Chemometrics and Intelligent Laboratory Systems 231 (2022) 104699

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

Optimal sample plans for multiresponse and multisubject experiments

J.M. Rodríguez-Díaz *

University of Salamanca, Spain

ARTICLE INFO

MSC:
62K05

Keywords:
Multiresponse model
Multisubject experiments
Optimal design
Two-way covariance
Time-depending experiments

ABSTRACT

In designed experiments quite often several variables are observed, which complicates the covariance structure of the data, especially when the objective is to observe the evolution of different characteristics along the time. Initial studies have been made assuming that the samples were taken on the same experimental unit, and thus some characteristics were assumed to keep constant for all the process. In the present work, a multisubject experiment is considered, which allows a variety of modifications of the previous paradigm: from considering different characteristics in the individuals, to equal or distinct models for the responses in each subject and the possibility to use several evolution models. Analytical results have been obtained, and an example were different bacteria are observed along the time has been analyzed and discussed.

1. Introduction

Multiresponse models have long been studied from the point of view of design of experiments, but usually with covariance structures close to the trivial ones, very often even assuming independent observations (see introductory section in [1]). However, as shown in that paper, in many situations a more complex covariance structure should be considered, or otherwise the designs employed when performing the corresponding experiments might be poorly informative. In that work, the study was focused in obtaining optimal designs for double-covariance structure models, with all the observations taken on the same experimental unit. The present study will enlarge and complete that research by considering multisubject models with different derivations, namely distinct individual characteristics, type of evolution model and equal or different response models for each variable in each subject.

1.1. Motivating example

Levan is a homopolysaccharide of fructose units usually produced by microorganisms, specially by *Zymomonas mobilis*, that is a Gram negative bacterium. These organisms use the enzyme levansucrase to break the sucrose bond between glucose and fructose, and then polymerize these fructoses. The fact that Levan is formed by furanoses units (fructose) instead of pyranoses as other biopolymers is what makes it very interesting for different applications in cosmetics, medicine or nanotechnology [2].

There are more bacteria that have been used for levan production, as *Bacillus subtilis* or new genera like *Hallomonas* or *Pseudomonas*, and a great number of works have been done for optimizing levan production. Most of them centered on the medium culture, but none of them analyzed the difference in levan production due to the type of bacteria, substrate and sucrose concentration till the work of González-Garcinuño et al. [3], that changed this tendency and studied the effect of the sucrose concentration on levan yield for types of bacteria that have not previously used for levan production, namely *Bacillus atrophaeus* and *Acinetobacter nectaris* (respectively Gram negative and Gram positive). In that study two variables, biomass growth and levan production, were measured five times every 8 h. Later on, Rodríguez-Díaz and Sánchez-León [1] propose different designs for these tests for one of the bacteria, *Bacillus atrophaeus*, in order to get better information about the parameters of the model. Now the study will be extended to multiple bacteria and/or cultures, in several ways.

1.2. Models and notation

Exact designs $\boldsymbol{\xi} = \{x_1, \ldots, x_n\}$, with x_i in the design space $\mathcal{X}$ and the size of the design, n, decided in advance, will be considered. In the general case x will be a vector containing the design conditions (controlled variables). For each response and subject linear models will be assumed, $y = \mathbf{f}(x)^T\boldsymbol{\beta} + u$, where $\boldsymbol{\beta}$ is the parameter vector of size m, u is the error term, and $\mathbf{f}(x) = (f_1(x), \ldots, f_m(x))^T$, with the $f_i(x)$ linearly independent in $\mathcal{X}$. Thus the one-response-one-subject model can be expressed as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U} \quad , \tag{1}$$

* Correspondence to: Department of Statistics-IUFFyM, Faculty of Science, Pl. de los Caídos s/n, 37008 Salamanca, Spain.
E-mail address: juanmrod@usal.es.

https://doi.org/10.1016/j.chemolab.2022.104699
Received 9 September 2022; Received in revised form 11 October 2022; Accepted 22 October 2022
Available online 7 November 2022

where $\mathbf{U} = (u_1, \dots, u_n)^T$ are the error terms, and $\mathbf{X} = (\mathbf{f}(x_1), \dots, \mathbf{f}(x_n))^T$ is the design matrix. For normal errors with 0 expectation and a covariance matrix $\boldsymbol{\Sigma} = Var(\mathbf{Y})$ describing the relation between the observations, the estimators of the model parameters is given by $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T \boldsymbol{\Sigma}^{-1}\mathbf{Y}$, with $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}} = Var(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T \Sigma^{-1}\mathbf{X})^{-1}$. The inverse of this covariance matrix,

$$\mathbf{M}(\xi) = \mathbf{X}^T \boldsymbol{\Sigma}^{-1}\mathbf{X} \quad , \tag{2}$$

is known as the *Information Matrix* of the design ξ for model (1). The objective is to find *optimal designs* $\boldsymbol{\xi}$ for the model, the ones that produce the most precise estimators of the model parameters, that is, those estimators that have minimum variance. In order to account for this, different 'variance measures' or *criterion functions* may be considered, most of them functions of the inverse of the information matrix (2). The most popular is D-optimality, that pays attention to the determinant of this inverse, while A-optimality focuses on the trace. Both of them have nice statistical properties: while D-optimality minimizes the volume of the confidence region of the estimators of the model parameters, A-optimality minimizes the average of the variances of these estimators. For non-linear models, the usual procedure is to linearize them by computing the derivatives with respect to the parameters. In this case the resulting information matrix (and thus the optimal designs) will depend on the unknown parameters (in fact it will depend only on the ones that appear 'non-linearly' in the model, Hill [4]), thus nominal values are needed for them, and the obtained designs will be *locally optimal*, good in an environment of those initial values. If no analytical expression of the model can be obtained, Rodríguez-Díaz and Sánchez-León [5] describe alternative methods for computing the derivatives. Some reference books on optimal design of experiments are Fedorov and Hackl [6], Pukelsheim [7] or Atkinson et al. [8].

When different responses $Y_1, \dots, Y_k$ should be observed, we talk about *multiresponse models*. In spite of the number of studies about optimal design of experiments in multiresponse models from different perspectives that can be found in literature (see the relation in [1]), in most of the cases the assumed correlation structure was quite simple or even null, assuming independent observations. However, that work showed that in many cases at least two types of correlation should be taken into account, namely *intra-correlation* between observations of different type taken on the same experimental unit at a specific design point x, and *inter-correlation* between observations of a specific variable taken on the same experimental unit at different design points $x_1, \dots, x_n$.

Let *sample* denote the set of values of the k variables observed on a experimental unit at a specific design point x, $\mathbf{y}(\mathbf{x}) = (y_1(x), \dots, y_k(x))^T$, and $\mathbf{S}(x) = cov[y_1(x), \dots, y_k(x)]$ its covariance matrix. It will be assumed that the relation between the different variables taken on the same subject will be similar for every x, thus in the following $\mathbf{S}$ will be considered non dependent on x. In a similar way, the intercovariance between the observations of each variable Y_i taken on the same experimental unit at different design points $x_1, \dots, x_n$ will be assumed similar for all the variables and thus the covariance matrix $cov[y_i(x_1), \dots, y_i(x_n)]$ will be denoted by $\mathbf{R}$ for all $i = 1, \dots, k$. This last assumption is not uncommon, especially when time is the only design variable. In that case it is usually assumed that the covariance between observations of the same variable taken on the same experimental unit at different design points x and x' will depend only on the distance between points, that is, $cov(y(x), y(x')) = \rho(|x - x'|)$, where ρ is a stationary covariance kernel (see for instance [9,10], Chapter 4). In the general case, with more than one controlled variable, it would be necessary to define a distance between design points (vectors), but the same idea could be applied. The assumption of the points $x_1, \dots, x_n$ being different avoids singular correlation matrices.

It will be assumed that the k variables will be observed at every point x_i (balanced designs). In other case, a procedure similar to the one used in Example 2 of [11] can be employed. Depending on the case, for the $k \times n$ observations taken for each experimental unit/subject sometimes will be more convenient to use the order $y_1(x_1), \dots, y_k(x_1), \dots, y_1(x_n), \dots, y_k(x_n)$, (design-point wise ordering), obtaining the global covariance matrix $\Sigma = R \otimes S$, and in other situations it will be more clever to use the order $y_1(x_1), \dots, y_1(x_n), \dots, y_k(x_1), \dots, y_k(x_n)$, (variable-type wise ordering), getting the covariance matrix $\Sigma = S \otimes R$.

Multiple subjects could be observed as well, always assuming that they will be independent of each other. They could share the same covariance characteristics (that is the same matrices $\mathbf{S}$ and $\mathbf{R}$ are valid for all of them), or specific ones. The former case would have the same effect than blocking, or repeating the experiment a number N of times, and in the following the same notation will be used for the model with similar blocks or with subjects with the same characteristics. However, when there are B subjects with different characteristics, it will be necessary to deal with several $\mathbf{S}_\mathbf{v}$ and $\mathbf{R}_\mathbf{v}$ and the computations become harder. The general model will have B different subject and N blocks. For each subject k different variables will be observed at the design points $x_1, \dots, x_n$, producing $N \times B \times k \times n$ observations. The observations vector will be

$$\mathbf{Y}_{(\mathbf{N})} = \left(\mathbf{Y}_\mathbf{1}^T, \dots, \mathbf{Y}_\mathbf{N}^T\right)^T \quad , \tag{3}$$

with $\mathbf{Y}_\mathbf{j} = \left(\mathbf{Y}^{\mathbf{j(1)}T}, \dots, \mathbf{Y}^{\mathbf{j(B)}T}\right)^T$, $\mathbf{Y}^{\mathbf{j(v)}} = \left(\mathbf{Y}_1^{\mathbf{j(v)}T}, \dots, \mathbf{Y}_k^{\mathbf{j(v)}T}\right)^T$, and $\mathbf{Y}_i^{\mathbf{j(v)}} = \left(y_{i1}^{j(v)}, \dots, y_{in}^{j(v)}\right)^T$, with $y_{il}^{j(v)}$ the observation of the ith variable taken of the vth subject at the lth design point and corresponding to the jth block. The global covariance matrix will be $\boldsymbol{\Sigma}_{(\boldsymbol{N})} = \mathbf{I}_\mathbf{N} \otimes \boldsymbol{\Sigma}$, with

$$\boldsymbol{\Sigma} = Diag\{\mathbf{S}_\mathbf{1} \otimes \mathbf{R}_\mathbf{1}, \dots, \mathbf{S}_\mathbf{B} \otimes \mathbf{R}_\mathbf{B}\} \quad , \tag{4}$$

the covariance matrix for each block, where $Diag\{a, b, \dots\}$ denotes the diagonal matrix with diagonal elements $a, b, \dots$ and $Diag_B\{a\}$ the diagonal matrix with B elements in the diagonal equal to a.

The core of the paper is contained in Section 2. After a brief summary of contents of previous works, quite a few new results are produced, specially for the model that allows different characteristics for the subjects. In Section 3 some of these analytical results will be applied to variations of the example described in Section 1.1. Section 4 discusses further extensions and derivations of the subject, and the final section contains the main conclusions.

2. Analytical results

Rodríguez-Díaz and Sánchez-León [1] obtain analytical results for the biresponse model observed in one subject, both for models without common parameters and for nested models. A result for k-response models is also shown, assuming constant covariance between any pair of responses and always for one experimental unit. Later on, Rodríguez-Díaz et al. [12] introduce the research for multiple subjects, assuming for all of them similar characteristics (namely the same correlation structures), as it would be the case when observing multiple variables on a cohort of students of similar age and belonging to the same classroom.

Now a generalization of previous studies will be presented. The situation we are interested in is measuring k variables for some experimental units at several design points $\boldsymbol{\xi} = \{x_1, \dots, x_n\}$. Initially (until Section 2.3) only one experiment (one block) will be considered, and thus in these sections for easiness of expressions the blocks notation will be removed from the models. The subjects may have different characteristics, especially inter and intra covariance. In the following, the example described in Section 1.1 will be used as common thread, that is, several responses may be observed in different cultures (treated as blocks, or subjects with similar characteristics) and/or bacteria (considered as subjects with different characteristics) along the time. Regarding this, [1] center on looking for the best observation times for several variables in **1 culture of 1 bacterium**. Possible extensions of

this case will follow, always for multiple responses and always having in mind obtaining D-optimal designs (sometimes A-optimal designs as well) for the response models. Let m be the number of parameters of the longest model. From now on m-parameter linear models will be assumed for all the responses, taking into account two possible situations:

Scenario I. — Each response variable adopts a particular model for each experimental unit

$$\hat{y}_i^{(v)}(x) = \mathbf{f}(x)^T \boldsymbol{\beta}_i^{(v)} \,, \quad \boldsymbol{\beta}_i^{(v)} = (\beta_{i1}^{(v)}, \dots, \beta_{im}^{(v)})^T \tag{5}$$

Scenario II. — Each response has the same model for all experimental units

$$\hat{y}_i^{(v)}(x) = \mathbf{f}(x)^T \boldsymbol{\beta}_i \,, \quad \boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{im})^T \tag{6}$$

where $i = 1, \dots, k$ denotes the response and $v = 1, \dots, B$ the experimental unit. Scenario I is the general case in which the subjects may be so different that the variables may have distinct models for them. It is clear that Scenario II is a particular case of Scenario I; however the results corresponding to both cases are so different (curiously Scenario II is more difficult to deal with) that it seems convenient to make this distinction. In both scenarios the covariance structure ($\mathbf{S}$, $\mathbf{R}$) may be different or not.

2.1. Observing several cultures of the same bacterium

In this case the different subjects are the cultures. It is sensible to assume the same intra and inter covariance for all the cultures of a specific bacterium, thus this situation is similar to that in [12], and some results obtained therein can be applied, namely:

Theorem 2.1. *Results for Scenario I:*

- *The D-optimal designs for the individual models of each variable in each experimental unit are as well D-optimal.*
- *The parameters of the individual models can be estimated independently and do not depend on* $\mathbf{S}$*, but their covariance matrix does depend on* $\mathbf{S}$*.*

Theorem 2.2. *Results for Scenario II:*

- *The D-optimal designs for the individual models of each variable in each experimental unit are as well D-optimal.*
- *The estimation of the parameters of each response is the average of the estimations for the different experimental units and does not depend on* $\mathbf{S}$*, but again their covariance matrix does depend on it.*

2.2. Observing 1 culture of each one of B different bacteria

In this setup the different subjects are the bacteria types, each one with its particular covariance characteristics, $\mathbf{S_v}$ and $\mathbf{R_v}$. We are interested in observing k variables in cultures of B different bacteria (one culture of each type of bacterium) at n different design points. Let $y_{ij}^{(v)}$ denote the observation of the i response variable taken on the subject v at the design point j, $i = 1, \dots, k$, $j = 1, \dots, n$, $v = 1, \dots, B$. The vector of the $B \times k \times n$ observations will be

$$\mathbf{Y} = \left(\mathbf{Y^{(1)}}^T, \dots, \mathbf{Y^{(B)}}^T\right)^T \quad ,$$

with $\mathbf{Y^{(v)}} = \left(\mathbf{Y_1^{(v)}}^T, \dots, \mathbf{Y_k^{(v)}}^T\right)^T$, and $\mathbf{Y_i^{(v)}} = \left(y_{i1}^{(v)}, \dots, y_{in}^{(v)}\right)^T$. The global covariance structure is given by (4), and the design matrix for the individual model of each variable in each subject, no matter whether the model is valid for all the subjects (Scenario II) or varying between subjects (Scenario I), is given by

$$\mathbf{X_0} = \left(\mathbf{f}(x_1)^T, \dots, \mathbf{f}(x_n)^T\right)^T \quad .$$

Scenario I

Let us begin with the first scenario described by (5). Regarding the optimal designs:

Theorem 2.3. *If the response models are given by (5)*

- *The D-optimal designs do not depend on the intra-covariances* S_v*, but the A-optimal designs do.*
- *If a design is D(A)-optimal for the individual models of each variable in each subject it will be globally D(A)-optimal*

The proof is deferred to Appendix. Now, paying attention to the estimation of the parameters:

Theorem 2.4. *Assuming the conditions of Scenario I (5), the parameters of the individual models can be estimated independently and do not depend on the intracovariance matrices* $\mathbf{S_v}$*, but their covariance matrix does depend on them.*

The proof is in Appendix.

Scenario II

Let us now assume the conditions given by Scenario II (6).

Theorem 2.5. *Assuming the conditions of the models (6) in Scenario II A- and D-optimal designs would in general depend on the intracovariances* $\mathbf{S_v}$*. However, if a design is D-optimal for the individual models of each variable in each subject it will be globally D-optimal. Furthermore:*

- *If* $\mathbf{S_v} = \mathbf{S}$ $\forall v$ *the A- and D-optimal designs do not depend on* $\mathbf{S}$*.*
- *If* $\mathbf{R_i} = \mathbf{R}$ $\forall i$ *the D- and A optimal designs do not depend on* $\mathbf{S_v}$ *and if a design is optimal for any response in any subject it will be globally optimal.*

The proof is deferred to Appendix.

The results about the estimation of the parameters are shown in the following theorem:

Theorem 2.6. *Assuming the conditions of the models (6) in Scenario II, the estimator of the model parameters can be computed as*

$$\hat{\tilde{\boldsymbol{\beta}}} = \left(\sum_{v=1}^{B} \mathbf{S_v}^{-1} \otimes \tilde{\mathbf{M}}_\mathbf{v}\right)^{-1} \left(\sum_{v=1}^{B} \left(\mathbf{S_v}^{-1} \otimes \mathbf{X_0}^T \mathbf{R_v}^{-1}\right) \mathbf{Y^{(v)}}\right) \quad .$$

Furthermore, if $\mathbf{S_v} = \mathbf{S}$ *for all* v*, then* $\hat{\tilde{\boldsymbol{\beta}}}$ *will not depend on the intracovariance* $\mathbf{S}$*. If, in addition,* $\mathbf{R_v} = \mathbf{R}$ *for all* v*, the estimation of the model parameters for each variable will be the average of the estimations for the B subjects, thus generalizing Theorem 2.2.*

The proof is in Appendix.

Corollary 2.1. *Under the conditions of Theorem 2.6, when the intracovariance remains specific for each subject,* $\mathbf{S_v}$*, but* $\mathbf{R_v} = \mathbf{R}$ *is the same for every subject, then the parameters estimation is a kind of weighted mean of the estimations in the subjects, where the 'weight' of each subject is proportional to the subject's intracovariance. Due to this relationship between samples of the different responses taken on the same subject at the same design point, this 'weight' makes that the estimations of all the responses in every subject take part in the final estimation of the model for each response.*

Proof. Regarding the last statement, let us now assume that $\mathbf{R_v} = \mathbf{R}$ for all v but the $\mathbf{S_v}$ are different. Then $\tilde{\mathbf{M}}_\mathbf{v} = \mathbf{X_0}^T \mathbf{R_v} \mathbf{X_0}$ is the same for every subject, let us say $\tilde{\mathbf{M}}_\mathbf{0}$, and thus from (12)

$$\begin{aligned}\hat{\tilde{\boldsymbol{\beta}}} &= \left[\left(\sum_{v=1}^{B} \mathbf{S_v}^{-1}\right)^{-1} \otimes \tilde{\mathbf{M}}_\mathbf{0}^{-1}\right] \left(\sum_{v=1}^{B} \left(\mathbf{S_v}^{-1} \otimes \mathbf{X_0}^T \mathbf{R}^{-1}\right) \mathbf{Y^{(v)}}\right) \\ &= \sum_{v=1}^{B} \left(\mathbf{S_v}^{-1}\right) \left(\sum_{i=1}^{B} \mathbf{S_i}^{-1}\right)^{-1} \left\{\tilde{\mathbf{M}}_\mathbf{0}^{-1} \mathbf{X_0}^T \mathbf{R}^{-1} \mathbf{Y_j^{(v)}}\right\}_{j=1,\dots,k} \quad .\end{aligned}$$

Then the estimation of the model for each response is a weighted average, with weights $\left(\mathbf{S_v}^{-1}\right)\left(\sum_{i=1}^{B} \mathbf{S_i}^{-1}\right)^{-1}$ for $v = 1, \dots, B$, of the estimations of the models in each subject, $\left\{\tilde{\mathbf{M}}_\mathbf{0}^{-1} \mathbf{X_0}^T \mathbf{R}^{-1} \mathbf{Y_j^{(v)}}\right\}_{j=1,\dots,k}$.

And the effect of using these weights is that actually the final estimation of the model for the jth response uses as well the estimation of every other response individual model. □

2.3. Observing N cultures of B different bacteria

Let us now add a new layer to the structure, that is consider the case in which N cultures of B different bacteria types are observed, taking in every culture n observations for each one of the k variables of interest along the time. That means a total of $N \times B \times k \times n$ observations, and again the objective is obtaining the optimal designs for the estimation of the model.

The problem may be approached as performing N experiments similar to the one described in Section 2.2. These experiments can be seen as N independent blocks under similar conditions, thus we have the general model (3), with the following results:

Theorem 2.7. *The information matrix for the experimental structure and models described in Section 2.3 is proportional to the information matrix for the models of Section 2.2, for both scenarios. Therefore the conclusions of Theorems 2.3 and 2.5 remain true for this case.*

The proof is shown in Appendix. And about the parameters estimators

Theorem 2.8. *Assuming the situation described in Section 2.3 in which N independent experiments are performed under similar conditions, the estimator of the model parameters is the average of the estimators obtained in the individual experiments.*

The proof is deferred to Appendix.

2.4. Biresponse model

Very often just two variables of interest are observed in a experiment (see for instance the relation in [1]). In that paper some results are obtained assuming homoskedasticity and constant covariance between different-type observations. Now some of those results will be generalized for general 2×2 intracovariance matrices

$$S = \begin{pmatrix} \sigma_1^2 & s \\ s & \sigma_2^2 \end{pmatrix} \quad . \tag{7}$$

It can be checked that the results in Sections 2.1 of Rodríguez-Díaz and Sánchez-León [1] remain true for this general matrix S. That is, if the models of the two variables have no common parameters, the expression of D- and A-optimality criterion functions for the bivariate model are functions of the corresponding expressions for the univariate models, and thus if a design is optimal for the individual models it will be globally optimal.

A similar result is obtained for nested models. In this case, for the sake of clarity more details are shown, since this situation fits with the example of Section 3. Similar notation than that in [1] will be used. The biresponse model is $y_1(x) = f_1(x)^T\theta_1 + u_1$, $y_2(x) = f(x)^T\theta + u_2$, depending on m parameters $\theta = (\theta_1^T, \theta_2^T)^T$, where $\theta_i = (\theta_{i1}, \dots, \theta_{im_i})^T$, $i = 1, 2$ and $m_2 = m - m_1$. That is, $y_1(x)$ depends on m_1 parameters, and $y_2(x)$ depends on these parameters plus some others.

Theorem 2.9. *Let us assume a bivariate model $y = (y_1, y_2)^T$ with linear individual responses depending on nested sets θ_1, $\theta_1 \amalg \theta_2$ of m_1 and $m_1 + m_2 = m$ parameters respectively as described above, and a intracovariance given by (7). Then the determinant of the information matrix of the bivariate model is proportional to the product of determinants of the information matrices of the univariate models $y_1(x)$ and $\tilde{y}_2(x) = f_2(x)^T\theta_2 + u$ and therefore the D-optimal designs will not depend on the intracovariance S. In particular, if a design $\xi = \{x_1, \dots, x_n\}$ (with $n \geq \max\{m_1, m_2\}$) is D-optimal for the individual responses $y_1(x)$, $\tilde{y}_2(x)$, it will be D-optimal as well for the bivariate model.*

The proof is in Appendix.

3. Application to the levan production example

Levan is a homopolysaccharide of fructose with industrial interest due to several potential applications [2]. González-Garcinuño et al. [3] study alternatives to the traditional levan producers paying attention to two different bacteria, *Bacillus atrophaeus* (*Ba*) and *Acinetobacter nectaris* (*An*), that are respectively Gram positive and Gram negative. An experiment was conducted in one culture of each bacteria observing two variables of interest every 8 h, namely biomass growth (b) and levan production (l), which make a biresponse model.

In the following, some of the results described in the previous section will be applied to this experiment. A first approach was made in [1], that study the obtaining of optimal designs for the biresponse model when observing one culture of *Ba*. Now both bacteria are considered, with different assumptions on the characteristics of each one. The first step will be to detail the models of both variables, following essentially the descriptions in [3] and some derivations obtained in [1]:

Microbial growth b can be estimated using the model

$$\frac{\partial b}{\partial t} = (\mu - D)b \quad ,$$

where b denotes the biomass concentration at a time t, μ is the growth rate and D is the dilution rate, which will be considered null in this case since the experiments took place at batch scale. Then, assuming null biomass at $t = 0$ it can be expressed as

$$b = \exp\left(\theta_1 t\right) \quad ,$$

where $\theta_1 = \mu$ will be the first parameter of the models.

Regarding the second variable of interest, levan production rate R_l can be computed as

$$R_l = \frac{1}{b}\frac{\partial l}{\partial t} \quad .$$

But levan production is as well related to microbial growth, and for this reason [3] describe this relation as $R_l = \beta_{bl}\mu + m_l$ using the Luedeking–Piret model, where β_{bl} is a constant coefficient that relates product with biomass and m_l is the maintenance coefficient. The parameter m_l can be considered constant when temperature and pH remain constant, but when pH varies linearly with t then a similar variation is produced in m_l [13]. Thus the model

$$\frac{\partial l}{\partial t} = \beta_{bl}\ b\ \mu + b\ m_l\ t \tag{8}$$

will be finally considered. Therefore levan production can be expressed as

$$l = e^{\theta_1 t}\left[\left(\theta_{21} - \frac{\theta_{22}}{\theta_1^2}\right) + \frac{\theta_{22}}{\theta_1}t\right] \quad , \tag{9}$$

where $\theta_{21} = \beta_{bl}$ and $\theta_{22} = m_l$. Linearizing the variables and removing the parts that do not depend on the parameters θ_1 and $\theta_2 = (\theta_{21}, \theta_{22})$ (and therefore have no influence in the optimal design) the equations

$$\begin{aligned} b &\approx te^{\theta_{1_0}t}\theta_1 \\ l &\approx e^{\theta_{1_0}t}\left[(-2\theta_{22_0}/\theta_{1_0}^3 + (\theta_{21_0} - 2\theta_{22_0}/\theta_{1_0}^2)t \right. \\ &\quad \left. + \theta_{22_0}t^2/\theta_{1_0})\theta_1 + \theta_{21} + ((t\theta_{1_0} - 1)/\theta_{1_0}^2)\theta_{22}\right] \end{aligned} \tag{10}$$

are obtained. Since the original models are non-linear, the linearized versions would in general depend on the parameters, thus initial values θ_0 are needed and the designs obtained will be locally-optimal. However [4], the linearized model will not depend on the parameters that *appear lineally* in the model, which in this case are θ_{21} and θ_{22}. Therefore just nominal values for θ_1 will be needed.

The above equations are valid for both bacteria. The differences come from the parameters, namely

- For *Ba* the mean is described by the Monod model,

$$\mu = \mu_{\max}\frac{c}{K_c + c} \quad ,$$

Table 1
2-point D-optimal designs and their efficiency for Ba and An bacteria, for models y_1, y_2 and global biresponse model in $\mathcal{X} = [0, 40]$. The upper extreme, 40, is always taken and thus only the first point is shown.

	Ba		*An*	
Model	D-opt	Eff(%)	D-opt	Eff(%)
y_1	37.740	69.4	37.747	69.6
y_2	30.370	99.6	30.495	99.7
$\{y_1, y_2\}$	31.419		31.510	

where $\mu_{\max}$ is the maximum growth rate, c is the sucrose concentration and K_c is the half velocity constant, that is, the value of the sucrose concentration for which the growth rate is $\mu_{\max}/2$. The values that will be used for these parameters will be $\mu_{\max} = 0.107\ \text{h}^{-1}$, $K_c = 5.48$ g/L and $c = 180$ g/L, that is the best sucrose concentration for Ba growth [3]. Then the nominal value $\theta_{1_0}(Ba)$ for the parameter θ_1 will be 0.104 .

- The biomass growth for An is assumed to follow a substrate inhibition kinetic, thus

$$\mu = \mu_{\max} \frac{c}{c^2/K_i + c + K_c} \quad ,$$

with values $\mu_{\max} = 3.207\ \text{h}^{-1}$, $K_c = 2068$ g/L, $K_i = 9.73$ g/L (inhibition constant) and $c = 120$ g/L. Then $\theta_{1_0}(An)$ will be 0.105 .

And it can be assumed that both β_{bl} and m_l will be specific for each bacterium, thus b and l will have the same structure for both bacteria, but with different parameters. On the other hand, when looking for the optimal designs for the bivariate parameter-nested model (10), using Theorem 2.9 the problem is equivalent to that for the responses

$$\begin{aligned} y_1 &= te^{\theta_{1_0} t}\theta_1 \\ y_2 &= e^{\theta_{1_0} t}\left[\theta_{21} + ((t\theta_{1_0} - 1)/\theta_{1_0}^2)\theta_{22}\right] \quad , \end{aligned} \tag{11}$$

simpler and without common parameters.

Also from Theorem 2.9 the D-optimal designs will not depend on the intra-covariance. The inter-covariance will be assumed to be the widely used exponential covariance $Cov[y_i(t), y_i(t+d)] = e^{-v\,d}$ [14] for both responses. In order to estimate the three parameters, at least a 2-sample design (4 observations) will be needed. Table 1 shows the 2-point D-optimal designs for the individual responses y_1, y_2 and for the biresponse model in $\mathcal{X} = [0, 40]$, as well as the efficiency of the optimal designs for the individual responses with respect to the respective global optimal design. It can be seen that the global optimal designs are ruled by the optimal designs for the second function y_2, the one with two parameters, which is very close to the global optimal. Fig. 1 shows the determinants of the information matrices of the three models y_1, y_2 and biresponse (y_1, y_2) for bacterium Ba as a function of the time of the first sample when the time of second one is fixed to be 40 h, the end of the observation period; the corresponding graphs for An are similar and thus they are not shown.

Table 2 shows optimal designs for both bacteria with different number of points $n = 3, \ldots, 5$. It can be checked that the designs are quite similar, nearly identical, giving almost a perfect efficiency when compared with the one of the opposite bacterium, and the same idea can be said for the 2-point designs of Table 1. Thus using the analytical results of Section 2 these designs can be considered globally optimal. The optimal design for any number of points should take one sample as late as possible, and the rest of the samples close to the end of the maximum culture period considered. Taking into account the number of points employed in the design, the most efficient is the 2-point design, decreasing the efficiency as the number of points increases (Fig. 2).

Finally, the uniform design employed in [3] loses a half of the efficiency when compared with the 2-point optimal design (efficiencies

Table 2
D-optimal n-point designs for Ba and An bacteria, for the biresponse model in $\mathcal{X} = [0, 40]$. The upper extreme, 40, is always taken and thus it is not shown.

n	3	4	5
Ba	{30.257, 35.733}	{28.580, 32.769, 37.480}	{27.106, 30.904, 34.584, 37.899}
An	{30.340, 35.734}	{28.670, 32.833, 37.484}	{27.192, 30.967, 34.613, 37.905}

52.72% and 52.47% for the respective bacteria Ba and An, Fig. 2) and about a quarter when compared with the 5-point optimal design (efficiencies 73.99% and 73.74% for Ba and An respectively). From all of these results it can be concluded that the samples taken at the beginning of the design space seem to be poorly informative for both bacteria, which practitioners explain because bacteria need an adaptation period previous to be able to produce levan.

4. Discussion

The ideas shown in this work can be extended to different models (not necessarily multiresponse models) that produce correlation structures similar to the ones studied above. For instance, let us assume the mixed model

$$y_i = \mathbf{f}^T(x_i)\boldsymbol{\beta} + \gamma + \varepsilon_i \quad ,$$

where x_i are the experimental conditions, $\boldsymbol{\beta}$ is the fix-effect parameter vector, γ is a random effect and ε_i is the residual. A common assumption ([15], Chapter 7) is that γ and ε_i are independent and normally distributed, with zero mean and variances σ_γ^2 and σ^2 respectively. Then

$$Var(y_i) = \sigma_\gamma^2 + \sigma^2 \quad \text{and} \quad Cov(y_i, y_j) = \sigma_\gamma^2 \quad .$$

Thus the covariance matrix becomes

$$\begin{pmatrix} \sigma_\gamma^2 + \sigma^2 & \sigma_\gamma^2 & \cdots & \sigma_\gamma^2 \\ \sigma_\gamma^2 & \sigma_\gamma^2 + \sigma^2 & \cdots & \sigma_\gamma^2 \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_\gamma^2 & \sigma_\gamma^2 & \cdots & \sigma_\gamma^2 + \sigma^2 \end{pmatrix} \quad ,$$

that is homoskedastic and with constant covariance, following the assumptions of the intracovariance $\mathbf{S}$ in [1]. Thus this set of observations (that we could think to be taken at a specific temporal point) could be sampled again at different times, assuming an intercorrelation $\mathbf{R}$ and producing a global covariance structure $\mathbf{R} \otimes \mathbf{S}$ or $\mathbf{S} \otimes \mathbf{R}$, depending on the order considered in the observations vector $\mathbf{Y}$. Therefore many of the results obtained here would be valid as well for this problem.

The number of observation increases with the complexity of the experiment, with every new layer that is added to the structure, and for this reason balanced designs may become more difficult to perform due to budged restrictions (time, money...). Regarding the bacteria example, and taking into account the cost of observing each variable in the distinct bacterium, culture, etc., it could be necessary to study different number of cultures for each bacterium, or not to observe some responses Y_i at every point. The treatment of unbalanced designs due to cost restrictions increases the complexity of obtaining optimal designs. Example 2 of [11] gives an insight of a procedure for analyzing unbalanced designs, and a method for introducing cost constraints in the model is discussed in the final Discussion of that work.

The knowledge of the covariance structure is a controversial issue as well. In case that the covariance parameters of $\mathbf{S_v}$ and/or $\mathbf{R_i}$ where not possible to be fixed (by experience of the practitioner, historical data, ...) and should be estimated as well, the information matrix would be greater [16]; [17] show an application for one variance parameter. Although that in that case, depending on the nature and number of the final set of parameters to be estimated, the results shown here may be not optimal, they are still valid and useful for obtaining good designs.

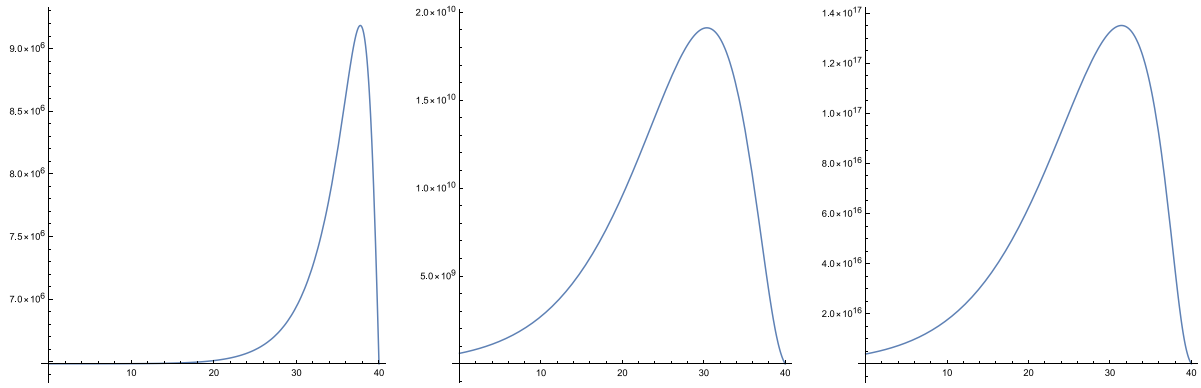

Fig. 1. Determinant of information matrix of 2-point designs as a function of the first observation (the second one fixed to be 40 h), for models y_1 (left), y_2 (center) and the biresponse model (right) for bacterium *Ba*.

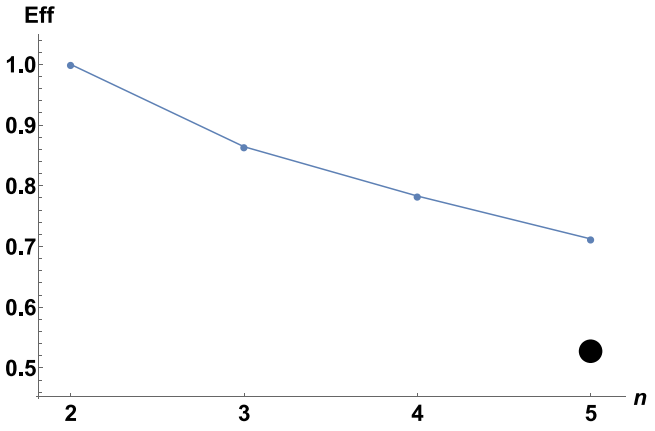

Fig. 2. Efficiency of *D*-optimal designs in n points for bacterium *Ba*. The big dot for $n = 5$ shows the efficiency of the 5-point uniform design employed in [3].

5. Conclusion

Analytical computation of Optimal Experimental Designs has been performed for complex data correlation structures. The different covariance structures studied here appear for instance when dealing with multiresponse models for different subjects, that may o may not share similar characteristics.

The analytical results shown above have been used to find optimal temporal points where samples of two variables of interest ('biomass growth' and 'levan production') should be taken, for different cases: 1 bacterium and 1 culture, 1 bacterium and several cultures (assumed to have similar characteristics), B bacteria (possibly with different characteristics) and 1 culture of each one, and B bacteria observing N cultures of each one; but the applications are countless. For instance, when the aim is estimating the quantity of radioactivity incorporated by workers in a facility dealing with radioactive material, in which two kind of tests (lungs measures and urine tests) are routinely made (biresponse model), and the need to move the worker to a distant special facility (with the corresponding costs) when making the tests. The problem was studied in [11] for one worker, but now it could be done for different workers with or without similar characteristics.

Rodríguez-Díaz et al. [12] computed optimal designs for checking the evolution of children's capability in solving mathematical problems while in primary school. Apart from this variable, a quite related one, 'linguistic comprehension' (capability of understanding the statement of the problem), was measured at the same time, again producing a biresponse model. Test were done for primary school children within the same classroom, thus assuming to share similar characteristics (**S**, **R**). But from the results obtained in this work now the study could be extended to different classrooms and/or school levels, organizing the procedure in order to make the tests in the school at the same time and therefore reducing the length and complexity of the global task.

Another interesting application is the study of optimal designs for compositional models. A compositional response in the simplex S^D can be identified with a multiresponse variable in R^{D-1}, and thus the theory above could be applied. Rodríguez-Díaz et al. [18] is the first approach to the subject of finding optimal designs for compositional models following this idea, and now more complex correlation structures could be studied in order to look for optimal designs.

In this work examples about time-depending experiments have been considered for the sake of clarity. However, the analytical results are general and they can be immediately extended to any design space where a convenient distance between points may produce a correlation structure between observations. Other different correlation structures, non based on distances between design points, could be explored, or even added to the studied ones, for instance spatial correlation (see [19], or [20] for an application to environmental studies). Apart from *D*- (sometimes also *A*-) optimality, different optimality criteria could be studied as well. All these ideas set the grounds of a future work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

Research was supported by the Spanish Ministry of Science, Innovation and Universities, and Junta de Castilla y León, Spain (Projects 'PID2021-125211OB-I00' and 'SA105P20' respectively).

Appendix

This section contains the proofs of the main results.

Theorem 2.3

Proof. There are B k individual models, all of them with the same design matrix $\mathbf{X_0}$, thus the global design matrix is

$$\mathbf{X} = \mathbf{I_{B\,k}} \otimes \mathbf{X_0} = \mathbf{I_B} \otimes \mathbf{I_k} \otimes \mathbf{X_0} = Diag\{\mathbf{I_k} \otimes \mathbf{X_0}, \ldots, \mathbf{I_k} \otimes \mathbf{X_0}\}$$
$$= Diag_B\{\mathbf{I_k} \otimes \mathbf{X_0}\} \quad .$$

Then the information matrix will be

$$\begin{aligned}\mathbf{M} &= \mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\\ &= Diag_B\{\mathbf{I_k}\otimes\mathbf{X_0^T}\}Diag\{\mathbf{S_1^{-1}}\otimes\mathbf{R_1^{-1}},\ldots,\mathbf{S_B^{-1}}\otimes\mathbf{R_B^{-1}}\}Diag_B\{\mathbf{I_k}\otimes\mathbf{X_0}\}\\ &= Diag\{\mathbf{S_1^{-1}}\otimes\tilde{\mathbf{M}}_1,\ldots,\mathbf{S_B^{-1}}\otimes\tilde{\mathbf{M}}_B\}\end{aligned}$$

with $\tilde{\mathbf{M}}_\upsilon = \mathbf{X_0^T}\mathbf{R_\upsilon^{-1}}\mathbf{X_0}$, the information matrix of each variable for the υth subject. Thus, taking into account the properties of Kronecker product (see for instance [21])

$$det(\mathbf{M}) = \prod_{\upsilon=1}^{B} det(\mathbf{S_\upsilon})^{-m}\ \prod_{\upsilon=1}^{B} det(\tilde{\mathbf{M}}_\upsilon)^k$$

and

$$\begin{aligned}tr\left(\mathbf{M}^{-1}\right) &= tr\left(Diag\{\mathbf{S_1}\otimes\tilde{\mathbf{M}}_1^{-1},\ldots,\mathbf{S_B}\otimes\tilde{\mathbf{M}}_B^{-1}\}\right)\\ &= \sum_{\upsilon} tr\left(\mathbf{S_\upsilon}\right) tr\left(\tilde{\mathbf{M}}_\upsilon^{-1}\right)\end{aligned}$$

where the matrices $\tilde{\mathbf{M}}_\upsilon$ do not depend on the intracovariance matrices $\mathbf{S_\upsilon}$. □

Theorem 2.4

Proof. Let us first compute

$$\begin{aligned}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y} &= Diag_B\{\mathbf{I_k}\otimes\mathbf{X_0^T}\}Diag\{\mathbf{S_\upsilon^{-1}}\otimes\mathbf{R_\upsilon^{-1}}\}_{\upsilon=1,\ldots,B}\\ &\quad\times\left(\mathbf{Y^{(1)^T}},\ldots,\mathbf{Y^{(B)^T}}\right)^T\\ &= Diag\{\mathbf{S_\upsilon^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\}_{\upsilon=1,\ldots,B}\left(\mathbf{Y^{(1)^T}},\ldots,\mathbf{Y^{(B)^T}}\right)^T\\ &= \{\left(\mathbf{S_\upsilon^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\}_{\upsilon=1,\ldots,B}\end{aligned}$$

Then

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \left(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y}\\ &= \mathbf{M}^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y}\\ &= Diag\{\mathbf{S_1}\otimes\tilde{\mathbf{M}}_1^{-1},\ldots,\mathbf{S_B}\otimes\tilde{\mathbf{M}}_B^{-1}\}\{\left(\mathbf{S_\upsilon^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\}_{\upsilon=1,\ldots,B}\\ &= \left\{\left(\mathbf{I_k}\otimes\left(\mathbf{X_0^TR_\upsilon^{-1}X_0}\right)^{-1}\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\right\}_{\upsilon=1,\ldots,B}\\ &= \left\{\left(\mathbf{X_0^TR_\upsilon^{-1}X_0}\right)^{-1}\mathbf{X_0^TR_\upsilon^{-1}Y_i^{(\upsilon)}}\right\}_{i=1,\ldots,k;\quad\upsilon=1,\ldots,B}\end{aligned}$$

and

$$Var\left(\hat{\boldsymbol{\beta}}\right) = \left(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\right)^{-1} = Diag\left\{\mathbf{S_1}\otimes\tilde{\mathbf{M}}_1^{-1},\ldots,\mathbf{S_B}\otimes\tilde{\mathbf{M}}_B^{-1}\right\}$$

Thus, taking into account that

$$\boldsymbol{\beta_i^{(\upsilon)}} = (\beta_1^{(\upsilon)},\ldots,\beta_m^{(\upsilon)})^T = \left(\mathbf{X_0^TR_\upsilon^{-1}X_0}\right)^{-1}\mathbf{X_0^TR_\upsilon^{-1}Y_i^{(\upsilon)}}$$

is the vector of parameter estimators of the ith response model for the υth subject, the theorem is proved. □

Theorem 2.5

Proof. When every individual has the same characteristics, that is, $\mathbf{S_\upsilon} = \mathbf{S}$ and $\mathbf{R_i} = \mathbf{R}$, it would be more convenient to use the variable-type wise ordering

$$\tilde{\mathbf{Y}} = \left(\mathbf{y_1^{(1)^T}},\ldots,\mathbf{y_1^{(N)^T}},\ldots,\mathbf{y_k^{(1)^T}},\ldots,\mathbf{y_k^{(N)^T}}\right)^T = \mathbf{PY}\quad,$$

where $\mathbf{P} = \mathbf{Q}\otimes\mathbf{I_n}$ is a suitable $B\ k\ n$ permutation matrix ($\mathbf{Q}$ is the $B\ k$ permutation matrix between the vectors $\mathbf{y_i^{(\upsilon)}}$, see [12]). But for subjects with different characteristics the resulting covariance matrix

$$\boldsymbol{\Sigma}_{\tilde{\mathbf{Y}}} = \boldsymbol{\Sigma}_{\mathbf{PY}} = \mathbf{P}\boldsymbol{\Sigma}_{\mathbf{Y}}\mathbf{P}^T = \left(\mathbf{Q}\otimes\mathbf{I_n}\right)Diag\{\mathbf{S_\upsilon}\otimes\mathbf{R_\upsilon}\}_{\upsilon=1,\ldots,B}\left(\mathbf{Q}^T\otimes\mathbf{I_n}\right)$$

is not easy to deal with, and thus the usual order

$$\mathbf{Y} = \left(\mathbf{Y_1^{(1)^T}},\ldots,\mathbf{Y_k^{(1)^T}},\ldots,\mathbf{Y_1^{(B)^T}},\ldots,\mathbf{Y_k^{(B)^T}}\right)^T$$

will be used. In this scenario the response vectors with the same sub-index share the same model (same parameters), therefore in this case the design matrix $\tilde{\mathbf{X}}$ will be

$$\tilde{\mathbf{X}} = \begin{pmatrix}\mathbf{X_0} & 0 & \ldots & 0\\ 0 & \mathbf{X_0} & \ldots & 0\\ \vdots & \vdots & \ldots & \vdots\\ 0 & 0 & \ldots & \mathbf{X_0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{X_0} & 0 & \ldots & 0\\ 0 & \mathbf{X_0} & \ldots & 0\\ \vdots & \vdots & \ldots & \vdots\\ 0 & 0 & \ldots & \mathbf{X_0}\end{pmatrix} = \mathbf{1_B}\otimes\mathbf{I_k}\otimes\mathbf{X_0}\quad,$$

where $\mathbf{1_B}$ denotes a column vector of '1's of length B. The model now becomes $\mathbf{Y} = \tilde{\mathbf{X}}\tilde{\boldsymbol{\beta}} + \mathbf{U}$, with

$$\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{\beta_1^T},\ldots,\boldsymbol{\beta_k^T}\right)^T = \left(\beta_{11},\ldots,\beta_{1m},\ldots,\beta_{k1},\ldots,\beta_{km}\right)^T$$

The information matrix will be

$$\begin{aligned}\tilde{\mathbf{M}} &= \tilde{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\tilde{\mathbf{X}}\\ &= \left(\mathbf{1_B^T}\otimes\mathbf{I_k}\otimes\mathbf{X_0^T}\right)Diag\{\mathbf{S_\upsilon^{-1}}\otimes\mathbf{R_\upsilon^{-1}}\}_{\upsilon=1,\ldots,B}\left(\mathbf{1_B}\otimes\mathbf{I_k}\otimes\mathbf{X_0}\right)\\ &= \left(\mathbf{I_k}\otimes\mathbf{X_0^T},\ldots,\mathbf{I_k}\otimes\mathbf{X_0^T}\right)Diag\{\mathbf{S_\upsilon^{-1}}\otimes\mathbf{R_\upsilon^{-1}}\}_{\upsilon=1,\ldots,B}\\ &\quad\times\left(\mathbf{I_k}\otimes\mathbf{X_0},\ldots,\mathbf{I_k}\otimes\mathbf{X_0}\right)^T\\ &= \sum_{\upsilon=1}^{B}\left(\mathbf{I_k}\otimes\mathbf{X_0^T}\right)\left(\mathbf{S_\upsilon^{-1}}\otimes\mathbf{R_\upsilon^{-1}}\right)\left(\mathbf{I_k}\otimes\mathbf{X_0}\right)\\ &= \sum_{\upsilon=1}^{B}\mathbf{S_\upsilon^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}X_0}\\ &= \sum_{\upsilon=1}^{B}\mathbf{S_\upsilon^{-1}}\otimes\tilde{\mathbf{M}}_\upsilon\end{aligned}$$

Thus, taking into account that the matrices $\tilde{\mathbf{M}}_\upsilon$ do not depend on the intracovariances $\mathbf{S_\upsilon}$, the theorem is proved. □

Theorem 2.6

Proof. Taking into account that

$$\begin{aligned}\tilde{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y} &= \left(\mathbf{S_1^{-1}}\otimes\mathbf{X_0^TR_1^{-1}},\ldots,\mathbf{S_B^{-1}}\otimes\mathbf{X_0^TR_B^{-1}}\right)\left(\mathbf{Y^{(1)^T}},\ldots,\mathbf{Y^{(B)^T}}\right)^T\\ &= \sum_{\upsilon=1}^{B}\left(\mathbf{S_\upsilon^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\end{aligned}$$

the estimator of the model parameters will be

$$\begin{aligned}\hat{\tilde{\boldsymbol{\beta}}} &= \left(\hat{\boldsymbol{\beta}}_1^T,\ldots,\hat{\boldsymbol{\beta}}_k^T\right)^T\\ &= \tilde{\mathbf{M}}^{-1}\tilde{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y}\\ &= \left(\sum_{\upsilon=1}^{B}\mathbf{S_\upsilon^{-1}}\otimes\tilde{\mathbf{M}}_\upsilon\right)^{-1}\left(\sum_{\upsilon=1}^{B}\left(\mathbf{S_\upsilon^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\right)\end{aligned}\tag{12}$$

with

$$Var\left(\hat{\tilde{\boldsymbol{\beta}}}\right) = \left(\sum_{\upsilon=1}^{B}\mathbf{S_\upsilon^{-1}}\otimes\tilde{\mathbf{M}}_\upsilon\right)^{-1}\quad.$$

From (12), if $\mathbf{S_\upsilon} = \mathbf{S}$ for all υ,

$$\begin{aligned}\hat{\tilde{\boldsymbol{\beta}}} &= \left(\mathbf{S^{-1}}\otimes\sum_{\upsilon=1}^{B}\tilde{\mathbf{M}}_\upsilon\right)^{-1}\left(\sum_{\upsilon=1}^{B}\left(\mathbf{S^{-1}}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\right)\\ &= \left[\mathbf{S}\otimes\left(\sum_{\upsilon=1}^{B}\tilde{\mathbf{M}}_\upsilon\right)^{-1}\right]\left(\mathbf{S^{-1}}\otimes\mathbf{I_m}\right)\sum_{\upsilon=1}^{B}\left(\mathbf{I_k}\otimes\mathbf{X_0^TR_\upsilon^{-1}}\right)\mathbf{Y^{(\upsilon)}}\\ &= \left[\mathbf{I_k}\otimes\left(\sum_{\upsilon=1}^{B}\tilde{\mathbf{M}}_\upsilon\right)^{-1}\right]\sum_{\upsilon=1}^{B}Diag_k\left\{\mathbf{X_0^TR_\upsilon^{-1}}\right\}\mathbf{Y^{(\upsilon)}}\end{aligned}$$

$$= Diag_k\left\{\left(\sum_{v=1}^{B}\tilde{\mathbf{M}}_{\mathbf{v}}\right)^{-1}\right\}\left(\sum_{v=1}^{B}\mathbf{X_0^T}\mathbf{R}_{\mathbf{v}}^{-1}\mathbf{Y}_{\mathbf{1}}^{(\mathbf{v})},\ldots,\sum_{v=1}^{B}\mathbf{X_0^T}\mathbf{R}_{\mathbf{v}}^{-1}\mathbf{Y}_{\mathbf{k}}^{(\mathbf{v})}\right)^T$$

$$= \left\{\left[\mathbf{X_0^T}\left(\sum_{v=1}^{B}\mathbf{R}_{\mathbf{v}}^{-1}\right)\mathbf{X_0}\right]^{-1}\mathbf{X_0^T}\left(\sum_{v=1}^{B}\mathbf{R}_{\mathbf{v}}^{-1}\mathbf{Y}_{\mathbf{j}}^{(\mathbf{v})}\right)\right\}_{j=1,\ldots,k}^{T},$$

that does not depend on $\mathbf{S}$. Now, if $\mathbf{R}_{\mathbf{v}} = \mathbf{R}$ for all v then

$$\hat{\boldsymbol{\beta}}_j = \left(B\mathbf{X_0^T}\mathbf{R}^{-1}\mathbf{X_0}\right)^{-1}\mathbf{X_0^T}\mathbf{R}^{-1}\left(\sum_{v=1}^{B}\mathbf{Y}_{\mathbf{j}}^{(\mathbf{v})}\right)$$

$$= \frac{1}{B}\sum_{v=1}^{B}\left(\mathbf{X_0^T}\mathbf{R}^{-1}\mathbf{X_0}\right)^{-1}\mathbf{X_0^T}\mathbf{R}^{-1}\mathbf{Y}_{\mathbf{j}}^{(\mathbf{v})}$$

$$= \frac{1}{B}\sum_{v=1}^{B}\hat{\boldsymbol{\beta}}_j^{(v)}$$

where $\hat{\boldsymbol{\beta}}_j^{(v)}$ is the parameter estimator of the model of the kth variable for the subject v. □

Theorem 2.7

Proof. It is sensible to assume that the model of each response will be the same for the N cultures of the same bacterium. Let $\check{\mathbf{X}}$ be the design matrix of each one of the N blocks, denoting both $\mathbf{X}$ in scenario I or $\tilde{\mathbf{X}}$ in scenario II of Section 2.2. The same rule will be used for the information matrix of each experiment, $\check{\mathbf{M}}$, that could refer to $\mathbf{M}$ or $\tilde{\mathbf{M}}$. The global design matrix $\mathbf{X}_{(\mathbf{N})}$ under the above assumptions can be expressed as $\mathbf{X}_{(\mathbf{N})} = \mathbf{1}_{\mathbf{N}} \otimes \check{\mathbf{X}}$, and the global information matrix can be computed as

$$\mathbf{M}_{(\mathbf{N})} = \mathbf{X}_{(\mathbf{N})}^T\boldsymbol{\Sigma}_{(\boldsymbol{N})}^{-1}\mathbf{X}_{(\mathbf{N})}$$
$$= \left(\mathbf{1}_{\mathbf{N}}^T \otimes \check{\mathbf{X}}^T\right)\left(\mathbf{I}_{\mathbf{N}} \otimes \boldsymbol{\Sigma}^{-1}\right)\left(\mathbf{1}_{\mathbf{N}} \otimes \check{\mathbf{X}}\right)$$
$$= N\,\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\check{\mathbf{X}}$$
$$= N\,\check{\mathbf{M}}$$

Thus, the conclusions shown in Theorems 2.3 and 2.5 keep for both scenarios. □

Theorem 2.8

Proof.

$$\mathbf{X}_{(\mathbf{N})}^T\boldsymbol{\Sigma}_{(\boldsymbol{N})}^{-1}\mathbf{Y}_{(\mathbf{N})} = \left(\mathbf{1}_{\mathbf{N}}^T \otimes \check{\mathbf{X}}^T\right)\left(\mathbf{I}_{\mathbf{N}} \otimes \boldsymbol{\Sigma}^{-1}\right)\left(\mathbf{Y}_{\mathbf{1}}^T,\ldots,\mathbf{Y}_{\mathbf{N}}^T\right)^T$$
$$= \left(\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1},\ldots,\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\right)\left(\mathbf{Y}_{\mathbf{1}}^T,\ldots,\mathbf{Y}_{\mathbf{N}}^T\right)^T$$
$$= \check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\left(\sum_{w=1}^{N}\mathbf{Y}_{\mathbf{w}}\right)$$

and then the estimator of the model parameters will be

$$\hat{\boldsymbol{\beta}}_{(\boldsymbol{N})} = \left(\mathbf{X}_{(\mathbf{N})}^T\boldsymbol{\Sigma}_{(\boldsymbol{N})}^{-1}\mathbf{X}_{(\mathbf{N})}\right)^{-1}\mathbf{X}_{(\mathbf{N})}^T\boldsymbol{\Sigma}_{(\boldsymbol{N})}^{-1}\mathbf{Y}_{(\mathbf{N})}$$
$$= \frac{1}{N}\sum_{w=1}^{N}\left(\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\check{\mathbf{X}}\right)^{-1}\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y}_{\mathbf{w}}$$

And the theorem is proved taking into account that

$$\left(\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\check{\mathbf{X}}\right)^{-1}\check{\mathbf{X}}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y}_{\mathbf{w}}$$

is the model estimators parameter for the wth Section 2.2-type experiment. □

Theorem 2.9

Proof. In order to use m-parameter functions for both models they will be expressed as $y_i(x) = g_i(x)^T\theta + u_i$, with $g_1(x)^T = (f_1(x)^T, \mathbf{0}_{\mathbf{m_2}}^T)$, $g_2(x)^T = f(x)^T = (f_3(x)^T, f_2(x)^T)$ such as $f(x)^T\theta = f_3(x)^T\theta_1 + f_2(x)^T\theta_2$. Then the design matrix for $n \geq \max\{m_1, m_2\}$ bi-samples will be the block-matrix

$$X = \begin{pmatrix} X_1 & \mathbf{0}_{\mathbf{nxm_2}} \\ X_3 & X_2 \end{pmatrix},$$

where $X_i^T = \left(f_i(x_1),\ldots,f_i(x_n)\right)$, $i = 1,2,3$, and the information matrix (2) with the covariance given by (7) is

$$M(\xi) = \frac{1}{\sigma_1^2\sigma_2^2 - s^2}\begin{pmatrix} \sigma_2^2M_1 - sM_{13}^T - sM_{13} + \sigma_1^2M_3 & -sM_{12} + \sigma_1^2M_{23}^T \\ -sM_{12}^T + \sigma_1^2M_{23} & \sigma_1^2M_2 \end{pmatrix},$$

where $M_{ij} = X_i^TR^{-1}X_j$ and $M_i = M_{ii}$. We thus have a block matrix with shape

$$W = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

If D is a non-singular square matrix then by [21] $det(W) = det(Q)\,det(D)$, with $Q = A - BD^{-1}C$. Using this property, after some algebra it can be checked that

$$\det[M(\xi)] \propto \det(M_1)\det(M_2),$$

which finishes the proof. □

References

[1] J.M. Rodríguez-Díaz, G. Sánchez-León, Optimal designs for multiresponse models with double covariance structure, Chemo. Intel. Lab. Syst. 189 (2019) 1–7b.

[2] S.K. Gupta, A.K. Pal, N.P. Sahu, R.S. Dalvi, M.S. Akhtar, A.K. Jha, S.K. Baruah, Dietary microbial levan enhances tolerance of *labeo rohita* (hamilton) juveniles to thermal stress, Aquaculture 306 (2010) 398–402.

[3] A. González-Garcinuño, A. Tabernero, J.M. Sánchez-Álvarez, M.A. Galán, E.M. Martin del Valle, Effect of bacteria type and sucrose concentration on levan yield and its molecular weight, Microbial. Cell Factories 16 (91) (2017) http://dx.doi.org/10.1186/s12934--017--0703--z.

[4] P.D.H. Hill, D-optimal designs for partially nonlinear regression models, Technometrics 22 (1980) 275–276.

[5] J.M. Rodríguez-Díaz, G. Sánchez-León, Design optimality for models defined by a system of ordinary differential equations, Biom. J. 56 (5) (2014) 886–900.

[6] V.V. Fedorov, P. Hackl, Model-Oriented Design of Experiments, Springer-Verlag, New York, 1997.

[7] F. Pukelsheim, Optimal Design of Experiments, SIAM, Philadelphia, PA, 2006.

[8] A.C. Atkinson, A.N. Donev, R.D. Tobias, Optimum Experimental Designs, with SAS, Oxford University Press, 2007.

[9] H. Dette, A. Pepelyshev, A. Zhigljavsky, Design for linear regression models with correlated errors, in: Handbook of Design and Analysis of Experiments, CRC Press, 2015, pp. 237–276.

[10] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[11] J.M. Rodríguez-Díaz, G. Sánchez-León, Efficient parameter estimation in multiresponse models measuring radioactivity retention, Radiat. Environ. Biophys. 58 (2) (2019) 167–182a.

[12] J.M. Rodríguez-Díaz, R.E. Pruneda, M. Rodríguez-Hernández, Design plan for an evolution study of related characteristics of a population, Mathematics 10 (2022) 792.

[13] R. Luedeking, E.L. Piret, A kinetic study of the lactic acid fermentation, J. Biochem. Microbiol. Tech. Eng. 1 (4) (1959) 393–412.

[14] N. Cressie, C.K. Wikle, Statistics for Spatio-Temporal Data, John Wiley & sons, New Jersey, 2011.

[15] P. Goos, B. Jones, Optimal design of experiments, in: A Case Study Approach, Wiley, U.K, 2011.

[16] A. Pázman, Criteria for optimal design for small-sample experiments with correlated observations, Kybernetika 43 (4) (2007) 453–462.

[17] J.M. Rodríguez-Díaz, M.T. Santos-Martín, H. Waldl, M. Stehlik, Filling and d-optimal designs for the correlated generalized exponential models, Chemometr. Intell. Lab. Syst. 114 (2012) 10–18.

[18] J.M. Rodríguez-Díaz, M.J. Rivas-López, M.T. Santos-Martín, I. Mariñas Collado, Optimal designs for a linear-model compositional response, Stoch. Environ. Res. Risk Assess. (34) (2020) 139–148, http://dx.doi.org/10.1007/s00477-019-01755-2.

[19] W.G. Müller, Collecting Spatial Data, third ed., Physica-Verlag, Heidelberg, 2007.

[20] D. Gumprecht, W. Müller, J.M. Rodríguez-Díaz, Optimal design for detecting spatial dependences, Geograp. Analy. 41 (2) (2009) 127–143.

[21] D.A. Harville, Matrix Algebra from a Statistician's Perspective, Springer-Verlag, 1997.