OPTIMAL DESIGNS FOR PARALLEL MODELS WITH CORRELATED RESPONSES

Mong-Na Lo Huang¹, Ray-Bing Chen², Chun-Sui Lin¹ and Weng Kee Wong³

¹National Sun Yat-sen University, ²National University of Kaohsiung and ³University of California at Los Angeles

Abstract: We consider a parallel linear model with correlated dual responses on a symmetric compact design region and construct locally $D$-optimal designs for estimating the unknown parameters in the model, and locally optimal designs for estimating the location shift parameter. The $D$-optimal designs for the additive model are invariant under linear transformation of the design space but locally optimal designs for estimating the location shift do not share this property. The latter optimal designs depend on the correlation between the dual responses in an interesting and sensitive way.

Key words and phrases: approximate design, biosassay, $D$-optimality, locally optimal design, location shift parameter.

1. Introduction

Consider a bioassay experiment that measures a response from different doses of the standard and test preparations. The interest is in estimating the potency of the test preparation relative to the standard, which by definition is the amount of the standard equivalent in effect to one unit of the test. Specifically, suppose that the dose interval of interest is $[a, b]$ and a dose from this interval is administered to an experimental unit. The response $y$ at this dose level, $d$, is measured and its expectation under the standard preparation is $E(y_1|d) = F_1(d), \forall d \in [a, b]$, where $F_1$ is some known functional with unknown parameters. Suppose, as is often the case in bioassay experiments, the expected response for the test preparation is $E(y_2|d) = F_2(d) = F_1(\tau d), \forall d \in [a, b]$, and $\tau$ is an unknown constant representing the relative potency between the standard and test preparations.

It is common practice to assume the regression function $F_1(d)$ is linearly related to $x = \log(d)$, see Finney (1978) for example. This implies

\begin{align*}
E(y_1|d) &= F_1(d) = \theta_0 + \theta_1 \log(d) = \theta_0 + \theta_1 x, \\
E(y_2|d) &= F_1(\tau d) = \theta_0 + \theta_1 (\log(d) + \log(\tau)) = \theta_0 + \theta_1 (x - \mu),
\end{align*}

where $\mu = -\log(\tau)$. Therefore, these two simple linear models are parallel with common slope $\theta_1$. The covariance matrix between the two responses from the
standard and test preparations is $\text{Cov}(y_1, y_2) = \Sigma = \sigma^2((1 - \rho)I_2 + \rho J_2)$, where $I_2$ is the $2 \times 2$ identity matrix, $J_2$ is the $2 \times 2$ matrix of one’s, and without loss of generality assume that $\sigma^2=1$. We also assume throughout that all models in the paper satisfy the parallelism assumption. Some test procedures for testing the hypothesis of parallelism are given in Smith and Choi (1982).

There is much research in bioassays, see for example Govindrajulu (2001) and Kshirsagar and Yuan (1992). Design papers for general bioassays are relatively scarce and they include Buonaccorsi (1986), Finney (1978), Kshirsagar and Yuan (1992), and Smith and Ridout (2003). Chai, Das and Dey (2001) and Kshirsagar and Yuan (1992) were among the few who addressed specific design issues for parallel line bioassays. Their interest, however, was in incomplete block designs, which is not the focus here.

This paper proposes optimal designs for a parallel line bioassay experiment when the responses from the standard and test preparations may be correlated. Such assumptions are realistic if observations come from the same litter, or observations are made from the same subjects under two experimental conditions. We provide closed form formulae for optimal designs for estimating model parameters, and optimal designs for estimating the relative potency.

We follow Kiefer’s approach and focus on continuous designs. A continuous design $\xi$ is a probability measure with a finite number of support points on a given compact design space. Throughout we assume that, after appropriate scaling, the design space is $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ and $\mathcal{X}_1 = \mathcal{X}_2 = [-1, 1]$. If the design has all its mass at the point $x$, we denote the design by $\delta_x$. A generic design on $m$ points is denoted by $\xi = p_1 \delta_{x_1} + p_2 \delta_{x_2} + \cdots + p_m \delta_{x_m}$, where each $x_i \in \mathcal{X}$, $\xi(x_i) = p_i > 0$, and $\sum_{i=1}^{m} p_i = 1$. A main advantage of continuous designs is that it can be readily verified if they are optimum among all designs on the design space $\mathcal{X}$ using equivalence theorems. Details of the continuous design framework and equivalence theorems are discussed in design monographs, see Pukelsheim (1993) or Fedorov (1972), for example.

In the next section we discuss optimal experimental design problems for estimating the unknown parameters in the model. Section 3 discusses optimal designs for estimating the logarithm of the relative potency. This parameter is important because it is widely used to measure the location-shift between the standard and test preparations in parallel line assays. Section 4 provides an application and a discussion.

2. Parameter Estimation

Throughout we focus on the parallel model with dose as the control variable, the dose level $x_1$ for the standard preparation and the dose level $x_2$ for the test
preparation may be different. Specifically, for each design point \( x = (x_1, x_2) \in \mathcal{X} = [-1, 1] \times [-1, 1] = [-1, 1]^2 \), we have

\[
\begin{align*}
E(y_1|x_1) &= \theta_{01} + \theta_1 x_1 \\
E(y_2|x_2) &= \theta_{02} + \theta_1 x_2.
\end{align*}
\] (1)

The model has three parameters and its mean function is given by \( F(x)^T = (I_2 X) \), where \( X = (x_1, x_2)^T \). Following convention, we measure the worth of a design \( \xi \) by its information matrix:

\[
M(\xi) = \int_{\mathcal{X}} F(x)\Sigma^{-1} F(x)^T d\xi(x),
\] (2)

where \( \Sigma \) is the covariance matrix for the dual responses.

Popular optimal designs used in practice are \( D \)-optimal designs and \( c \)-optimal designs. \( D \)-optimal designs are useful for parameter estimation because they minimize the generalized variance and therefore have the smallest volume in the confidence ellipsoid for the model parameters. Mathematically, a design \( \xi^* \) is called a \( D \)-optimal design if \( |M(\xi^*)| \geq |M(\xi)| \) for all \( \xi \) defined on the design space \( \mathcal{X} \).

If interest is centered on estimating a given function of the model parameters, say \( c(\theta) \), \( c \)-optimal designs are used because they provide the smallest asymptotic variance of the estimate. The variance of the estimated function has the form \( \hat{c}(\theta)^T M(\xi)^{-1} \hat{c}(\theta) \), where \( \hat{c}(\theta) \) is the derivative function of \( c \) with respect to \( \theta \). Because the optimality criterion contains parameters that need to be estimated, our optimal designs are locally optimal. This means that for all our problems the parameter \( \rho \) is assumed to be fixed and known. This assumption is plausible when there are previous studies or expert opinions about the possible value of the parameter \( \rho \). The following notation will be used repeatedly when we have two factors: \( x_1 \) and \( x_2 \) are defined on the design space \( \mathcal{X} \); for a given design \( \xi \) on \( \mathcal{X} \), \( c_i = \int_{\mathcal{X}} x_i^1 d\xi \), \( d_i = \int_{\mathcal{X}} x_i^2 d\xi \), \( i = 1, 2 \), and \( \gamma = \int_{\mathcal{X}} x_1 x_2 d\xi \).

**Theorem 2.1.** Suppose (11) holds. The \( D \)-optimal design on \( \mathcal{X} \) for estimating parameters is (i) \( \xi^* = (1/2)\delta_{(-1,1)} + (1/2)\delta_{(1,-1)} \) if \( \rho > 0 \), and (ii) \( \xi^* = (1/2)\delta_{(-1,-1)} + (1/2)\delta_{(1,1)} \) if \( \rho < 0 \).

**Proof.** It is straightforward to verify that

\[
M(\xi) = \frac{1}{1 - \rho^2} \begin{pmatrix}
1 & -\rho & c_1 - \rho d_1 \\
-\rho & 1 & d_1 - \rho c_1 \\
(c_1 - \rho d_1) & (d_1 - \rho c_1) & c_2 + d_2 - 2\rho \gamma
\end{pmatrix}.
\] (3)
Using the formula for the determinant of a partitioned matrix, we have
\[
\det(M(\xi)) = \frac{c_2 + d_2 - 2\rho\gamma}{1 - \rho^2} - \frac{(c_1 d_1)\Sigma^{-1}(c_1 d_1)^T}{(1 - \rho^2)}
\]
\[
\leq \frac{c_2 + d_2 - 2\rho\gamma}{(1 - \rho^2)^2} \leq 2 - 2\rho\gamma.
\]

The above inequality shows that we may restrict attention to designs with
\[c_2 = d_2 = 1\]. Accordingly, we may consider designs with supports on the four corner
points \((1,1), (1,-1), (-1,1), (-1,-1)\) with weight \(\alpha_1, \alpha_2, \alpha_3, \alpha_4\),
respectively, and each \(\alpha_i \geq 0\) and \(\sum_{i=1}^{4} \alpha_i = 1\). If we let \(\gamma = \alpha_1 + \alpha_4 - \alpha_2 - \alpha_3\),
we observe that maximizing the determinant over all such designs is the same as
minimizing \(\rho\gamma\). Therefore, when \(\rho \geq 0\), the optimal design has to satisfy \(\gamma = -1\) and this means the design is equally supported on \((-1,1), (1,-1)\). When \(\rho < 0\),
the optimal design has to satisfy \(\gamma = 1\) and this implies the design is equally
supported on \((-1,-1), (1,1)\).

The parallel model may be generalized to include \(p\) control variables, \(p \geq 2\), and the dose levels for the two preparations may be different. Let \(x_{i,p} = (x_{i1}, \ldots, x_{ip})^T \in \mathcal{X}_{i}^p = [-1,1]^p, i = 1, 2\), and let \(x^T = (x_{1,p}^T, x_{2,p}^T)\) be a design
point defined on \(\mathcal{X}^p = \mathcal{X}_1^p \times \mathcal{X}_2^p\). The model is
\[
E(y_i | x_{i,p}) = \theta_{0i} + \sum_{k=1}^{p} \theta_k x_{ik}, \quad i = 1, 2, \quad (4)
\]
with covariance matrix \(\Sigma\) for the dual responses. In the following theorem, we
let \(\xi^*_k\) be a \(D\)-optimal design for the model \(E(y_{ik} | x_{ik}) = \theta_{0i} + \theta_k x_{ik}, i = 1, 2,\) and \(k = 1, \ldots, p\).

**Theorem 2.2.** Consider (4) on \(\mathcal{X}^p\) with \(p \geq 2\). The product design \(\xi^* =
\xi_1^* \otimes \cdots \otimes \xi_p^*\) is \(D\)-optimal for estimating \(\theta = (\theta_{01}, \theta_{02}, \theta_{11}, \ldots, \theta_{p})^T\).

**Proof.** Under (4), we have \(F(x)^T = (I_2 \ X)\) where \(X = (x_{1,p} \ x_{2,p})^T\). Let
\[
U = \begin{pmatrix} I_2 - \int_{\mathcal{X}^p} X d\xi \\ 0 \end{pmatrix}
\]
and let \(\tilde{M}(\xi) = U^T M(\xi) U\). It can be shown that
\[
\max_{\xi} |M(\xi)| = \max_{\xi} |\tilde{M}(\xi)| = \max_{\xi} \int_{\mathcal{X}^p} X^T \Sigma^{-1} X d\xi - \left( \int_{\mathcal{X}^p} X^T d\xi \right) \Sigma^{-1} \left( \int_{\mathcal{X}^p} X d\xi \right).
\]
Let \(\tilde{X} = X - \int_{\mathcal{X}^p} X d\xi\), and let \(\tilde{x}_{ij} = (\tilde{x}_{1j}, \tilde{x}_{2j})^T, j = 1, \ldots, p\). It follows that
\[
\int_{\mathcal{X}^p} X^T \Sigma^{-1} X d\xi - \left( \int_{\mathcal{X}^p} X^T d\xi \right) \Sigma^{-1} \left( \int_{\mathcal{X}^p} X d\xi \right) = \int_{\mathcal{X}^p} \tilde{X}^T \Sigma^{-1} \tilde{X} d\xi.
\]
Because the determinant of a positive-semidefinite matrix is less than or equal to the absolute value of the product of the diagonal elements, it follows that
\[
|\int_{X^p} \tilde{X}^T \Sigma^{-1} \tilde{X} d\xi| \leq (\int_{X^p} \tilde{x}_i^T \Sigma^{-1} \tilde{x}_i d\xi) \cdots (\int_{X^p} \tilde{x}_p^T \Sigma^{-1} \tilde{x}_p d\xi),
\]
with equality if \( \int_{X^p} \tilde{x}_i^T \Sigma^{-1} \tilde{x}_i d\xi = 0, \forall i \neq j \). Hence the product design \( \xi^* = \xi_1^* \times \cdots \times \xi_p^* \) is a \( D \)-optimal design for model (1).

**Example 1.** Suppose \( \Xi \) holds with \( p = 2 \) on \( \mathcal{X}^2 \), and that \( \rho \geq 0 \). The \( D \)-optimal design on each factor space is \( \xi_1^* \) and \( \xi_2^* \) where both are equally supported at points \((-1,1)\) and \((1,-1)\). Theorem 2.2 implies that the design \( \xi_1^* \otimes \xi_2^* = (1/4)\delta_{(-1,-1,1,1)} + (1/4)\delta_{(-1,1,1,-1)} + (1/4)\delta_{(1,-1,-1,1)} + (1/4)\delta_{(1,1,-1,-1)} \) is \( D \)-optimal for estimating \( \theta = (\theta_{01}, \theta_{02}, \theta_1, \theta_2)^T \).

**3. Location-Shift Parameter**

In this section, we consider optimal designs for estimating the location-shift parameter \( \mu \) in model (1), that is,
\[
\begin{align*}
E(y_1|x_1) &= \theta_{01} + \theta_1 x_1 \\
E(y_2|x_2) &= \theta_{01} + \theta_1 (x_2 - \mu).
\end{align*}
\]
(5)
The location-shift parameter \( \mu \) can be expressed as
\[
\mu = \frac{\theta_{01} - \theta_{02}}{\theta_1} = \frac{l_1^T \theta}{l_2^T \theta} = \frac{\beta_1}{\beta_2},
\]
where \( \theta = (\theta_{01}, \theta_{02}, \theta_1)^T \), \( l_1 = (1,-1,0)^T \), \( l_2 = (0,0,1)^T \), \( \beta_1 = l_1^T \theta \), and \( \beta_2 = l_2^T \theta \). We have
\[
\text{Cov}(\hat{\beta}) \propto \begin{pmatrix} l_1^T M(\xi)^{-1} l_1 & l_1^T M(\xi)^{-1} l_2 \\ l_2^T M(\xi)^{-1} l_1 & l_2^T M(\xi)^{-1} l_2 \end{pmatrix} = LM(\xi)^{-1} L^T,
\]
where \( \beta = (\beta_1, \beta_2)^T \) and \( L^T = (l_1, l_2) \). By McDonald and Studden (1990), the approximate variance of the ratio of the two estimated parameters is
\[
\text{Var}(\hat{\beta}_1/\hat{\beta}_2) \propto (h_1, h_2) \dot{L}(\xi)^{-1} L^T(h_1, h_2)^T,
\]
where \( h(\beta_1, \beta_2) = \beta_1/\beta_2 \), and \( h_i = \partial h/\partial \beta_i, i = 1, 2 \). In our case, we have
\[
\text{Var}(\hat{\beta}_1/\hat{\beta}_2) \propto (1,-\beta_1/\beta_2) \dot{L}(\xi)^{-1} L^T(1,-\beta_1/\beta_2)^T = c^T M(\xi)^{-1} c,
\]
with \( c = (1,-1,-\mu)^T \). This means the best design for estimating \( \mu \) is a locally \( c \)-optimal design. Here and throughout the rest of the paper, we construct locally \( c \)-optimal designs on the design space \( \mathcal{X} \), and note that the design set-up requires the dosage levels for both preparations be on the logarithmic scale and appropriately standardized. Optimal designs for estimating \( \mu \) on other design spaces will have to be reconstructed because, unlike \( D \)-optimal designs, these designs are not invariant under linear transformation on the design space.
When \( \rho = 0 \), the two responses are uncorrelated and they do not have to be observed in pairs. We may thus relax our designs to include different numbers of observations, say \( n_1, n_2 \) for the two responses, respectively. In this case, designs \( \xi \) for such a set-up can be expressed as \( \xi = p_1 \xi_1 + p_2 \xi_2 \), where \( p_1 = n_1/n, p_2 = n_2/n, n = n_1 + n_2 \), and \( \xi_i^{(n)} \), \( i = 1, 2 \), represents the design for the \( i \)th response on \([-1,1]\). Theorems 3.1 and 3.2 below present locally optimal designs for estimating \( \mu \) when the two responses are uncorrelated. Many of the optimal designs are found by first restricting attention to a subclass of designs and, among these designs, determining the smallest non-trivial lower bound for the determinant of the inverse of the information matrix, or the variance of the estimate of interest. The optimal design is then found by constructing a design that attains the lower bound. Recall that \( c_i = \int_{X_1} x^i d\xi_1 \) and \( d_i = \int_{X_2} x^i d\xi_2 \).

**Theorem 3.1.** Suppose \( |\mu| \leq 2 \), and \( \xi_1 \) and \( \xi_2 \) are two designs supported on \([-1,1]\). The design \( \xi^* = (1/2)\xi_1 + (1/2)\xi_2 \) is a locally optimal design for estimating \( \mu \) provided \( d_1 - c_1 = \mu \).

**Proof.** When \( \rho = 0 \), a direct calculation shows

\[
M(\xi) = \begin{pmatrix} p_1 & 0 & p_1 c_1 \\
0 & p_2 & p_2 d_1 \\
p_1 c_1 & p_2 d_1 & p_1 c_2 + p_2 d_2 \end{pmatrix}.
\]

Recalling that \( c = (1, -1, -\mu)^T \), we have

\[
c^T M(\xi)^{-1} c = \frac{1}{p_1} + \frac{1}{p_2} + \frac{(\mu - (d_1 - c_1))^2}{p_1 c_2 + p_2 d_2 - p_1 c_1^2 - p_2 d_1^2}. \tag{6}
\]

For any two designs \( \xi_1 \) and \( \xi_2 \) on \([-1,1]\), we have \( |d_1 - c_1| \leq 2 \) and \( 1/p_1 + 1/p_2 \geq 4 \). This means that, from (6), we can find a design \( \xi^* \) such that \( p_1 = p_2 = 1/2 \) and \( \mu = d_1 - c_1 \).

**Theorem 3.2.** Suppose \( |\mu| > 2 \). If designs \( \xi_1^* \) and \( \xi_2^* \) are supported on \([-1,1]\) and if \( \xi^* = p_1 \xi_1^{(n)} + p_2 \xi_2^{(n)} \) satisfies (i) \( 1/|\mu| < p_1 < 1 - 1/|\mu| \), and (ii) \(-\mu c_1 p_1 = \mu d_1 p_2 = 1 \), where \( p_2 = 1 - p_1 \), then \( \xi^* \) is a locally optimal design for estimating \( \mu \). Moreover, \( c^T M(\xi^*)^{-1} c = \mu^2 \).

**Proof.** It is straightforward to verify from (6) that

\[
c^T M(\xi)^{-1} c \geq \frac{1}{p_1} + \frac{1}{p_2} + \frac{(\mu - (d_1 - c_1))^2}{1 - p_1 c_1^2 - p_2 d_1^2},
\]

with equality if the design \( \xi \) is supported on \([-1,1]\). In particular, equality is attained for the optimal designs \( \xi_1^* \) and \( \xi_2^* \). For \( 1/|\mu| < p_1 < 1 - 1/|\mu| \), define

\[
h(p_1, c_1, d_1) = \frac{1}{p_1} + \frac{1}{p_2} + \frac{(\mu - (d_1 - c_1))^2}{1 - p_1 c_1^2 - p_2 d_1^2}.
\]
If we take partial derivatives of the function $h$ with respect to $c_1$ and $d_1$ and set them equal to 0, we have

$$\mu = \frac{[d_1(p_1c_1 + p_2d_1) - 1]}{(p_1c_1)} = \frac{[1 - c_1(p_1c_1 + p_2d_1)]}{(p_2d_1)}.$$ 

It follows that $p_1c_1 + p_2d_1 = (p_1c_1 + p_2d_1)(p_1c_1 + p_2d_1)$ and, because $p_1c_1 + p_2d_1 \neq 1$, we must have $p_1c_1 + p_2d_1 = 0$. By assumption, it follows that $-\mu p_1c_1 = 1 = \mu p_2d_1$, and the optimal design $\xi^*$ satisfies $c^T M(\xi^*)^{-1}c = h(p_1, -1/(\mu p_1), 1/(\mu p_2)) = \mu^2$.

**Example 2.** Suppose (1) holds, $\rho = 0$, and $\mu = 3$. If we take an equal number of observations from the test and standard preparations, i.e., $p_1 = p_2 = 1/2$, and use designs $\xi^*_1 = (5/6)\delta_{-1} + (1/6)\delta_1$ for the standard preparation, $\xi^*_2 = (1/6)\delta_{-1} + (5/6)\delta_1$ for the test preparation, we have $1/3 < p_1 < 2/3$, $c_1 = -2/3$, $d_1 = 2/3$, and condition (ii) of the theorem holds. It follows that the average of these two designs, the design equally supported at $\pm 1$, is locally optimal for estimating $\mu$.

Table 1 and Table 2 display selected optimal designs constructed from Theorems 3.1 and 3.2. For example, in Table 1, the third row shows that when $\mu = -0.5$, the designs for the two preparations are $\xi^*_1 = \delta_1$ and $\xi^*_2 = 0.25\delta_{-1} + 0.75\delta_1$. In addition, they have the property that $d_1 - c_1 = 0.5 - 1 = -0.5 = \mu$, and consequently, the design $\xi^* = (1/2)\xi^*_1 + (1/2)\xi^*_2$ is locally optimal for estimating $\mu$. Alternatively, if we take $\xi^*_1 = 0.75\delta_{-1} + 0.25\delta_1$ and $\xi^*_2 = \delta_{-1}$ as shown in the fourth row, the design $\xi^* = (1/2)\xi^*_1 + (1/2)\xi^*_2$ also satisfies $d_1 - c_1 = -1 - (-0.5) = -0.5 = \mu$, and hence is also locally optimal for estimating $\mu$.

Table 1. Designs for constructing optimal designs for model (5) with $\Sigma = I_2$ and a given $\mu, |\mu| \leq 2$, using Theorem 3.1.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>Design points of $\xi^*_1$</th>
<th>Design points of $\xi^*_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.5</td>
<td>0.000 0.500 0.125 0.375</td>
<td>0.375 0.125 0.500 0.000</td>
</tr>
<tr>
<td>0.0</td>
<td>0.500 0.000 0.500 0.000</td>
<td>0.000 0.500 0.000 0.500</td>
</tr>
<tr>
<td>0.5</td>
<td>0.500 0.000 0.375 0.125</td>
<td>0.125 0.375 0.000 0.500</td>
</tr>
<tr>
<td>1.0</td>
<td>0.500 0.000 0.250 0.250</td>
<td>0.250 0.250 0.000 0.500</td>
</tr>
<tr>
<td>1.5</td>
<td>0.500 0.000 0.125 0.375</td>
<td>0.375 0.125 0.000 0.500</td>
</tr>
<tr>
<td>2.0</td>
<td>0.500 0.000 0.000 0.500</td>
<td>0.500 0.000 0.000 0.500</td>
</tr>
</tbody>
</table>
Table 2. Designs for constructing optimal designs for model (5) with $\Sigma = I_2$ and a given $\mu, |\mu| > 2$, using Theorem 3.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>Design points of $\xi_1^*$</th>
<th>Design points of $\xi_2^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0.500 0.100 0.000 0.400</td>
<td>0.400 0.000 0.100 0.500</td>
</tr>
<tr>
<td>3.0</td>
<td>0.500 0.167 0.000 0.333</td>
<td>0.333 0.000 0.167 0.500</td>
</tr>
<tr>
<td>4.0</td>
<td>0.500 0.250 0.000 0.250</td>
<td>0.250 0.000 0.250 0.500</td>
</tr>
<tr>
<td>5.0</td>
<td>0.500 0.300 0.000 0.200</td>
<td>0.200 0.000 0.300 0.500</td>
</tr>
</tbody>
</table>

The next three results concern correlated responses from the test and standard preparations with $\rho \neq 0$ and $|\rho| < 1$.

**Theorem 3.3.** Suppose (5) holds, $0 < |\rho| < 1$, and $|\mu| \leq 2$. If a design $\xi^*$ satisfies $d_1 - c_1 = \mu$, $\xi^*$ is a locally optimal design for estimating $\mu$.

**Proof.** From (3), it is straightforward to calculate that
\[
\begin{align*}
    c^T M(\xi)^{-1} c &= 2(1 - \rho) + \frac{(1 - \rho^2)(\mu - (d_1 - c_1))^2}{c_2 + d_2 - 2\rho\gamma - c_1^2 - d_1^2 + 2c_1d_1\rho}.
\end{align*}
\] (7)

If $|\mu| \leq 2$, we observe that
\[
\begin{align*}
    (c_2 + d_2 - 2\rho\gamma) &- (c_1^2 + d_1^2 - 2c_1d_1\rho) \\
    &= \int (x_1 x_2)^T \Sigma^{-1}(x_1 x_2)^T d\xi - (c_1 d_1)^T \int (x_1 x_2)^T \Sigma^{-1}(c_1 d_1)^T \\
    &= \int (\tilde{x}_1 \tilde{x}_2)^T \Sigma^{-1}(\tilde{x}_1 \tilde{x}_2)^T d\xi > 0,
\end{align*}
\]

where $\tilde{x}_i = x_i - \int x_i d\xi$, $i = 1, 2$. It follows that $c^T M(\xi)^{-1} c \geq 2(1 - \rho)$, and equality holds if $d_1 - c_1 = \mu$. The desired result follows.

**Theorem 3.4.** Suppose (5) holds, $0 < \rho < 1$, and $|\mu| > 2$. The design $\xi^* = (1/2 + 1/\mu)\delta_{(1,-1)} + (1/2 - 1/\mu)\delta_{(-1,1)}$ is a locally optimal design for estimating $\mu$.

**Proof.** From the general expression of $c^T M(\xi)^{-1} c$ in (7), we have
\[
\begin{align*}
    c^T M(\xi)^{-1} c &\geq 2(1 - \rho) + (1 - \rho^2) + \frac{(\mu - (d_1 - c_1))^2}{2 - 2\rho\gamma - c_1^2 - d_1^2 + 2c_1d_1\rho} \\
    &= 2(1 - \rho) + (1 - \rho^2)g(c_1, d_1, \gamma),
\end{align*}
\]
with equality if $\xi$ is supported on $(-1, -1)$, $(-1, 1)$, $(1, -1)$, and $(1, 1)$. Now we want to find designs with $c_1, d_1$ that minimize $g(c_1, d_1, \gamma)$. For fixed $\gamma$ and $\rho$, we first take partial derivatives of $g(c_1, d_1, \gamma)$ with respect to $c_1, d_1$, and set them to 0. A straightforward argument shows the optimal design must have $c_1 = -d_1$. Under this constraint, let

$$h(d_1, \gamma) = g(-d_1, d_1, \gamma) = \frac{(\mu - 2d_1)^2}{2 - 2\rho\gamma - (2d_1^2 + 2\rho d_1^2)}$$

verify directly that $d_1^* = [2(1 - \rho\gamma)/[\mu(1 + \rho)]$ minimizes the function $h(d_1, \gamma)$ because $\partial^2 h/\partial d_1^2 = [(1 + \rho)^2\mu^2]/[(1 - \rho\gamma)^2(-4 + 4\rho\gamma + \mu^2 + \rho\mu^2)] > 0$ when $\gamma \geq -1 > [4 - \mu^2(1 + \rho)]/(4\rho)$. Hence, with the additional condition that $\gamma = -1$, $h(d_1^*, \gamma)$ attains its minimum value. Consequently the locally optimal design for estimating $\mu$ is $\xi^* = (1/2 + 1/\mu)\delta(-1, -1) + (1/2 - 1/\mu)\delta(-1, 1)$, because it has the property that $c_1 = -d_1 = -2/\mu$ and $\gamma = -1$.

The next result allows us to construct locally optimal design when $|\mu| > 2$ and $-1 < \rho < 0$. The proof is more complicated and is deferred to the Appendix.

**Theorem 3.5.** Suppose (5) holds, $-1 < \rho < 0$, and $|\mu| > 2$. Consider a design of the form $\xi = p_1\delta(-1, -1) + p_2\delta(-1, 1) + p_3\delta(1, -1) + p_4\delta(1, 1)$. The design $\xi^*$ is a locally optimal design for estimating $\mu$ if

(i) $p_1 = p_4 = (\mu - 2)/[2(\mu + \mu\rho - 2\rho)], p_2 = 1 - 2p_1$ and $p_3 = 0$, provided $2 < \mu \leq 2 - 2/\rho$,

(ii) $p_1 = p_4 = (\mu + 2)/[2(\mu + \mu\rho + 2\rho)], p_2 = 0$ and $p_3 = 1 - 2p_1$, provided $-2 + 2/\rho \leq \mu < -2$,

(iii) $p_1 = p_4 = 1/2$ and $p_2 = p_3 = 0$, provided $|\mu| > 2 - 2/\rho$.

### 4. An Application and Discussion

Darby (1980) analyzed a data set on the assay of the antibiotic tobramycin where the same levels of dose were used in both the standard and test preparations. The range of the variable $x$ (logdose) in the study was between $-1.8$ and $-3$, which is not symmetric about 0. However, the $c$-optimal design on the interval $[-3, -1.8]$ for estimating $\mu$ can still be found by applying results in Section 3. In this assay, there exist constants $c_1$ and $d_1$, both inside the range $[-3, -1.8]$, such that $-1.2 \leq c_1 - d_1 \leq 1.2$. Then, if it is known from prior experience that the location shift parameter $\mu$ is approximately zero, we would be interested in designs such that the design points on the test and standard preparations are the same, and that $c_1 - d_1 = 0$ approximately. Such designs are optimal or nearly optimal for estimating $\mu$, by Theorem 3.1. Moreover, as long as $\mu$ is inside the interval $[-1.2, 1.2]$, any design that satisfies $c_1 - d_1 = \mu$ is $c$-optimal. If $\rho \neq 0$ and $\mu$ exceeds the maximum possible values of $c_1 - d_1$, the design problem will have
to be specifically worked out. This is a drawback of designs that lack invariance
under a linear change of the design space. D-optimal designs have the invariance
property and so they can be constructed on any interval once the optimal design
is worked out on the interval $[-1,1]$.

We note that the information matrices for the optimal designs in Theorem
2.1 are actually non-singular even though they have only two support points. This is because, under the given bivariate structure, both responses are observed
at two levels $-1$ and $1$ of the dose variables $x_1$ and $x_2$, and the common slope
parameter for the parallel model can be estimated with information from either
response. The nonsingularity of the other information matrices of the optimal
designs could be similarly explained.

There are other design issues for the parallel line model not yet addressed
here. First, we focused only on symmetric design spaces; occasionally a non-
symmetrical design space is used, see Kent-Jones and Meiklejohn (1994) for example. Second, we have assumed the variances of the responses from both
preparations are equal. If these variances are unequal, the locally optimal de-
signs found here may not apply. Third, if the researcher is primarily interested
in $\mu$ and, at the same time, wishes to examine the parallelism assumption, we
may resort to multiple objective designs, see Cook and Wong (1994) and the
many references in Wong (1999) for more details. In this case, we can construct
a multiple-objective optimal design that incorporates the T-optimal design cri-
terion for discriminating between the two rival multiresponse models discussed
in Ucinski and Bogacka (2005). Under our model, the results for the T-optimal
design criterion are relatively simple and therefore are not discussed here. For
more complicated situations such as when there are more than two responses, we
would have to resort to algorithms to search for the optimal designs for estimating
relative potencies. The algorithms in Mueller and Pazman (1999) or Ucinski and
Atkinson (2004) for finding optimal designs in problems with correlated errors
may be helpful.

Acknowledgement

This research was partially supported by the National Science Council of
Taiwan, R.O.C. under grant NSC 91-2118-M-110-002.

Appendix

Lemmas 1 to 3 are needed for the proof of Theorem 3.5, which deals with
the case when the dual responses are negatively correlated and $\mu$ is large in
magnitude. It is helpful to recall from Theorem 3.4 that an optimal design for
estimating $\mu$ on $\mathcal{X}$ must satisfy $c_1 = -d_1$. Accordingly, we focus on designs of
the form $\xi = p_1 \delta_{(-1,-1)} + p_2 \delta_{(-1,1)} + p_3 \delta_{(1,-1)} + p_4 \delta_{(1,1)}$. 
**Lemma 1.** Suppose (5) holds, $-1 < \rho < 0$, and $\mu > 2$. If the design $\xi = p_1\delta_{(-1,-1)} + p_2\delta_{(-1,1)} + p_3\delta_{(1,-1)} + p_1\delta_{(1,1)}$ satisfies $p_2 + p_3 = \alpha$, where $\alpha$ is a fixed constant, and $0 \leq \alpha < (2 - 2\rho)/(\mu + \mu\rho - 4\rho)$, the design $\xi^* = p_1\delta_{(-1,-1)} + \alpha\delta_{(-1,-1)} + p_1\delta_{(1,1)}$ with $p_1 = (1 - \alpha)/2$ minimizes $c^T M(\xi)^{-1}c$.

**Proof.** Since $\alpha$ is fixed, we have $d_1 = \alpha - 2p_3$ and $\gamma = 1 - 2\alpha$. The function $h(d_1, \gamma)$ in (5) can be rewritten as

$$h_1(p_3; \alpha) = h(\alpha - 2p_3, 1 - 2\alpha) = \frac{(\mu - 2\alpha + 4p_3)^2}{2 - 2\rho(1 - 2\alpha) - 2(1 + \rho)(\alpha - 2p_3)^2}. \quad (9)$$

It is easy to verify that the derivative of $h_1(p_3; \alpha)$ with respect to $p_3$ is

$$\dot{h}_1(p_3; \alpha) = \frac{2(\mu + 4p_3 - 2\alpha)(2 - \alpha\mu + 2\mu(1 + \rho + 4\alpha - 2\rho - 4\alpha\rho + 4\rho^2 - \alpha\mu + 2\mu\rho - \alpha\mu)}{(-1 + \alpha^2 - 4\alpha p_3 + 4p_3^2 + \rho - 2\alpha p - \alpha^2 p - 4\alpha\rho + 4\rho^2 - \alpha\mu + 2\mu\rho - \alpha\mu\rho)} > 0$$

for $0 \leq p_3 \leq \alpha$ and $0 \leq \alpha < (2 - 2\rho)/(\mu + \mu\rho - 4\rho)$. Thus $h_1(p_3; \alpha)$ is increasing in $[0, \alpha]$ and the minimum of $h_1(p_3; \alpha)$ occurs when $p_3 = 0$.

**Lemma 2.** Suppose (5) holds, $-1 < \rho < 0$, and $\mu > 2$. Suppose $\xi = p_1\delta_{(-1,-1)} + p_2\delta_{(-1,1)} + p_1\delta_{(1,1)}$ satisfies $0 \leq p_2 \leq (2 - 2\rho)/(\mu + \mu\rho - 4\rho)$ and $p_1 = (1 - p_2)/2$.

(i) If $2 < \mu \leq 2 - 2/\rho$, the design $\xi^*$ with $p_1 = (\mu - 2)/[2(\mu + \mu\rho - 2\rho)]$ and $p_2 = (\mu\rho - 2\rho + 2)/(\mu + \mu\rho - 2\rho)$ minimizes $c^T M(\xi)^{-1}c$.

(ii) If $\mu > 2 - 2/\rho$, the design $\xi^*$ with $p_1 = 1/2$ and $p_2 = 0$ minimizes $c^T M(\xi)^{-1}c$.

**Proof.** Consider the design $\xi = p_1\delta_{(-1,-1)} + p_2\delta_{(-1,1)} + p_1\delta_{(1,1)}$ with $d_1 = p_2$ and $\gamma = 1 - 2p_2$. The function $h(d_1, \gamma)$ in (5) becomes

$$h_2(p_2) = h(p_2, 1 - 2p_2) = \frac{(\mu - 2p_2)^2}{2 - 2\rho(1 - 2p_2) - 2(1 + \rho)p_2^2}, \quad (10)$$

and it is straightforward to verify that $p_2^* = (\mu - 2\rho + 2)/[\mu + \mu\rho - 2\rho]$ is a critical number of $h_2(p_2)$ in interval $[0, (2 - 2\rho)/(\mu + \mu\rho - 4\rho)]$ such that the second derivative $\ddot{h}_2(p_2^*) = [(\mu - 2\rho + \mu\rho)^4]/[(\mu - 2)(2 + \mu - 2\rho + \mu\rho)]$ is positive. The first part of the theorem is proved. The second part of the theorem follows because when $\mu > 2 - 2/\rho$, the derivative of $h_2(p_2)$ is positive for all $p_2 \in [0, (2 - 2\rho)/(\mu + \mu\rho - 4\rho)]$. Therefore, $p_2^* = 0$ minimizes $h_2(p_2)$.

**Lemma 3.** Suppose (5) holds, $-1 < \rho < 0$, and $\mu > 2$. Suppose the design $\xi = p_1\delta_{(-1,-1)} + p_2\delta_{(-1,1)} + p_3\delta_{(1,-1)} + p_1\delta_{(1,1)}$ satisfies $p_2 + p_3 = \alpha$, $\alpha$ is a fixed constant and $(2 - 2\rho)/(\mu + \mu\rho - 4\rho) \leq \alpha \leq 1$. Then the design $\xi^* = p_1\delta_{(-1,-1)} + p_2\delta_{(-1,1)} + p_3\delta_{(1,-1)} + p_1\delta_{(1,1)}$
(α – p3)δ(1,1) + p3δ(1,1) = p1 with p3 = [μ(1 + ρ)α – 4ρα – 2ρ]/[2μ(1 + ρ)] and p1 = (1 – α)/2 minimizes cTM(ξ)c.

**Proof.** Direct calculus shows that the restriction α > (2 – 2ρ)/(μ + μρ – 4ρ) on the derivative of h1(p3; α) in (10) implies that p3 = [μ(1 + ρ)α – 4ρα – 2ρ]/[2μ(1 + ρ)] and satisfies h1(p3; α) = 0 and h1(p3; α) = [4μ(1 + ρ)2]/[(1 – ρ + 2ρ)2(μ2 + 4ρ – 8ρ + μ2ρ – 4ρ)] > 0. The lemma is proved.

**Proof of Theorem 3.5.** Consider μ > 2. By Lemmas 1, 2 and 3, we only need to show that h1(p3; α) in (9) and h2(p3) in (10) satisfy h1(p3; α) > h2(p3) for all α in the range (2 – 2ρ)/(μ + μρ – 4ρ) ≤ α ≤ 1. Additional calculation shows that if 2 < μ ≤ 2 – 2/ρ,

\[
\frac{h_1(p_3^*; \alpha) - h_2(p_2^*)}{(1 + \rho)(1 - \rho + 2\alpha\rho)} - \frac{1}{2}(\mu - 2)(2 + \mu - 2\rho + \mu\rho) > 0,
\]

and if μ > 2 – 2/ρ, we have

\[
\frac{h_1(p_3^*; \alpha) - h_2(p_2^*)}{2(1 + \rho)(1 - \rho + 2\alpha\rho)} - \frac{\mu^2}{2 - 2\rho} > \frac{2(-\mu\rho + \rho + 1)}{(1 - \rho)(1 + \rho)} > 0.
\]

Hence inequality h1(p3; α) > h2(p3) holds for μ > 2. Thus parts (i) and (ii) of the theorem are proved. The remaining parts of the theorem can be proved analogously by considering the case when μ is less than –2.

**References**


Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan R.O.C.
E-mail: lomn@math.nsysu.edu.tw

Institute of Statistics, National University of Kaohsiung, Kaohsiung 811, Taiwan, R.O.C.
E-mail: rbchen@nuk.edu.tw

Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan R.O.C.
E-mail: cslin@math.nsysu.edu.tw

Department of Biostatistics, University of California at Los Angeles, Los Angeles, CA 90095, U.S.A.
E-mail: wkwong@ucla.edu

(Received June 2004; accepted February 2005)