Optimal statistical designs for the accurate estimation of the parameters of a growth rate model for *Listeria monocytogenes*

Jean-Pierre Gauchi
National Institute of Agricultural Research (INRA)
Applied Mathematics and Computational Sciences
Jouy-en-Josas, France
Email: jean-pierre.gauchi@jouy.inra.fr

June 7, 2005

Abstract

We propose optimal statistical designs for the accurate determination of the parameters of a growth rate model, function of the temperature, for Listeria monocytogenes (LM). This model is generally called a $\mu_{\max}$ model in the predictive microbiology field. It is called a secondary model because $\mu_{\max}$ is a parameter of a primary model for describing the growth as a function of time. There are many models for $\mu_{\max}$ (functions of temperature, pH, water activity, ...) but the behaviour of LM when temperature is low needs a particular $\mu_{\max}$ model. In this paper, beyond the determination of useful optimal designs, we want to emphasize on the advantage of an experimental methodology based on such statistical tools. Indeed, we observed that this kind of efficient experimental approach is not sufficiently known and put into practice in the predictive microbiology field, whereas it is well known in other scientific fields as applied chemical research for instance.

Contents

1 Introduction **2**

2 The statistical context **4**

3 The CFT model **6**

4 The optimal designs **9**
4.1 The local $D-$optimal design 9
4.1.1 Definition 9
4.1.2 Computation 9
4.1.3 Comment 11
4.2 The local D_S-optimal design 14
4.2.1 Definition 14
4.2.2 Computation 14
4.2.3 Comment 15
4.3 The $ELD-$optimal design 15
4.3.1 Definition 15
4.3.2 Computation 15
4.3.3 Comment 15

5 Comparison of designs **17**
5.1 Naive designs 17
5.2 Montecarlo simulation and computed statistics 18
5.3 Results of the comparison 19

6 Discussion **21**
6.1 Four major principles 22
6.1.1 Sequential approach 22
6.1.2 $D-$optimality family 22
6.1.3 Designs with $N_S = p$ and $N > N_S$ 23
6.1.4 The nature of the error variance 24
6.2 Specific designs for the nonlinear models 24

7 Conclusion **25**

8 Appendix A: The confidence ellipsoïd and its volume **25**

9 Appendix B: The exact X confidence region **26**

10 References **27**

1 Introduction

This paper is devoted to a proposition of an experimental methodology based on the use of optimal statistical designs in a view to reach a very accurate

estimation of the parameters of a growth rate model where temperature is the explanatory variable. It is called a $\mu_{\max}$ secondary model in the predictive microbiology area, and applied here to Listeria monocytogenes (LM) data. Typically, $\mu_{\max}$ is the slope of the primary growth models. Notably the accurate estimation of the parameter corresponding to the minimal growth temperature of LM is of major interest. The statistical theoretical aspects on the optimal designs we propose are not developped here (however numerous and useful references are indicated in a pedagogical point of view), but only evocated. The practical point of view, with the goal to be useful to the scientists of the predictive microbiology area, is emphasized.

For about thirty years many analytic models have been proposed for the modelling of a bacterian population growth, as a function of time, by means of said primary models, see, e.g., Baranyi and Roberts (1994, 1995), Rosso (1995), Rosso and Flandrois (1995). Also, was proposed the modelling of the parameters of these primary models, that lead to the secondary models when these parameters appear themselves as functions of environmental factors (temperature, pH, water activity, ...), see, e.g., Ratkovsky et al. (1982), Rosso et al. (1993), Augustin and Carlier (2000), Le Marc et al. (2002), Bajard et al. (1996), Charles-Bajard et al. (2003). For instance this is the case for the parameter $\mu_{\max}$ when it is considered as a function of the temperature. A very well known model of $\mu_{\max}$ is the pioneer model called the Ratkovsky square model (Ratkovsky et al. 1982). Then, other models were exhibited over these last years, see, e.g., the Rosso's cardinal model (Rosso et al., 1993), and also the Charles-Bajard-Flandrois-Tomassone model (Charles-Bajard et al., 2003), referred as the CFT model further in this paper. This CFT model is particularly relevant to show the anomaly of *LM* (Bajard et al., 1996). The CFT model contains two parameters and seems to be really efficient to model $\mu_{\max}$ as a function of the temperature, that has been clearly shown by the authors. It is completely defined in Section 3. Our main objective is to show, via the CFT model, the advantage of the following approach when a large accuracy is needed. When an information is previously available due to a first experimental step (generally non optimal) and simultaneously a first model is assumed, then optimal designs can be advantageously undertaken in a second experimental step to reach an efficient estimation of the model parameters. This second step is particularly advantageous if the first step showed that the experimental variance (error variance) is heterogenous over the experimental domain as it was the case in the CFT problem. From an additionnal point of view, there is a need for such an accuracy if we want to manage a risk assessment in the food industry.

In this paper, beyond the determination of real practical optimal designs for a real model, we want to emphasize on the usefulness of an experimental

methodology based on such statistical tools whatever the postulated model. Indeed, we observed last years that this kind of efficient experimental approach is not sufficiently known and put into practice in the predictive microbiology field. However, notice that some good papers were published with a similar approach but based on another optimality criteria and a rather different point of view (Versyk et al., 1999; Bernaerts et al., 2000).
This paper is divided in seven sections. After this first introductive section, Section 2 introduces to the statistical context and notations, Section 3 presents the CFT model, Section 4 describes the optimal designs proposed (definition and computation), Section 5 shows a montecarlo simulation comparison of the optimal designs with naive (non optimal) designs. Section 6 gives some major principles for helping to construct an efficient experimental methodology that can be applied in a general way, and also underlining that another criteria exist for the more difficult cases. Section 7 is a conclusion to close the paper.

2 The statistical context

We consider simultaneously the classical parametric nonlinear regression model

$$y_{ik} = \eta_{\xi}\left(x_i, \theta^*\right) + \varepsilon_{ik} \quad ; \; i = 1, \ldots, N_S \; ; \; k = 1, \ldots, r_i \tag{1}$$

and the design

$$\xi_{N,N_S,\{r\}} = \left\{ \begin{array}{cccccc} x_1 & \ldots & x_i & \ldots & x_{N_S} \\ w\left(x_1\right) & \ldots & w\left(x_i\right) & \ldots & w\left(x_{N_S}\right) \\ r_1 & \ldots & r_i & \ldots & r_{N_S} \end{array} \right\} \tag{2}$$

where

- y_{ik} is the k^{th} observation collected on the i^{th} support point x_i defined hereafter,
- x_i is a $m-$dimensional support point ($m-$vector) whom each component x_{il} is a controlled level of the explanatory variable X_l, $l = 1, \ldots, m$; we have $x_i \in \Xi \subset \mathbb{R}^m$, $i = 1, \ldots, N_S$, Ξ being the experimental design, compact subset of $\mathbb{R}^m$,
- θ^* is the $p-$dimensional vector of p unknown parameters to be estimated, they will be considered as certain or random parameters depending on the optimal design chosen as it will be shown further; we have $\theta^* \in \Theta \subset \mathbb{R}^p$, Θ being the parametric domain, a Borelian subset of $\mathbb{R}^p$,

- $\eta_\xi(.)$ is a continuous scalar function defined on $\Xi \times \Theta$, valued in $\mathbb{R}$, doubly differentiable relatively to the parameters and the explanatory variables,
- N is the total number of observations y_{ik} collected on the N_S support points with the replication pattern $\{r\} = r_1, \ldots r_{N_S}$, $r_i \geq 1$, $\forall i$, $\sum_{i=1}^{N_S} r_i = N$,
- ε_{ik} is the error attached to y_{ik}; in this paper we will assume that it is normally distributed as $\mathcal{N}(0, \sigma_i^2)$, with mean equals to zero and variance σ_i^2; the errors are assumed independent,
- $w(x_i) = 1/\sigma_i^2$ is the weight of each observation y_{ik}, $\forall k$, this weight dependent on x_i, because we assume a heterogenous variance over the experimental domain Ξ.

Moreover let the $(N \times 1)$ vector y of the N observations, we have also:

$$Var(y) = \Sigma_\xi = diag\left\{\overbrace{\sigma_1^2 \ldots \sigma_1^2}^{r_1} \ldots\ldots \overbrace{\sigma_{N_S}^2 \ldots \sigma_{N_S}^2}^{r_{N_S}}\right\} \quad (3)$$

$$W_\xi = \Sigma_\xi^{-1} = diag\left\{\overbrace{w(x_1) \ldots w(x_1)}^{r_1} \ldots\ldots \overbrace{w(x_{N_S}) \ldots w(x_{N_S})}^{r_{N_S}}\right\} \quad (4)$$

Note that generally the σ_i^2 are unknown. For the determination of the optimal designs proposed here it is necessary to have some estimates of these σ_i^2. These estimates, referred as $\hat{\sigma}_i^2$, must have been computed from anterior available data. For instance there are sometimes calculated by means of a parametric model as it is the case for the CFT model (see equation (7)).
Because the errors are assumed independent and gaussian we can define the Fisher information matrix (see for instance Atkinson and Doneev (1992) for useful details on this well known matrix) as the $(p \times p)$ matrix

$$M(\xi, \theta^*) = J(\xi, \theta^*)^T W_\xi J(\xi, \theta^*) \quad (5)$$

where $J(\xi, \theta^*)$ is the $(N \times p)$ jacobian matrix, whom the general term is $\left\{\frac{\partial \eta(x_{ik}, \theta^*)}{\partial \theta_j}\right\}$, $i = 1, \ldots, N_S$, $k = r_1, \ldots, r_{N_S}$, $j = 1, \ldots, p$. We assume that these first derivatives exist and they are continuous over $\Xi \times \Theta$, and they are continously differentiable relatively to the explanatory variables and the parameters. Finally, for estimating θ^* from the data we will use the weigthted least squares estimate, referred as $\hat{\theta}_{WLS}$, but more simply referred as $\hat{\theta}$ in this paper.

3 The CFT model

The CFT model contains two parameters and seems to be really efficient to model $\mu_{\max}$ as a function of the temperature, that has been clearly shown by the authors (Charles-Bajard et al., 2003).
The analytical function of this model is

$$\eta_{\xi}(\theta) = \eta_0(\theta)\left[\eta_1(\theta) + \eta_2(\theta)\,\eta_3(\theta)\,\eta_4(\theta)\right]^2 \tag{6}$$

with $\eta_0(\theta) = k_1\theta_1 + k_2\theta_2$; $\eta_1(\theta) = (T-\theta_1)/\theta_2$; $\eta_2(\theta) = \theta_1/(\theta_2(\theta_1-\theta_2))$; $\eta_3(\theta) = 1+\exp(\theta_1-\theta_2)$; $\eta_4(\theta) = \ln((1+\exp(\theta_2-T))/(1+\exp(\theta_2-\theta_1)))$, where θ_1 and θ_2 are the parameters to be estimated, and T is the explanatory factor that is the temperature expressed in Celsius degrees; k_1 and k_2 are two constants given in Charles-Bajard et al. (2003): $k_1 = 0.004$; $k_2 = 0.0149$. For this model we have $m = 1$, $p = 2$, $\Theta = [-7\,;-3]\times[8\,;13]$, and $\Xi = [-5°C\,;+30°C]$.
It has been proved by the authors that the model (6), notably under its square root form, takes very well into account the break in the curve, typical anomaly of the Listeria behaviour shown by Bajard et al (1996). However, in this paper we want to deal with the heterogeneity of the error variance, and then it is not recommended to take the square root of (6). This transformation is not always optimal for stabilizing the error variance and moreover it destroys the true nature of the error variance. In Charles-Bajard et al. (2003) it is given twenty pairs of data (x_i, y_i) and also the corresponding twenty values of the $\mu_{\max}$ variance, that are reliable enough because each of these twenty values are based on six repeated experiments. These values clearly show that the variance is not homogenous. To take into account this heterogeneity we established the model :

$$var(\mu_{\max}) = \exp(-12.80 + 0.27T) \tag{7}$$

with these data. We obtained also the estimates $\hat{\theta}_1 = -5.33$ and $\hat{\theta}_2 = 11.27$ by means of the weighted nonlinear regression (NLIN procedure) of the SAS/STAT software (version 8.1). The weights are defined in section 2. We computed the 95% marginal asymptotic confidence intervals for θ_1 and θ_2:

$$\theta_1 : [-5.86\ ;\ -4.82]\ \ ;\ \theta_2 : [10.80\ ;\ 11.85] \tag{8}$$

from the 95% confidence ellipsoïd (see Fig.1), that is here only an approximate confidence region, generally called asymptotic confidence region in classical textbooks (see for instance Bates and Watts, (1988)).

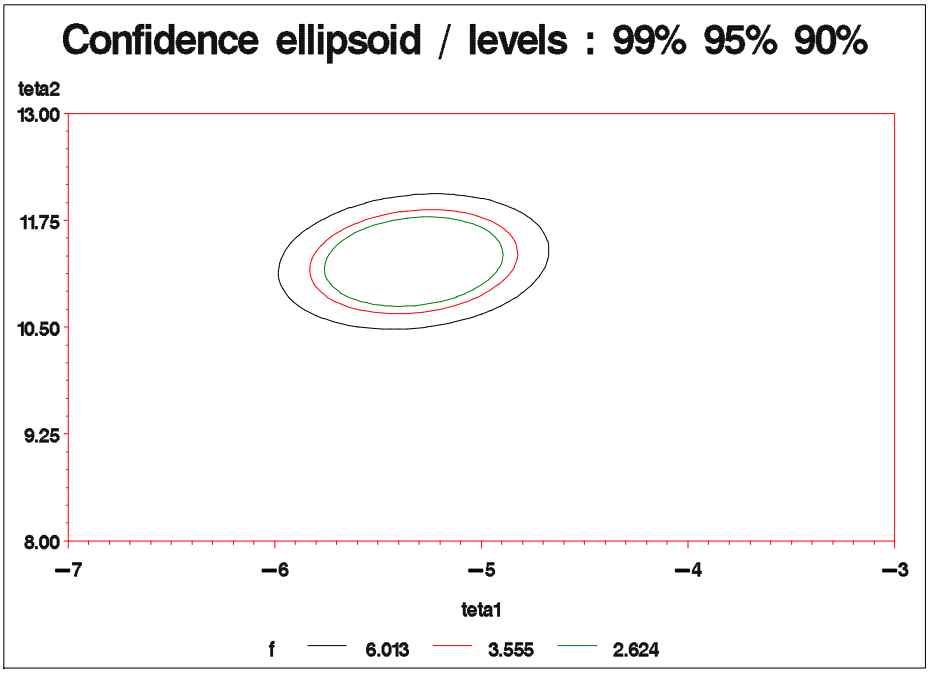

Figure 1: Plot of parametric confidence ellipsoids, at three levels, for the CFT model, in the conditions and the data of the paper of Charles-Bajard et al. (2003).

The volume of this ellipsoïd equals to 0.95. We recall some useful details about this ellipsoid in Appendix A. On the Fig. 2 appears an exact − exact means here that its covering probabilty of θ^* is exactly 95% − confidence region, called in the following the X confidence region. Its volume equals to 8.47 and the X (exact) marginal confidence intervals obtained from it are for θ_1 and θ_2:

$$\theta_1 : [-6.78 \ ; \ -3] \quad ; \quad \theta_2 : [9.75 \ ; \ 13] \tag{9}$$

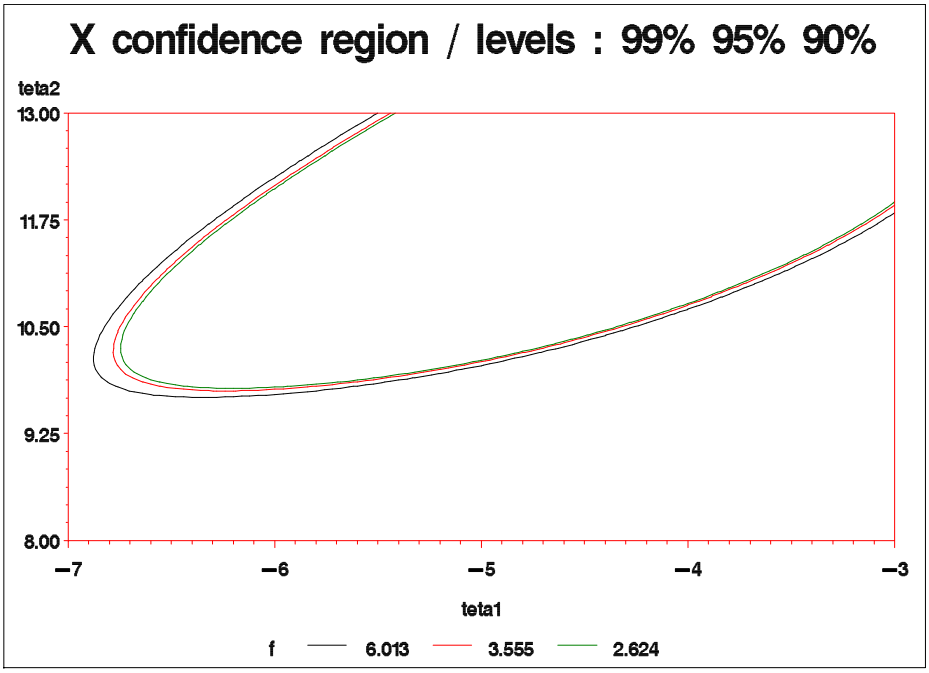

Figure 2: Plot of the X (exact) parametric confidence regions, at three levels, for the CFT model, in the conditions and the data of the paper of Charles-Bajard et al. (2003).

The X confidence region belongs to the Halperin family (Halperin, 1963) and has been completely defined by Hamilton and Watts (1985) and Vila (1985, 1986). We give a brief definition of this region in Appendix B. Comparing Fig. 1 and 2 it is clear here that the ellipsoïd seems to be a very bad approximation (too flaterring and false) of an exact region, and we show in the next section that the main reason is due to the nonoptimal design used in the paper of Charles-Bajard et al. (2003).

4 The optimal designs

The optimal design methodology is only a subset included in the huge field of the statistical and rational designing of experiments. Of course it is not our purpose to give here a review on this matter. We want just to say that numerous criteria exist to build optimal designs and several books are devoted on this matter, see, e.g., Atkinson and Doneev (1992) for a useful introduction on this matter, Pázman (1986) and Pukelsheim (1993) for advanced readers involved in the linear models, and Pázman (1993) and Gallant (1987) for advanced readers involved in the nonlinear models. Some criteria deal with linear models, others are specialized to nonlinear models. In this paper we show how three different optimal designs coming fron this litterature can solve our microbiological question, even if those criteria are not very recent. In this section we recall the definition of the criteria used, and why they are chosen, and we give the results i.e. the optimal designs determined with these criteria.

4.1 The local $D-$optimal design

4.1.1 Definition

If we consider that all the parameters of the postulated model are parameters of interest, then their estimates must be obtained with an equivalent precision, and we must choose a design wich enables to reach this aim. It is the aim of the $D-$optimality criterion. Indeed this criterion leads to minimize the volume (the square root in fact) of the usual parametric confidence ellipsoïd (see Appendix A). So the $D-$optimal design is defined (with the notations of section 2) as:

$$\xi^{D}_{N,N_S=p,\{r\}} = Arg\left\{\max_{\xi\in\Xi} \det\left(M\left(\xi,\theta_0\right)\right)\right\} \tag{10}$$

Three important remarks must be made now: i) Note that this design depends on a prior value θ_0 of θ^*, so it is in fact a *local* $D-$optimal design, and for the designs shown in this section θ_0 will be set to the estimate $\hat{\theta}$ given in section 3, ii) Minimize the ellipsoïd volume is mathematically equivalent to maximize the determinant of the Fisher matrix information, iii) We want here a design based on only p support points, and it is explained why below.

4.1.2 Computation

First of all the model (7) leads to the predictions of the local variances $\hat{\sigma}^2_i$, and then to the weights $\hat{w}\left(x_i\right) = 1/\hat{\sigma}^2_i$. These predictions will be the diagonal

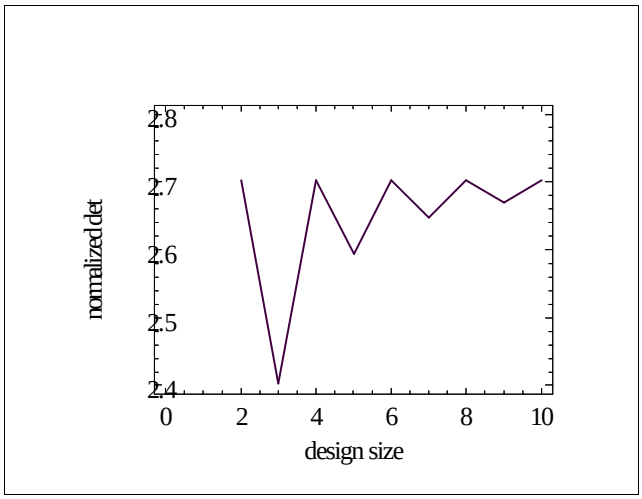

Figure 3: Evolution of the normalized determinant (see text, section 4.1.2) versus the design size.

terms of the W_ξ matrix. Now a minimal discrete (local) $D-$optimal design ($N = N_S = p$) can be easily obtained by the very well-known Fedorov double exchange algorithm (Fedorov, 1972), implemented in the OPTEX procedure of the SAS/QC software. We obtained:

$$\xi^D_{2,2,\{1,1\}} = \left\{ \begin{array}{ll} T_1 = -1.12°C & T_2 = +15.40°C \\ \hat{w}\left(T_1\right) = 490117 & \hat{w}\left(T_2\right) = 5665 \\ r_1 = 1 & r_2 = 1 \end{array} \right\} \tag{11}$$

If we want replications the same algorithm leads to the graph on Fig. 3.

The replicated designs correspond to the peaks at the two-multiples on Fig. 3. That means that the replicated designs are made of balanced replications of the same minimal ($N_S = 2$)$-$design. The best designs are those that are equireplicated on the two N_S support points defined by (11). Now, let $\Delta = \det\left(M\left(\xi, \theta_0\right)\right)$ and $\Delta_n = \Delta / N^p$ the normalized determinant. We found the maximum normalized determinant $\Delta^*_n = \max\left(\Delta_n\right) = 2.70$. The General Equivalence Theorem (Kiefer and Wolfowitz, 1960, Pukelsheim, 1993) enables us to confirm this value is a global maximum on the experimental domain Ξ, and then this minimal $D-$optimal design is said 100% $D-$efficient. Moreover, the Vila condition of the replication in $D-$optimality (Vila, 1991) is checked here and then indicates that the best designs are those that are

equireplicated on the support (11). This concusion can be visualized in Table 1.

Balanced replications	Δ	Δ_n	Unbalanced replications	Δ	Δ_n
1 ; 1	10.81	2.70	1 ; 1	10.81	2.70
2 ; 2	43.24	2.70	1 ; 2 or 2 ; 1	21.62	2.40
3 ; 3	97.30	2.70	1 ; 3 or 3 ; 1	32.43	2.03
4 ; 4	172.97	2.70	1 ; 4 or 4 ; 1	43.24	1.73
5 ; 5	270.27	2.70	1 ; 5 or 5 ; 1	54.05	1.50
10 ; 10	1080	2.70	1 ; 10 or 10 ; 1	108	0.89

Table 1 : $D-$optimality criterion: evolution of the determinant and the normalized determinant versus the number of balanced and unbalanced replications, always on the same support points given in (9).
To close this subsection we can say that a practical and useful (local) $D-$optimal design that can be put into practice at the laboratory is the design

$$\xi^{D}_{6,2,\{3,3\}} = \left\{ \begin{array}{ll} T_1 = -1.12°C & T_2 = +15.40°C \\ \hat{w}\left(T_1\right) = 490117 & \hat{w}\left(T_2\right) = 5665 \\ r_1 = 3 & r_2 = 3 \end{array} \right\} \qquad (12)$$

4.1.3 Comment

In fact, if the postulated model is linear relatively to the parameters, and also if we are sure that is the right model to analyze the data, then it is well know that a $D-$optimal design is an adequate design to estimate the parameters of the model with the best joint precision (small variance) as possible. However, if the postulated model is nonlinear relatively to the parameters, and even if the model is absolutely true, then it is known also that the $D-$optimality criterion is only an approximate criterion (Atkinson and Doneev, 1992). Indeed, in the nonlinear case the $D-$optimality criterion is based on the first-order approximation of the model $\eta\left(.\right)$. However, in certain cases the nonlinearity of the model is small, i.e. its intrinsic and parametric components (see Bates and Watts (1988, chapter 6) for a good introduction on this matter, and also our discussion in 6.1.3) are small. In such cases we can accept this approximation. Several methods enable us to quantify this nonlinearity, but two among them are easy to apply, *before to collect* the observations y_{ik} , i.e. in our experimental design situation. The first method consists in the computation of the curvature measures proposed by Bates

and Watts (1980, 1988, chapter 6). The second method is the comparison (relatively to the shape and the volume) of the confidence ellipsoïd with an exact region, for instance the X confidence region (see Appendix B) that has been intensively studied by Vila (1985, 1986), Gauchi (1999), and Gauchi and Vila (2004). Here, for the design (12) we obtain a quadratic mean of the parametric nonlinearity equals to 0.44, that is a significant value because it overtakes the threshold of 0.38, and otherwise the intrinsic linearity is zero because $N_S = p$. This latter consequence is particularly interesting as we will explain in 6.1.3; one can find a full and elegant geometrical proof of this phenomenon in Pázman (1993). We show in Fig. 4 and 5 the confidence ellipsoïd and the (expected) X confidence region for the design (12), and in Table 2 the 95% marginal confidence intervals are given.

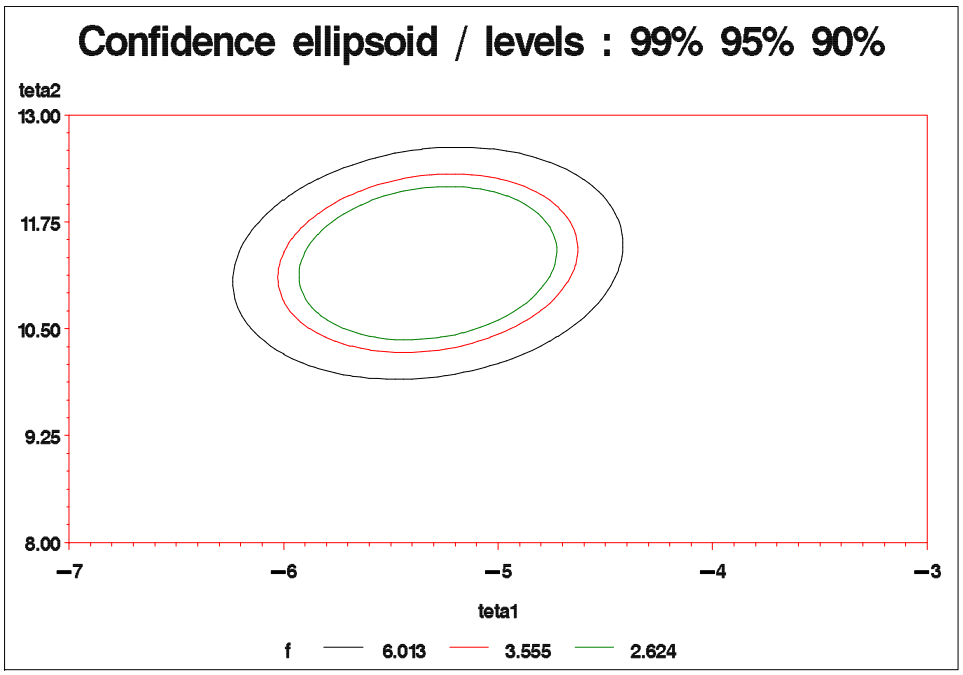

Figure 4: Plot of parametric confidence ellipsoids, at three levels, for the local $D-$optimal described in (12).

These regions and intervals look like, respectively, on the contrary of those shown in Fig.1 and 2. This likeness heavily confirms the usefulness of the optimal design approach.
In addition, in Table 2 we put the 95% marginal asymptotic confidence intervals and X (exact) confidence intervals for a $D-$optimal design with same

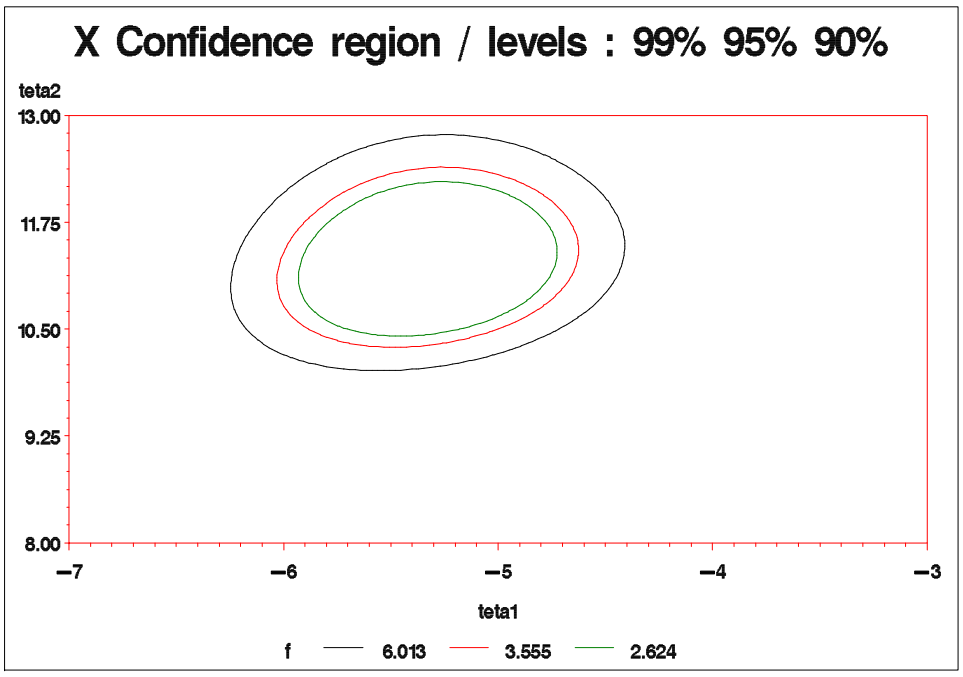

Figure 5: Plot of X (exact) parametric confidence regions, at three levels, for the local $D-$optimal described in (12).

support points but with the replication pattern $\{r_1 = 10\ ;\ r_2 = 10\}$ in a view to compare these intervals with those of section 3. We observe that even with twenty experiments and even if the volumes are almost the same, the asymptotic marginals intervals are always rather wrong.
In conclusion of this paragraph we can say that this local $D-$optimal design defined by (12) is not the best one but it can be a fairly acceptable design, notably because the ellipsoidal confidence region is very close to the exact X confidence region. We have to add now that it exist other very well adapted optimality criteria for the nonlinear case, they are evocated in section 6.2 of this paper.

4.2 The local D_S-optimal design

4.2.1 Definition

Now we suppose that we are more particularly interested in estimating θ_1 corresponding to the growth minimal temperature, i.e. the second parameter θ_2 is considered here as a nuisance parameter. Then we can choose an optimal design that takes into account this constraint. The (local) D_S-optimal, referred as $\xi^{D_S(\theta_1)}_{N,N_S=p,\{r\}}$, enables us to reach this goal, and it is defined here (see, e.g., Silvey (1980) for a general presentation) as:

$$\xi^{D_S(\theta_1)}_{N,N_S=p,\{r\}} = Arg\left\{\max_{\xi\in\Xi} \det\left(M^{[1,1]}\right)\right\} \tag{13}$$

where

$$M^{[1,1]} = M^{[1,1]} - M^{[1,2]}\left(M^{[2,2]}\right)^{-1} M^{[1,2]T} \tag{14}$$

with $M^{[i,j]}$ the $(i,j)-$element of (5).

4.2.2 Computation

If we compute and plot the determinant of (14) at each node of a two-dimensional grid over Ξ we can determine easily the maximum and the corresponding T values. These are $T_1 = -1.15°C$ and $T_2 = 11.97°C$. A D_S-optimal design (100% D_S-efficient) could be then made of balanced repetitions on these T values that are the support points. We refer a practical D_S-optimal design as:

$$\xi^{D_S(\theta_1)}_{6,2,\{3,3\}} = \left\{\begin{array}{ll} T_1 = -1.15°C & T_2 = +11.97°C \\ \hat{w}\left(T_1\right) = 495746 & \hat{w}\left(T_2\right) = 13396 \\ r_1 = 3 & r_2 = 3 \end{array}\right\} \tag{15}$$

4.2.3 Comment

First of all note that the second support point is closer to the first support point with this design (compare with (12)). The ellipsoid and exact confidence regions look like each other and also look like to those of the $D-$optimal. We give only the 95% marginal intervals in Table 2.

4.3 The $ELD-$optimal design

4.3.1 Definition

Now both parameters are of interest but we want to be more robust relatively to a misspecification of the prior value θ_0. Then instead of a single value θ_0 we propose a prior distribution for the unknown parameter (that is now considered as a random parameter), for instance a gaussian distribution $p(\theta_0) = \mathcal{N}(\theta_0, \Sigma_{\theta_0})$. This distribution is then introduced in the optimality criterion. The $ELD-$optimal design (Chaloner and Verdinelli, 1995) is based on this approach. It is defined as:

$$\begin{aligned} \xi^{ELD}_{N,N_S=p,\{r\}} &= Arg\left\{\max_{\xi\in\Xi} E\left[\log\det\left(M(\xi,\theta_0)\right)\right]\right\} \\ &= Arg\left\{\max_{\xi\in\Xi}\left\{\int_\Theta \log\det(M(\xi,\theta_0))p(\theta_0)\,d\theta_0\right\}\right\} \end{aligned} \tag{16}$$

4.3.2 Computation

Firtsly, we determine

$$\Sigma_{\theta_0} = \begin{pmatrix} 0.03567 & 0.006843 \\ 0.006843 & 0.0522 \end{pmatrix} \tag{17}$$

with the Charles-Bajard paper data. Secondly, with the same graphical way as for the preceeding design, we obtain now the maximum of (16) at the T values: $T_1 = -0.92°C$ and $T_2 = 16.58°C$. A practical $ELD-$optimal design could be then:

$$\xi^{ELD}_{6,2,\{3,3\}} = \begin{Bmatrix} T_1 = -0.92°C & T_2 = +16.58°C \\ \hat{w}(T_1) = 464353 & \hat{w}(T_2) = 4119 \\ r_1 = 3 & r_2 = 3 \end{Bmatrix} \tag{18}$$

4.3.3 Comment

This design is particularly well adapted here in a sequential approach because we have available data to compute (17), see section 6.1.1 for a discussion.

Otherwise we observe that the second support point is farther than to the first support point (compare with (12)). The ellipsoid and exact confidence regions look like each other and also look like to those of the $D-$optimal. We give only the 95% marginal intervals in Table 2.

$D-$optimal design with $\{r_1 = 3; r_2 = 3\}$
Anticipated (asymptotic) confidence intervals ; vol $(E) = 2.26$
$\theta_1 : [-6.02 \ ; \ -4.63] \implies L = 1.40$
$\theta_2 : [+10.22 \ ; \ +12.31] \implies L = 2.09$
Expected X (exact) confidence intervals ; vol$(X) = 2.29$
$\theta_1 : [-6.03 \ ; \ -4.63] \implies L = 1.41$
$\theta_2 : [+10.29 \ ; \ +12.40] \implies L = 2.11$
$D-$optimal design with $\{r_1 = 10; r_2 = 10\}$
Anticipated (asymptotic) confidence intervals (vol $(E) = 0.68$)
$\theta_1 : [-5.71 \ ; \ -4.95] \implies L = 0.77$
$\theta_2 : [+10.70 \ ; \ +11.84] \implies L = 1.14$
Expected X (exact) confidence intervals (vol $X = 0.68$)
$\theta_1 : [-5.71 \ ; \ -4.94] \implies L = 0.77$
$\theta_2 : [10.72 \ ; \ 11.86] \implies L = 1.15$
D_S-optimal design
Anticipated (asymptotic) confidence intervals ; vol$(E) = 2.89$
$\theta_1 : [-6.02 \ ; \ -4.63 \ ] \implies L = 1.40$
$\theta_2 : [+9.95 \ ; \ +12.59] \implies L = 2.64$
Expected X (exact) confidence intervals ; vol$(X) = 3.08$
$\theta_1 : [-6.03 \ ; \ -4.62] \implies L = 1.40$
$\theta_2 : [+10.13 \ ; \ +13.00] \implies L = 2.85$
$ELD-$optimal design
Anticipated (asymptotic) confidence intervals ; vol$(E) = 2.29$
$\theta_1 : [-6.03 \ ; \ -4.63 \ ] \implies L = 1.40$
$\theta_2 : [+10.22 \ ; \ +12.32] \implies L = 2.10$
Expected X (exact) confidence intervals ; vol$(X) = 2.31$
$\theta_1 : [-6.04 \ ; \ -4.63] \implies L = 1.41$
$\theta_2 : [+10.28 \ ; \ +12.40] \implies L = 2.12$

Table 2 : Anticipated (asymptotic) confidence intervals and expected X (exact) confidence intervals for the optimal designs given in section 4. L is the length of the interval, and vol(E) and vol(X) mean volume of the ellipsoid and the X region, respectively. All the intervals are given at the 95% level.

5 Comparison of designs

The objective of this section is to compare the three optimal designs we gave in the preceeding section, between each other but also with naive designs, designs that we define hereafter.

5.1 Naive designs

We define now a naive (nonoptimal) design whom the goal is to simulate the usual way to design experiments in predictive microbiology. We propose such a design after attending numerous meetings of microbiologists, these latter years. Of course this only a personnal point of view. Firstly, if there are only two parameters in the model the number N_S of support points chosen by the experimenter is never be only two, but let us say a number between 5 and 10. We chose here $N_S = 6$. Morever, almost instinctively, these support points are equidistantly set on the experimental design. Secondly, the repetition of experiments seems not always very relevant, because that means more work, and also the link between the quality of the parameter estimation and the number of repetitions is not easily understood. So, we consider a naive design without repetitions and other two with two repetitions at each of the six support points. Finally, the heterogeneity of the error variance is not taken into account when designing experiments, the model is rather transformed for instance by taking square root of the response. Here in a view to compare with the optimal designs we will consider the two situations: taking into account or not this heterogeneity with the model (7). Then, the naive designs we consider are defined as:

$$\xi^{naive1}_{6,6,\{1,\ldots,1\}} = \left\{ \begin{array}{l} T_1 = -5°C,\ T_{k+1} = T_k + 7 \\ \hat{w}\left(T_k\right) = 1 \\ r_k = 1 \end{array} \right\} \quad ; \quad k = 1, \ldots, 5 \qquad (19)$$

$$\xi^{naive2}_{12,6,\{2,\ldots,2\}} = \left\{ \begin{array}{l} T_1 = -5°C,\ T_{k+1} = T_k + 7 \\ \hat{w}\left(T_k\right) = 1 \\ r_k = 2 \end{array} \right\} \quad ; \quad k = 1, \ldots, 5 \qquad (20)$$

$$\xi^{naive3}_{12,6,\{2,\ldots,2\}} = \left\{ \begin{array}{l} T_1 = -5°C,\ T_{k+1} = T_k + 7 \\ \hat{w}\left(T_k\right) = 1/var\left(\mu_{\max}\right)_k \\ r_k = 2 \end{array} \right\} \quad ; \quad k = 1, \ldots, 5 \qquad (21)$$

So, the simulation of the observation sets for the naive designs (19) and (20) is performed with a uniform $var\left(\mu_{\max}\right)_S = 0.0008612$ on Ξ (obtained

by averaging the twenty values of $var\,(\mu_{\max})$ given in the Charles-Bajard paper), and consequently $\hat{w}\,(T_k) = 1, \forall k$. For the naive design (21) we use the model (7). At last we consider a $D-$optimal design, referred as $\xi^{D_0}_{6,2\{3,3\}}$,where $var\,(\mu_{\max})$ is assumed homogenous (then equals to 0.0008612). This design was obtained with the same computing means as for (11), it will be very useful for the comparison study as we will see in section 5.3. It is defined by

$$\xi^{D_0}_{6,2,\{3,3\}} = \left\{ \begin{array}{ll} T_1 = +1.76°C & T_2 = +30°C \\ \hat{w}\,(T_1) = 1 & \hat{w}\,(T_2) = 1 \\ r_1 = 3 & r_2 = 3 \end{array} \right\} \tag{22}$$

5.2 Montecarlo simulation and computed statistics

To compare the relative efficiency of the optimal and naive designs we must compare the quality of parameter estimates obtained with each design. To reach this goal two steps are necessary:

- In a first step, we simulate by a classical montecarlo method a large number G (typically G equals to 5000) of observation sets y_{ik}, $i = 1, \ldots, N_S$, $k = 1, \ldots, r_i$. For this simulation we use the CFT model assuming that $\hat{\theta}_j$, $j = 1, 2$, given in section 3, are the true value, and with an homogenous or heterogenous error distribution according to the design studied. For each observation set u ($u = 1, \ldots, G$) the parameter estimates $\hat{\theta}^{(u)}_j$ are computed by nonlinear regresssion.
- In a second step, with these G estimates we compute simulation estimates $\hat{\theta}^S_j$, and some relevant statistics relative to these $\hat{\theta}^S_j$.

These statistics enable us to compare efficiently the designs. These are, for j = 1,2, the following:

- $\hat{\theta}^S_j$, the simulation estimate, computed as

$$\hat{\theta}^S_j = \frac{1}{G} \sum_{u=1}^{G} \hat{\theta}^{(u)}_j \tag{23}$$

- $\hat{\sigma}(\hat{\theta}^S_j)$, the estimated standard error of the simulation estimate $\hat{\theta}^S_j$, computed as

$$\hat{\sigma}(\hat{\theta}^S_j) = \left[\frac{1}{G} \sum_{u=1}^{G} \left(\hat{\theta}^{(u)}_j - \hat{\theta}^S_j \right)^2 \right]^{1/2} \tag{24}$$

- $\hat{b}_{\%}(\hat{\theta}_j^S)$, the estimated bias pourcentage of $\hat{\theta}_j^S$, computed as

$$\hat{b}_{\%}(\hat{\theta}_i^S) = 100 \times \frac{\widehat{E\left(\hat{\theta}_j^S\right)} - \hat{\theta}_j}{\hat{\theta}_j^S} \tag{25}$$

  where $\widehat{E\left(\hat{\theta}_j^S\right)} = \hat{\theta}_j^S$, and where $\hat{\theta}_i$ is assumed to be the true value of the unknown parameter during the simulation,

- $\bar{b}_{\%}$, the average of the two $\hat{b}_{\%}(\hat{\theta}_j^S)$,

- $\widehat{CV}_{\%}\left(\hat{\theta}_j^S\right)$, the estimated variation coefficient, computed as

$$\widehat{CV}_{\%}\left(\hat{\theta}_j^S\right) = 100 \times \frac{\hat{\sigma}(\hat{\theta}_j^S)}{\left|\hat{\theta}_j^S\right|} \tag{26}$$

- $\widehat{EQM}_j$, the estimated mean quadratic error of the estimate $\hat{\theta}_j^S$, computed as

$$\widehat{EQM}_j = \left[\hat{b}_{\%}(\hat{\theta}_j^S)/100\right]^2 + \hat{\sigma}^2(\hat{\theta}_j^S) \tag{27}$$

- $\overline{EQM}$, the average of the two $\widehat{EQM}_j$.

5.3 Results of the comparison

In Table 3 are given these results. Important comments are necessary about this Table 3.

Simulation Statistics	Designs						
	$\xi^{D}_{6,2\{3,3\}}$	$\xi^{D_S}_{6,2\{3,3\}}$	$\xi^{ELD}_{6,2\{3;3\}}$	$\xi^{D_0}_{6,2\{3,3\}}$	$\xi^{naive1}_{6,6\{1,...,1\}}$	$\xi^{naive2}_{12,6\{2,...,2\}}$	$\xi^{naive3}_{12,6\{2,...,2\}}$
	hetero	hetero	hetero	homo	homo	homo	hetero
$\hat{\theta}_1^S$	-5.3296	-5.3314	-5.3301	-5.1662	-7	-3.3089	-5.3308
$\hat{\theta}_2^S$	11.2683	11.2925	11.2679	11.5354	10.76	12.0982	11.2654
$\hat{\sigma}(\hat{\theta}_1^S)$	0.2682	0.2682	0.2687	1.7324	NC	0.4430	0.4097
$\hat{\sigma}(\hat{\theta}_2^S)$	0.3982	0.5159	0.4010	0.9326	NC	1.0147	0.3561
$\hat{b}_{\%}(\hat{\theta}_1^S)$	0.042	0.074	0.052	3.12	NC	61	0.063
$\hat{b}_{\%}(\hat{\theta}_2^S)$	0.022	0.236	0.018	2.34	NC	7	0.004
$\bar{b}_{\%}$	0.032	0.15	0.035	2.73	NC	34	0.034
$\widehat{CV}_{\%}\left(\hat{\theta}_1^S\right)$	5.03	5.03	5.04	33	NC	13.39	7.68
$\widehat{CV}_{\%}\left(\hat{\theta}_2^S\right)$	3.53	4.57	3.56	8	NC	8.39	3.16
$\widehat{EQM}_1$	0.07	0.07	0.07	3	NC	0.57	0.17
$\widehat{EQM}_2$	0.16	0.27	0.16	0.9	NC	1.03	0.13
$\overline{EQM}$	0.11	0.17	0.12	1.94	NC	0.80	0.15
Γ_{int}	0	0	0	0	7.45(0.38)	5.27(0.49)	5.27(0.49)

Table 3 : Simulated statistics for comparing the designs: the three $D-$, D_S-, $ELD-$, optimal designs (see section 4 for details), the D_0-optimal design (homogeneous error variance), and the three naive designs (see section 5.1 for details). Hetero means heterogenous variance, homo means homogenous variance. NC means non computable.

First of all, we can see on the last line of Table 3 the quadratic mean of the intrinsic curvature, called Γ_{int}. This measure, and others, was proposed by Bates and Watts (1980) and they are particularly well explained in their book (Bates and Watts (1988, chapter 6)). We refer the reader to this latter one, and here we just recall how using Γ_{int}. This measure can be zero or positive. If it is zero that means that the expectation surface S_η is a plane, it is the case if $N_S = p$. Otherwise, if Γ_{int} is not zero, this surface S_η is not a plane, it can be a very complicated surface for some models especially if Γ_{int} is greater than a certain threshold (this one is given between parenthesis in Table 3). This surface S_η, included in the observation space $\mathbb{R}^{N_S}$, is the $p-$dimensionnal surface described by $\eta_\xi(x,\theta)$ when θ takes all its possible values in Θ. In nonlinear regression we are very concerned by S_η because the orthogonal (or the $W_\xi-$orthogonal) projection of the observation vector is performed on it. If it is a plane the projection is unic and the solution, i.e. the estimate $\hat{\theta}$, is unic: notice that this is always the case in linear regression.

If S_η is not a plane the projection can be not unic, and consequently it can appear more than one estimate solution. This phenomenon can be increased if the heterogeneity of the error variance is not taken into account: it occurs here with the naive design number 1, and we can see in Table 3 that even the very efficient classical algorithm (Levenberg-Marquardt algorithm) cannot avoid to lead to a nonsense solution for $\hat{\theta}$ with this naive design number 1. In comparison we observe that, even if (22) does not take into account the heterogeneity of the error variance, the estimate obtained is reasonably correct. The reason is based on the fact that $N_S = p$. Of course the estimated standard errors are unacceptable, they should lead to very inflated parametric confidence intervals, because the heterogeneity of the error variance is forgotten.
With the naive design number 2, N is now twice, and it is well known that Γ_{int} decreased when N increases. Then with this design the estimate is not very good, but it is better than for the preceeding naive design.
At last the naive design number 3 takes into account the heterogeneity of the error variance, and we can see that the estimates and their standard errors are quiet good. However, twelve experiments were necessary to reach this goal instead of only six with the three optimal designs. Moreover, its rather large intrinsic nonlinearity could cause large imprecision in the estimates if a mistaken prior θ_0 would be chosen, in a further analysis: indeed, it is known that the robustness of the estimates relatively to the prior θ_0 is greater if, notably, the intrinsic nonlinearity is large.
Now, if we compare the three optimal designs we can see that they are close each other relatively to the simulated statistics shown in Table 3. That is often observed when the parametric nonlinearity is reasonable that is the case here: we recall that these three optimal designs are all based on the first-order approximation of the analytic form of the CFT model.

6 Discussion

Now, we want to present and discuss four major principles that are the basis of an efficient experimental approach using optimal designs in the nonlinear situation. These principles have been applied above in the context of the CFT model, they are: the sequential approach, the $D-$optimality family, designs with $N_S = p$ and $N > N_S$, and taking into account the nature of the error variance.

6.1 Four major principles

6.1.1 Sequential approach

First of all, even if it seems obvious, we want to underline that it is a nonsense attitude to try to determine an optimal design at the beginning of a new experimental problem, is to say when none practical information is available. Typically in this case, the experimenter has no idea at all about the feasible range of the explanatory variables, the borders of the parametric domain, the eventual positiveness of the parameters, the experimental reproducibility, and so on. To undertake a first optimal design it is necessary to have some information, especially on the aspects just above evocated. Then, we recommend a sequential approach, at least based on two steps: the first experimental step cannot be optimal, but it provides (in general) some useful information because the experimenter has a scientific knowledge on his problem, and in a second step this information helps to determine a first optimal design that can be experimentally achieved. These two steps were applied here to the CFT model. The optimal design can be based on an approximate criterion or an exact criterion depending on several aspects such as the level on the nonlinearities of the postulated model (see the following subsections). If very accurate results are necessary further sequential optimal designs can be achieved.

6.1.2 $D-$optimality family

By the $D-$optimality family we mean the family of all the criteria based on the minimization of a scalar function of the determinant of an approximation of the Fisher information matrix. We know that this matrix tends to be the true variance matrix of the parameter estimate when N increases to infinity. In this sense the criteria used above, the local $D-$optimality, the D_S-optimality, and the $ELD-$optimality belong to this family. Other criteria exist in this family: they were proposed by Pronzato (1986), Pronzato and Walter (1988), and they common goal is to be robust relatively to a missspecification of θ_0. This design family is based on an approximate criteria for the nonlinear case, but we can check if it is an acceptable approximation for instance by plotting and comparing an X (exact) parametric confidence region (defined in Appendix B) to the classical ellipsoid. If p is too large the plotting is not possible but we can compare the volume of the anticipated ellipsoid (see Appendix A) and the expected volume of the X region by computing it with a classical montecarlo method. If this comparison shows a too large difference then specific optimal designs must be preferred, even if they are much complicated to compute. Note that a design of this $D-$optimality

family is often a good initial design, from a computational point of view, for the algorithms used for determining those specific designs. See the subsection 6.2 for a brief information on these speific designs.

6.1.3 Designs with $N_S = p$ and $N > N_S$

A major principle is to determine optimal designs with $N_S = p$, and with repetitions of experiments on these p support points, i.e. $N > N_S$. Let us explain the deep reasons of this principle, some aspects already being indicated in 5.3. First of all in this case the intrinsic nonlinearity, i.e. the curvature of S_η is zero, and consequently the projection $\hat{y}$ is always unic and also the minimum of the residual sum of squares is unic and global. This situation, especially if the error variance is heterogenous, is then a strong advantage, particularly from a computing point of view. We encountered a big difficulty with the naive designs (19) and (20) where $N_S \neq p$. Of course, it depends also on the analytic form of the postulated model, and the case $N_S \neq p$ does not always lead to multiple minima. Secondly, if we want to find a local $D-$optimal design, it exists a necessary and sufficient condition (NSC) proved by Vila (1991) that indicates that, if this NSC is true, the $D-$optimal design with N experiments is the repetition of the minimum $D-$optimal design with $Ns = p$. So, more experiments on another support points are not necessary. Notice that this condition does not exist for all the criteria. However, we can always add a few additionnal experiments in a view to check the predictive ability of the model in a second time. Of course, these additionnal experiments must not be used for the computation of the estimates, but only for the prediction. In fact these additionnal experiments should be set in spots where the predictive ability is clearly shown, is to say, in spots that could enable the validity of an alternative model. This is an open problem: where these additionnal experiments must be set in Ξ to discriminate two (or more) alternative models ? Atkinson (1972, 1975), Atkinson and Cox (1974), Atkinson and Fedorov (1975), Mathieu (1981), gave the solution if the models are linear and nested, but there is not yet a solution if the models are nonlinear and not nested. Otherwise, we want just to underline the very important aspect of performing true repetitions ($N > N_S = p$), because this the only one way to obtain a pure error. This latter enables to compute an unbiased estimation of the true error variance on one hand and to check its eventual heterogeneity on the other hand.

6.1.4 The nature of the error variance

About the nature of the error variance we want to underline that the best way is to take into account the eventual heterogeneity of the error variance by means of an approximate parametric model of this latter, and then using this model for the construction of the criterion and the optimal design. We showed how it is possible to do that with the CFT problem in section 2. Typically we can see in Charles-Bajard et al. (2003) that two minima occured when the error variance was assumed homogenous, whereas a single minimum occurs when this heterogeneity is not forgotten and modellized as we shown before. Otherwise, we recommend strongly this approach instead of trying to transform the response as with the square root transformation very often encountered. Indeed, this transformation is generally not optimal notably relatively to the predictive ability of the model, and moreover destroys the nature of the error variance and then can lead to very erroneous (generally inflated) parametric confidence intervals.

6.2 Specific designs for the nonlinear models

If the intrinsic and parametric nonlinearities of the model are large, and simultaneously the number N is small, especially if the error variance is large also, it can be a very bad approach to use designs from the $D-$optimality family. Also other families specific to the linear models, such as the $E-$optimality family (see for example Atkinson and Doneev (1992) for an introduction), is not a better approximation if those constraints yet exist.
As specific designs for nonlinear models we mean designs based on criteria that manage in some way the intrinsic and parametric nonlinearities of the model. In other words the key point is that a first-order approximation of the analytic form of the model is not achieved. There is not a lot of such specific criteria for constructing these specific designs. In this paper we want to point to two important approaches. The first specific approach is based on the exact parametric confidence regions, for instance kinds of regions defined in Appendix B. Then the criterion consists to minimize the expected volume of such regions. This approach was firstly proposed by Vila (1985, 1986) and also intensively analyzed by Gauchi (1999) and Gauchi and Vila (2004). The second specific approach is very different: it takes into account the exact probability density function (thus for N given) of the nonlinear least squares estimate $\hat{\theta}$ (Pázman, 1993). It was firstly proposed by Pázman and Pronzato (1992), and was extended and improved by Gauchi and Pázman (2003, 2004). The criteria involved in this second approach are based on scalar functions of the mean square error matrix of the nonlinear

least squares estimate $\hat{\theta}$, simultaneously with its exact density. Of course, this approach is highly technic and difficult to comput, but now the available computer power enables us to determine efficient designs as shown in Gauchi and Pázman (2003, 2004), especially if a stochastic minimization procedure is used.
A few words now on the computation of optimal designs. Of course, in the linear case, for many years we have been using specific algorithms (as the well-known double-exchange algorithm (Fedorov, 1972) that we used in section 4.1), or the classical deterministic optimization procedures for computing the optimal designs. However for the specific optimal designs for the nonlinear case we recommend, notably if p is large (let us say > 4), to use stochastic minimization procedures, see, e.g., Ermoliev (1990), Kushner and Yin (1997). Indeed, these procedures present a strong advantage because it is not necessary to express the analytic forms of the first and second order derivatives of the objective function (the criterion). This latter task is obligatory for instance when using deterministic quasi newton methods of optimization,that is almost untractable if p is large and the model is quiet sophisticated. In Gauchi and Pázman (2003, 2004) we give details of a very efficient method of stochastic optimization (MSO) improved from Vila (1990).

7 Conclusion

The results given in this paper show the advantages of optimal designs. These advantages more increase when the models contain a lot of parameters, see, e.g., the cardinal model with 10 parameters (Augustin, 1999). In this paper our intention was to be essentially pedagogical. However, the optimal designs proposed are good designs for the CFT model, designs we recommend to use at the laboratory. Our intention was also to underline major principles detailed in the discussion section. These principles can contribute to build an efficient and optimal experimental methodology in the future of the predictive microbiology. We plan in a further paper to show application to sophisticated predictive microbiology models (with large p), computation achieved by means of our MSO.

8 Appendix A: The confidence ellipsoïd and its volume

We recall that the confidence ellipsoïd of θ^* for a regression model is defined by:

$$\mathcal{R}_E = \left\{ \theta \in \Theta : \left(\theta - \hat{\theta}\right)^T M^{-1}\left(x, \hat{\theta}\right)\left(\theta - \hat{\theta}\right) \leq p s_\nu^2 \mathcal{F}(\alpha; p, \nu) \right\} \quad (28)$$

where s_ν^2 is an independant estimation of the unknown true error variance σ^2, based on ν degrees of freedom (in this paper we took $\nu = 18$, value coming from the Charles-Bajard paper), and $F(\alpha; p, \nu)$ is the $\alpha-$quantile of a Fisher distribution at p and ν degrees of freedom. It is an approximate (only asymptotically exact) confidence region if the model is nonlinear relatively to its parameters, as it is the case for the CFT model.
The volume of this ellipsoïd, determined with N data, is given by the formula (Gallant, 1976):

$$V_E = \frac{2\Gamma(N/2)\left[2\pi\sigma^2 \mathcal{F}(\alpha; p, N-p)/(N-p)\right]^{p/2}}{\Gamma(p/2)\,\Gamma((N-p)/2)} \times \left| M\left(x, \hat{\theta}\right)\right|^{-1/2} \quad (29)$$

where Γ is the eulerian gamma function. Notice that i) In the design situation where $\hat{\theta}$ is unknown we have to replace $\hat{\theta}$ by a prior θ_0 in these two formulas. This volume becomes an anticipated volume, it is the approach we used in this paper for the figures shown. ii) In the CFT problem σ^2 is equals to one because it is taken into account by means of the model (7) leading to the diagonal weights of the matrix W_ξ.

9 Appendix B: The exact X confidence region

In the classical statistical context an exact confidence region for a (vectorial) parameter (or a function of it) is a region whom its recovering probability of the true (unknown) θ^* parameter is exactly $1-\alpha$, where α is the risk level of the associated test connected to the region, see Lehman (1986) for a full and rigorous theory on this aspect. Halperin (1963) proposed an elegant region family based on the decomposition of the error vector in the observation space (sampling space). Later, Hamilton and Watts (1985) improved the test power connected to this family by introducing a new exact region. This region has been intensively analyzed by Vila (1985, 1986). We call this region as the X region and it is defined with our notations, at a $(1-\alpha)\,\%-$level, by:

$$\mathcal{R}_X = \left\{ \theta \in \Theta : R_X \leq p s_\nu^2 F(\alpha; p, \nu) \right\} \quad (30)$$

where

$$R_X = \left(y - \eta_\xi(x, \theta)\right)^T J(x, \theta_0)\, M^{-1}(x, \theta_0)\, J^T(x, \theta_0)\, W_\xi \left(y - \eta_\xi(x, \theta)\right) \quad (31)$$

The plots of Fig. 2 and 5 are based on the equation (**??**). The volume of this region and its marginal confidence intervals have been obtained by a usual montecarlo method.

10 References

Atkinson, A.C., 1972. Planning experiments to detect inadequate regression models. Biometrika, 29, 2, 275.

Atkinson, A.C., 1975. Planning experiments for model testing and discrimination. Math. Operationsforsch. u. Statist., 6, 2, 253-267

Atkinson, A.C., Cox, D. R., 1974. Planning experiments for discriminating between models. J. of Roy. Statist. Soc, B36, 3, 321-348.

Atkinson, A.C., Doneev, A.N., 1992. Optimum experimental designs. Oxford, Clarendon Press.

Atkinson, A.C., Fedorov, V.V., 1975. The design of experiments for discriminating between two rival models. Biometrika, 62, 1, 57.

Augustin, J.-C. 1999. Modélisation de la dynamique de croissance des populations de Listeria monocytogenes dans les aliments. Thesis. Université Claude Bernard - Lyon I, Lyon. France.

Augustin, J.-C., Carlier, V., 2000. Modelling the growth rate of Listria monocytogenes with a multiplicative type model including interactions environmental between environmental factors. International Journal of Food Microbiology, 56, 1, 53-70.

Bajard, S., Rosso, L., Fardel, G., Flandrois, J.P., 1996. The particular behaviour of Listeria monocytogenes under sub-optimal conditions. International Journal of Food Microbiology, 29, 201-211.

Baranyi, J., Roberts, T.A., 1994. A dynamic approach to predicting bacterial growth in food. International Journal of Food Microbiology, 23, 277-294.

Baranyi, J., Roberts, T.A., 1995. Mathematics of predictive food microbiology. International Journal of Food Microbiology, 26, 199-218.

Bates, D.M., Watts, D.G., 1980. Relative curvatures measures of nonlinearity. Journal of Royal Statistical Society, B42, 1-25.

Bates, D.M., Watts, D.G., 1988. Nonlinear Regression Analysis and its Applications. Wiley, New York.

Bernaerts, K., Versyck, K.J., Van Impe, J.F., 2000. On the design of optimal dynamic experiments for parameter estimation of a Ratkowsky-type growth kinetics at suboptimal temperatures. International Journal of Food Microbiology, 54, 27-38.

Chaloner, K., Verdinelli, 1995. Bayesian experimental design: a review. Statistical Science, 10, 3, 273-304.

Charles-Bajard, S., Flandrois, J.-P., Tomassone, R., 2003. Quelques problèmes statistiques rencontrés dans l'estimation de la température minimale de croissance de Listeria monocytogenes. Revue de Statistique Appliquée, vol LI, 3, 59-71.
Ermoliev, Y., 1990. Stochastic quasigradient methods. In Numerical techniques for stochastic optimization, chap. 6, Ermoliev & Wets, Scientific Editors. Springer-Verlag.
Fedorov, V.V., 1972. Theory of optimal experiments. Academic Press. New-York.
Gallant, A.R., 1976. Confidence regions for the parameters of a nonlinear regression model. Institute of Statistics Mimeograph Series n°1077. North Carolina State University. USA.
Gallant, A.R., 1987. Nonlinear Statistical Models. Wiley. New York.
Gauchi, J.-P., 1999. Contribution à l'étude et au calcul de critères de plans d'expériences optimaux pour des modèles de régression non linéaire. Thèse de doctorat, CNAM, Paris.
Gauchi, J.-P., Vila, J.-P., 2004. Planification expérimentale optimale pour des modèles de régression non linéaire: Approche basée sur des régions de confiance paramétriques exactes. Technical Report 2004-2. National Institute of Agricultural Research (INRA), MIA Unit, Jouy-en-Josas. France.
Gauchi, J.-P., Pázman, A., 2003. Distribution of LSE, stochastic optimization and optimum designs in nonlinear models. Technical Report 2003-7, National Institute of Agricultural Research (INRA), MIA Unit, Jouy-en-Josas, France.
Gauchi, J.-P., Pázman, A., 2005. Designs in nonlinear regression by stochastic minimization of functionals of the mean square error matrix. *In Press* in J. of Statistical Planning and Inference.
Halperin, M., 1963. Confidence interval estimation in non-linear regression. J. of Roy. Statist. Soc, B25, 2, 330-333.
Hamilton, D.C., Watts, D.G.,1985. A quadratic design criterion for precise estimation in nonlinear regression models. Technometrics, 27, 3.
Kiefer, J., Wolfowitz, J., 1960. The equivalence of two extremum problems. Canadian Journal of Mathematics, 12, 275-285.
Kushner, H.J., Yin, G.G., 1997. Stochastic approximation algorithms and applications. Springer. Heidelberg.
Lehman, E.L., 1986. Testing Statistical Hypotheses. Wiley. New-York.
Le Marc, Y., Huchet, V., Bourgeois, C.M., Guyonnet, J.P., Mafart, P., Thuault, D., 2002. Combined effects of pH and organic aids on the growth rate of Listeria inocua. Int. J. Food Microbiol., 73, 2-3, 219-237.
Mathieu, D., 1981. Contribution de la méthodologie de la recherche expérimentale à l'étude des relations structure-activité. Thèse. Université d'Aix-

Marseille. France.
Pázman, A., 1986. Foundations of optimum experimental designs. Dordrecht, The Netherlands, Reidel.
Pázman, A., 1993. Nonlinear statistical models. Dordrecht, Kluwer.
Pázman, A., Pronzato, L., 1992. Nonlinear experimental design based on the distribution of estimators. J. of Statist. Planning and Inference 33, 385-402.
Pronzato, L., 1986. Synthèse d'expériences robustes pour modèles à paramètres incertains. Thèse. Université d'Orsay. France.
Pronzato, L., Walter, E., 1988. Robust experiment design for nonlinear regression models, in V.V. Fedorov and E. Laüter (Eds), Model Oriented Data Analysis, Proc. Eisenach GDR, 1987. Lectures Notes in Economics and Mathematical Systems, 297, 77-86. Berlin. Springer.
Pukelsheim, F., 1993. Optimal designs of experiments. Wiley.
Ratkovsky, D.A., Olley, J., Ball, A., 1982. Relationship between temperature and growth rate of bacterial cultures. J. of Bacteriology, 149,1, 1-5.
Rosso, L. , Lobry, J.R., Flandrois, J.-P., 1993. An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model. J. of Theoretical Biology, 162, 447-463.
Rosso, L., 1995. Modélisation et microbiologie prévisionnelle: élaboration d'un nouveau outil pour l'Agroalimentaire. Thèse. Lyon.
Rosso, L. , Flandrois, J.-P., 1995. Application de la modélisation à la description et la prévision de la croissance des flores de micro-organismes. Bull. Soc. Fr. Microbiol., 10, 3.
Silvey, S.D., 1980. Optimal designs: an introduction to the theory for parameter estimation. Chapman and Hall. London.
Versyk, K.J., Bernaerts, K., Geeraerd, A.H., Van Impe, J.F., 1999. Introducing optimal experimental design in predictive microbiology: A motivating example. International Journal of Food Microbiology, 51, 39-51.
Vila, J.P., 1985. Etudes et comparaisons de critères de plans d'expériences optimaux pour l'estimation des paramètres d'un modèle de régression non linéaire. Thèse de Mathématiques. Université d'Orsay (n°696). France.
Vila, J.P., 1986. Optimal designs for parameters estimations in nonlinear regression models : exact criteria. Invited Lecture XIIIth, Int. Biometrics Conf., Seattle, USA.
Vila, J.P., 1990. Exact experimental designs via stochastic optimization for nonlinear regression models. Compstat, 1990, 291-296.
Vila, J.P., 1991. Local optimality of replications from a minimal $D-$optimal design in regression: a sufficient and a quasi-necessary condition. J. of Statistical Planning and Inference, 29, 261-277.