}) \sim \text{Wishart}(\mathbf{I}_k, h)$  and  $\mathbf{L}$  is a  $d \times k$  matrix with rank  $d$ , by the property of Wishart distribution

TABLE 5: Each critical value is generated by 100,000 simulations.

$\nu$	$h$	$k$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
2	2	3	361.2	69.40	33.57
3	2	3	101.7	32.31	18.99
4	2	3	58.37	23.10	14.80
5	2	3	40.90	18.71	12.71
6	2	3	33.96	16.22	11.47
7	2	3	27.87	14.77	10.58
8	2	3	25.05	13.72	10.01
9	2	3	23.13	13.04	9.65
10	2	3	21.18	12.45	9.37
20	2	3	17.14	10.53	8.18
30	2	3	15.04	10.06	7.75
40	2	3	14.16	9.26	7.28
50	2	3	13.78	9.18	7.53
60	2	3	13.33	9.09	7.28
70	2	3	13.42	9.09	7.29
80	2	3	12.84	8.95	7.22
90	2	3	12.90	8.98	7.18
100	2	3	12.76	9.00	7.29
2	2	4	473.9	95.48	46.42
3	2	4	142.5	45.26	26.65
4	2	4	74.17	30.73	20.11
5	2	4	53.50	24.96	17.17
6	2	4	43.52	21.75	15.41
7	2	4	35.82	19.17	14.10
8	2	4	31.84	18.08	13.40
9	2	4	29.91	17.17	12.88
10	2	4	25.82	15.81	12.02
20	2	4	19.22	13.10	10.54
30	2	4	18.27	12.30	10.06
40	2	4	16.58	11.87	9.75
50	2	4	16.23	11.62	9.54
60	2	4	16.01	11.54	9.48
70	2	4	15.44	11.44	9.49
80	2	4	16.05	11.50	9.43
90	2	4	15.24	11.26	9.28
100	2	4	15.61	11.15	9.39
2	3	4	513.8	98.84	46.93
3	3	4	139.7	44.78	26.45
4	3	4	76.92	31.03	20.09
5	3	4	53.05	24.64	16.92
6	3	4	42.91	21.71	15.38
7	3	4	36.41	19.55	14.20
8	3	4	32.46	18.06	13.42
9	3	4	29.49	16.94	12.84
10	3	4	26.53	16.05	12.28
20	3	4	20.14	13.28	10.70
30	3	4	17.63	12.28	9.96
40	3	4	16.88	12.06	9.94
50	3	4	16.85	11.87	9.56
60	3	4	15.92	11.39	9.35

TABLE 5: Continued.

$\nu$	$h$	$k$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
70	3	4	16.00	11.66	9.50
80	3	4	16.16	11.33	9.55
90	3	4	15.70	11.24	9.37
100	3	4	15.07	11.09	9.17

(see the theorem in Rencher 1998 [23, page 56]),  $\mathbf{L}\mathbf{G}(\mathbf{z})'\mathbf{G}(\mathbf{z})\mathbf{L}' \sim \text{Wishart}(\mathbf{L}\mathbf{L}', h)$ . Note that  $\text{rank}(\mathbf{L}'\mathbf{L}) = \text{rank}((\mathbf{L}'\mathbf{L})^{-1}) = \text{rank}(\mathbf{L}) = d$ . Since  $(\mathbf{L}\mathbf{L}')^{-1}$  is symmetric and positive definite, then  $(\mathbf{L}\mathbf{L}')^{-1/2}$  is also symmetric and positive definite, and its rank is  $d$ . Hence,  $\mathbf{A} \sim \text{Wishart}(\mathbf{I}_d, h)$ .

Part 2. For the intercept case:

The intercept case can be proved using the same arguments as indicated by Miller ([13, page 113]). Note that, in the intercept case, the rank of  $(\mathbf{L}\mathbf{L}')^{-1/2}$  is  $d + 1$ . Hence,  $\mathbf{A} \sim \text{Wishart}(\mathbf{I}_{d+1}, h)$ .

**B. Critical Values Based on (10)**

See Table 5.

**C. Proof of (13)**

Note that

$$\max_{\psi \in C_\psi} \max_{\{\mathbf{x}:\mathbf{M}(\mathbf{x})\psi=0\}} Q(\mathbf{x}; \hat{\Psi}) = \max_{\{(\mathbf{x}, \psi):\mathbf{M}(\mathbf{x})\psi=0, \psi \in C_\psi\}} Q(\mathbf{x}; \hat{\Psi}). \tag{C.1}$$

Adapting the proof of Theorem 2.1 in Peterson et al. [25], it follows that for any critical value,  $b_\alpha$ ,

$$C_x(b_\alpha) = \{\mathbf{x} : \mathbf{M}(\mathbf{x})\psi = 0, \psi \in C_\psi(b_\alpha)\}. \tag{C.2}$$

Since  $Q(\mathbf{x}; \hat{\Psi})$  is not a function of  $\psi$ , it follows directly that

$$\max_{\{(\mathbf{x}, \psi):\mathbf{M}(\mathbf{x})\psi=0, \psi \in C_\psi\}} Q(\mathbf{x}; \hat{\Psi}) = \max_{\mathbf{x} \in C_x(b_\alpha)} Q(\mathbf{x}; \hat{\Psi}). \tag{C.3}$$

So (13) follows directly from (C.1) and (C.3).

**D. The Proof That  $C_x(c_\alpha) = C_{\mathcal{L}}(c_\alpha)$**

Recall that  $C_x(c_\alpha) = \{\mathbf{x} : Q(\mathbf{x}; \hat{\Psi}) \leq c_\alpha\}$ ,  $C_{\mathcal{L}}(c_\alpha) = \{\mathcal{L} : Q(\mathbf{x}; \hat{\Psi}) \leq c_\alpha \text{ for all } \mathbf{x} \in \mathcal{L}\}$ . By definition,  $\mathbf{x} \in C_{\mathcal{L}}(c_\alpha)$  implies that  $\mathbf{x} \in C_x(c_\alpha)$ . Next, we will show that if  $\mathbf{x} \in C_x(c_\alpha)$ , then  $\mathbf{x} \in C_{\mathcal{L}}(c_\alpha)$ . This can be proved by contradiction. Suppose that there exist some  $\mathbf{x}'$ s such that  $\mathbf{x} \in C_x(c_\alpha)$ , but  $\mathbf{x} \notin C_{\mathcal{L}}(c_\alpha)$ . Then there exists at least one  $\mathbf{x}$  point, say  $\mathbf{x}^*$ , in  $C_x(c_\alpha)$  such that there is no linear subspace that satisfies the two conditions: (1) it contains  $\mathbf{x}^*$ , (2) for every point  $\mathbf{x}$  in this subspace  $Q(\mathbf{x}; \hat{\Psi}) \leq c_\alpha$ . Define  $C_\psi(c_\alpha) = \{\psi : (\hat{\Psi} - \psi)' \hat{\mathbf{V}}_\psi^{-1} (\hat{\Psi} - \psi) \leq c_\alpha\}$  and consider the set,  $C_x^* = \{\mathbf{x} : \mathbf{M}(\mathbf{x})\psi = 0, \text{ for all } \psi \in C_\psi(c_\alpha)\}$ . It follows using a proof analogous to that in Theorem 2.1 in Peterson et al. [25] that

$C_x(c_\alpha) = C_x^*$ . Therefore,  $\mathbf{x}^* \in C_x^*$ . This then implies that there exists a  $\psi^* \in C_\psi(c_\alpha)$  such that  $\mathbf{M}(\mathbf{x}^*)\psi^* = 0$ . If  $h < k$ , then there exists a line or hyperplane  $\mathcal{L}^*$  such that all the  $\mathbf{x}'$ s that satisfy  $\mathbf{M}(\mathbf{x})\psi^* = 0$  are in  $\mathcal{L}^*$ . Hence,  $\mathbf{x}^* \in \mathcal{L}^*$ . Again using the proof of Theorem 2.1 in Peterson et al. [25], it follows that for any given point  $\mathbf{x}$ ,  $Q(\mathbf{x}; \hat{\Psi}) \leq c_\alpha$  if and only if  $\mathbf{M}(\mathbf{x})\psi = 0$  for some  $\psi \in C_\psi(c_\alpha)$ . Therefore,  $Q(\mathbf{x}; \hat{\Psi}) \leq c_\alpha$  for all the  $\mathbf{x}'$ s in  $\mathcal{L}^*$ . In other words,  $\mathcal{L}^*$  satisfies the above two conditions, which is a contradiction.

**References**

- [1] V. N. Nair, "Taguchi's parameter design: a panel discussion," *Technometrics*, vol. 34, no. 2, pp. 127–161, 1992.
- [2] T. Robinson, C. Borrer, and R. H. Myers, "Robust parameter design: a review," *Quality and Reliability Engineering International*, vol. 20, no. 1, pp. 81–101, 2004.
- [3] G. Taguchi, *Introduction to Quality Engineering*, UNIPUB/Kraus International, White Plains, NY, USA, 1986.
- [4] R. H. Myers, A. I. Khuri, and G. G. Vining, "Response surface alternatives to the taguchi robust parameter design approach," *The American Statistician*, vol. 46, no. 2, pp. 131–139, 1992.
- [5] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, *Response Surface Methodology—Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, New York, NY, USA, 3rd edition, 2009.
- [6] R. H. Myers, Y. Kim, and K. L. Griffiths, "Response surface methods and the use of noise variables," *Journal of Quality Technology*, vol. 29, no. 4, pp. 429–440, 1997.
- [7] J. Kunert, C. Auer, M. Erdbrügge, and R. Ewers, "An experiment to compare Taguchi's product array and the combined array," *Journal of Quality Technology*, vol. 39, no. 1, pp. 17–34, 2007.
- [8] T. J. Robinson, S. S. Wulff, D. C. Montgomery, and A. I. Khuri, "Robust parameter design using generalized linear mixed models," *Journal of Quality Technology*, vol. 38, no. 1, pp. 65–75, 2006.
- [9] H. B. Goldfarb, C. M. Borrer, D. C. Montgomery, and C. M. Anderson-Cook, "Using genetic algorithms to generate mixture-process experimental designs involving control and noise variables," *Journal of Quality Technology*, vol. 37, no. 1, pp. 60–74, 2005.
- [10] E. Del Castillo and S. Cahya, "A tool for computing confidence regions on the stationary point of a response surface," *The American Statistician*, vol. 55, no. 4, pp. 358–365, 2001.
- [11] G. E. P. Box and S. Jones, "Design products that are robust to environment," Tech. Rep. 56, Center for Quality and Productivity Improvement, University of Wisconsin, Madison, WI, USA, 1990.
- [12] J. M. Lucas, "How to achieve a robust process using response surface methodology," *Journal of Quality Technology*, vol. 26, no. 4, pp. 248–260, 1994.

- [13] R. G. Miller, *Simultaneous Statistical Inference*, Springer, New York, NY, USA, 2nd edition, 1981.
- [14] N. H. Timm, *Applied Multivariate Analysis*, Springer, New York, NY, USA, 2002.
- [15] A. Cheng, "Confidence Regions for Optimal Control Variables for Robust Parameter Design Experiments," Ph.D.dissertation, Temple University, Department of Statistics, 2011.
- [16] G. P. Y. Clarke, "Approximate confidence limits for a parameter function in nonlinear regression," *Journal of the American Statistical Association*, vol. 82, no. 397, pp. 221–230, 1987.
- [17] S. Kotz and N. L. Johnson, "Multivariate  $t$ -distribution," in *Encyclopedia of Statistical Sciences*, vol. 6, pp. 129–130, John Wiley & Sons, Hoboken, NJ, USA, 1982.
- [18] D. C. Montgomery, *Design and Analysis of Experiments*, John Wiley & Sons, New York, NY, USA, 7th edition, 2009.
- [19] J. B. Kristensen, X. Xu, and H. Mu, "Process optimization using response surface design and pilot plant production of dietary diacylglycerols by lipase-catalyzed glycerolysis," *Journal of Agricultural and Food Chemistry*, vol. 53, no. 18, pp. 7059–7066, 2005.
- [20] J. Wu and M. Hamada, *Experiments-Planning, Analysis, and Parameter Design Optimization*, John Wiley & Sons, New York, NY, USA, 2000.
- [21] J. Singh, D. D. Frey, N. Soderborg, and R. Jugulum, "Compound noise: evaluation as a robust design method," *Quality and Reliability Engineering International*, vol. 23, no. 3, pp. 387–398, 2007.
- [22] X. S. Hou, "On the use of compound noise factor in parameter design experiments," *Applied Stochastic Models in Business and Industry*, vol. 18, no. 3, pp. 225–243, 2002.
- [23] A. C. Rencher, *Multivariate Statistical Inference and Applications*, John Wiley & Sons, New York, NY, USA, 1998.
- [24] C. R. Rao, *Linear Statistical Inference and its Application*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1973.
- [25] J. J. Peterson, S. Cahya, and E. Del Castillo, "A general approach to confidence regions for optimal factor levels of response surfaces," *Biometrics*, vol. 58, no. 2, pp. 422–431, 2002.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

