The Annals of Statistics
2010, Vol. 38, No. 4, 2499–2524
DOI: 10.1214/09-AOS787

ON THE DE LA GARZA PHENOMENON

BY MIN YANG[1]

University of Missouri

Deriving optimal designs for nonlinear models is, in general, challenging. One crucial step is to determine the number of support points needed. Current tools handle this on a case-by-case basis. Each combination of model, optimality criterion and objective requires its own proof. The celebrated de la Garza Phenomenon states that under a $(p-1)$th-degree polynomial regression model, any optimal design can be based on at most p design points, the minimum number of support points such that all parameters are estimable. Does this conclusion also hold for nonlinear models? If the answer is yes, it would be relatively easy to derive any optimal design, analytically or numerically. In this paper, a novel approach is developed to address this question. Using this new approach, it can be easily shown that the de la Garza phenomenon exists for many commonly studied nonlinear models, such as the Emax model, exponential model, three- and four-parameter log-linear models, Emax-PK1 model, as well as many classical polynomial regression models. The proposed approach unifies and extends many well-known results in the optimal design literature. It has four advantages over current tools: (i) it can be applied to many forms of nonlinear models; to continuous or discrete data; to data with homogeneous or nonhomogeneous errors; (ii) it can be applied to any design region; (iii) it can be applied to multiple-stage optimal design and (iv) it can be easily implemented.

1. Introduction.

The usefulness and popularity of nonlinear models have spurred a large literature on data analysis, but research on design selection has not kept pace. One complication in studying optimal designs for nonlinear models is that information matrices and optimal designs depend on unknown parameters. A common approach to solve this dilemma is to use locally optimal designs, which are based on one's best guess of the unknown parameters. While a good guess may not always be available, this approach remains of value to obtain benchmarks for all designs Ford, Torsney and Wu (1992). In fact, most available results are under the context of locally optimal designs. (Hereafter, the word "locally" is omitted for simplicity.)

There is a vast literature on identifying good designs for a wide variety of linear models, but the problem is much more difficult and not nearly as well understood

Received September 2009; revised December 2009.
[1]Supported by NSF Grants DMS-07-07013 and DMS-07-48409.
AMS 2000 subject classifications. Primary 62K05; secondary 62J12.
Key words and phrases. Locally optimal, Loewner ordering, support points.

for nonlinear models. Relevant references will be provided in later sections in this paper.

In the field of optimal designs, there exist no general approaches for identifying good designs for nonlinear models. There are three main reasons for this significant research gap. First, in nonlinear models the mathematics tends to become more difficult, which makes proving optimality of designs a more intricate problem. Current available tools are mainly based on the geometric approach by Elfving (1952) or the equivalence approach by Kiefer and Wolfowitz (1960). This typically means that results can only be obtained on a case-by-case basis. Each combination of model, optimality criterion and objective requires its own proof. It is not feasible to derive a general solution. Second, while linear models are all of the form $E(y) = X\beta$, there is no simple canonical form for nonlinear models. Coupled with the first challenge, this means it is very difficult to establish unifying and overarching results for nonlinear models. Again, this means that individual consideration is typically needed for different models, different optimality criteria, and different objectives. Third, when considering the important practical problem of multi-stage experiments, the search for optimal designs becomes even more complicated because one needs to add design points on top of an existing design.

Is there a practical way to overcome these challenges and derive a general approach for finding optimal designs for nonlinear models? One feasible strategy is to identify a subclass of designs with a simple format, so that one can restrict considerations to this subclass for any optimality problem. With a simple format, it would be relatively easy to derive an optimal design, analytically or numerically.

To make this strategy meaningful, the number of support points for designs in the subclass should be as small as possible. By Carathéodory's theorem, we can always restrict our consideration to at most $p(p+1)/2$ design points (where p is the number of parameters). On the other hand, if we want all parameters to be estimable, the minimum number of support points should be at least p. Thus, the ideal situation is that the designs in the subclass have no more than p points. This reminds one of de la Garza's (1954) result, which was discussed in detail by Pukelsheim (2006) under the concept of "admissibility." This result was named the celebrated de la Garza Phenomenon by Khuri et al. (2006).

The de la Garza phenomenon can be explained as follows: suppose we consider a $(p-1)$th-degree polynomial regression model (p parameters in total) with i.i.d. random errors. For any n point design where $n > p$, there exists a design with exactly p support points such that the information matrix of the latter one is not inferior to that of the former one under Loewner ordering. Does this phenomenon also exist for other models? For nonlinear models with two parameters, Yang and Stufken (2009) provided an approach to identify the subclass of designs:

for any design ξ which does not belong to this class, there is a design in the class with an information matrix that dominates ξ in the Loewner ordering. By applying this approach, they showed that many commonly studied models, such as logistic and probit models, are based on two design points. This result unifies and extends most available optimality results for binary response models. However, a limitation exists since it can only be applied to nonlinear models with one or two parameters.

The purpose of this paper is to generalize Yang and Stufken (2009) to nonlinear models with an arbitrary number of parameters. The proposed approach makes it relatively easy to prove the de la Garza Phenomenon for many nonlinear models. In fact, for many commonly studied nonlinear models, including the Emax model, exponential model, three- and four-parameter log-linear models, Emax-PK1 model, as well as many classical polynomial regression models, it can be shown that for any given design ξ, there exists a design ξ^* with at most p (number of parameters) points, where the information matrix under ξ^* is not inferior to that of ξ under Loewner ordering. Thus, when searching for an optimal design, one can restrict consideration to this subclass of designs, both for one-stage and multi-stage problems. Here, the optimal design can be for arbitrary parameter functions under any information matrix based-optimality criterion, including the commonly used A-, D-, E-, Φ_p-, etc. criteria as well as standardized optimality criteria proposed by Dette (1997). Refer to Yang and Stufken (2009) for more details on the significance of these flexibilities.

This paper is organized as follows. In Section 2, we introduce the strategy. Main results are presented in Section 3. Applications to many commonly studied nonlinear models are presented in Section 4. Section 5 is a short discussion. Most proofs are included in the Appendix.

2. The strategy. Suppose we have a nonlinear regression model for which at each point x the experimenter observes a response y. We assume that the y's are independent and follow some exponential distribution G with mean $\eta(x,\theta)$, where θ is $p \times 1$ parameters vector. Typically, the optimal nonlinear designs are studied under approximate theory, that is, instead of exact sample sizes for design points, design weights are used. An approximate design ξ can be written as $\xi = \{(x_i, \omega_i), i = 1, \dots, n\}$, where $\omega_i > 0$ is the design weight for design point x_i and $\sum_{i=1}^n \omega_i = 1$. It is more convenient to rewrite ξ as $\xi = \{(c_i, \omega_i), i = 1, \dots, n\}$, where $c_i \in [A, B]$ may depend on θ and is one-to-one map of $x_i \in [U, V]$. Typically, the information matrix for θ under design ξ can be written as

$$I_\xi(\theta) = P(\theta)\left(\sum_{i=1}^n \omega_i C(\theta, c_i)\right)(P(\theta))^T, \tag{2.1}$$

where

$$C(\theta, c_i) = \begin{pmatrix} \Psi_{11}(c_i) & \Psi_{12}(c_i) & \cdots & \Psi_{1p}(c_i) \\ \Psi_{12}(c_i) & \Psi_{22}(c_i) & \cdots & \Psi_{2p}(c_i) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_{1p}(c_i) & \Psi_{2p}(c_i) & \cdots & \Psi_{pp}(c_i) \end{pmatrix}. \tag{2.2}$$

Here, $P(\theta)$ is a $p \times p$ nonsingular matrix that depends on the value of θ only. Notice that while $I_\xi(\theta)$ is fixed for given θ and ξ, there is flexibility on $P(\theta)$ and $C(\theta, c_i)$. For many models, we can adjust $P(\theta)$ so that all Ψ_{lt}'s in (2.2) are free of ξ and θ. Some examples of (2.1) and (2.2) are given in Section 4.

Under locally optimality context, for two given designs $\xi = \{(c_i, \omega_i), i = 1, \dots, n\}$ and $\xi^* = \{(\tilde{c}_j, \tilde{\omega}_j), j = 1, \dots, \tilde{n}\}$, $I_\xi(\theta) \le I_{\xi^*}(\theta)$ is equivalent to $\sum_{i=1}^n \omega_i C(\theta, c_i) \le \sum_{j=1}^{\tilde{n}} \tilde{\omega}_j C(\theta, \tilde{c}_j)$ (here and elsewhere in this paper, matrix inequalities are under the Loewner ordering). One strategy to show $I_\xi(\theta) \le I_{\xi^*}(\theta)$ is to prove that the following equations hold:

$$\sum_{i=1}^{n} \omega_i \Psi_{lt}(c_i) = \sum_{j=1}^{\tilde{n}} \tilde{\omega}_j \Psi_{lt}(\tilde{c}_j) \tag{2.3}$$

for $1 \le l \le t \le p$ except for some $l = t$ (one or more)

$$\sum_{i=1}^{n} \omega_i \Psi_{ll}(c_i) \le \sum_{j=1}^{\tilde{n}} \tilde{\omega}_j \Psi_{ll}(\tilde{c}_j). \tag{2.4}$$

The development of the new tool is based on this strategy. Notice that Yang and Stufken (2009) used the same strategy for the $p = 2$ case. However, the picture for a general p is completely different. This is because when $p = 2$, the existence of ξ^* can be based on the existence of one $\tilde{c}$ and one $\tilde{\omega}$ satisfying two nonlinear equations, which can be solved explicitly. For an arbitrary p, it is unlikely to derive such explicit expressions for ξ^* since the existence of ξ^* is based on the existence of multiple [approximately $p(p+1)/4$] $\tilde{c}$'s and $\tilde{\omega}$'s satisfying multiple [approximately $p(p+1)/2$] nonlinear equations. Alternative approaches must be employed. In the next section, some new algebra results will be provided to address these needs.

3. The approach. In this section, we shall show that, under certain conditions, for a general nonlinear model, there exists a subclass of designs such that for any given design ξ, there exists a design $\tilde{\xi}$ in this subclass such that $I_\xi(\theta) \le I_{\xi^*}(\theta)$. We first introduce some new algebra results.

3.1. *Algebra results.* Let $\Psi_1, \dots, \Psi_k$ be k functions defined on $[A, B]$. Throughout this paper, we have the following assumptions:

(i) $\Psi_1, \dots, \Psi_k$ are infinity differentiable;
(ii) $f_{l,l}$ has no zero value on $[A, B]$.

Here, $f_{l,t}$, $1 \le t \le k$; $t \le l \le k$ are defined as follows:

$$f_{l,t}(c) = \begin{cases} \Psi_l'(c), & t = 1, l = 1, \dots, k, \\ \left(\dfrac{f_{l,t-1}(c)}{f_{t-1,t-1}(c)}\right)', & 2 \le t \le k, t \le l \le k. \end{cases} \tag{3.1}$$

The structure of computations of $f_{l,t}$ can be viewed as the following lower triangular matrix:

$$\begin{pmatrix} f_{1,1} = \Psi_1' & & & & \\ f_{2,1} = \Psi_2' & f_{2,2} = \left(\dfrac{f_{2,1}}{f_{1,1}}\right)' & & & \\ f_{3,1} = \Psi_3' & f_{3,2} = \left(\dfrac{f_{3,1}}{f_{1,1}}\right)' & f_{3,3} = \left(\dfrac{f_{3,2}}{f_{2,2}}\right)' & & \\ f_{4,1} = \Psi_4' & f_{4,2} = \left(\dfrac{f_{4,1}}{f_{1,1}}\right)' & f_{4,3} = \left(\dfrac{f_{4,2}}{f_{2,2}}\right)' & f_{4,4} = \left(\dfrac{f_{4,3}}{f_{3,3}}\right)' & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.2}$$

The $(t+1)$th column is obtained from the tth column. The lth ($l \ge t+1$) element of the $(t+1)$th column is the derivative of the ratio between the lth and the tth element of the tth column.

LEMMA 1. *Let $\Psi_1, \dots, \Psi_k$ be k functions defined on $[A, B]$. Assume that $f_{l,l}(c) > 0$, $c \in [A, B]$, $l = 1, \dots, k$. Then we have following conclusions*:

(a) *When $k = 2n - 1$. For any given $A \le \tilde{c}_0 < c_1 < \cdots < c_n \le B$ and $\omega_i > 0$, $i = 1, \dots, n$, there exist n pairs $(\tilde{c}_j, \tilde{\omega}_j)$, $j = 0, \dots, n-1$, where $\tilde{c}_0 < c_1 < \tilde{c}_1 < c_2 < \cdots < \tilde{c}_{n-1} < c_n$ and $\tilde{\omega}_j > 0$, such that* (3.3) *and* (3.4) *hold, and* (3.5) > 0.
(b) *When $k = 2n - 1$. For any given $A \le c_1 < \cdots < c_n < \tilde{c}_n \le B$ and $\omega_i > 0$, $i = 1, \dots, n$, there exist n pairs $(\tilde{c}_j, \tilde{\omega}_j)$, $j = 1, \dots, n$, where $A \le c_1 < \tilde{c}_1 < c_2 < \cdots < \tilde{c}_{n-1} < c_n < \tilde{c}_n \le B$ and $\tilde{\omega}_j > 0$, such that* (3.3) *and* (3.4) *hold, and* (3.5) < 0.
(c) *When $k = 2n$. For any given $A \le \tilde{c}_0 < c_1 < \cdots < c_n < \tilde{c}_n \le B$ and $\omega_i > 0$, $i = 1, \dots, n$, there exist $n+1$ pairs $(\tilde{c}_j, \tilde{\omega}_j)$, $j = 0, \dots, n$, where $A \le \tilde{c}_0 < c_1 < \tilde{c}_1 < \cdots < c_n < \tilde{c}_n \le B$ and $\tilde{\omega}_j > 0$, such that* (3.3) *and* (3.4) *hold, and* (3.5) < 0.
(d) *When $k = 2n$. For any given $A \le c_1 < \cdots < c_{n+1} \le B$ and $\omega_i > 0$, $i = 1, \dots, n+1$, there exist n pairs $(\tilde{c}_j, \tilde{\omega}_j)$, $j = 1, \dots, n$, where $A \le c_1 < \tilde{c}_1 < \cdots < c_n < \tilde{c}_n < c_{n+1} \le B$ and $\tilde{\omega}_j > 0$, such that* (3.3) *and* (3.4) *hold, and* (3.5) > 0.

Here,

$$\sum_i \omega_i = \sum_j \tilde{\omega}_j; \tag{3.3}$$

$$\sum_i \omega_i \Psi_l(c_i) = \sum_j \tilde{\omega}_j \Psi_l(\tilde{c}_j), \qquad l = 1, \dots, k-1; \tag{3.4}$$

$$\sum_i \omega_i \Psi_k(c_i) - \sum_j \tilde{\omega}_j \Psi_k(\tilde{c}_j). \tag{3.5}$$

Yang and Stufken (2009) have proven Lemma 1 for $k = 2$ and 3. For arbitrary k, the proof is rather complicated (see the Appendix). Lemma 1 requires that $f_{l,l}(c) > 0$ for every $l = 1, \dots, k$, which is very demanding. In fact, such strict conditions are not required. Suppose there are some $f_{l,l}(c) < 0$, we can consider $-\Psi_l(c)$ instead of $\Psi_l(c)$ depending on the situation, such that the corresponding $f_{l,l}(c) > 0$ for every $l = 1, \dots, k$. Notice that (3.4) is invariant to such transformation and the sign of the inequality (3.5) may need to be reversed. Thus we can have similar results to Lemma 1 with a relaxed condition. We are ready to present our first main theorem.

THEOREM 1. *Let $\Psi_1, \dots, \Psi_k$ be k functions defined on $[A, B]$. Let $F(c) = \prod_{l=1}^k f_{l,l}(c)$. For any given N pairs (c_i, ω_i), where $c_i \in [A, B]$ and $\omega_i > 0$, $i = 1, \dots, N$, there exists a set of pairs $(\tilde{c}_j, \tilde{\omega}_j)$, where $\tilde{c}_j \in [A, B]$ and $\tilde{\omega}_j > 0$, such that* (3.3) *and* (3.4) *hold, and* (3.5) < 0. *Specifically*:

(a) *when $k = 2n - 1$, $N \ge n$ and $F(c) < 0$ for $c \in [A, B]$, there are n pairs $(\tilde{c}_j, \tilde{\omega}_j)$ in the set and one of $\tilde{c}_j$'s is A*;
(b) *when $k = 2n - 1$, $N \ge n$ and $F(c) > 0$ for $c \in [A, B]$, there are n pairs $(\tilde{c}_j, \tilde{\omega}_j)$ in the set and one of $\tilde{c}_j$'s is B*;
(c) *when $k = 2n$, $N \ge n$ and $F(c) > 0$ for $c \in [A, B]$, there are $n + 1$ pairs $(\tilde{c}_j, \tilde{\omega}_j)$ in the set and two of $\tilde{c}_j$'s are A and B*;
(d) *when $k = 2n$, $N \ge n + 1$ and $F(c) < 0$ for $c \in [A, B]$, there are n pairs $(\tilde{c}_j, \tilde{\omega}_j)$ in the set.*

PROOF. The proofs for the above four cases are completely analogous. Here we will provide the proof of Theorem 1(a). First, we prove the conclusion holds when $N = n$. From (3.2), it is easy to verify that if we change only one $\Psi_l(c)$ to $-\Psi_l(c)$, say, $l = l_0$, and keep all other $\Psi_l(c)$'s the same, then all $f_{l,l}(x)$ will maintain their original signs with two exceptions: (i) $f_{l_0,l_0}(c)$ and $f_{l_0+1,l_0+1}(c)$ reverse the sign when $l_0 < k$ or (ii) $f_{k,k}(c)$ reverse the sign when $l_0 = k$. Among all $f_{l,l}$, $l = 1, \dots, k$, suppose a of them are negative, say $f_{l_1,l_1}, \dots, f_{l_a,l_a}$. Here, $l_1 < \cdots < l_a$ and a must be an odd number.

When $l_{2b-1} \le l < l_{2b}$ $[1 \le b \le (a-1)/2]$ or $l \ge l_a$, $\widetilde{\Psi}_l(c)$ is defined as $-\Psi_l(c)$. Otherwise, $\widetilde{\Psi}_l(c) = \Psi_l(c)$. We can verify that the corresponding $\widetilde{f_{l,l}}(c) > 0, l = 1, \dots, k$ by repeatedly using the argument for the change of signs of $f_{l,l}(c)$'s when we change only one $\Psi_l(c)$ to $-\Psi_l(c)$ each time. Now let $\tilde{c}_0 = A$, and notice that $\widetilde{\Psi}_k(c) = -\Psi_k(c)$, by Lemma 1(a), the conclusion follows.

Assume that Lemma 1(a) holds for $n \le N \le M$. Now we consider $N = M + 1$. Following this assumption, for the M pairs (c_i, ω_i), $1 \le i \le M$, there exist n pairs $(\overline{c}_j, \overline{\omega}_j)$, $j = 0, \dots, n-1$, where $\overline{c}_0 = A$, such that

$$\begin{aligned} &\sum_{i=1}^{M} \omega_i = \sum_{j=0}^{n-1} \overline{\omega}_j, \\ &\sum_{i=1}^{M} \omega_i \Psi_l(c_i) = \sum_{j=0}^{n-1} \overline{\omega}_j \Psi_l(\overline{c}_j), \qquad l = 1, \dots, k-1, \\ &\sum_{i=1}^{M} \omega_i \Psi_k(c_i) < \sum_{j=0}^{n-1} \overline{\omega}_j \Psi_k(\overline{c}_j). \end{aligned} \tag{3.6}$$

Consider the $n-1$ pairs $(\overline{c}_j, \overline{\omega}_j)$, $j = 1, \dots, n-1$ and (c_{M+1}, ω_{M+1}). Apply (a) when $N = n$, there exist n pairs $(\tilde{c}_j, \tilde{\omega}_j)$, $j = 0, \dots, n-1$ where $\tilde{c}_0 = A$, such that

$$\begin{aligned} &\omega_{M+1} + \sum_{j=1}^{n-1} \overline{\omega}_j = \sum_{j=0}^{n-1} \tilde{\omega}_j, \\ &\omega_{M+1} \Psi_l(c_{M+1}) + \sum_{j=1}^{n-1} \overline{\omega}_j \Psi_l(\overline{c}_j) = \sum_{j=0}^{n-1} \tilde{\omega}_j \Psi_l(\tilde{c}_j), \qquad l = 1, \dots, k-1, \\ &\omega_{M+1} \Psi_k(c_{M+1}) + \sum_{j=1}^{n-1} \overline{\omega}_j \Psi_k(\overline{c}_j) < \sum_{j=0}^{n-1} \tilde{\omega}_j \Psi_k(\tilde{c}_j). \end{aligned} \tag{3.7}$$

Combining (3.6) and (3.7), we establish Lemma 1(a) when $N = M + 1$. By mathematical induction, the conclusion follows. $\square$

3.2. *The main tools.* We are now ready to present our main tools.

THEOREM 2. *For a nonlinear regression model, suppose the information matrix can be written as* (2.1) *and* $c_i \in [A, B]$. *Rename all distinct* Ψ_{lt}, $1 \le l \le t \le p$ *to* $\Psi_1, \dots, \Psi_k$ *such that* (i) Ψ_k *is one of* Ψ_{ll}, $1 \le l \le p$ *and* (ii) *there is no* $\Psi_{lt} = \Psi_k$ *for* $l < t$. *Let* $F(c) = \prod_{l=1}^{k} f_{l,l}(c)$, $c \in [A, B]$. *For any given design* ξ, *there exists a design* $\tilde{\xi}$, *such that* $I_\xi(\theta) \le I_{\tilde{\xi}}(\theta)$. *Here*, $\tilde{\xi}$ *depends on different situations.*

(a) *When* k *is odd and* $F(c) < 0$, $\tilde{\xi}$ *is based on at most* $(k+1)/2$ *points including point* A.

(b) *When k is odd and $F(c) > 0$, $\tilde{\xi}$ is based on at most $(k+1)/2$ points including point B.*
(c) *When k is even and $F(c) > 0$, $\tilde{\xi}$ is based on at most $k/2 + 1$ points including points A and B.*
(d) *When k is even and $F(c) < 0$, $\tilde{\xi}$ is based on at most $k/2$ points.*

PROOF. The proof for the four cases are completely analogous. Here we provide the proof of Theorem 2(a).

By (2.1) and the fact that $P(\theta)$ depends on θ only, it is sufficient to show that $C_\xi(\theta) \le C_{\tilde{\xi}}(\theta)$. Let $\xi = \{(c_i, \omega_i), i = 1, \dots, N\}$. If $N < n$, then we can just take $\tilde{\xi} = \xi$. If $N \ge n$, by (a) of Theorem 1, there exist n paris $(\tilde{c}_j, \tilde{\omega}_j)$, $j = 0, \dots, n-1$, where $\tilde{c}_0 = A$, such that (3.3), (3.4) and (3.5) < 0 hold. Let $\tilde{\xi} = \{(\tilde{c}_j, \tilde{\omega}_j), j = 0, \dots, n-1\}$. Direct computation shows that the diagonal elements of $C_{\tilde{\xi}}(\theta) - C_\xi(\theta)$ are either 0 or greater than 0, and the off-diagonal elements are all 0. Thus the conclusion follows. □

REMARK 1. For cases (a), (b) and (d) of Theorem 2, the conclusions stay the same if the interval $[A, B]$ is not finite. For case (a), $[A, B]$ can be replaced by $[A, \infty)$. In this situation, for any given design ξ, we can choose $B = \mathrm{Max}_{1 \le i \le N} c_i$ and the same conclusion follows. Similarly, $[A, B]$ can be replaced by $(-\infty, B]$ in case (b) or $(-\infty, \infty)$ in case (d).

REMARK 2. There are many different ways to rename all distinct Ψ_{lt}, $1 \le l \le t \le p$ to $\Psi_1, \dots, \Psi_k$. Not all orders can satisfy the requirements in Theorem 2. However, as long as there exists one order of $\Psi_1, \dots, \Psi_k$ such that these requirements can be satisfied, the conclusion holds. Notice that Ψ_k must be one of Ψ_{ll}, $1 \le l \le p$.

4. Applications. Theorem 2 can be applied to many commonly studied statistical models. In fact, as we demonstrate next, for many models, any optimal design can be based on the minimum number of support points, that is, number of support points such as all parameters are estimable. As we discussed earlier, this makes it much easier to study an optimal design. Theorem 2 works on the information matrix directly. It is very general. It can be applied to any models, continuous or discrete data with homogeneous or nonhomogeneous error, as long as the information matrix can be written as (2.1). Here, we demonstrate its applications for the model

$$Y_{ij} = \eta(x_i, \theta) + \varepsilon_{ij}, \tag{4.1}$$

where ε_{ij}'s are i.i.d. $N(0, \sigma^2)$ with known σ^2, $x_i \in [L, U]$ is the design variable, and θ is a $p \times 1$ parameter vector. Most commonly studied models can be written as (4.1). For a given design $\xi = \{(x_i, \omega_i), i = 1, \dots, N\}$, the corresponding information matrix for θ can be written as

$$I_\xi(\theta) = \sum_{i=1}^{N} \omega_i \frac{\partial \eta(x_i, \theta)}{\partial \theta} \left(\frac{\partial \eta(x_i, \eta)}{\partial \theta} \right)^T. \tag{4.2}$$

Next, we apply Theorem 2 for some popular choices of $\eta(x, \theta)$. Notice that Theorem 2 is not limited to this model format. It can be applied to many other models. Some examples will be shown in Sections 4.3 and 4.4.

4.1. *Models with three parameters.* Dette et al. (2008) studied $E_{\max}$, exponential, and log-linear models. These models can be written in the form of (4.1) with

$$\eta(x, \theta) = \begin{cases} \theta_0 + \dfrac{\theta_1 x}{x + \theta_2}, & E_{\max}, \\ \theta_0 + \theta_1 \exp(x/\theta_2), & \text{exponential}, \\ \theta_0 + \theta_1 \log(x + \theta_2), & \text{log-linear}. \end{cases} \tag{4.3}$$

Here, $x_i \in [L, U] \subset (0, \infty)$, $\theta_1 > 0$, and $\theta_2 > 0$. They showed that local MED-optimal designs (MED is defined as the smallest dose producing a practically relevant response) are either a two points design with low end point L, or a three points design with two end points L and U. In fact, as the following theorem shows, any optimal design can be based on three points including one or two end points.

THEOREM 3. *Under model* (4.3), *for an arbitrary design ξ, there exists a design ξ^* with at most three support points such that $I_\xi(\theta) \le I_{\xi^*}(\theta)$. Specifically, the three points include the two end points L and U for the $E_{\max}$ model, the upper end point U for the exponential model and the two end points L and U for the log-linear model.*

PROOF. We first consider the $E_{\max}$ model. By some routine algebra, it can be shown that the information matrix can be written as in the form of (2.1) with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ \dfrac{1}{\theta_2} & -\dfrac{1}{\theta_2} & 0 \\ \dfrac{1}{\theta_2^2} & -\dfrac{1}{\theta_2^2} & \dfrac{1}{\theta_1\theta_2} \end{pmatrix}^{-1} \quad \text{and} \quad C(\theta, c_i) = \begin{pmatrix} 1 & c_i & c_i^2 \\ c_i & c_i^2 & c_i^3 \\ c_i^2 & c_i^3 & c_i^4 \end{pmatrix}, \tag{4.4}$$

where $c_i = 1/(x_i + \theta_2)$. Let $\Psi_1(c) = c$, $\Psi_2(c) = c^2$, $\Psi_3(c) = c^3$ and $\Psi_4(c) = c^4$, and we can verify that the corresponding $f_{1,1} = 1$, $f_{2,2} = 2$, $f_{3,3} = 3$ and $f_{4,4} = 4$. By case (c) of Theorem 2, the conclusion follows.

The proofs for exponential and log-linear models are similar with different $P(\theta)$ and $C_\xi(\theta)$. For the exponential model,

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\dfrac{\theta_2}{\theta_1} \end{pmatrix}^{-1} \quad \text{and} \quad C(\theta, c_i) = \begin{pmatrix} 1 & e^{c_i} & c_i e^{c_i} \\ e^{c_i} & e^{2c_i} & c_i e^{2c_i} \\ c_i e^{c_i} & c_i e^{2c_i} & c_i^2 e^{2c_i} \end{pmatrix}, \tag{4.5}$$

where $c_i = x_i/\theta_2$. Let $\Psi_1(c) = e^c$, $\Psi_2(c) = ce^c$, $\Psi_3(c) = e^{2c}$, $\Psi_4(c) = ce^{2c}$ and $\Psi_5(c) = c^2 e^{2c}$, and we can verify that the corresponding $f_{1,1} = e^c$, $f_{2,2} = 1$, $f_{3,3} = 2e^c$, $f_{4,4} = 1$ and $f_{5,5} = 2$. By case (b) of Theorem 2, the conclusion follows.

For the log-linear model,

$$\begin{gathered} P(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \dfrac{1}{\theta_1} \end{pmatrix}^{-1} \quad \text{and} \\ C(\theta, c_i) = \begin{pmatrix} 1 & \log(c_i) & c_i \\ \log(c_i) & \log^2(c_i) & c_i \log(c_i) \\ c_i & c_i \log(c_i) & c_i^2 \end{pmatrix}, \end{gathered} \tag{4.6}$$

where $c_i = 1/(x_i + \theta_2)$. Let $\Psi_1(c) = \log(c)$, $\Psi_2(c) = c$, $\Psi_3(c) = c\log(c)$ and $\Psi_4(c) = \log^2(c)$ or c^2, and we can verify that the corresponding $f_{1,1} = 1/c$, $f_{2,2} = 1$, $f_{3,3} = 1/c$ and $f_{4,4} = 2/c^2$ or 4 when $\Psi_4(c) = \log^2(c)$ or c^2, respectively. Apply case (c) of Theorem 1, and for any design we can find a design with three points including end points L and U, such that the off-diagonal elements are the same and diagonal elements are either the same or larger. Thus the conclusion follows. $\square$

REMARK 3. Han and Chaloner (2003) studied D- and c-optimal design under a slightly different exponential model where $\eta(x, \theta) = \theta_0 + \theta_1 \exp(-\theta_2 x)$. By applying a similar approach as used for the exponential model in Theorem 3, we can show that for an arbitrary design ξ, there exists a design ξ^* with three support points including low end point L such that $I_\xi(\theta) \le I_{\xi^*}(\theta)$. This confirms and extends results in Han and Chaloner (2003) for this model, while Han and Chaloner (2003) showed that the D- and c-optimal designs are based on three points including two end points L and U.

Dette, Melas and Wong (2005) studied another version of $E_{\max}$ model, which can be written in the form of (4.1) with

$$\eta(x, \theta) = \frac{\theta_0 x^{\theta_2}}{\theta_1 + x^{\theta_2}}, \tag{4.7}$$

where $x \in [0, T]$, $\theta_0 > 0$, $\theta_1 > 0$ and $\theta_2 \neq 0$. They showed that D- and D_1-optimal designs are based on three points including end point T. The next theorem shows that we can restrict ourselves with three points designs for any optimal designs.

THEOREM 4. *Under model (4.7), for an arbitrary design ξ, there exists a design ξ^* with at most three support points such that $I_\xi(\theta) \leq I_{\xi^*}(\theta)$.*

PROOF. By some routine algebra, it can be shown that the information matrix can be written in the form of (2.1) with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \dfrac{\theta_1}{\theta_0} & 0 \\ 0 & -\dfrac{\theta_1 \log(\theta_1)}{\theta_0} & -\dfrac{\theta_2}{\theta_0} \end{pmatrix}^{-1} \quad \text{and}$$

$$C(\theta, c_i) = \begin{pmatrix} \dfrac{1}{(1+c_i)^2} & \dfrac{1}{(1+c_i)^3} & \dfrac{c_i \log(c_i)}{(1+c_i)^3} \\ \dfrac{1}{(1+c_i)^3} & \dfrac{1}{(1+c_i)^4} & \dfrac{c_i \log(c_i)}{(1+c_i)^4} \\ \dfrac{c_i \log(c_i)}{(1+c_i)^3} & \dfrac{c_i \log(c_i)}{(1+c_i)^4} & \dfrac{c_i^2 \log^2(c_i)}{(1+c_i)^4} \end{pmatrix}, \tag{4.8}$$

where $c_i = \theta_1 x_i^{-\theta_2}$. Let $\Psi_1(c) = \frac{1}{(1+c)^4}$, $\Psi_2(c) = \frac{1}{(1+c)^3}$, $\Psi_3(c) = \frac{c \log(c)}{(1+c)^4}$ and $\Psi_4(c) = \frac{1}{(1+c)^2}$, $\Psi_5(c) = \frac{c \log(c)}{(1+c)^3}$, and $\Psi_6(c) = \frac{c^2 \log^2(c)}{(1+c)^4}$. We can verify that the corresponding $f_{1,1} = -\frac{4}{(1+c)^5}$, $f_{2,2} = 3/4$, $f_{3,3} = \frac{3c+1}{3c^2}$, $f_{4,4} = \frac{4c(3c+2)}{(3c+1)^2}$, $f_{5,5} = \frac{9c^3+15c^2+7c+1}{(3c+2)^2 c^2}$ and $f_{6,6} = \frac{18c^2+15c+2}{c(9c^2+6c+1)}$. Notice that $c > 0$, and this implies that $F(c) < 0$. By case (d) of Theorem 2, the conclusion follows. □

REMARK 4. Li and Majumdar (2008) studied D-optimal design for a three-parameter logistic model where $\eta(x, \theta) = \frac{\theta_0}{1+\theta_1 \exp(\theta_2 x)}$. It can be shown that the information matrix can be written in the form of (4.8) with $c_i = \theta_1 \exp(\theta_2 x_i)$. Thus for any arbitrary design ξ, there exists a design ξ^* with three support points such that $I_\xi(\theta) \leq I_{\xi^*}(\theta)$. This confirms and extends Li and Majumdar's (2008) results.

Han and Chaloner (2003) studied D- and c-optimal designs for a model which can be written in the form of (4.1) with

$$\eta(x, \theta) = \log\bigl(\theta_0 + \theta_1 \exp(-\theta_2 x)\bigr), \tag{4.9}$$

where $x \in [L, U] \subset (0, \infty)$, $\theta_0 > 0$, $\theta_1 > 0$ and $\theta_2 > 0$. They showed that D- and c-optimal designs are based on three points including end points L and U. In fact, any optimal design based on the information matrix can be restricted to a three-point design including lower end point L.

THEOREM 5. *Under* (4.9), *for any arbitrary design* ξ, *there exists a design* ξ^* *with at most three support points including lower end point* L *such that* $I_\xi(\theta) \le I_{\xi^*}(\theta)$.

PROOF. It can be shown that the information matrix can be written in the form of (2.1) with

$$P(\theta) = \begin{pmatrix} \theta_0 & \theta_1 & 0 \\ \theta_0 & 0 & 0 \\ 0 & \theta_1 \log\left(\dfrac{\theta_1}{\theta_0}\right) & \theta_2 \end{pmatrix}^{-1} \quad \text{and}$$

$$C(\theta, c_i) = \begin{pmatrix} 1 & \dfrac{1}{1+c_i} & \dfrac{c_i \log(c_i)}{1+c_i} \\ \dfrac{1}{1+c_i} & \dfrac{1}{(1+c_i)^2} & \dfrac{c_i \log(c_i)}{(1+c_i)^2} \\ \dfrac{c_i \log(c_i)}{1+c_i} & \dfrac{c_i \log(c_i)}{(1+c_i)^2} & \dfrac{c_i^2 \log^2(c_i)}{(1+c_i)^2} \end{pmatrix}, \tag{4.10}$$

where $c_i = \frac{\theta_1}{\theta_0}\exp(-\theta_2 x_i)$. Let $\Psi_1(c) = \frac{1}{(1+c)^2}$, $\Psi_2(c) = \frac{1}{1+c}$, $\Psi_3(c) = \frac{c\log(c)}{(1+c)^2}$ and $\Psi_4(c) = \frac{c\log(c)}{1+c}$, and $\Psi_5(c) = \frac{c^2\log^2(c)}{(1+c)^2}$. We can verify that the corresponding $f_{1,1} = -\frac{2}{(1+c)^3}$, $f_{2,2} = \frac{1}{2}$, $f_{3,3} = \frac{1+c}{c^2}$, $f_{4,4} = -2$, and $f_{5,5} = \frac{2}{c}$. Notice that $c > 0$, which implies that $F(c) > 0$. By case (b) of Theorem 2 and $c_i = \frac{\theta_1}{\theta_0}\exp(-\theta_2 x_i)$, the conclusion follows. □

4.2. *Models with four parameters.* Fang and Hedayat (2008) studied a composed $E_{\max}$-PK1 model, which can be written in the form of (4.1) with

$$\eta(x, \theta) = \theta_0 + \frac{\theta_1 D}{D + \theta_2 \exp(\theta_3 x)}. \tag{4.11}$$

Here, D is a positive constant, $x_i \in [0, U]$ and $\theta_i > 0$, $i = 0, 1, 2, 3$. Fang and Hedayat (2008) showed that local D-optimal designs are based on four points including end points 0 and U. The next theorem tells us that any optimal designs can be based on four points designs.

THEOREM 6. *Under model* (4.11), *for any arbitrary design* ξ, *there exists a design* ξ^* *with at most four support points such that* $I_\xi(\theta) \le I_{\xi^*}(\theta)$.

PROOF. It can be shown that the information matrix can be written in the form of (2.1) with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & \dfrac{\theta_2}{\theta_1} & 0 \\ 0 & 0 & -\dfrac{\theta_2}{\theta_1}\log\left(\dfrac{\theta_2}{D}\right) & -\dfrac{\theta_3}{\theta_1} \end{pmatrix}^{-1} \quad \text{and} \tag{4.12}$$

$$C(\theta, c_i) = \begin{pmatrix} 1 & \dfrac{1}{1+c_i} & \dfrac{1}{(1+c_i)^2} & \dfrac{c_i \log(c_i)}{(1+c_i)^2} \\ \dfrac{1}{1+c_i} & \dfrac{1}{(1+c_i)^2} & \dfrac{1}{(1+c_i)^3} & \dfrac{c_i \log(c_i)}{(1+c_i)^3} \\ \dfrac{1}{(1+c_i)^2} & \dfrac{1}{(1+c_i)^3} & \dfrac{1}{(1+c_i)^4} & \dfrac{c_i \log(c_i)}{(1+c_i)^4} \\ \dfrac{c_i \log(c_i)}{(1+c_i)^2} & \dfrac{c_i \log(c_i)}{(1+c_i)^3} & \dfrac{c_i \log(c_i)}{(1+c_i)^4} & \dfrac{c_i^2 \log^2(c_i)}{(1+c_i)^4} \end{pmatrix},$$

where $c_i = \frac{\theta_2}{D} \exp(\theta_3 x_i)$. Let $\Psi_1(c) = \frac{1}{(1+c)^4}$, $\Psi_2(c) = \frac{1}{(1+c)^3}$, $\Psi_3(c) = \frac{c \log(c)}{(1+c)^4}$ and $\Psi_4(c) = \frac{1}{(1+c)^2}$, $\Psi_5(c) = \frac{c \log(c)}{(1+c)^3}$, $\Psi_6(c) = \frac{1}{1+c}$, $\Psi_7(c) = \frac{c \log(c)}{(1+c)^2}$ and $\Psi_8(c) = \frac{c^2 \log^2(c)}{(1+c)^4}$. We can verify that the corresponding $f_{1,1} = -\frac{4}{(1+c)^5}$, $f_{2,2} = \frac{3}{4}$, $f_{3,3} = \frac{3c+1}{3c^2}$, $f_{4,4} = \frac{4c(3c+2)}{(3c+1)^2}$, $f_{5,5} = \frac{9c^3+15c^2+7c+1}{c^2(3c+2)^2}$, $f_{6,6} = \frac{9c(3c+2)}{9c^2+6c+1}$, $f_{7,7} = \frac{3c+1}{3c^2}$ and $f_{8,8} = \frac{2}{3c^2}$. Notice that $c > 0$, which implies that $F(c) < 0$. By applying case (d) of Theorem 2, the conclusion follows. $\square$

Dette et al. (2008) studied a four-parameter logistic model, which can be written in the form of (4.1) with

$$\eta(x, \theta) = \theta_0 + \frac{\theta_1}{1 + \exp((\theta_2 - x)/\theta_3)}. \tag{4.13}$$

Here, $x_i \in [L, U] \subset (0, \infty)$, $\theta_1 > 0$, $\theta_2 > 0$ and $\theta_3 > 0$. Although they did not provide an analytical solution, their numerical solution shows that local MED-optimal designs are based on four points including the end point L. In fact, any optimal design under (4.13) can be based on a four-point design. Notice that the information matrix of Model (4.13) can be written as (4.12) except that $c_i = \exp(\frac{\theta_2 - x_i}{\theta_3})$ and

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & \dfrac{\theta_3}{\theta_1} & 0 \\ 0 & 0 & 0 & \dfrac{\theta_3}{\theta_1} \end{pmatrix}^{-1}. \tag{4.14}$$

Immediately, we have the following theorem.

THEOREM 7. *Under model (4.13), for any arbitrary design ξ, there exists a design ξ^* with at most four support points such that $I_\xi(\theta) \leq I_{\xi^*}(\theta)$.*

Theorem 7 confirms and extends Dette et al.'s (2008) numerical results.

REMARK 5. Li and Majumdar (2008) studied D-optimal design for a different version of four-parameter logistic model where

$$\eta(x,\theta) = \theta_0 + \frac{\theta_1}{1+\exp(\theta_2+\theta_3 x)}.$$

It can be shown that the information matrix can be written in the form of (4.12) with $c_i = \exp(\theta_2 + \theta_3 x_i)$ and

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & \dfrac{1}{\theta_1} & 0 \\ 0 & 0 & -\dfrac{\theta_2}{\theta_1} & -\dfrac{\theta_3}{\theta_1} \end{pmatrix}^{-1}. \tag{4.15}$$

Thus for an arbitrary design ξ, there exists a design ξ^* with at most four support points such that $I_\xi(\theta) \le I_{\xi^*}(\theta)$. This confirms and extends the results in Li and Majumdar (2008) for this model.

4.3. *Models with $p+1$ parameters.* Theorem 2 can be applied to many classical polynomial regression models. De la Garza (1954) studied a pth-degree polynomial regression model, which can be written in the form of (4.1) with

$$\eta(x,\theta) = \theta_0 + \sum_{i=1}^{p} \theta_i x^i, \tag{4.16}$$

where $x \in [-1, 1]$. De la Garza (1954) proved that any optimal design can be based on at most $p+1$ points including end points -1 and 1. Here, we provide an alternative way to prove this result under a general design region $[L, U]$.

THEOREM 8 (de la Garza Phenomenon). *Under model* (4.16), *for any arbitrary design ξ, there exists a design ξ^* with at most $p+1$ support points including end points L and U such that $I_\xi(\theta) \le I_{\xi^*}(\theta)$.*

PROOF. The information matrix can be written in the form of (2.1) with $P(\theta) = I_{(p+1)\times(p+1)}$ and

$$C(\theta, c_i) = \begin{pmatrix} 1 & c_i & \cdots & c_i^p \\ c_i & c_i^2 & \cdots & c_i^{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_i^p & c_i^{p+1} & \cdots & c_i^{2p} \end{pmatrix}, \tag{4.17}$$

where $c_i = x_i$. Let $\Psi_l(c) = c^l$, $l = 1, \dots, 2p$. We can check that the corresponding $f_{l,l} = l$ for $l = 1, \dots, 2p$. By applying case (c) of Theorem 2, we can draw the desired conclusion. $\square$

Weighted polynomial regression is an extension of Model (4.16), where the error terms ε_{ij}'s are i.i.d. $N(0, \sigma^2/\lambda(x))$ (σ^2 is known). Both Karlin and Studden (1966) and Dette, Haines and Imhof (1999) studied D-optimal designs under various choices of $\lambda(x)$ and design regions. Their results show that the number of support points of D-optimal designs is $p+1$ [except Lemma 2.2 of Dette, Haines and Imhof (1999) which has at most $p+2$ points]. By applying Theorem 2, we can extend their conclusions to any optimal designs. The results are summarized below:

THEOREM 9. *Under model* (4.16), *where the error terms* ε_{ij}*'s are i.i.d* $N(0, \sigma^2/\lambda(x))$ (σ^2 *is known*), *for an arbitrary design* ξ, *there exists a design* ξ^* *such that* $I_\xi(\theta) \le I_{\xi^*}(\theta)$. *Here*, ξ^* *is defined as follows*:

(i) ξ^* *is based on at most* $p+1$ *points when* $\lambda(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}$, $x \in [-1, 1]$, $\alpha + 1 > 0$ *and* $\beta + 1 > 0$.

(ii) ξ^* *is based on at most* $p+1$ *points including point* 0 *when* $\lambda(x) = \exp(-x)$ *and* $x \ge 0$.

(iii) ξ^* *is based on at most* $p+1$ *points when* $\lambda(x) = x^{\alpha+1}\exp(-x)$, $x \ge 0$, *and* $\alpha + 1 > 0$.

(iv) ξ^* *is based on at most* $p+1$ *points when* $\lambda(x) = \exp(-x^2)$.

(v) ξ^* *is based on at most* $p+1$ *points when* $\lambda(x) = (1+x^2)^{-n}$ *and* $p \le n$.

(vi) ξ^* *is based on at most* $p+2$ *points including either lower end point* L *or upper end point* U *when* $\lambda(x) = (1+x^2)^{-n}$ *and* $p > n$.

PROOF. The information matrix can be written in the form of (2.1) with $P(\theta) = I_{(p+1)\times(p+1)}$ and

$$C(\theta, c_i) = \begin{pmatrix} \lambda(c_i) & \lambda(c_i)c_i & \cdots & \lambda(c_i)c_i^p \\ \lambda(c_i)c_i & \lambda(c_i)c_i^2 & \cdots & \lambda(c_i)c_i^{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda(c_i)c_i^p & \lambda(c_i)c_i^{p+1} & \cdots & \lambda(c_i)c_i^{2p} \end{pmatrix}, \tag{4.18}$$

where $c_i = x_i$. The proofs are similar for all cases except for (ii), which can be proven with a similar approach as in Theorem 8 proof. Here we give the proof of Theorem 9(i). Let $\Psi_1(c) = -\int_0^c (1-t)^\alpha (1+t)^\beta dt$ and $\Psi_l(c) = (1-c)^{\alpha+1}(1+c)^{\beta+1}c^{l-2}$, $l = 2, \dots, 2p+2$. Notice that $\Psi_1(c)$ is not one of the elements in (4.18). We simply choose its value here for computation convenience. We can check that the corresponding $f_{1,1} < 0$ and $f_{l,l} > 0$ for $l = 2, \dots, 2p+2$. By applying case (d) of Theorem 2, the conclusion follows. □

4.4. *Loglinear model with quadratic term.* Theorem 2 is not limited to the model format (4.1). It can be applied to any nonlinear model, as long as the information matrix can be written in the form of (2.1). Here we give one such example.

Wang et al. (2006) studied D-optimal designs for loglinear models with a quadratic term, where

$$y_i \sim \text{Poisson}(\mu_i) \quad \text{and} \quad \log(\mu_i) = \theta_0 + \theta_1 x_i + \theta_2 x_i^2. \tag{4.19}$$

Here, $x_i \in [L, U]$. They showed that D-optimal designs are based on three points for some selected parameters by numerical searching. Their conclusion can be verified with the following theorem.

THEOREM 10. *Under model* (4.19), *for any arbitrary design* ξ, *there exists a design* ξ^* *such that* $I_\xi(\theta) \le I_{\xi^*}(\theta)$. *Here, when* $\theta_2 < 0$, ξ^* *is based on at most three points*; *when* $\theta_2 > 0$, ξ^* *is based on at most four points including the end points* L *and* U.

PROOF. It can be shown that the information matrix can be written in the form of (2.1) with

$$P(\theta) = \exp\left(\frac{4\theta_2\theta_0 - \theta_1^2}{8\theta_2}\right) \times \begin{pmatrix} 1 & 0 & 0 \\ \dfrac{\theta_1\sqrt{|\theta_2|}}{2\theta_2} & \sqrt{|\theta_2|} & 0 \\ \text{sign}(\theta_2)\dfrac{\theta_1^2}{4\theta_2} & \text{sign}(\theta_2)\theta_1 & \text{sign}(\theta_2)\theta_2 \end{pmatrix}^{-1} \quad \text{and} \tag{4.20}$$

$$C(\theta, c_i) = \begin{pmatrix} e^{\text{sign}(\theta_2)c_i^2} & c_i e^{\text{sign}(\theta_2)c_i^2} & c_i^2 e^{\text{sign}(\theta_2)c_i^2} \\ c_i e^{\text{sign}(\theta_2)c_i^2} & c_i^2 e^{\text{sign}(\theta_2)c_i^2} & c_i^3 e^{\text{sign}(\theta_2)c_i^2} \\ c_i^2 e^{\text{sign}(\theta_2)c_i^2} & c_i^3 e^{\text{sign}(\theta_2)c_i^2} & c_i^4 e^{\text{sign}(\theta_2)c_i^2} \end{pmatrix},$$

where $c_i = \sqrt{|\theta_2|}x_i + \frac{\theta_1\sqrt{|\theta_2|}}{2\theta_2}$. Let $\Psi_1(c) = \text{sign}(\theta_2)\int_0^c e^{\text{sign}(\theta_2)t^2}\,dt$ and $\Psi_l(c) = c^{l-2}e^{\text{sign}(\theta_2)c^2}$, $l = 2, \dots, 6$. We can verify that the corresponding (i) $f_{1,1} < 0$ when $\theta_2 < 0$ or $f_{1,1} > 0$ when $\theta_2 > 0$; (ii) $f_{l,l} > 0$ for $l = 2, \dots, 6$. By Theorem 2, the conclusion follows. $\square$

REMARK 6. Notice that the first derivative of $e^{\text{sign}(\theta_2)c^2}$ is 0 when $c = 0$. So we cannot apply Theorem 2 if c ranges from a negative to a positive number and $\Psi_1(c) = e^{\text{sign}(\theta_2)c^2}$. To avoid this situation, a specific $\Psi_1(c)$ is chosen although it is not among the functions in (4.20). This is the general strategy to handle such situations. The disadvantage of this strategy is that it could increase the number of support points unnecessarily.

5. Discussion. Deriving optimal designs for nonlinear models is complicated. Currently, the main tools are Elfving's geometric approach and Kiefer's equivalence theorem. Although these two approaches have been proven to be powerful tools, the results have to be derived on a case-by-case basis and some optimal designs are difficult to derive. In contrast, the proposed approach in this paper can yield very general results. As we have illustrated in the last section, for many commonly studied nonlinear models, this approach gives some simple structures based on which any optimal design can be found. As a result, it is a relatively easy to find an optimal design since one only needs to consider these simple structures. Many practical models have a moderate number of parameters (see Sections 4.1 and 4.2). For those models, any optimal designs can be derived readily. At a minimum, numerical search is feasible with the algorithm proposed by Stufken and Yang (2010).

The well-known Carathéodory theorem gives $p(p+1)/2$ as a upper bound for the number of support points in optimal designs. Examples in Section 4 show that the upper bound can be as small as p, the minimum number of support points such that all parameters are estimable. Although this may not be true for arbitrary nonlinear models, the proposed approach can be used to improve the upper bound. On the other hand, this approach gives an alternative way to prove the de la Garza Phenomenon with little effort. Furthermore, this phenomenon is extended for more general weighted polynomial regression models.

The proposed approach offers a lot of flexibility. It can be applied to multistage design, an important feature for locally optimal design. It works for any design region. The conditions are mild and can be easily verified using symbolic computational software packages, such as Maple or Mathematica.

While the results of this paper are already far reaching, we believe that there is potential to extend them further. In general, this approach can be applied to any nonlinear model as long as the corresponding functions are differentiable. One possible obstacle is that some $f_{l,l}$ may take the value 0. In this situation, the proposed approach may not be applied directly. One way to handle it is to introduce some new functions. For example, refer to the proof of Theorem 10. This, however, may increase the number of support points unnecessarily. How to handle this situation remains an open question for future research.

APPENDIX

This Appendix contains an outline of the proof of Lemma 1. A step-by-step complete proof, which contains a series of propositions and their respective proofs, is available in the full version of this paper, which can be found at http://arXiv.org.

First, we introduce two useful propositions which can support the proof of Lemma 1. The proofs of Propositions 1 and 2 are given in the complete proof at the above website.

PROPOSITION 1. *Let $\Psi_1, \dots, \Psi_k$ be $k = 2n$ functions defined on $[A, B]$. Assume that $f_{l,l}(c) > 0$, $c \in [A, B]$, $l = 1, \dots, k$. Suppose for some $c_1, \dots, c_n$ and $\tilde{c}_0, \dots, \tilde{c}_n$, where $A \le \tilde{c}_0 < c_1 < \tilde{c}_1 < c_2 < \tilde{c}_2 < \cdots < c_n < \tilde{c}_n \le B$, there exist*

$\omega_1, \ldots, \omega_n$ and $\tilde{\omega}_0, \ldots, \tilde{\omega}_n$, such that $\sum_{i=1}^n \omega_i = \sum_{j=0}^n \tilde{\omega}_j$ and the following $k-1$ equations hold:

$$\sum_{i=1}^{n} \omega_i \Psi_l(c_i) = \sum_{j=0}^{n} \tilde{\omega}_j \Psi_l(\tilde{c}_j), \qquad l = 1, \ldots, k-1. \tag{A.1}$$

If at least one of $\omega_1, \ldots, \omega_n$ and $\tilde{\omega}_0, \ldots, \tilde{\omega}_n$ is positive, then all of them should be positive. Under this situation,

$$\sum_{i=1}^{n} \omega_i \Psi_k(c_i) < \sum_{j=0}^{n} \tilde{\omega}_j \Psi_k(\tilde{c}_j). \tag{A.2}$$

PROPOSITION 2. *Let $\Psi_1, \ldots, \Psi_k$ be $k = 2n-1$ functions defined on $[A, B]$. Assume that $f_{l,l}(c) > 0$, $c \in [A, B]$, $l = 1, \ldots, k$. Suppose for some $c_1, \ldots, c_n$ and $\tilde{c}_1, \ldots, \tilde{c}_n$, where $A \le c_1 < \tilde{c}_1 < c_2 < \tilde{c}_2 < \cdots < c_n < \tilde{c}_n \le B$, there exist $\omega_1, \ldots, \omega_n$ and $\tilde{\omega}_1, \ldots, \tilde{\omega}_n$, such that $\sum_{i=1}^n \omega_i = \sum_{j=1}^n \tilde{\omega}_j$ and the following $k-1$ equations hold*:

$$\sum_{i=1}^{n} \omega_i \Psi_l(c_i) = \sum_{j=1}^{n} \tilde{\omega}_j \Psi_l(\tilde{c}_j), \qquad l = 1, \ldots, k-1. \tag{A.3}$$

If at least one number in $\omega_1, \ldots, \omega_n$ and $\tilde{\omega}_1, \ldots, \tilde{\omega}_n$ is positive, then all of them should be positive. Under this situation,

$$\sum_{i=1}^{n} \omega_i \Psi_k(c_i) < \sum_{j=1}^{n} \tilde{\omega}_j \Psi_k(\tilde{c}_j). \tag{A.4}$$

With the above two propositions, one can show that, if (3.3) and (3.4) of Lemma 1 hold, then inequalities in (3.5) will hold. Now we are one step away from proof of Lemma 1. We only need to prove that there exist $(\tilde{c}_j, \tilde{\omega}_j)$'s such that (3.3) and (3.4) hold. To do this, we need Proposition 3, which give some basic properties of $\tilde{c}_j$'s assuming Lemma 1 holds.

PROPOSITION 3. *Assume that Lemma 1 holds for $k \le K$. Let $(\tilde{c}_j, \tilde{\omega}_j)$'s be the solution set for given (c_i, ω_i)'s, and $\tilde{c}_0$ or $\tilde{c}_n$ (if applicable). Let ω_i^m be a sequence of bounded positive number for each i and $(\tilde{c}_j^m, \tilde{\omega}_j^m)$'s be the solution set with ω_i's being replaced by ω_i^m's, and all other values are fixed including $\tilde{c}_0$ or $\tilde{c}_n$ (if applicable). Then we have*:

(i) *$\tilde{c}_j$'s are unique.*

(ii) *If one of the ω_i sequences, say ω_{i_1} is an increasing sequence, and all other given values are the same (including $\tilde{c}_0$ or $\tilde{c}_n$ if applicable), then the $\tilde{c}_j$, $j < i_1$, are increasing sequences, and $\tilde{c}_j$, $j \ge i_1$, are decreasing sequences. On the other hand, if ω_{i_1} is an decreasing sequence and all other given values are the same, then $\tilde{c}_j$, $j < i_1$, are decreasing sequences, and $\tilde{c}_j$, $j \ge i_1$, are increasing sequences.*

(iii) *If* $\omega_i^m \to \omega_i$ *for all* i*'s, then* $\tilde{c}_j^m \to \tilde{c}_j$ *for all* j*'s.*

(iv) *If* $\omega_{i_1}^m \to 0$ *and* $\omega_{i_2}^m \nrightarrow 0$*, then either* $\underline{\lim}|\tilde{c}_{i_2-1}^m - c_{i_2}| = 0$ *or* $\underline{\lim}|\tilde{c}_{i_2}^m - c_{i_2}| = 0$.

(v) *Suppose that* $\omega_{i_1}^m \to 0$. *If there exists* $i_2 > i_1$, *such that* $\underline{\lim}\,\omega_i^m > 0$ *for* $i \geq i_2$, *then* $\tilde{c}_j^m \to c_{j+1}$ *for all* $j \geq i_2 - 1$. *If there exist* $i_3 < i_1$, *such that* $\underline{\lim}\,\omega_i^m > 0$ *for* $i \leq i_3$, *then* $\tilde{c}_j^m \to c_j$ *for all* $j \leq i_3$.

(vi) *If* $\underline{\lim}|\tilde{c}_{j_1}^m - c_{j_1+1}| = 0$, *then there exists a subsequence* $\{m_1\}$ *and* $i_1(\leq j_1)$, *such that* $\lim \omega_{i_1}^{m_1} = 0$ *and* $\lim|\tilde{c}_j^{m_1} - c_{j+1}| = 0$ *for* $i_1 \leq j \leq j_1$. *Similarly, if* $\underline{\lim}|\tilde{c}_{j_2}^m - c_{j_2}| = 0$, *then there exists a subsequence* $\{m_2\}$ *and* $i_2(> j_2)$, *such that* $\lim \omega_{i_2}^{m_2} = 0$ *and* $\lim|\tilde{c}_j^{m_2} - c_j| = 0$ *for* $j_2 \leq j < i_2$.

(vii) *Suppose that* $\omega_i^m < \omega_i$ *when* $i \leq i_1$ *and* $\omega_i^m = \omega_i$ *otherwise. If* $\tilde{c}_{j_1}^m \to \tilde{c}_{j_1}$, *for some* $j_1 \geq i_1$, *then* $\omega_i^m \to \omega_i$ *for all* i*'s and* $\tilde{c}_j^m \to \tilde{c}_j$ *for all* j*'s.*

(viii) *Let* $\tilde{c}_0^m$ *be a sequence numbers between* $\tilde{c}_0$ *and* c_1, *and suppose that* $\omega_i^m < \omega_i$ *when* $i \leq i_1$ *and* $\omega_i^m = \omega_i$ *otherwise. Let* $(\tilde{c}_j^m, \tilde{\omega}_j^m)$ *be the solution set in case* (a) *for given* (c_i, ω_i^m), $i = 1, \ldots, n$ *and* $\tilde{c}_0^m$. *If* $\tilde{c}_{n-1}^m \to \tilde{c}_{n-1}$, *then we must have* $\tilde{c}_0^m \to \tilde{c}_0$ *and* $\omega_i^m \to \omega_i$, $i = 1, \ldots, n$.

We will use mathematical induction to prove Lemma 1. When $k = 2$ and 3, Lemma 1 has been proven by Yang and Stufken (2009). We use the following two propositions to prove Lemma 1 for arbitrary k, that is, (i) assume Lemma 1 holds when $k \leq 2n - 2$, then show that it also holds for $k = 2n - 1$, and (ii) assume Lemma 1 holds when $k \leq 2n - 1$, then show that it also holds for $k = 2n$. Notice that once we can show there exist $(\tilde{c}_j, \tilde{\omega}_j)$'s such that (3.3) and (3.4) are satisfied, then by either Propositions 1 or 2, the inequality in (3.5) holds. To prove the existence of $(\tilde{c}_j, \tilde{\omega}_j)$'s, the strategy is similar for each of the four cases.

For the ease of presentation, we define $(C, \Omega) = \{(c_i, \omega_i)\}$, where $c_i < c_{i+1}$ and $\omega_i > 0$; $(\widetilde{C}, \widetilde{\Omega}) = \{(\tilde{c}_j, \tilde{\omega}_j)\}$, where $\tilde{c}_j < \tilde{c}_{j+1}$ and $\tilde{\omega}_j > 0$; $G_l(C, \Omega, \widetilde{C}, \widetilde{\Omega}) = \sum_i \omega_i \Psi_l(c_i) - \sum_j \tilde{\omega}_j \Psi_l(\tilde{c}_j)$. For given (C, Ω) with appropriate cardinality, let:

(i) $S_j^I(C, \Omega, \tilde{c}_0) = \tilde{c}_j$, where $\tilde{c}_j$'s are given under case (a);

(ii) $S_j^{II}(C, \Omega, \tilde{c}_n) = \tilde{c}_j$, where $\tilde{c}_j$'s are given under case (b);

(iii) $S_j^{III}(C, \Omega, \tilde{c}_0, \tilde{c}_n) = \tilde{c}_j$, where $\tilde{c}_j$'s are given under case (c);

(iv) $S_j^{IV}(C, \Omega) = \tilde{c}_j$, where $\tilde{c}_j$'s are given under case (d).

PROPOSITION 4. *If Lemma* 1 *holds for* $k \leq 2n - 2$, *then it will also hold for* $k = 2n - 1$.

PROOF. If case (a) holds, we can consider a new function set $\widetilde{\Psi}_1, \ldots, \widetilde{\Psi}_k$ on $[-B, -A]$. Here $\widetilde{\Psi}_i(c) = -\Psi_i(-c)$ when i is odd and $\widetilde{\Psi}_i(c) = \Psi_i(-c)$ when i is even. For the new function set, we can verify that the corresponding $f_{l,l} > 0$, $c \in [-B, -A]$, $l = 1, \ldots, k$. Let $C^- = \{-c_i, i = 1, \ldots, n\}$ and $\tilde{c}_0^- = -\tilde{c}_n$. Apply case (a) to the new function set $\widetilde{\Psi}_1, \dots, \widetilde{\Psi}_k$ with C^- and $\tilde{c}_0^-$, we obtain the solution set $\widetilde{C}^- = \{\tilde{c}_j^-, j = 0, \dots, n-1\}$. Let $\widetilde{C} = \{\tilde{c}_j = -\tilde{c}_{n-j}^-, j = 1, \dots, n\}$, then case (b) follows by replacing $\widetilde{\Psi}_i$ with Ψ_i. So we only need to prove case (a).

In this case, $(C, \Omega) = \{(c_i, \omega_i), i = 1, \dots, n\}$ and $\tilde{c}_0$ are given. It is sufficient to show that there exists a solution set $(\widetilde{C}, \widetilde{\Omega}) = \{(\tilde{c}_j, \tilde{\omega}_j), j = 0, \dots, n-1\}$ that satisfy (3.3) and (3.4) of Lemma 1.

Let $D = \{d_1, \dots, d_{n-1}, \omega_n\}$, where $d_i \in (0, \omega_i)$, $i = 1, \dots, n-1$. Define $\Omega^{\overline{D}} = \{\omega_1 - d_1, \dots, \omega_{n-1} - d_{n-1}\}$ and $C^{-n} = \{c_1, \dots, c_{n-1}\}$. We are going to show that for any given $d_{n-1} \in (0, \omega_{n-1})$, there exist d_i, $i = 1, \dots, n-2$, and $\tilde{c}_{n-1} \in (c_{n-1}, c_n)$, such that

$$S_j^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) = S_j^{IV}(C, D) \tag{A.5}$$

for $j = 1, \dots, n-1$. Once we show that (A.5) holds, we can let $\tilde{\omega}_j'$ be the corresponding weight of $S_j^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1})$, $j = 0, \dots, n-1$ and $\tilde{d}_j$ be the corresponding weight of $S_j^{IV}(C, D)$, $j = 1, \dots, n-1$. Define $\widetilde{C} = \{\tilde{c}_0, S_j^{IV}(C, D), j = 1, \dots, n-1\}$ and $\widetilde{\Omega} = \{\tilde{\omega}_0', \tilde{\omega}_j' + \tilde{d}_j, j = 1, \dots, n-1\}$, then

$$G_l(C, \Omega, \widetilde{C}, \widetilde{\Omega}) = 0, \qquad l = 1, \dots, 2n-3. \tag{A.6}$$

It can be shown that $G_{2n-2}(C, \Omega, \widetilde{C}, \widetilde{\Omega})$ is a continuous function of d_{n-1}. If we further show that $G_{2n-2}(C, \Omega, \widetilde{C}, \widetilde{\Omega})$ has different signs when $d_{n-1} \downarrow 0$ and $d_{n-1} \uparrow \omega_{n-1}$, then there must exist a $d_{n-1} \in (0, \omega_{n-1})$, such that $G_{2n-2}(C, \Omega, \widetilde{C}, \widetilde{\Omega}) = 0$. Then our conclusion follows.

This strategy will be achieved in three steps: (i) for any given $d_{n-1} \in (0, \omega_{n-1})$ and $\tilde{c}_{n-1} \in (c_{n-1}, c_n)$, there exists d_i, $i = 1, \dots, n-2$, such that (A.5) holds for $j = 1, \dots, n-2$; (ii) for any given $d_{n-1} \in (0, \omega_{n-1})$, there exist $\tilde{c}_{n-1} \in (c_{n-1}, c_n)$ and d_i, $i = 1, \dots, n-2$, such that (A.5) holds for $j = 1, \dots, n-1$; (iii) $G_{2n-2}(C, D, \widetilde{C}, \widetilde{\Omega})$ has different signs when $d_{n-1} \downarrow 0$ and $d_{n-1} \uparrow \omega_{n-1}$.

Step (i) can be proven by mathematical induction. We first show that for any given $d_i \in (0, \omega_i)$, $i = 2, \dots, n-1$ and $\tilde{c}_{n-1} \in (c_{n-1}, c_n)$, there exists $d_1 \in (0, \omega_1)$, such that (A.5) holds when $j = 1$. This is because when $d_1 \uparrow \omega_1$, we have

$$S_1^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) \to c_2 \quad \text{and} \quad S_1^{IV}(C, D) \to S_1^{IV}(C, D') < c_2, \tag{A.7}$$

where $D' = \{\omega_1, d_2, \dots, d_{n-1}, \omega_n\}$. This is due to (v) and (iii) of Proposition 3, respectively. When $d_1 \downarrow 0$, we have

$$\begin{gathered} S_1^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) \to S_1^{III}(C^{-n}, \Omega', \tilde{c}_0, \tilde{c}_{n-1}) < c_2 \quad \text{and} \\ S_1^{IV}(C, D) \to c_2, \end{gathered} \tag{A.8}$$

where $\Omega' = \{\omega_1, \omega_2 - d_2, \dots, \omega_{n-1} - d_{n-1}\}$. This is due to (iii) and (v) of Proposition 3, respectively. By (A.7) and (A.8), it is clear that $S_1^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) -$

$S_1^{IV}(C, D)$ is positive when $d_1 \uparrow \omega_1$ and negative when $d_1 \downarrow 0$. It can be shown that $S_1^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - S_1^{IV}(C, D)$ is a continuous function of d_1. Thus there exists a point d_1 such that (A.5) holds when $j = 1$. Notice that d_1 depends on d_i's, $i = 2, \dots, n-1$ and $\tilde{c}_{n-1}$.

Now, we assume that for any given $d_i \in (0, \omega_i)$, $i = p(\le n-2), \dots, n-1$, and $\tilde{c}_{n-1} \in (c_{n-1}, c_n)$, there exists $d_i \in (0, \omega_i)$, $i = 1, \dots, p-1$, such that (A.5) holds when $j = 1, \dots, p-1$. Consider any given $d_i \in (0, \omega_i)$, $i = p+1, \dots, n-1$ and $\tilde{c}_{n-1}$. By assumption, for any $d_p \in (0, \omega_p)$, there exists $d_i \in (0, \omega_i)$, $i = 1, \dots, p-1$ such that (A.5) holds when $j = 1, \dots, p-1$. When $d_p \downarrow 0$, by (v) of Proposition 3, we have

$$S_p^{IV}(C, D) \to c_{p+1}. \tag{A.9}$$

Next, we are going to show that $d_i \to 0$, $i = 1, \dots, p-1$ when $d_p \downarrow 0$. Suppose there exists some $i(<p)$, such that $d_i \nrightarrow 0$. Let i_1 be the smallest i that satisfies this condition. If $i_1 = 1$, then we have $\underline{\lim}|S_1^{IV}(C, D) - c_1| = 0$ by (iv) of Proposition 3. This implies that $\underline{\lim}|S_1^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - c_{i_1}| = 0$ since (A.5) holds for $j = 1, \dots, p-1$. If $i_1 > 1$, then by (iv) of Proposition 3, we have either $\underline{\lim}|S_{i_1-1}^{IV}(C, D) - c_{i_1}| = 0$ or $\underline{\lim}|S_{i_1}^{IV}(C, D) - c_{i_1}| = 0$. Suppose that $\underline{\lim}|S_{i_1-1}^{IV}(C, D) - c_{i_1}| = 0$, and then we have $\underline{\lim}|S_{i_1-1}^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - c_{i_1}| = 0$ by (A.5) again. By (vi) of Proposition 3, $\underline{\lim}(\omega_i - d_i) = 0$ for some $i \le i_1 - 1$. By the definition of i_1, we have $d_i \to 0$ for $i < i_1$. This is a contradiction. So we must have $\underline{\lim}|S_{i_1}^{IV}(C, D) - c_{i_1}| = 0$, which implies that $\underline{\lim}|S_{i_1}^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - c_{i_1}| = 0$ by (A.5).

By (vi) of Proposition 3, there exists a subsequence of $\{d_p \downarrow 0\}$ and $i_2 > i_1$ such that $\lim(\omega_{i_2} - d_{i_2}) = 0$ and $\lim |S_i^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - c_i| = 0$ for $i_1 \le i < i_2$. For this subsequence, by (iv) of Proposition 3 and the fact that $\lim d_{i_2} = \omega_{i_2}$, we have either $\underline{\lim}|S_{i_2-1}^{IV}(C, D) - c_{i_2}| = 0$ or $\underline{\lim}|S_{i_2}^{IV}(C, D) - c_{i_2}| = 0$. However, $\lim S_{i_2-1}^{IV}(C, D) = \lim S_{i_2-1}^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) = c_{i_2-1}$. Thus we must have $\underline{\lim}|S_{i_2}^{IV}(C, D) - c_{i_2}| = 0$, this also implies that $\underline{\lim}|S_{i_2}^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - c_{i_2}| = 0$.

By the exact same argument, there must exist $i_3 > i_2$ and a subsequence, such that $\lim(\omega_{i_3} - d_{i_3}) = 0$ and $\underline{\lim}|S_{i_3}^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) - c_{i_3}| = 0$. Repeat this argument again, and we can find strictly increasing numbers $i_4, i_5, \dots,$ and each of them has the same property as i_2 and i_3. Since p is finite, one of $\{i_2, i_3, i_4, \dots\}$ must be greater than or equal to p. This leads to a contradiction since all $d_i(<\omega_i)$, $i > p$ are fixed and $d_p \downarrow 0$. Thus we have $d_i \to 0$, $i = 1, \dots, p-1$ when $d_p \downarrow 0$. This implies that

$$S_p^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) \to S_p^{III}(C^{-n}, \Omega', \tilde{c}_0, \tilde{c}_{n-1}) < c_{p+1}, \tag{A.10}$$

where $\Omega' = \{\omega_1, \dots, \omega_p, \omega_{p+1} - d_{p+1}, \dots, \omega_{n-1} - d_{n-1}\}$. (A.10) is due to (iii) of Proposition 3. By (A.9) and (A.10), we have

$$S_p^{IV}(C, D) - S_p^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) > 0, \tag{A.11}$$

when $d_p \downarrow 0$.

On the other hand, notice that $d_p \uparrow \omega_p$ is equivalent to $\omega_p - d_p \downarrow 0$. We can show that the inequality sign in (A.11) will reverse using an analogous approach as used in the case of $d_p \downarrow 0$. Due to space limit, we will just give the outline of the proof here. First, we have

$$S_p^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) \to c_{p+1}. \tag{A.12}$$

Next, we are going to show that $d_i \to \omega_i$, $i = 1, \dots, p-1$ when $d_p \uparrow \omega_p$. Suppose there exists some $i(<p)$, such that $d_i \nrightarrow \omega_i$. Let i_1 be the smallest one. We can show that $\underline{\lim}|S_{i_1}^{IV}(C, D) - c_{i_1}| = 0$. This implies that there exists a subsequence of $\{d_p \uparrow \omega_p\}$ and $i_2 > i_1$ such that $\lim d_{i_2} = 0$ and $\underline{\lim}|S_{i_2}^{IV}(C, D) - c_{i_2}| = 0$. Repeat this argument, and we can find strictly increasing numbers $i_3, i_4, \dots$, and each of them has the same property as i_2. Since p is finite, one of $\{i_2, i_3, i_4, \dots\}$ must be greater than or equal to p. This leads to a contradiction since all d_i, $i > p$ are fixed and $d_p \uparrow \omega_p > 0$. Thus we have $d_i \to \omega_i$, $i = 1, \dots, p-1$ when $d_p \uparrow \omega_p$. This implies that

$$S_p^{IV}(C, D) \to S_p^{IV}(C, D') < c_{p+1}, \tag{A.13}$$

where $D' = \{\omega_1, \dots, \omega_p, d_{p+1}, \dots, d_{n-1}, \omega_n\}$. By (A.12) and (A.13), we have

$$S_p^{IV}(C, D) - S_p^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) < 0, \tag{A.14}$$

when $d_p \uparrow \omega_p$. It can be shown that $S_p^{IV}(C, D) - S_p^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1})$ is a continuous function of d_p. By (A.11) and (A.14), there must exist d_p, such that

$$S_p^{IV}(C, D) = S_p^{III}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}). \tag{A.15}$$

By mathematical induction, we have shown step (i).

Now we are going to prove step (ii). Let $\Omega_{d_{n-1}} = \{\omega_1, \dots, \omega_{n-2}, d_{n-1}, \omega_n\}$. Applying Lemma 1 when $k = 2n-2$, we have $\tilde{c}^*(d_{n-1}) \in (c_{n-1}, c_n)$, where $\tilde{c}^*(d_{n-1}) = S_{n-1}^{IV}(C, \Omega_{d_{n-1}})$. From step (i), we know that for any given d_{n-1} and $\tilde{c}_{n-1} \in (\tilde{c}^*(d_{n-1}), c_n)$, there exists d_i, $i = 1, \dots, n-2$, such that (A.5) holds for $j = 1, \dots, n-2$. It can be shown that $S_{n-1}^{IV}(C, D)$ is a continuous function of $\tilde{c}_{n-1}$. Then it is sufficient to show that $\underline{\lim}\, S_{n-1}^{IV}(C, D) > \tilde{c}_{n-1}$ when $\tilde{c}_{n-1} \downarrow \tilde{c}^*(d_{n-1})$ and $\overline{\lim} S_{n-1}^{IV}(C, D) < \tilde{c}_{n-1}$ when $\tilde{c}_{n-1} \uparrow c_n$.

Suppose that $\underline{\lim}\, S_{n-1}^{IV}(C, D) \le \tilde{c}_{n-1}$ when $\tilde{c}_{n-1} \downarrow \tilde{c}^*(d_{n-1})$. There exists a subsequence of $\tilde{c}_{n-1} \downarrow \tilde{c}^*(d_{n-1})$, such that $\lim S_{n-1}^{IV}(C, D) \le \tilde{c}_{n-1}$. Since $d_i < \omega_i$, $i = 1, \dots, n-2$, by (ii) of Proposition 3, we must have $S_{n-1}^{IV}(C, D) >$

$S^{IV}_{n-1}(C, \Omega_{d_{n-1}}) = \tilde{c}^*(d_{n-1})$. This implies that for the subsequence of $\tilde{c}_{n-1} \downarrow \tilde{c}^*(d_{n-1})$, $\lim S^{IV}_{n-1}(C, D) = \tilde{c}^*(d_{n-1})$. By (vii) of Proposition 3 and the fact that $d_i < \omega_i$, $i = 1, \dots, n-2$, we have $d_i \to \omega_i$, $i = 1, \dots, n-2$. Consequently, we have

$$\begin{aligned} &S^{IV}_{n-2}(C, D) \to S^{IV}_{n-2}(C, \Omega_{d_{n-1}}) < c_{n-1} \quad \text{and} \\ &S^{III}_{n-2}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) \to c_{n-1} \end{aligned} \tag{A.16}$$

by (iii) and (v) of Proposition 3, respectively. This is a contradiction to (A.5) when $j = n-2$.

Suppose that $\overline{\lim} S^{IV}_{n-1}(C, D) \geq \tilde{c}_{n-1}$ when $\tilde{c}_{n-1} \uparrow c_n$. There exists a subsequence of $\tilde{c}_{n-1} \uparrow c_n$, such that $\lim S^{IV}_{n-1}(C, D) \geq \tilde{c}_{n-1}$. By the assumption that Lemma 1 holds when $k = 2n-2$, we have $S^{IV}_{n-1}(C, D) < c_n$. This implies that for the subsequence of $\tilde{c}_{n-1} \uparrow c_n$, $S^{IV}_{n-1}(C, D) \to c_n$. By (vi) of Proposition 3, there exists a sub-subsequence of $\tilde{c}_{n-1} \uparrow c_n$ and $i_1 (< n-1)$ (notice that d_{n-1} is fixed), such that $\lim d_{i_1} = 0$ and $\lim S^{IV}_j(C, D) = c_{j+1}$ for $i_1 \leq j \leq n-1$. From the proof of step (i), $\lim d_{i_1} = 0$ means $\lim d_i = 0$ for $i \leq i_1$. On the other hand, we have $S^{III}_{i_1}(C^{-n}, \Omega^{\overline{D}}, \tilde{c}_0, \tilde{c}_{n-1}) = c_{i_1+1}$ by (A.5) holds for $j = i_1$. By (vi) of Proposition 3, there exists a sub-sub-subsequence of $\tilde{c}_{n-1} \uparrow c_n$ and $i_2 (\leq i_1)$, such that $\lim d_{i_2} = \omega_{i_2}$. This is a contradiction to $\lim d_{i_2} = 0$. This proves step (ii).

Now we are going to prove step (iii). By similar arguments as used in the proof of $\lim d_p = 0$ in step (i), which implies $\lim d_i = 0$ for $i \leq p$, we can show that $\lim d_i = 0$ for $i = 1, \dots, n-1$ when $d_{n-1} \downarrow 0$. Recall the definition of $(\widetilde{C}, \widetilde{\Omega})$ at the beginning of the proof,

$$G_{2n-2}(C, \Omega, \widetilde{C}, \widetilde{\Omega}) \to G_{2n-2}(C^{-n}, \Omega^{-n}, \widetilde{C}^{-n}, \widetilde{\Omega}^{-n}) < 0. \tag{A.17}$$

Here, $(\widetilde{C}^{-n}, \widetilde{\Omega}^{-n})$ is the solution set of $(C^{-n}, \Omega^{-n}) = \{(c_i, \omega_i), i = 1, \dots, n-1\}$ with given $\tilde{c}_0$ and $\tilde{c}_n (= c_n)$ under case (c) of Lemma 1 when $k = 2n-2$.

Similarly, we can show that $\lim d_i = \omega_i$ for $i = 1, \dots, n-1$ when $d_{n-1} \uparrow \omega_{n-1}$. Therefore, we have

$$G_{2n-2}(C, \Omega, \widetilde{C}, \widetilde{\Omega}) \to G_{2n-2}(C, \Omega, \widetilde{C}', \widetilde{\Omega}') > 0. \tag{A.18}$$

Here, $(\widetilde{C}', \widetilde{\Omega}')$ is the solution set of (C, Ω) under case (d) of Lemma 1 while $k = 2n-2$. (A.17) and (A.18) give the proof of step (iii). This completes the proof of Proposition 4. $\square$

PROPOSITION 5. *If Lemma 1 holds when $k \leq 2n-1$, then it also holds when $k = 2n$.*

PROOF. We first prove that case (c) holds. In this case, $(C, \Omega) = \{(c_i, \omega_i), i = 1, \dots, n\}$, $\tilde{c}_0$ and $\tilde{c}_n$ are given. It is sufficient to show that there exists a solution set $(\widetilde{C}, \widetilde{\Omega}) = \{(\tilde{c}_j, \tilde{\omega}_j), j = 0, \dots, n\}$ which satisfies (3.3) and (3.4) of Lemma 1.

The proof is similar to that in Proposition 4. Here we will provide an outline of the proof.

Define $D = \{d_i \in (0, \omega_i), i = 1, \dots, n\}$ and $\Omega^{\overline{D}} = \{\omega_i - d_i, i = 1, \dots, n\}$. We are going to show that for any given $d_n \in (0, \omega_n)$, there exist $d_i \in (0, \omega_i)$, $i = 1, \dots, n-1$, such that

$$S_j^I(C, \Omega^{\overline{D}}, \tilde{c}_0) = S_j^{II}(C, D, \tilde{c}_n) \tag{A.19}$$

for $j = 1, \dots, n-1$. Once we show that (A.19) holds, we can let $\tilde{\omega}_j'$ be the corresponding weight of $S_j^I(C, \Omega^{\overline{D}}, \tilde{c}_0)$, $j = 0, \dots, n-1$ and $\tilde{d}_j$ be the corresponding weight of $S_j^{II}(C, D, \tilde{c}_n)$, $j = 1, \dots, n$. Then define $\widetilde{C} = \{\tilde{c}_0, S_j^I(C, \Omega^{\overline{D}}, \tilde{c}_0), j = 1, \dots, n-1, \tilde{c}_n\}$ and $\widetilde{\Omega} = \{\tilde{\omega}_0', \tilde{\omega}_j' + \tilde{d}_j, j = 1, \dots, n-1, \tilde{d}_n\}$. Then we have

$$G_l(C, D, \widetilde{C}, \widetilde{\Omega}) = 0, \qquad l = 1, \dots, 2n-2. \tag{A.20}$$

If we further show that $G_{2n-1}(C, D, \widetilde{C}, \widetilde{\Omega})$ has different signs when $d_n \downarrow 0$ and $d_n \uparrow \omega_n$, then there must exist a $d_n \in (0, \omega_n)$, such that $G_{2n-1}(C, D, \widetilde{C}, \widetilde{\Omega}) = 0$. Then our conclusion follows.

This strategy will be achieved in two steps: (i) for any given $d_n \in (0, \omega_n)$, there exist d_i, $i = 1, \dots, n-1$, such that (A.19) holds for $j = 1, \dots, n-1$; and (ii) $G_{2n-1}(C, D, \widetilde{C}, \widetilde{\Omega})$ has different signs when $d_n \downarrow 0$ and $d_n \uparrow \omega_n$. The two steps can be proven similarly as in steps (i) and (iii) of Proposition 4.

Now we shall show that case (d) holds. In this case, $(C, \Omega) = \{(c_i, \omega_i), i = 1, \dots, n+1\}$. The proof is similar to the proof of case (a) in Proposition 4.

Let $D = \{d_2, \dots, d_n, \omega_{n+1}\}$, where $d_i \in (0, \omega_i)$, $i = 2, \dots, n$. Define $\Omega^{\overline{D}} = \{\omega_1, \omega_2 - d_2, \dots, \omega_n - d_n\}$, $C^{-1} = \{c_2, \dots, c_{n+1}\}$ and $C^{-(n+1)} = \{c_1, \dots, c_n\}$. We are going to show that for any given $d_n \in (0, \omega_n)$, there exists d_i, $i = 2, \dots, n-1$, $\tilde{c}_1$, and $\tilde{c}_n$ such that

$$S_j^{II}(C^{-(n+1)}, \Omega^{\overline{D}}, \tilde{c}_n) = S_{j-1}^I(C^{-1}, D, \tilde{c}_1) \tag{A.21}$$

for $j = 1, \dots, n$. Once we show that (A.21) holds, we can define

$$\widetilde{C} = \{S_j^{II}(C^{-(n+1)}, \Omega^{\overline{D}}, \tilde{c}_n), j = 1, \dots, n\}$$

with appropriate $\widetilde{\Omega}$ [similar as that of case (a)]. Then we have

$$G_l(C, \Omega, \widetilde{C}, \widetilde{\Omega}) = 0, \qquad l = 1, \dots, 2n-2. \tag{A.22}$$

If we further show that $G_{2n-1}(C, \Omega, \widetilde{C}, \widetilde{\Omega})$ has different signs when $d_n \downarrow 0$ and $d_n \uparrow \omega_n$, then there must exist a $d_n \in (0, \omega_n)$, such that $G_{2n-1}(C, \Omega, \widetilde{C}, \widetilde{\Omega}) = 0$. Thus our conclusion follows.

This strategy will be achieved with the following three steps: (i) for any given $d_n \in (0, \omega_n)$ and $\tilde{c}_n \in (c_n, c_{n+1})$, there exists $\tilde{c}_1 \in (c_1, c_2)$, d_i, $i = 2, \dots, n-1$, such that (A.21) holds for $j = 1, \dots, n-1$; (ii) for any given $d_n \in (0, \omega_n)$, there exists $\tilde{c}_n \in (c_n, c_{n+1})$, $\tilde{c}_1 \in (c_1, c_2)$, and d_i, $i = 2, \dots, n-1$, such that (A.21) holds for $j = 1, \dots, n$; (iii) $G_{2n-1}(C, D, \widetilde{C}, \widetilde{\Omega})$ has different signs when $d_n \downarrow 0$ and $d_n \uparrow \omega_n$.

Define $\tilde{c}_1 = S_j^{II}(C^{-(n+1)}, \Omega^{\overline{D}}, \tilde{c}_n)$ for given $d_2, \dots, d_n$ and $\tilde{c}_n$. Thus, we have $\tilde{c}_1 \in (c_1, c_2)$ and (A.21) holds $j = 1$.

The proof of steps (i), (ii) and (iii) are almost exactly the same as that of case (a). One only needs to change the notation and make two modifications in step (ii): first, $\tilde{c}^*(d_n) = S_{n-1}^I(C^{-1}, \Omega_{d_n}, c_1)$ where $\Omega_{d_n} = \{\omega_2, \dots, \omega_{n-1}, d_n, \omega_{n+1}\}$; second, use (viii) instead of (vii) of Proposition 3. This completes the proof of Proposition 5. □

REFERENCES

DE LA GARZA, A. (1954). Spacing of information in polynomial regression. *Ann. Math. Statist.* **25** 123–130. MR0060777

DETTE, H. (1997). Designing experiments with respect to “Standardized” optimality criteria. *J. Roy. Statist. Soc. Ser. B* **59** 97–110. MR1436556

DETTE, H., BRETZ, F., PEPELYSHEV, A. and PINHEIRO, J. (2008). Optimal designs for dose-finding studies. *J. Amer. Statist. Assoc.* **103** 1225–1237. MR2462895

DETTE, H., HAINES, L. M. and IMHOF, L. (1999). Optimal designs for rational models and weighted polynomial regression. *Ann. Statist.* **27** 1272–1293. MR1735997

DETTE, H., MELAS, V. B. and WONG, W. K. (2005). Optimal design for goodness-of-fit of the Michaelis–Menten enzyme kinetic function. *J. Amer. Statist. Assoc.* **100** 1370–1381. MR2236448

ELFVING, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23** 255–262. MR0047998

FANG, X. and HEDAYAT, A. S. (2008). Locally D-optimal designs based on a class of composed models resulted from blending Emax and one-compartment models. *Ann. Statist.* **36** 428–444. MR2387978

FORD, I., TORSNEY, B. and WU, C. F. J. (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *J. Roy. Statist. Soc. Ser. B* **54** 569–583. MR1160483

HAN, C. and CHALONER, K. (2003). D- and c-optimal designs for exponential regression models used in viral dynamics and other applications. *J. Statist. Plann. Inference* **115** 585–601. MR1985885

KARLIN, S. and STUDDEN, W. J. (1966). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815. MR0196871

KHURI, A. I., MUKHERJEE, B., SINHA, B. K. and GHOSH, M. (2006). Design issues for generalized linear models: A review. *Statist. Sci.* **21** 376–399. MR2339137

KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12** 363–366. MR0117842

LI, G. and MAJUMDAR, D. (2008). D-optimal designs for logistic models with three and four parameters. *J. Statist. Plann. Inference* **138** 1950–1959. MR2406416

PUKELSHEIM, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia, PA. MR2224698

STUFKEN, J. and YANG, M. (2010). On locally optimal designs for generalized linear models with group effects. Technical report, Dept. Statistics, Univ. Missouri.

WANG, Y., MYERS, R. H., SMITH, E. P. and YE, K. (2006). D-optimal designs for Poisson regression models. *J. Statist. Plann. Inference* **136** 2831–2845. MR2279838

YANG, M. and STUFKEN, J. (2009). Support points of locally optimal designs for nonlinear models with two parameters. *Ann. Statist.* **37** 518–541. MR2488361

DEPARTMENT OF STATISTICS
UNIVERSITY OF MISSOURI
COLUMBIA, MISSOURI 65211
USA
E-MAIL: yangmi@missouri.edu