A Comparison of Three Approaches for Constructing Robust Experimental Designs

Vincent Agboto, Department of Family and Community Medicine

Meharry Medical College, Nashville, TN 37208

Email:vagboto@mmc.edu

William Li and Christopher Nachtsheim

Carlson School of Management

University of Minnesota

Minneapolis, MN 55455

wli@umn.edu

nacht001@umn.edu

Abstract

While optimal designs are commonly used in the design of experiments, the optimality of those designs frequently depends on the form of an assumed model. Several useful criteria have been proposed to reduce such dependence, and efficient designs have been then constructed based on the criteria, often algorithmically. In the model robust design paradigm, a space of possible models is specified and designs are sought that are efficient for all models in the space. The Bayesian criterion given by DuMouchel and Jones (1994), posits a single model that contains both primary and potential terms. In this article we propose a new Bayesian model robustness criterion that combines aspects of both of these approaches. We then evaluate the efficacy of these three alternatives empirically. We conclude that the model robust criteria generally lead to improved robustness; however, the increased robustness can come at a significant cost in terms of computing requirements.

KEY WORDS: Bayesian designs; Model-robust design; D-optimality; Supersaturated design.

1 Introduction

A challenge in the domain of experimental design is the identification of statistical/mathematical criteria that adequately reflect the various goals of a study. These goals may concern (among other things) the ability to detect and estimate active effects, the ability to predict future observations, and/or the ability to assess experimental error. In order to assess the efficacy of the design, the investigator is usually asked to identify the form of the regression model that is to be employed in the analysis. An experimental design is then obtained that is optimal, in some sense, for the specified model and the goals of the experiment. Of course, if the model is wrong, the design will generally not be optimal.

In most applications, the true model is not necessarily known, but can be assumed to be an unknown element of a known set of models. For example, in a one-factor experiment, it may be reasonable to assume that either a linear or a quadratic model will provide a sufficient approximation to the true model. Then the model space consists of two models—the single-factor linear regression model and the single-factor quadratic regression model—and a design can be derived with the requirement that it will be efficient for both models. This is the general approach used to construct model robust designs. Examples of previous work in this realm include Läuter (1974) Cook and Nachtsheim, (1982), Sun (1993), Li and Nachtsheim (2000), Bingham and Li (2002), and Heredia-Langner et al. (2004).

In a pioneering paper, a Bayesian approach to model robustness was introduced by DuMouchel and Jones (1994, "DJ" herein). The DJ approach is based on a single model that contains what DuMouchel and Jones refer to as *primary* and *potential* terms. They derived a simple modification of the Bayesian D-optimality criterion for a single model, and showed how it could be used to construct designs that are generally efficient for estimation of the primary effects and, where possible, provide added facility for estimation of potential effects. This approach was also used by Jones, Lin, and Nachtsheim (2008) to construct efficient supersaturated designs. A key advantage of the DuMouchel and Jones (1994) approach, in contrast with model robust approaches, is ease of computation. The DJ-approach is based on a single model matrix. In contrast, the model robust approach generally requires the computation of information matrices and criterion values for all possible models.

While both approaches have found application in practical settings, the relative advantages and disadvantages have not been assessed. Our purpose was originally to explore the efficacy of these two somewhat disparate approaches to design robustness. In the process of conducting this study, it became apparent that a third approach, which we call Bayesian Model Robust design (BMR) could be suggested. As we show in

Section 2, the BMR criterion can be viewed as a compromise between the Bayesian approach of DJ and the frequentist, model robust design approaches described above. As a result, this paper explores the relative efficacy of these three approaches for design robustness.

The paper is organized as follows. In Section 2, we review in greater detail the three approaches to robustness discussed in this paper. In Section 3, we carry out empirical comparisons for some common design scenarios. Concluding remarks are provided in Section 4.

2 Three approaches to design robustness

In this section we provide greater detail on our implementations of the model robust design approach and the DJ approach, and we propose a Bayesian model robust (BMR) design criterion as an intermediate approach. In what follows, we employ the following notation. A design d having n runs and m factors is represented by the $n \times m$ design matrix $\mathbf{X}(d) = [\mathbf{x}_1, \ldots, \mathbf{x}_n]'$, where the ith row of $\mathbf{X}(d)$ is $(x_{i1}, \ldots, x_{im})$. Let $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_L\}$ denote the finite set of L models under consideration. The $n \times p_l$ model matrix $\mathbf{X}$ for model l is:

$$\mathbf{X}_l(d) = [\mathbf{f}_l(\mathbf{x}_1), \ldots, \mathbf{f}_l(\mathbf{x}_n)]'$$

where the functional $\mathbf{f}_l$ indicates which effects are present in model l. For example, if model l consists of all m main effects and all $\binom{m}{2}$ two-factor interactions, we have:

$$\mathbf{f}_l'(\mathbf{x}_i) = (1, x_{i1}, \ldots, x_{im}, x_{i1}x_{i2}, x_{i1}x_{i3}, \ldots, x_{i,m-1}x_{im})$$

Throughout, we assume that the standard linear model assumptions are valid for at least one model in $\mathcal{F}$:

$$\mathbf{y} = \mathbf{X}_T\boldsymbol{\beta}_T + \boldsymbol{\varepsilon}$$

where T corresponds to the index of the "true" model in $\mathcal{F}$, $\boldsymbol{\beta}_T$ is a $p_T \times 1$ vector of unknown parameters and the error vector $\boldsymbol{\varepsilon}$ has variance-covariance matrix $\sigma^2\mathbf{I}$.

2.1 Model robust design: criteria and model spaces

A design d_{mr} is model-robust if it optimizes some measure of efficacy over the model space $\mathcal{F}$. For example, Li and Nachtsheim (2000) advocate maximization of the average efficiency:

$$\bar{E}_C(d) = \sum_{l=1}^{L} w_l E_l^C(d) \tag{2.1}$$

where $E_l^C(d)$ is the efficiency of design d for model l and design criterion C. Läuter (1974) proposed the S-optimality criterion for approximate designs, ξ. A design ξ_S is S-optimal if it maximizes the average log determinant:

$$\xi_S = \text{argmax}_\xi \sum_{l=1}^{L} w_l \log \mid M_l(\xi) \mid$$

where $M_l(\xi) = \int_\chi \boldsymbol{f}(x)\boldsymbol{f}'(x)d\xi(x)$ is the information matrix for model l, design ξ, and design space χ. Thibodeau (1977) constructed approximate designs that maximize the minimum G-efficiency over $\mathcal{F}$. Srivastava (1975) defined search designs as being those that allow estimation of all effects in one set and up to k elements in another set. Extending Läuter's approach, Cook and Nachtsheim (1982) developed the notion of model robust linear optimal designs. Specifically, they proposed minimization of the average inefficiency for linear criteria of optimality. More recently, Sun (1993) established a methodology for assessing the capability of fractional factorial designs for the estimation of models having main effects and interactions. In doing so, he proposed *Estimation Capacity* (EC) and *Information Capacity* (IC) criteria. The estimation capacity of a design d is defined as the fraction of models in the model space that are estimable for that design. In (2.1), we have $e_l(d) = 1$, if model l is estimable, and zero otherwise. As discussed in Li and Nachtsheim (2000), information capacity is closely related to the Läuter (1974) criterion mentioned above. For IC, we take $e_l(d) = |\boldsymbol{X}_l|/n^{p_L}$ in (2.1). Li and Nachtsheim (2000) proposed a combined (EC,IC)-optimality criterion. A design is optimal by this criterion if it maximizes the IC criterion among all designs having maximum estimation capacity. Bingham and Li (2002) further extended the methodology and constructed (EC,IC)-optimal designs in the context of robust parameter designs.

An experimenter's choice of model space will depend on the application. However it is frequently the case that experimenters have an over-riding interest in a standard model, such as a first-order (screening model), a main-effects plus two-factor interactions model, or a second-order (response surface) model, but they would also prefer that the design has some capability to estimate models with containing a few terms of higher order. In the empirical comparisons to follow, we will focus on the MEPI_g model space of Li and Nachtsheim (2000), and a model space, denoted as $\text{SS}_{n,m}$, that is appropriate for constructing n-run supersaturated designs for m factors. The MEPI_g is defined by:

$$\text{MEPI}_g = \{\text{models with all } m \text{ main effects and any } g \text{ two-factor interactions}\}. \tag{2.2}$$

This model space was motivated by the work of Srivastava (1975), who classified factorial effects into three categories: (i) effects certain to be negligible, (ii) required effects, and (iii) remaining effects which may or may not be active. The model space for Srivastava's search designs consisted of all models involving all effects of type (ii) plus g effects of type (iii). MEPI_g is clearly a special case. MEPI_g has been employed recently by Sun (1993), Li and Nachtsheim (2000), Jones, Li, Nachtsheim, and Ye (2007), and Li (2006) in the construction of model-robust and model-discriminating designs. The number of models in MEPI_g is $n_m = \binom{\binom{m}{2}}{g}$. As described in Section 3.2, the supersaturated model space $\text{SS}_{n,m}$ is defined by:

$$\text{SS}_{n,m} = \{\text{all possible models having } n-1 \text{ or fewer main effects}\}, \tag{2.3}$$

where $n \leq m$. The number of models in $\text{SS}_{n,m}$ is $1 + \sum_{g=1}^{n-1} \binom{m}{g}$. We note that this model space is slightly different from that used in Jones, Li, Nachtsheim, and Ye (2009), who constructed a class of (EC_g,IC_g)-optimal designs, assuming that exactly g out of m factors are active.

2.2 The modified Bayesian approach of DuMouchel and Jones

As noted earlier, the DuMouchel and Jones (1994) approach to robustness focuses on a single model. Their objective was to allow the precise estimation of all primary terms, while providing some estimability for the potential terms. For convenience assume that the model may be written:

$$\boldsymbol{f}'(\boldsymbol{x}) = (\boldsymbol{f}'_{pri}(\boldsymbol{x}), \boldsymbol{f}'_{pot}(\boldsymbol{x})),$$

where $\boldsymbol{f}'_{pri}(\boldsymbol{x})$ and $\boldsymbol{f}'_{pri}(\boldsymbol{x})$ are comprised of the p and q primary and potential model terms, respectively. We similarly partition the model matrix and $\boldsymbol{\beta}$ as:

$$\boldsymbol{X}(d) = [\boldsymbol{X}_{pri}(d), \boldsymbol{X}_{pot}(d)] \qquad \text{and} \qquad \boldsymbol{\beta} = (\boldsymbol{\beta}'_{pri}, \boldsymbol{\beta}'_{pot})'$$

The coefficients of the primary regression coefficients are assumed to have a noninformative prior distribution, with an arbitrary prior mean and a prior variance tending to infinity. By definition, the absolute values of the regression coefficients for potential terms are unlikely to be large, and so they are assumed to have a prior mean of 0 and finite variance $\tau^2\mathbf{I}$. Let $\mathbf{K}$ be the $(p+q)\times(p+q)$ diagonal matrix whose first p diagonal elements are equal to 0 and whose last q elements are 1. Without loss of generality, DJ take $\sigma = 1$ so that $\mathbf{Y} \mid \boldsymbol{\beta} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{I})$. By Bayes's rule for conjugate normal distributions, $\boldsymbol{\beta} \mid \mathbf{Y} \sim N(\mathbf{b}, [\mathbf{X}'\mathbf{X} + \mathbf{K}/\tau^2]^{-1})$ where $\mathbf{b} = [\mathbf{X}'\mathbf{X}+\mathbf{K}/\tau^2]^{-1}\mathbf{X}'\mathbf{Y}$. For a design d, the DJ Bayesian D-optimality criterion can then be defined as:

$$\Phi_{DJ}(d) = \mid \boldsymbol{X}_d'\boldsymbol{X}_d + \mathbf{K}/\tau^2 \mid . \tag{2.4}$$

Clearly, as the prior variance $\tau^2 \to \infty$, the DJ criterion approaches the standard D criterion for a single model comprised of all primary and potential terms. Herein, we use $\tau = 1$ as suggested by DJ.

2.3 Bayesian model robust design

The model robustness criterion (2.1) requires specification of weights, w_i, for each element of the model space. These weights have been loosely motivated as prior model probabilities in the past (see, e.g., Cook and Nachtsheim (1982)) although the model robustness approach is not Bayesian. Agboto (2006) developed a Bayesian model robust approach that, as we shall see, can be viewed as a compromise between the frequentist model robust approach and the modified Bayesian approach of DJ.

Let $p(l)$ denote the prior probability of model $\boldsymbol{f}_l$. Following DJ, we take $\sigma^2 = 1$ and assume that, given model l is the true model, $\boldsymbol{\beta} = \boldsymbol{\beta}_l \sim \mathrm{N}(\mathbf{0}, \boldsymbol{R}_l^{-1})$. Then it is straightforward to show (Appendix 1) that the Bayesian model robust design satisfies:

$$d_{bmr} = \mathrm{argmax}_d \sum_{l=1}^{L} p(l) \log \mid \boldsymbol{X}_l(d)'\boldsymbol{X}_l(d) + \boldsymbol{R}_l \mid . \tag{2.5}$$

If noninformative prior distributions are assumed for the prior distribution of $\boldsymbol{\beta} \mid l$, then the term $\boldsymbol{R}_l$ in (2.5) drops out and the criterion is nearly identical to IC criterion with $p(l) = w_l, l = 1, \ldots, L$. The only distinction is that in the former the criterion corresponds to a weighted average of the logs of the determinants, whereas in the latter, the criterion corresponds to a weighted average of normalized determinants. Also, for a noninformative prior distribution, criterion (2.5) is identical to the S-optimality criterion of Läuter (1974). As noted by Chaloner and Verdinelli (1995), (for a single model) there is no real advantage to the Bayesian approach in the presence of a noninformative prior, and Bayesian D-optimality is equivalent to standard D-optimality in this instance. There is another disadvantage to the use of a noninformative prior here. If, for any design d, the estimation capacity is less than 1, then $d_{bmr} = -\infty$. For this reason, we employ $\boldsymbol{R} = \boldsymbol{I}/c$ with $c > 0$. For numerical work to be reported here, we employ the DJ rationale and fix $c = 1$.

We note that one could further leverage the DJ paradigm within the Bayesian model robust design approach through choice of prior variance covariance matrices $\boldsymbol{R}_l^{-1}$. Partition each of the models in $\mathcal{F}$ into primary and potential terms as follows. Given model l, let $\boldsymbol{R}_{l,pri}^{-1} = \sigma_{l,pri}^2 \boldsymbol{I}_{l,pot}$ and $\boldsymbol{R}_{l,pot}^{-1} = \sigma_{l,pot}^2 \boldsymbol{I}_{l,pot}$, with $\sigma_{l,pri}^2 > \sigma_{l,pot}^2$, so that:

$$\boldsymbol{R}_l^{-1} = \begin{bmatrix} \sigma_{\beta,pri}^2 \boldsymbol{I}_{l,pri} & \mathbf{0} \\ \mathbf{0} & \sigma_{\beta,pot}^2 \boldsymbol{I}_{l,pot} \end{bmatrix}.$$

Although we do not report results here, our limited experience with this criterion suggests that it often leads to designs that were similar to those produced by the DJ approach.

3 Empirical comparisons

In this section we construct robust designs using the three approaches for a variety of design problems and compare the performances. In Section 3.1, we study MEPI$_g$ models spaces, and in Section 3.2 we consider supersaturated model spaces. For simplicity, only two-level designs are considered in this article, although the proposed approach is also applicable to constructing designs whenever model uncertainty is present.

In constructing optimal designs, the coordinate-exchange algorithm of Meyer and Nachtsheim (1995) was used. The algorithm can be summarized as follows: Start with a random n-run design with m factors. The design matrix is updated iteratively by changing rows. For each row, change the sign of the coordinate whose sign change will result in the largest gain in the criterion values among the m coordinates. Update the design matrix accordingly. The procedure continues until no further improvement can be made.

The above procedure was used to obtain optimal designs for all three criteria: BMR of (2.5), DJ of (2.4), and (EC,IC) (which sequentially maximizes EC and IC). In the coordinate-exchange algorithm, there is no

Table 1: Comparison of designs for $n = 8$, $m = 5$, and $g = 1$ (BMR values are for $(c_1, c_2) = (1, 1)$)

Design	EC	IC	BMR	DJ	balance
FFD	.40	.40	14.44	18.14	yes
BMR-optimal	1.00	.84	14.49	18.13	no
(EC,IC)-optimal	1.00	.84	14.49	18.13	no
DJ-optimal	.40	.40	14.44	18.14	yes

guarantee that using a random starting design will lead to convergence to a global optimum. Thus, several starting designs should be used. The "optimal" designs reported in this section were constructed using 50 random starting designs, unless specified otherwise.

3.1 MEPI$_g$ model space

We first consider the MEPI$_g$ space of (2.2), and we construct optimal designs using the three criteria: DJ, BMR, and (EC,IC). As the DJ criterion requires far less computing time than BMR or (EC,IC), an important research question arises: Can the DJ-approach generally produce designs that are just as efficient over all candidate models in the model space as those constructed by the BMR or (EC,IC) approach? Our research results show that this may not necessarily be the case.

We start with a simple 8-run example, in which $m = 5$ factors are considered. Assume that $g = 1$ two-factor interaction is included in the model, and the model space MEPI$_1$ of (2.2) is used. Four types of designs are considered in Table 1: The first design, denoted by FFD, is the fractional factorial design obtained by the generators: $D = AB$, and $E = AC$, where five factors are denoted by A, B, C, D, and E, respectively. It is the maximum resolution design. However, it is not efficient in dealing with the model uncertainty over MEPI$_1$, with EC=40%. Among the three optimal designs, the DJ-optimal design has the same criterion values as the FFD. By comparison, both BMR-optimal and (EC,IC)-optimal designs have EC=100%. It can be argued that the latter designs are preferable from the perspective of model robustness, as they guarantee that all $L = 15$ possible models in the model space are estimable.

We also compared optimal designs generated from the three criteria (DJ, BMR, and (EC,IC)) for a variety

Table 2: Comparison of designs using BMR and DJ criteria (BMR values are for $(c_1, c_2) = (1, 1)$)

(n, m, g)	Design criterion	EC	IC	BMR	DJ
(8,4,2)	BMR	.80	.80	15.07	18.90
	(EC,IC)	1.00	.85	14.59	17.72
	DJ	.80	.80	15.07	18.90
(8,5,2)	BMR	.82	.60	15.96	18.13
	(EC,IC)	.82	.60	15.96	18.13
	DJ	.09	.09	15.63	18.14
(12,8,2)	BMR	.89	.73	26.46	34.13
	(EC,IC)	.97	.76	26.23	33.64
	DJ	.89	.73	26.46	34.13
(16,8,2)	BMR	.89	.89	30.92	54.17
	(EC,IC)	1.00	.88	29.99	50.80
	DJ	.89	.89	30.92	54.17

of design scenarios. Some selected results are reported in Table 2. We first note that, for all the examples considered in Table 2, the BMR-optimal designs converge to either DJ-optimal designs or (EC,IC)-optimal design. For instance, in case of $n = 8$ and $m = 5$, if $g = 2$ two-factor interactions are contained in the models over the MEPI$_2$ model space, then the BMR-optimal design has (EC,IC)=(.82,.60), the same as that of the (EC,IC)-optimal design. In comparison, the DJ-optimal design has a much smaller estimation capacity value (EC=.09). In another example, in which $(n, m, g) = (8, 4, 2)$, however, the BMR-optimal design has the same criterion value as the DJ-optimal design.

The last example considered in Table 2 was also used by DuMouchel and Jones (1994, Example 4 on Page 42). This design can be used as a screening experiment for $m = 8$ factors with $n = 16$ runs. Suppose that up to $g = 2$ two-factor interactions are contained in the model. As pointed out by DuMouchel and Jones (1994), the DJ-optimal design actually converges to the resolution IV 2^{8-4} design. Interestingly, the corresponding BMR-optimal design also converges to the same design, for which EC= 89%. In comparison, the (EC,IC)-optimal design has EC= 100%.

The results in Table 2 show that the the added computational effort required by the model robust approach can lead to substantive gains in robustness as measured by estimation capacity. As noted by DuMouchel and Jones (1994), when the primary terms are comprised of all main effects and potential terms are comprised of all two-factor interactions, the DJ-approach tends to favor resolution IV designs (and, in a more general sense, maximum resolution designs). However, it has been shown previously (see, for instance, Li and Nachtsheim (2000)) that the maximum resolution design may not perform well in terms of model robust criteria such as EC and IC. The BMR and (EC,IC) approaches, which explicitly consider the efficacy of all possible models, may lead to designs having better model robustness properties.

3.2 Supersaturated design model space $\mathrm{SS}_{n,m}$

We now compare use of the BMR, (EC,IC), and DJ criteria for constructing supersaturated designs. In a supersaturated design with n runs and m factor, we have $n \leq m$. Thus it is not possible to estimate all m main effects. A supersaturated design can be useful at the screening stage, with the assumption that only a few of the m factors are active. In Table 3 we compare the three classes of SS designs over the model space $\mathrm{SS}_{n,m}$ of (2.3). For simplicity and consistency we continue to use DJ to denote the criterion of Jones, Lin, and Nachtsheim (2008), who advocated the use of the DJ criterion for constructing Bayesian SS designs. We can draw similar conclusions from Table 3 as those from Table 2 for the MEPI model space. Model robust criteria such as (EC,IC) generally lead to better robustness properties. Among the six examples for various (n, m) values, all (EC,IC)-optimal designs have a higher estimation capacity value than the corresponding DJ-optimal design. For instance, in the first example of $(n, m) = (6, 8)$, the (EC,IC)-optimal design has EC = .9862, which is significantly higher than the EC of .7385 for the DJ-optimal design. The difference between the DJ-criterion values of the two designs appear to be small (7.8883 vs. 8.1630). Note that both DJ and BMR criteria values are dependent on a tuning parameter value c. It was noted in Jones, Lin, and Nachtsheim (2008) that optimal designs are relatively insensitive to the choice of prior variance τ^2 and thus $\tau^2 = 5$ was employed. We observed similar phenomena and used $c = 1/\tau^2 = .2$ for all calculations in Table 3.

The results in Table 3 show that the BMR criterion, as expected, led to designs that perform between (EC,IC) and DJ criteria. For the six examples considered in the table, the BMR-optimal designs are closer to DJ-optimal designs than to (EC,IC)-optimal designs. More specifically, BMR-optimal designs have the same EC values as the corresponding DJ-optimal designs in three cases: $(n, m) = (6, 10)$, (8,10), and (10,15). In the other three cases, they have higher EC values than DJ-optimal designs.

Table 3: Comparison of designs using BMR and DJ criteria ($c = .2$ for both BMR and DJ)

(n, m)	Design criterion	EC	IC	BMR	DJ
(6,8)	BMR	.9312	.3521	7.5054	8.1341
	(EC,IC)	.9862	.3627	7.3342	7.8883
	DJ	.7385	.2821	7.3025	8.1630
(6,10)	BMR	.8352	.2840	7.9436	6.2517
	(EC,IC)	.9686	.3134	7.6332	5.9326
	DJ	.8352	.2840	7.9436	6.2517
(8,10)	BMR	.9535	.3276	11.1821	14.0843
	(EC,IC)	.9990	.3216	10.6364	13.0006
	DJ	.9535	.3276	11.1821	14.0843
(8,12)	BMR	.9727	.2940	11.7981	12.1850
	(EC,IC)	.9897	.2900	11.4679	11.6356
	DJ	.9476	.2872	11.7963	12.2285
(10,12)	BMR	.9674	.3094	15.0353	20.5695
	(EC,IC)	1.0000	.3047	15.8523	19.5698
	DJ	.9589	.3048	15.0137	20.6209
(10,15)	BMR	.9904	.2700	16.3986	17.8790
	(EC,IC)	.9960	.2611	15.8523	16.9605
	DJ	.9904	.2700	16.3986	17.8790

4 Conclusion

In this article, we have empirically evaluated the performance of three approaches for constructing robust experimental designs in the presence of model uncertainty. Model spaces employed included the MEPI_g space (2.2) and the "supersaturated" model space $\text{SS}_{n,m}$ defined in (2.3). We conclude that use of the model robust paradigm, as advocated by Li and Nachtsheim (2000) can lead, at times, to substantial improvements in estimation capacity when compared to the modified Bayesian approach of DuMouchel and Jones (1994). Nonetheless, in many instances, these seemingly disparate approaches to robustness produce the same designs. The new BMR approach introduced herein, appears to provide a compromise approach. The DJ approach has a clear advantage in terms of computational efficiency. Times required to construct model robust designs—using either the (EC,IC) or the BMR criteria—can be two to three orders of magnitude larger than those required by the DJ approach for problems considered here. We therefore recommend use of model robust designs when computing time is not an issue, or when the need to achieve estimation capacity near 100% is paramount. In other circumstances, the DJ approach provides a very fast and effective alternative.

5 References

Agboto, V., 2006. *Bayesian approaches to model robust and model discrimination designs*, unpublished Ph.D. dissertation, School of Statistics, University of Minnesota.

Bingham, D., Li, W., 2002. A class of optimal robust parameter designs. *Journal of Quality and Technology* 34, 244-259.

Chaloner, K., Verdinelli, I., 1995. Bayesian experimental design: a review. *Statist. Sci.* 10, 273-304.

Cook, R. D., Nachtsheim, C. J., 1982. Model robust, linear-optimal designs. *Technometrics* 24, 49-54.

DuMouchel, W., Jones, B., 1994, A simple Bayesian modification of D-optimal designs to reduce dependence on an assumed model. *Technometrics* 36, 37-47.

Heredia-Langner, A., Montgomery, D. C., Carlyle, W. M., Borror, C. M., 2004. Model-robust optimal designs: A genetic algorithm approach. *Journal of Quality Technology* 36, 263-279.

Jones, B., Lin, D. K. J., Nachtsheim, C., J., 2008. Bayesian D-optimal supersaturated designs. *Journal of Statistical Planning and Inference* 138, 86-92.

Jones, B., Li, W., and Nachtsheim, C. J., and Ye, K., 2007. Model Discrimination - Another Perspective of Model-Robust Designs. *Journal of Statistical Planning and Inference* 137, 1576-1583.

Jones, B., Li, W., Nachtsheim, C. J., Ye, K., 2009. Model-robust supersaturated and partially supersaturated designs. *Journal of Statistical Planning and Inference* 139, 45-53.

Läuter, E., 1974. Experimental design in a class of models. *Math. Operationsforsch. Statist.* 5, 379-396.

Li, W., 2006, Screening Designs for Model Selection. In *Screening: Methods for Experimentation in Industry, Drug Discovery and Genetics*, Editors: Dean, A. M. and Lewis, S. M, Chapter 10, Springer Verlag.

Li, W., Nachtsheim, C. J., 2000. Model-robust factorial designs. *Technometrics* 42, 379-396.

Meyer R. K., Nachtsheim C. J., 1995. The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* 37, 6069.

Sun, D. X., 1993. *Estimation Capacity and Related Topics in Experimental Designs*, unpublished PhD dissertation, Department of Statistics and Actuarial Science, University of Waterloo.

Srivastava, J. N., 1975. Designs for Searching Non-negligible Effects, in *A Survey of Statistical Designs and Linear Models* (ed. J.N. Srivastava), Amsterdam: North Holland, pp. 507–519.

Thibodeau, L.A., 1977. *Robust design for regression problems*, unpublished PhD dissertation, Department of Applied Statistics, University of Minnesota.

Appendix A: Bayesian Model Robust Design Criterion

The prior distribution of the parameter $\boldsymbol{\beta}$ given model i is given by $\boldsymbol{\beta} \mid i \sim N(\boldsymbol{\beta}_{0i},\ \sigma^2\mathbf{R}_i^{-1})$, where $\boldsymbol{\beta}_{0i}$ is a $k_i \times 1$ vector, and the $k_i \times k_i$ matrix $\mathbf{R}_i$ is known. The posterior distribution of $\boldsymbol{\beta}$ given response $\boldsymbol{y}$, model i, and design d, is $\boldsymbol{\beta} \mid \boldsymbol{y}, i, d \sim N(\boldsymbol{\beta}_i^*, \sigma^2\mathbf{D}_i)$ where $\boldsymbol{\beta}_i^* = (\mathbf{X}'_{i,d}\mathbf{X}_{i,d} + \mathbf{R}_i)^{-1}(\mathbf{X}'_i\mathbf{y} + \mathbf{R}_i\boldsymbol{\beta}_{0i})$ and $\sigma^2\mathbf{D}_i = \sigma^2(\mathbf{X}'_{i,d}\mathbf{X}_{i,d} + \mathbf{R}_i)^{-1}$. We have the requirement that $\sum_{l=1}^{d} w_l = 1$. Let μ be a counting measure associated with the variable l. If we employ the expected Kullback divergence measure as in the standard Bayesian design approach (Chaloner and Verdinelli, 1995), the expected utility for the class of linear models considered is:

$$\begin{aligned}
U(d) &= \int\int\int \log p(\boldsymbol{\beta} \mid \mathbf{y}, i, d) p(\mathbf{y}, \boldsymbol{\beta}, i \mid d) d\boldsymbol{\beta} d\mathbf{y} \mu(di) \\
&= \int\int\int \log p(\boldsymbol{\beta} \mid \mathbf{y}, i, d) p(\mathbf{y}, \boldsymbol{\beta} \mid i, d) p(i) d\boldsymbol{\beta} d\mathbf{y} \mu(di) \\
&= \sum_{i=1}^{d} p(i) \int\int \log p(\boldsymbol{\beta} \mid \mathbf{y}, i, d) p(\mathbf{y}, \boldsymbol{\beta} \mid i, d) d\boldsymbol{\beta} d\mathbf{y} \\
&= \sum_{i=1}^{d} p(i) \{-\frac{k_i}{2}\log(2\pi) - \frac{k_i}{2} + \frac{1}{2}\log\det(\sigma^{-2}(\mathbf{X}'_{i,d}\mathbf{X}_{i,d} + \mathbf{R}_i))\} \\
&= \sum_{i=1}^{d} p(i) \log\{\det(\sigma^{-2}(\mathbf{X}'_{i,d}\mathbf{X}_{i,d} + \mathbf{R}_i))\} + \sum_{i=1}^{d} p(i) \{\frac{k_i}{2} + \frac{k_i}{2}\log(2\pi)\}.
\end{aligned}$$

The second expression of the equation is a constant, therefore the criterion is reduced then to maximizing the function:

$$\phi_{\mathrm{BMR}}(d) = \sum_{i=1}^{d} p(i) \log \mid \sigma^{-2}(\mathbf{X}'_{i,d}\mathbf{X}_{i,d} + \mathbf{R}_i) \mid . \tag{5.1}$$