



Optimal Design of Experiments for Hybrid Nonlinear Models, with Applications to Extended Michaelis–Menten Kinetics

Yuanzhi HUANG, Steven G. GILMOUR , Kalliopi MYLONA, and Peter GOOS

Biochemical mechanism studies often assume statistical models derived from Michaelis–Menten kinetics, which are used to approximate initial reaction rate data given the concentration level of a single substrate. In experiments dealing with industrial applications, however, there are typically a wide range of kinetic profiles where more than one factor is controlled. We focus on optimal design of such experiments requiring the use of multifactor hybrid nonlinear models, which presents a considerable computational challenge. We examine three different candidate models and search for tailor-made D- or weighted-A-optimal designs that can ensure the efficiency of nonlinear least squares estimation. We also study a compound design criterion for discriminating between two candidate models, which we recommend for design of advanced kinetic studies.

Supplementary materials accompanying this paper appear on-line

Key Words: Biochemistry; Compound criterion; D-optimality; Exchange algorithm; Weighted-A-optimality.

1. INTRODUCTION

Both empirical experimental research and physical and chemical theory play important roles in the applied biological sciences. Mechanistic models built from widely accepted and validated theories can help understanding and optimization of processes, though experimental data are often still required to validate the model and/or to estimate some unknown parameters in the models. Such experiments should be designed carefully to meet the research objectives, and optimal design of experiments (DoE) has an important role to play in ensuring the experiments produce informative data. The purely mechanistic models typically represent the effects of a single treatment factor on the response and are usually nonlinear.

Yuanzhi Huang, School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, UK. Steven G. Gilmour (✉) and Kalliopi Mylona Department of Mathematics, King's College London, Strand, London, UK (E-mail: steven.gilmour@kcl.ac.uk). Peter Goos, Department of Biosystems, KU Leuven, Leuven, Belgium.

© 2020 The Author(s)

Journal of Agricultural, Biological, and Environmental Statistics, Volume 25, Number 4, Pages 601–616
<https://doi.org/10.1007/s13253-020-00405-3>

When theory does not suggest a specific model, on the other hand, purely empirical models are used. This is often the case in experiments with multiple treatment factors, whose effects on the response, including interactions, need to be understood. In such experiments, the use of polynomial linear regression models to approximate the response surface is widespread. In this context, DoE is essential not only for parameter estimation, as with mechanistic models, but also for model selection and testing.

In reality, the distinction between mechanistic and empirical models is not so clear-cut. If theory gives the model form for the effect of a factor on the response, this model would be used in a single-factor experiment. However, it might be that this factor is only one among several which are being studied and there is no theory for the combined effects of all of the factors. Rather than abandoning the theory and using a purely empirical model, we should use a *hybrid model*, which reduces to the mechanistic model when the levels of all other factors are fixed, but uses an empirical form for the effects of these other factors on the response. Such hybrid models are used fairly extensively, though often rather informally, but we believe they should probably be used much more. Like for any other models, it is important to design experiments that allow the parameters of these hybrid models to be estimated efficiently and facilitate the selection and testing of the empirically chosen model parts. This paper develops optimal designs for such experiments.

Here, we focus on experiments on biochemical mechanisms involving kinetics leading to the Michaelis–Menten (M–M) model and its extensions. Such models are often seen in applied biosciences in areas including agriculture, environmental science, chemical engineering and drug development (PK/PD modeling). The nonlinear M–M model is derived as the stable solution of a set of theoretical differential equations with respect to the time course of a biochemical reaction. It describes the mechanistic association between the initial reaction rate v and the (initial) substrate (i.e. the reaction reagent) concentration S . With the substrate concentration chosen in n experimental runs, the M–M model is written as

$$v_i = \frac{v_{\max} S_i}{k + S_i} + \varepsilon_i, \text{ for } i = 1, 2, \dots, n, \quad (1)$$

where v_i is the initial reaction rate at observation i , S_i is the substrate concentration used at that observation, ε_i denotes the corresponding random error, v_{\max} is a parameter giving the maximal reaction rate, and k is a parameter called the Michaelis constant. Despite the one-factor structure, the two nonnegative parameter values of the M–M model depend on conditions such as the temperature and reaction time, different values of which could result in different kinetic profiles.

In general, a nonlinear regression model takes the form

$$Y_i = f(\mathbf{x}_i; \boldsymbol{\theta}) + \varepsilon_i, \quad (2)$$

for observations $i = 1, \dots, n$. In this expression, Y_i is the response at observation i , \mathbf{x}_i is the vector of factor settings at that observation and $\boldsymbol{\theta}$ is the vector of model parameters $\theta_1, \theta_2, \dots, \theta_p$. For nonlinear least squares (NLS) estimation, the errors ε_i are assumed to be uncorrelated and have constant variance σ^2 . The covariance matrix of the parameter estimators, $\sigma^2(\mathbf{F}^T \mathbf{F})^{-1}$, can then be written in terms of the model matrix \mathbf{F} , the (i, j) th element of which is given by

$$F_{ij} = \frac{\partial f(\mathbf{x}_i; \boldsymbol{\theta})}{\partial \theta_j}.$$

Depending on both the unknown “true” parameter values and the chosen factor levels of the design \mathbf{X} , $\mathbf{F}^T \mathbf{F}$ is identical to the Fisher information matrix, which measures the informativeness of the design.

While the M–M model gives a reliable approximation to the simplest biochemical mechanisms (e.g. for two-step reactions), more advanced kinetic studies require more complex nonlinear models. Such kinetic models, derived from a number of differential equations and reaction rate laws, are often written in terms of rational functions in the initial substrate concentration; e.g. see [Wong \(1975\)](#) for the biochemical theory and [Wang et al. \(2019\)](#) for recent work on parameter estimation. We can treat (1), a first-order rational function, as a special case of the higher-order kinetic models. In model-oriented studies, we need to build one or several reasonable tentative models based on relevant theories. The data collected would be used to estimate the treatment effects and unknown kinetic constants, requiring careful allocation of factor levels to different runs of the experiment.

Both standard and optimal design approaches can improve statistical results expected in treatment comparison, parameter estimation, statistical inference and model discrimination ([Cornish-Bowden 2014](#)). Standard multifactor designs are appropriate for purely empirical modeling. Apart from basic full factorial designs (such as 2^q or 3^q designs), central composite designs and Box–Behnken designs are commonly used for kinetic studies and second-order linear models are assumed to analyze the data. Nevertheless, in the event that a mechanistic model developed from (1) is assumed, we should construct and use a tailor-made optimal experimental design (OED) because the model is nonlinear in the kinetic constants. OEDs are highly flexible in that they can deal with any model structure and any constraints on the factor levels. Several authors have obtained approximate optimal designs for two-factor experiments with models based on the M–M equation, e.g. [Bogacka et al. \(2011\)](#), [Chen et al. \(2017\)](#) and [Schorning et al. \(2018\)](#). In this paper, we focus on exact designs, since they are more easily generalizable to more than two factors.

In Sect. 2, we introduce three nonlinear candidate models that might describe a two-factor M–M mechanism. DoE for such models is of considerable value but has received little attention in the literature. One interesting question we try to answer is to what extent the partially linear structure of each model can be exploited to simplify the resulting OED. We will use data from an experimental investigation of a similar M–M mechanism. Our results in Sects. 3 and 4 are based on the assumed model(s) and the parameter priors we calculate using these reference data. The same idea should fit various other applications in agriculture, environmental science, chemical engineering and medicine, as we discuss in Sect. 5.

2. MULTIFACTOR NONLINEAR KINETIC MODELS

2.1. ALTERNATIVE MODELS

While the M–M model can be used in theoretical studies of one-factor mechanisms, we focus on OED for multifactor M–M mechanisms requiring an extension of (1). In advanced

kinetics studies, often in an industrial context, there could be several factors to investigate but no biochemical theory explaining how these factors affect the reaction rate simultaneously. As factors such as temperature and pH influence M-M kinetics, a multifactor model should be used to describe the extended reaction mechanism. Common practice is to assume an empirical model that is, for instance, a second-order polynomial. However, since one of the factors is the substrate concentration defined in (1), when all other factors are held constant, the M-M model should hold as well. Therefore, it makes sense to assume a “hybrid model,” which could contain both M-M kinetics and some empirical model elements, as the best approximation to the kinetic data at hand. Similar models for multifactor response surfaces can be seen in Bogacka et al. (2017), Moyano et al. (2018), Strouwen et al. (2019) and Weilandt and Hatzimanikatis (2019).

Let x be a controlled variable in addition to the substrate concentration S . Then we can revise (1) to get, for instance, the *Additive Model*

$$v_i = \frac{(a_0 + a_1x_i + a_2x_i^2)S_i}{k + S_i} + \varepsilon_i, \text{ for } i = 1, 2, \dots, n. \quad (3)$$

The Additive Model is considered to be partially linear, as it involves a second-order linear predictor in x , which assumes a linear (in the parameters) relationship between x and the reaction rate v when the substrate concentration is fixed, but reduces to the M-M model when x is fixed. The model parameters are k , a_0 , a_1 and a_2 . This linear predictor in x can be modified before we decide on the most suitable model structure. It is also possible to substitute k with another linear function in x , though that could lead to an undesirably complex model and, in many biochemical mechanisms, factors such as the temperature and pH affect the response, but not the Michaelis constant.

In the Additive Model, we substitute the maximum reaction rate v_{\max} (a nonnegative parameter) with a linear function in x which can be negative. As an alternative, we can transform the linear function to derive the *Exponential Model*

$$v_i = \frac{\exp(a_0 + a_1x_i + a_2x_i^2)S_i}{k + S_i} + \varepsilon_i, \text{ for } i = 1, 2, \dots, n. \quad (4)$$

Again, when x_i is fixed, we obtain the M-M equation, but, this time, fixing S gives a simple nonlinear model in x .

This transformation also makes assumptions about the distribution of the errors ε , the structure of which can be additive, multiplicative, or something else. We have to speculate about the unknown error structure and that might indicate a different model too. For instance, Cornish-Bowden (2004, chap.140) conceived a M-M model of the form

$$v_i = \frac{v_{\max}S_i}{k + S_i}(1 + \varepsilon_i),$$

where ε_i is a normal random variable as part of the multiplicative error function. When the error term is not additive, the classical assumptions (e.g. constant variance) for NLS would be invalid. In this case, a transform-both-sides (TBS) model (Carroll and Ruppert 1984; Ruppert et al. 1989) can be used to stabilize the errors and better approximate the data.

In (4), supposing $\varepsilon_i^{(1)}$ indicates the additive error of the linear function in x , we assume an error function $\exp(\varepsilon_i^{(2)})$ and link it with M–M kinetics (this function is sensible given that the error has a potential to increase with the substrate concentration). As we take account of both error sources, the initial rate model can be written as

$$v_i = \frac{\exp(a_0 + a_1x_i + a_2x_i^2 + \varepsilon_i^{(1)})S_i}{k + S_i} \times \exp(\varepsilon_i^{(2)}),$$

where $\varepsilon_i^{(1)}$ and $\varepsilon_i^{(2)}$ are independent, with $\varepsilon_i^{(1)} \sim \mathcal{N}(0, \sigma_1^2)$ and $\varepsilon_i^{(2)} \sim \mathcal{N}(0, \sigma_2^2)$. To derive the *Transformed Model* that justifies the use of NLS, both sides of the above initial rate model should be transformed such that

$$v'_i = \log(v_i) = \log\left(\frac{S_i}{k + S_i}\right) + a_0 + a_1x_i + a_2x_i^2 + \varepsilon_i, \quad (5)$$

where $\varepsilon = \varepsilon^{(1)} + \varepsilon^{(2)} \sim \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$. In this case fixing S_i gives a linear (in the parameters) model in x_i , while fixing x_i produces the M–M model, but with multiplicative errors. Given the assumed multiplicative error structure, the Transformed Model should fit the data better than (4). When the true error structure is not known, as is usual in practice, one has to examine each candidate model to find the most appropriate one.

2.2. ILLUSTRATIVE APPLICATION: A MICHAELIS–MENTEN MECHANISM

Martins et al. (1999) studied a biochemical reaction mechanism that can be linked to pharmaceutical applications. In one of their kinetic studies, glyoxalase II was used as a protein catalyst (i.e. the enzyme), and S-D-lactoylglutathione (i.e. substrate) concentration levels S were set within the range [0.15, 3] mM (millimolars, the unit of concentration) in a 2-ml (milliliters) mixture. The measured response was the initial rate ν of formation of GSH, a form of the antioxidant glutathione (the desirable product), assuming M–M kinetics. Protein weight $E \in [0.02, 0.12]$ mg (milligram) was another controlled factor, in which case M–M model (1) can be written as

$$v_i = \frac{a_1 E_i S_i}{k + S_i} + \varepsilon_i, \quad (6)$$

where k and a_1 are unknown constants. As an alternative to (6), which is a pure mechanistic model, we assume one of models (3)–(5) that can be generalized to other case studies. Table 1 shows the 30-run reference design, which we approximated from the graphical illustration in Martins et al. (1999), since the raw data are not available, a common problem recently noted by Halling et al. (2018). We use this experiment as a proof of concept for two reasons. First, model (6) is one of the simplest models one can assume for the kind of data in Table 1. Second, Martins et al. (1999) presented two fitted models: fitted M–M model (1), based on the 15 runs of the reference design with the protein weight fixed at 0.03, and a first-order linear model in E , based on the other 15 runs with substrate concentration fixed at 1.5 mM. The benefit of hybrid modeling is to combine the separate models, so that we can use all the data to estimate fewer parameters and make optimal use of all available information.

Table 1. Data of 30-run reference experimental design.

E	S	100ν	E	S	100ν	E	S	100ν
0.023	1.5	0.0425	0.03	0.9	0.051	0.045	1.5	0.075
0.023	1.5	0.0475	0.03	0.9	0.0535	0.0675	1.5	0.13
0.023	1.5	0.0475	0.03	1.5	0.057	0.0675	1.5	0.135
0.03	0.15	0.023	0.03	1.5	0.06	0.0675	1.5	0.1375
0.03	0.15	0.024	0.03	1.5	0.063	0.089	1.5	0.17
0.03	0.15	0.025	0.03	3	0.07	0.089	1.5	0.18
0.03	0.3	0.0355	0.03	3	0.07	0.089	1.5	0.19
0.03	0.3	0.0365	0.03	3	0.07	0.112	1.5	0.225
0.03	0.3	0.037	0.045	1.5	0.07	0.112	1.5	0.225
0.03	0.9	0.0485	0.045	1.5	0.075	0.112	1.5	0.225

If we fit mechanistic model (6) to the data in Table 1, the NLS estimates are $\tilde{a}_1 = 2.422 (\pm 0.098) \times 10^{-2}$ and $\tilde{k} = 0.3290 (\pm 0.0696)$. To minimize the variances of the two estimators, we should use an OED, a design \mathbf{X} of n runs that, for instance, maximizes the D-criterion function $\phi_0 = \log|\mathbf{F}^T\mathbf{F}|$, the logarithm of the determinant of the Fisher information matrix $\mathbf{F}^T\mathbf{F}$. The criterion function involves the unknown parameter values, which we should replace with prior values (e.g. the NLS estimates based on reference data). For M–M model (6),

$$\mathbf{F} = \begin{bmatrix} \frac{E_1 S_1}{k+S_1} - \frac{a_1 E_1 S_1}{(k+S_1)^2} \\ \vdots \\ \frac{E_1 S_n}{k+S_n} - \frac{a_1 E_n S_n}{(k+S_n)^2} \end{bmatrix}$$

and the D-optimal design (at \tilde{a}_1 and \tilde{k}) can be derived from the General Equivalence Theorem, as shown in Dette and Biedermann (2003) and Matthews and Allcock (2004). In this case, this gives the 30-run D-optimal design

$$\mathbf{X} = \begin{pmatrix} (0.12, 0.2698) & (0.12, 3) \\ 15 & 15 \end{pmatrix}, \quad (7)$$

where the two optimal factor level combinations (E, S) are shown in the first row and the corresponding numbers of replicates are shown in the second row. Overall, it seems much easier to use the OED than the reference design and it also suggests that we can consider reducing the total number of runs to, for instance, 10 with five replicate runs for each factor level combination, although this depends on the required precision of the estimators. Dette and Biedermann (2003) showed that the optimal design does not depend on a_1 , but does depend on k . While the protein weight is fixed at its maximum (0.12 mg), we should use the estimated value \tilde{k} to determine the first substrate concentration at 0.2698. If design (7) is used, although we cannot fit a model with more than two unknown parameters, it is the most efficient for estimating (6).

When we define a linear transformation $x = (E - 0.07)/0.05 \in [-1, 1]$, as done throughout this paper, and assume one of multifactor models (3)–(5), the OEDs cannot be derived from theorems and it is necessary to optimize the experimental designs numerically. We use a MATLAB implementation of the multiphase OED algorithm developed by Huang et al. (2019), where the levels of the two controlled factors and the number of factor level combinations are optimized iteratively according to the D-criterion for each nonlinear model. The multiphase OED algorithm involves a combination of the traditional Fedorov exchange procedure (Fedorov 1972), a constrained Nelder–Mead method for continuous optimization and a modified iterative procedure for factor level correction. For the designs we report in the current paper, we executed the multiphase OED algorithm 30 times. All designs are based on one prior value of the model parameters. Hence, we use the locally optimal design approach to deal with the fact that OEDs depend on the unknown parameters.

In addition to the D-criterion, we also define and use a weighted-A-criterion to minimize a weighted sum of the variances of the parameter estimators,

$$\phi_1 = \text{Trace}(\mathbf{W}(\mathbf{F}^T \mathbf{F})^{-1}),$$

where the weight matrix \mathbf{W} is diagonal. We set the diagonal elements of \mathbf{W} to be the reciprocals of the variances calculated based on the D-optimal design. This approach gives a reasonable default weighting, assuming equal interest in all parameters—see Gilmour and Trinca (2012) for more details. In general, it is more difficult to optimize the weighted-A-criterion with the Fedorov exchange procedure, but this criterion gives a more natural summary of the parameter estimation properties of the design. To find weighted-A-optimal designs, we adapted the multiphase OED algorithm of Huang et al. (2019).

For D-, weighted-A- and other optimality criteria, it is natural to define the efficiency of a design with respect to the optimal design. The D-efficiency of a design is given by

$$\left(\frac{|\mathbf{F}^T \mathbf{F}|}{|\mathbf{F}^{*T} \mathbf{F}^*|} \right)^{\frac{1}{p}} = \frac{\exp\left(\frac{\phi_0}{p}\right)}{\exp\left(\frac{\phi_0^*}{p}\right)},$$

and the weighted-A-efficiency of a design is given by ϕ_1^*/ϕ_1 , where \mathbf{F}^* , ϕ_0^* and ϕ_1^* are obtained from the corresponding optimal design.

3. OPTIMAL DESIGN OF EXPERIMENTS

3.1. MODEL (3): THE ADDITIVE MODEL

From the data in Table 1, the reference estimates of $\boldsymbol{\theta} = \{k, a_0, a_1, a_2\}$ of (3) are $\tilde{\boldsymbol{\theta}} = \{0.3191, 0.0016, 0.0012, 0.0001\}$, which we take as the prior values for evaluating the D- and weighted-A-criterion functions. As a benchmark, the D-criterion function value of the reference DoE in Table 1 is $\phi_0 = -9.0594$ and the weighted-A-criterion function value is $\phi_1 = 40.3671$. A natural alternative benchmark design would be five replicates of the 3×2 factorial design, with levels 0.02, 0.07 and 0.12 for x and 0.2698 and 3 for S , which

has $\phi_0 = -4.9392$ and $\phi_1 = 5.8285$. For an OED of 30 runs, we find four factor level combinations that give the D-optimal design for (3):

$$\mathbf{X}_D = \begin{pmatrix} (0.02, 3) & (0.07, 3) & (0.12, 0.26) & (0.12, 3) \\ 8 & 8 & 7 & 7 \end{pmatrix}. \quad (8)$$

The D-criterion function value of the D-optimal design \mathbf{X}_D is -3.6864 , so that the reference design is $\exp(-9.0594/4)/\exp(-3.6864/4) \approx 26.10\%$ D-efficient with respect to (8) and the 3×2 factorial design is $\exp(-4.9392/4)/\exp(-3.6864/4) \approx 73.11\%$ D-efficient with respect to (8). The number of factor level combinations is equal to the number of unknown parameters, in which case the four numbers of replicate runs should be close to each other under the D-criterion (due to the factorization of the information matrix). The factor level combinations in the D-optimal design were obtained by optimizing the D-criterion function over the continuous factor space $\mathcal{X} = [0.02, 0.12] \times [0.15, 3]$.

The optimal protein weight levels are 0.02, 0.07 and 0.12 in (8): the first and third of these correspond to the minimum and maximum levels for this factor, whereas 0.07 is the central level. These three levels are often seen in conventional OEDs for second-order linear models. That these levels are optimal here too is not surprising, given the partially linear form of the model and the inclusion of a quadratic effect of the protein weight. The two optimal substrate concentration levels are 0.26 and 3. These levels are close to those in (7). Consequently, design (8) is somewhat similar to a one-factor-at-a-time OED. On the one hand, the 1st, 2nd and 4th factor level combinations contribute most to the estimation of a_0 , a_1 and a_2 , with 23 observations in total. On the other hand, the 3rd and 4th factor level combinations are included in the design to estimate the parameter k of the mechanistic function, with 14 observations in total.

With the Fisher information matrix and the parameter prior $\boldsymbol{\theta} = \{k, a_0, a_1, a_2\}$ obtained from the reference experimental design, we calculate the 4×4 covariance matrix, for illustrative purpose, for the 3×2 factorial design as

$$\mathbf{V}_F = \sigma^2 \mathbf{M}_0^{-1} = \sigma^2 \begin{pmatrix} 91800 & 79.432 & 60.267 & 5.0372 \\ 79.432 & 0.2513 & 0.0521 & -0.1782 \\ 60.267 & 0.0521 & 0.1309 & 0.0033 \\ 5.0372 & -0.1782 & 0.0033 & 0.2741 \end{pmatrix},$$

where σ^2 is an unknown constant that should be estimated from the experimental data. For the D-optimal design \mathbf{X}_D , the covariance matrix is

$$\mathbf{V}_D = \sigma^2 \mathbf{M}_1^{-1} = \sigma^2 \begin{pmatrix} 48606 & 24.037 & 38.744 & 22.031 \\ 24.037 & 0.1649 & 0.0192 & -0.1421 \\ 38.744 & 0.0192 & 0.1042 & 0.0144 \\ 22.031 & -0.1421 & 0.0144 & 0.2363 \end{pmatrix}.$$

In the latter covariance matrix, all the variances of the four parameter estimators are substantially smaller than in the former. Therefore, the D-optimal design is substantially more efficient than the 3×2 factorial design.

Table 2. Relative efficiencies (%) of experimental designs with respect to the optimal design.

Design	Model (equation number) and criterion						
	(3)	(4)		(5)		(4) and (16)	
	D	WA	D	WA	D	WA	Compound
(8)	100	95.69	93.20	81.40	70.42	53.94	–
(9)	97.65	100	91.87	76.97	73.36	60.87	–
(12)	88.38	78.73	100	92.52	63.96	50.89	44.24
(13)	85.57	75.83	94.14	100	68.26	52.91	47.71
(14)	72.77	71.92	68.59	59.57	100	96.17	–
(15)	76.03	79.63	71.64	63.52	97.98	100	–
(17)	89.11	85.73	91.24	82.66	64.11	52.09	100

The efficiencies in bold are for the criterion under which the corresponding design was obtained

Instead of calculating a D-optimal design, we can also compute a weighted-A-optimal design and minimize ϕ_1 instead of maximize ϕ_0 . The weighted-A-optimal design is

$$\mathbf{X}_{WA} = \begin{pmatrix} (0.02, 3) & (0.071, 3) & (0.12, 0.21) & (0.12, 3) \\ 6 & 10 & 7 & 7 \end{pmatrix}, \quad (9)$$

the weighted-A-criterion value of which is 3.8276. This value is smaller than that of the D-optimal design given in (8): the D-optimal design is $3.8276/4 = 95.69\%$ weighted-A-efficient (with the weighted-A-criterion value being exactly 4 because of the way in which the weight matrix \mathbf{W} is based on (8)). The reference design is only $3.8276/40.3671 \approx 9.48\%$ weighted-A-efficient, whereas the 3×2 factorial design is $3.8276/5.8285 \approx 65.67\%$ weighted-A-efficient. The latter two designs are thus clearly inferior to the optimal designs. The factor levels obtained in (9) are similar to those in (8), but the numbers of replicate runs are more unequal under the weighted-A-criterion. On the other hand, the D-criterion value of this weighted-A-optimal design (9) is -3.7815 , so that it is 97.65% D-efficient with respect to (8). Table 2 summarizes the efficiencies for most of the designs we show in this paper.

The true D- and weighted-A-efficiencies of OEDs for nonlinear models are dependent on the deviation of the prior values $\tilde{\theta}$ from the true parameter values of the assumed model. As an illustration, we examine the sensitivities of design (8) to this unknown deviation (i.e., to variation in the true parameter values) and quantify to what extent an inappropriate parameter prior would affect the D- or weighted-A-criterion for the OED. We consider 16 different scenarios $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_{16}$ as indicated in Table 3. The 16 scenarios correspond to the 2^4 combinations of two possible values for each of the four model parameters, labeled $-$ and $+$. The parameter values used are 0.1058 and 0.5324 for k , 0.0014 and 0.0019 for a_0 , 0.0011 and 0.0014 for a_1 and -0.0001 and 0.0003 for a_2 . Under each of scenarios $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_8$, for which the parameter k acts at its low level, the design

$$\mathbf{X}_{D-} = \begin{pmatrix} (0.02, 3) & (0.07, 3) & (0.12, 0.15) & (0.12, 3) \\ 8 & 8 & 7 & 7 \end{pmatrix} \quad (10)$$

Table 3. Assumed true parameter values under each of the 16 postulated scenarios.

	k	a_0	a_1	a_2		k	a_0	a_1	a_2
\mathfrak{S}_1	-	-	-	-	\mathfrak{S}_9	+	-	-	-
\mathfrak{S}_2	-	+	-	-	\mathfrak{S}_{10}	+	+	-	-
\mathfrak{S}_3	-	-	+	-	\mathfrak{S}_{11}	+	-	+	-
\mathfrak{S}_4	-	+	+	-	\mathfrak{S}_{12}	+	+	+	-
\mathfrak{S}_5	-	-	-	+	\mathfrak{S}_{13}	+	-	-	+
\mathfrak{S}_6	-	+	-	+	\mathfrak{S}_{14}	+	+	-	+
\mathfrak{S}_7	-	-	+	+	\mathfrak{S}_{15}	+	-	+	+
\mathfrak{S}_8	-	+	+	+	\mathfrak{S}_{16}	+	+	+	+

is optimal under the D-criterion. In this design, the substrate concentration level in the third factor level combination is its minimum value, 0.15. Under scenarios $\mathfrak{S}_9, \mathfrak{S}_{10}, \dots, \mathfrak{S}_{16}$, for which the parameter k acts at its high level, the D-optimal design is

$$\mathbf{X}_{D+} = \begin{pmatrix} (0.02, 3) & (0.07, 3) & (0.12, 0.39) & (0.12, 3) \\ 8 & 8 & 7 & 7 \end{pmatrix}. \tag{11}$$

Both designs are similar in structure to (8). This demonstrates the robustness of design (8) across different scenarios. The three designs in (8), (10) and (11) differ only in substrate concentration level in the third factor level combination. The three other factor level combinations, the numbers of replicates and the first entry of the third support point are the same for all three designs. The third factor level combination of design (8) lies between the corresponding factor level combinations of the two other designs (10) and (11). So, design (8) can be viewed as a compromise. The differing results for scenarios $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_8$, on the one hand, and scenarios $\mathfrak{S}_9, \mathfrak{S}_{10}, \dots, \mathfrak{S}_{16}$, on the other hand, suggest that only the value of k has an influence on the optimal design. This rather high robustness might be due to the partially linear nature of the assumed model, since, in the case of linear (in the parameters) models, the parameter prior has no impact on the OED.

To demonstrate that, in the present model, the optimal designs really do depend on parameters other than k , we also report the D-optimal design obtained when changing the prior value of a_2 by a factor of 100 to $100\tilde{a}_2 = 0.01$:

$$\mathbf{X}_{D,a_2} = \begin{pmatrix} (0.02, 0.37) & (0.02, 3) & (0.07, 3) & (0.12, 0.27) & (0.12, 3) \\ 1 & 7 & 8 & 7 & 7 \end{pmatrix}.$$

Unlike initial D-optimal design (8), the new one consists of five instead of four factor level combinations.

3.2. MODEL (4): THE EXPONENTIAL MODEL

Model (4) exhibits stronger nonlinear behavior than (3). The NLS estimate for that model is $\tilde{\theta} = \{\tilde{k}, \tilde{a}_0, \tilde{a}_1, \tilde{a}_2\} = \{0.3122, -6.4086, 0.8383, -0.2861\}$. We use this vector as

the prior parameter vector for (4), where \tilde{k} is close to the estimates for k in models (3) and (6). With the D-criterion, the optimal design is found to be

$$\mathbf{X}_D = \begin{pmatrix} (0.025, 3) & (0.086, 3) & (0.097, 0.33) & (0.12, 0.26) & (0.12, 3) \\ 8 & 7 & 1 & 7 & 7 \end{pmatrix}. \quad (12)$$

The D-criterion value of this design is $\phi_0 = -43.0242$. Compared with design (8) for model (3), the new design involves an extra treatment which is used just once. The factor level combination (0.12, 0.26) is used in both the D-optimal design for model (3) and in the new D-optimal design for model (4). With respect to design (12), the reference design (which has $\phi_0 = -48.2255$ and $\phi_1 = 34.8335$ for model (4)) is 27.24% D-efficient and the 3×2 factorial design (which has $\phi_0 = -44.5141$ and $\phi_1 = 6.4432$ for model (4)) is 68.90% D-efficient.

For model (4), the weighted-A-optimal design is

$$\mathbf{X}_{WA} = \begin{pmatrix} (0.023, 3) & (0.085, 3) & (0.098, 0.25) & (0.12, 0.21) & (0.12, 3) \\ 10 & 8 & 2 & 6 & 4 \end{pmatrix}, \quad (13)$$

the weighted-A-criterion value of which is $\phi_1 = 3.7008$. This implies that the reference design is 10.62% weighted-A-efficient and the 3×2 factorial design is 57.44% weighted-A-efficient. The D-criterion value of design (13) is $\phi_0 = -43.2658$ for model (4), so that it is 94.14% D-efficient with respect to design (12) (other comparisons of D- and weighted-A-efficiencies can be found in Table 2). While weighted-A-optimal design (13) involves factor level combinations that are similar to those of D-optimal design (12), its numbers of replicates are quite different.

3.3. MODEL (5): THE TRANSFORMED MODEL

Transform-both-sides model (5) involves a polynomial linear regression model in x and is only nonlinear in the parameter k . For that model, the information matrix and, hence, the optimal designs do not depend on the parameters a_0 , a_1 and a_2 . In other words, both the D-criterion and the weighted-A-criterion are independent of the values of a_0 , a_1 and a_2 . Fitting model (5) to the reference data results in the NLS estimate $\tilde{\theta} = \{\tilde{k}, \tilde{a}_0, \tilde{a}_1, \tilde{a}_2\} = \{0.2838, -6.4406, 0.8420, -0.2561\}$, which does not deviate much from that for untransformed model (4). When the prior value for k is $\tilde{k} = 0.2838$, the D-optimal design for model (5) is the equireplicated 3×2 design

$$\mathbf{X}_D = \begin{pmatrix} (0.02, 0.15) & (0.02, 3) & (0.07, 0.15) & (0.07, 3) & (0.12, 0.15) & (0.12, 3) \\ 5 & 5 & 5 & 5 & 5 & 5 \end{pmatrix}, \quad (14)$$

with a D-criterion value equal to $\phi_0 = 11.6960$. Hence, the reference design (which has $\phi_0 = 8.5739$ and $\phi_1 = 10.6548$ for model (5)) is 45.82% D-efficient. In this case, where we have a close-to-linear model, the D-optimal design thus is a 3×2 factorial design, which is essentially a standard design. A similar behavior has been observed for continuous optimal designs for linear models with noninteracting factors (Schwabe and Wierich 1995). The

benchmark 3×2 factorial design we defined in Sect. 3.1 and have been using throughout this paper involves an x value of 0.2698 instead of 0.15 for half of its runs, and has $\phi_0 = 11.3852$ and $\phi_1 = 4.2698$ for model (5). It is therefore 92.52% D-efficient.

Likewise, we can use the weighted-A-criterion for model (5) and find that the unequally replicated 3×2 factorial design

$$\mathbf{X}_{\text{WA}} = \begin{pmatrix} (0.02, 0.15) & (0.02, 3) & (0.07, 0.15) & (0.07, 3) & (0.12, 0.15) & (0.12, 3) \\ 4 & 5 & 4 & 8 & 4 & 5 \end{pmatrix} \quad (15)$$

is weighted-A-optimal. That design's weighted-A-criterion value is $\phi_1 = 3.8466$. The reference design is 36.10% weighted-A-efficient for model (5), while the benchmark 3×2 factorial design is 89.52% weighted-A-optimal. The D-criterion value of weighted-A-optimal design (15) is $\phi_0 = 11.6143$, so it is 97.98% D-efficient with respect to D-optimal design (14). Note that the factor level combinations are identical in the D-optimal and weighted-A-optimal designs, so that the designs only differ in the factor level combinations' replications.

4. A COMPOUND CRITERION

To construct an OED, we must focus on an assumed model that suits the reaction mechanism in question. We can start from one of the candidate models given in Sect. 2 or look for an even better model that reduces to the M–M model when all factors other than the substrate concentration are held constant. In some studies, however, there could be more than one kinetic model to be discriminated between as one of the aims of the experiment and the modeling (Cornish-Bowden 2014). The candidate models are often nonlinear and may involve more than one factor. Therefore, an efficient design should be robust to the unknown parameter values, as well as to several candidate models.

To discriminate between non-nested rival models, one can use a T-criterion (Atkinson and Fedorov 1975; Atkinson 2008) for selecting a design. In our kinetics application, however, the challenge is to determine a suitable polynomial model in x in for the empirical part of the model. As low-order polynomial models are nested within high-order polynomial models, we can make use of a compound criterion for our examples.

To illustrate this approach, we focus on the D-criterion and on an alternative for model (4). More specifically, as an alternative model, we consider

$$v_i = \frac{\exp(a_0 + a_1x_i + a_2x_i^2 + a_3x_i^3)S_i}{k + S_i} + \varepsilon_i, \text{ for } i = 1, 2, \dots, n. \quad (16)$$

Since, in polynomial models, third-order terms are often unimportant, it is reasonable to assume that the value of the parameter a_3 will be near zero.

To discriminate between two nested models (4) and (16), for instance using a two-tailed t test, we should certainly ensure that the parameter a_3 is estimated with a small variance. This would result in a high power for the t test. Therefore, in total, there are three criterion functions we should combine to obtain a suitable compound criterion: (i) a D_s -criterion

function ϕ_s to maximize the power of the t test for a_3 , (ii) the D-criterion function for model (4) (ϕ_{02} , of which \mathbf{M}_2 is the information matrix) and (iii) the D-criterion function for model (16) (ϕ_{01} , of which \mathbf{M}_1 is the information matrix).

The criterion ϕ_s can be formulated in terms of the ratio of the determinants of the information matrices \mathbf{M}_1 and \mathbf{M}_2 . If we denote by w_1 , w_2 and w_3 the weights for the three criterion functions ϕ_{01} , ϕ_s and ϕ_{02} , the compound criterion function to be maximized is

$$\begin{aligned}\phi_c &= w_1\phi_{01} + w_2\phi_s + w_3\phi_{02} = w_1\log|\mathbf{M}_1| + w_2\log\frac{|\mathbf{M}_1|}{|\mathbf{M}_2|} + w_3\log|\mathbf{M}_2| \\ &= (w_1 + w_2)\log|\mathbf{M}_1| + (w_3 - w_2)\log|\mathbf{M}_2|.\end{aligned}$$

When $w_1 = w_3 = 0$, the compound criterion concentrates on the t test and the discrimination between nested models (4) and (16). When $w_2 = w_3 = 0$ or $w_1 = w_2 = 0$, the compound criterion concentrates on one candidate model and the function is reduced to ϕ_{01} or ϕ_{02} . For the sake of illustration, we choose equal weights $w_1 = w_2 = w_3$. Since, for this choice, $w_2 = w_3$, the compound criterion reduces to the D-criterion for extended model (16).

When using the NLS estimate $\tilde{\theta} = \{\tilde{k}, \tilde{a}_0, \tilde{a}_1, \tilde{a}_2, \tilde{a}_3\} = \{0.3148, -6.4151, 0.8959, -0.2696, -0.0928\}$ as the prior parameter vector, we find that the design

$$\mathbf{X}_c = \begin{pmatrix} (0.02, 3) & (0.062, 3) & (0.1, 3) & (0.104, 0.3) & (0.12, 0.27) & (0.12, 3) \\ 6 & 6 & 6 & 2 & 4 & 6 \end{pmatrix}, \quad (17)$$

maximizes the compound criterion value, with $\phi_c = -56.3837$. Compared with design (12) for model (4), design (17) involves more levels for x , as a result of which it is suitable for a model involving a third-order polynomial in x . The new design also involves relatively few replicates of the factor level combination (0.12, 0.27), implying that this design pays less attention to the estimation of the parameter k than does design (12).

For alternative model (4), this design is $\exp(-43.3911/4)/\exp(-43.0242/4) = 91.24\%$ D-efficient with respect to design (12) which is D-optimal for model (4). More efficiencies for the design that maximizes the compound criterion are reported in Table 2. Note that several of the designs constructed in this do not allow model (16) to be fitted, so there are some blank entries (indicating a zero efficiency) in the table. This is due to the fact that these designs only involve three levels for x and at least four levels are required to fit a third-order polynomial model.

5. DISCUSSION

In this paper, our focus was on optimal designs for hybrid models, combining mechanistic models and empirical models for multi-factor experiments. For the sake of illustration, we focused on extensions of Michaelis–Menten kinetic models. Nevertheless, the ideas and methods we introduce here can also be applied to a variety of other mechanistic models in the applied biological sciences.

It is vital that experimental data contain sufficient information for mechanistic studies and allow for powerful statistical inference and correct conclusions. As noted in [Cornish-Bowden](#)

(2014), an efficient experimental design facilitates the collection of highly informative data. When compared to standard designs (e.g., 3^q factorials), all the OEDs we find in this paper are quite simple, in spite of the complex nonlinear structure of the assumed models. As a matter of fact, the OEDs are no more complex than standard DoEs for second-order linear models and are therefore applicable in practice. In our demonstrations, the OEDs were obtained using the multiphase optimization algorithm of Huang et al. (2019).

A kinetic reaction can be modeled in detail if we are able to derive the correct chemical equations, based on reaction steps or phases and the corresponding simultaneous differential rate laws. The kinetic model may eventually result from the closed-form solution of a set of differential equations. In less simple cases, as in Atkinson and Bogacka (2002), it is necessary to use a direct method to obtain an alternative numerical solution to the differential equations as well as to approximate the Fisher information of the resulting model.

To describe a non-Michaelis–Menten mechanism, the initial reaction rate is often modeled using a rational function in terms of the initial substrate concentration; e.g. see Wong (1975). The mechanistic model is then written in a form such as

$$v_i = \frac{a_1 + a_2 S_i + a_3 S_i^2 + a_4 S_i^3}{a_5 + a_6 S_i + a_7 S_i^2 + S_i^3} + \varepsilon_i.$$

These models should then be modified as well to incorporate the impact of experimental factors other than the substrate concentration. Because the most suitable rational function is generally unknown and the ideal way in which to incorporate the effect of extra experimental factors is usually not clear before data collection, compound optimality criteria for selecting suitable experimental designs are key for non-Michaelis–Menten models.

Kinetic model discrimination is quite difficult at times when data do not “place the models in jeopardy” (Box and Hunter 1965). In these cases, we cannot accept or reject candidate models because multiple models fit the observed data well. However, competing models may have complex structures, with some involving more kinetic constants than needed. To tackle difficulties in model selection, we recommend experimenters to start with kinetic models with fewer parameters (e.g., the three candidate models in Sect. 2, which assume M–M kinetics and a second-order linear function) in sequential mechanism studies. An advanced model can be examined on those occasions where there is solid prior evidence for such a postulated model. Otherwise, experimenters should be asked to simplify both the mechanistic and empirical components of the hybrid model under consideration, so as to reduce the total number of parameters. Although we can use a compound criterion, the number of candidate models should be minimized in advance if at all possible. This is to allow experimenters to better utilize their resources and to focus on the most realistic models.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

[Received November 2019. Accepted June 2020. Published Online July 2020.]

REFERENCES

- Atkinson, A. C. (2008). DT-optimum designs for model discrimination and parameter estimation. *Journal of Statistical Planning and Inference*, 138:56–64.
- Atkinson, A. C. and Bogacka, B. (2002). Compound and other optimum designs for systems of nonlinear differential equations arising in chemical kinetics. *Chemometrics and Intelligent Laboratory Systems*, 61:17–33.
- Atkinson, A. C. and Fedorov, V. V. (1975). The design of experiments for discriminating between two rival models. *Biometrika*, 62:57–70.
- Bogacka, B., Latif, A. H. M. M., Gilmour, S., and Youdim, K. (2017). Optimum designs for non-linear mixed effects models in the presence of covariates. *Biometrics*, 73:927–937.
- Bogacka, B., Patan, M., Johnson, P. J., Youdim, K., and Atkinson, A. C. (2011). Optimum design of experiments for enzyme inhibition kinetic models. *Journal of Biopharmaceutical Statistics*, 21:555–572.
- Box, G. E. P. and Hunter, W. G. (1965). The experimental study of physical mechanisms. *Technometrics*, 7:23–42.
- Carroll, R. J. and Ruppert, D. (1984). Power-transformations when fitting theoretical-models to data. *Journal of the American Statistical Association*, 79:321–328.
- Chen, P.-Y., Chen, R.-B., Tung, H.-C., and Wong, W. K. (2017). Standardized maximin D-optimal designs for enzyme kinetic inhibition models. *Chemometrics and Intelligent Laboratory Systems*, 86:79–86.
- Cornish-Bowden, A. (2004). *Fundamentals of Enzyme Kinetics*. Portland, London, 3rd edition.
- Cornish-Bowden, A. (2014). Analysis and interpretation of enzyme kinetic data. *Perspectives in Science*, 1:121–125.
- Dette, H. and Biedermann, S. (2003). Robust and efficient designs for the Michaelis-Menten model. *Journal of the American Statistical Association*, 98:679–686.
- Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.
- Gilmour, S. G. and Trinca, L. A. (2012). Bayesian L-optimal exact design of experiments for biological kinetic models. *Applied Statistics*, 61:237–251.
- Halling, P., Fitzpatrick, P. F., Raushel, F. M., Rohwer, J., Schnell, S., Wittig, U., Wohlgemuth, R., and Kettner, C. (2018). An empirical analysis of enzyme function reporting for experimental reproducibility: missing/incomplete information in published papers. *Biophysical Chemistry*, 242:22–27.
- Huang, Y., Gilmour, S., Mylona, K., and Goos, P. (2019). Optimal design of experiments for non-linear response surface models. *Applied Statistics*, 68:623–640.
- Martins, A. M., Cordeiro, C., and Freire, A. P. (1999). Glyoxalase II in *saccharomyces cerevisiae*: In situ kinetics using the 5,5'-dithiobis(2-nitrobenzoic acid) assay. *Archives of Biochemistry and Biophysics*, 366:15–20.
- Matthews, J. N. S. and Allcock, G. C. (2004). Optimal designs for Michaelis-Menten kinetic studies. *Statistics in Medicine*, 23:477–491.
- Moyano, F. E., Vasilyeva, N., and Menichetti, L. (2018). Diffusion limitations and Michaelis-Menten kinetics as drivers of combined temperature and moisture effects on carbon fluxes of mineral soils. *Biogeosciences*, 15:5031–5045.
- Ruppert, D., Cressie, N., and Carroll, R. J. (1989). A transformation/weighting model for estimating Michaelis-Menten parameters. *Biometrics*, 45:637–656.
- Schorning, K., Dette, H., Kettelhake, K., and Möller, T. (2018). Optimal designs for non-competitive enzyme inhibition kinetic models. *Statistics*, 52:1359–1378.
- Schwabe, R. and Wierich, W. (1995). D-optimal designs of experiments with non-interacting factors. *Journal of Statistical Planning and Inference*, 44:371–384.
- Strouwen, A., Nicolai, B. M., and Goos, P. (2019). Optimizing oxygen input profiles for efficient estimation of Michaelis-Menten respiration models. *Food and Bioprocess Technology*, 12:769–780.

- Wang, M., Jacobsen, E. W., Chotteau, V., and Hjalmarsson, H. (2019). A multi-step least squares method for nonlinear rational models. *American Control Conference*, pages 4509–4514.
- Weilandt, D. R. and Hatzimanikatis, V. (2019). Particle-based simulation reveals macromolecular crowding effects on the Michaelis-Menten mechanism. *Biophysical Journal*, 117:355–368.
- Wong, J. T. (1975). *Kinetics of Enzyme Mechanisms*. Academic Press, London.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.