Optimal Experimental Designs in the Flow Rate of Particles

Mariano AMO-SALAS, Elvira DELGADO-MÁRQUEZ, and Jesús LÓPEZ-FIDALGO

Department of Mathematics
University of Castilla-La Mancha
Ciudad Real, Spain
(Mariano.Amo@uclm.es; Elvira.Delgado@uclm.es; Jesus.LopezFidalgo@uclm.es)

This article focuses on analyzing the process of jam formation during the discharge by gravity of granular material stored in a two-dimensional silo. The aim of the article is two-fold. First, optimal experimental designs are computed, in which four approaches are considered: D-optimality, a combination of D-optimality and a cost/gain function, Bayesian D-optimality, and sequential designing. These results reveal that the efficiency of the design used by the experimenters can be improved dramatically. A sensitivity analysis with respect to the most important parameter is also performed. Second, estimation of the unknown parameters is done using least squares, that is, assuming normality, and also via maximum likelihood assuming the exponential distribution. Simulations for the designs considered in this article show that the variance, the mean squared error, and the bias of the estimators using maximum likelihood are in most cases lower than those using least squares. Supplementary materials for this article are available online.

KEY WORDS: Bayesian design; Compound criterion; D-optimal design; Granular material; Jam formation; Sensitivity analysis; Sequential design.

1. INTRODUCTION

Material in granular form is a collection of macroscopic solid similar particles, which interact with each other by forces of contact. The size of the granular material ranges from millimeters or centimeters (e.g., grains of sand or seeds) to meters (e.g., rocks). This granular material is usually stored in silos (see Figure 1 as an illustrative model in two dimension). A real-world (three-dimensional) example is the transport of granular material in the mining industry through a vertical tube. During the discharging process caused by gravity, the flow of granular material can be interrupted due to the formation of an arch at the outlet (Figure 1). An arch is a structure where the particles are mutually stable. The formation of an arch causes a jam. If one of the particles that form an arch is removed, the arch collapses under the effect of gravity. After the formation of an arch, an input of energy (blowing, shaking, or tapping) is necessary to break the blocking arch and restart the flow. Then the grains fall again until a new arch is formed. The event of breaking the arches may be dangerous, expensive, or just difficult, as happens in the case of the mining industry. An avalanche can be defined as the amount of granular material that falls between two successive jamming events. The waiting time between two successive jamming events can be used to measure the extent of an avalanche.

In this work, we focus on analyzing the process of jam formation and creation of avalanches during the discharge by gravity of granular material stored in a two-dimensional silo. When the silo is filled with granular material, grains pour freely from the outlet due to gravity (see Figure 1). In this article, the average size of avalanches, $S$, is measured as the waiting time between two successive jamming events. Following (Janda et al. 2008), we shall assume that $S$ has an exponential distribution with mean and variance

$$E(S) = \frac{A}{(\phi_C - \phi)}^\gamma, \quad \text{var}(S) = \left(\frac{A}{(\phi_C - \phi)}\right)^2,$$

where $\phi$ is the diameter of the outlet and $\phi_C$ is the theoretical critical diameter above which jamming is not possible. The parameter $A$ corresponds to the value of $S$ when $\phi = \phi_C - 1$ and $\gamma$ is the rate of exponential decline. The vector of unknown parameters, $\theta = (\gamma, A, \phi_C)$, has to be estimated from the data collected on $S$.

In the literature, least-square estimators (LSE) are computed for this problem as if the data were Gaussian, neglecting the exponential distribution of the data. In this article, maximum likelihood estimators (MLE) for this distribution are computed and compared to LSE (MLE assuming a Gaussian distribution). We need data points $(\phi_i, S_i), i = 1, \ldots, n$ to perform the estimation. To generate such data, we need to choose the input points $\phi$ within an interval $[\phi_L, \phi_U]$. Earlier, this has been done in a rudimentary form (e.g., Janda et al. 2008). In this article, we propose a methodology for optimal selection of these points with the objective of achieving maximum possible efficiency in parameter estimation.

This article has been organized as follows. In Section 2, an introduction to relevant aspects of Optimal Experimental Design theory is provided. In Section 3, the best sample points to run experiments for estimating the vector of unknown parameters of the model given by (1) are computed using the D-optimality criterion. In Section 4, a compound criterion is considered.
to include the cost of the experimentation in the problem. In Section 5, a sensitivity analysis is performed in two ways. On the one hand, the relative efficiency of the design obtained is evaluated for possible changes in the most important parameter of the model. On the other hand, a comparative study of the LSEs and MLEs is made. Finally, in Section 6, Bayesian and sequential approaches are considered to compute D-optimal designs.

2. OPTIMAL EXPERIMENTAL DESIGN THEORY

An exact experimental design of size \( n \) is a collection of values of the explanatory variable, \( \phi_1, \ldots, \phi_n \), in a defined design space \( \mathcal{X}_\phi = [\phi_L, \phi_U] \), where some values may be repeated. In the example considered in this article, \( \phi_L \) and \( \phi_U \) are the lower and upper extremes of the possible outlet diameters of the silo, satisfying \( \rho < \phi_L < \phi_U < \phi_C \) where \( \rho \) is the diameter of the granular material. Assuming there are just \( k \) different points in the design and they are, without loss of generality, \( \phi_1, \ldots, \phi_k \), then a probability measure can be defined with \( 0 \leq p_i \leq 1 \) being the proportion of times \( \phi_i \) appears in the design (\( i = 1, \ldots, k \)).

Thus, the design can be denoted as

\[
\xi = \left\{ \phi_1, \ldots, \phi_k \right\} /_{p_1, \ldots, p_k},
\]

where \( \sum_{i=1}^{k} p_i = 1 \).

This suggests a generalization of the concept of experimental design to any discrete probability measure with finite support on the design space, where \( \xi \) is the probability mass function associated. This is called an approximate design.

A major advantage of the optimal experimental design (OED) is the cost reduction of experimentation by allowing statistical models to be estimated with fewer experimental runs. For computing optimal designs, the Fisher information matrix (FIM), or simply information matrix, is used because its inverse is asymptotically proportional to the covariance matrix of the estimators of the parameters under some regularity conditions.

The information matrix for a design \( \xi \) is defined as

\[
M(\xi, \theta) = \sum_{i=1}^{k} \xi(\phi_i)I(\phi_i, \theta),
\]

where \( \xi(\phi_i) \) is the proportion of experiments at \( \phi_i \), \( I(\phi_i, \theta) = -E(\frac{\partial^2 L(\theta)}{\partial \theta^2}) \) is the FIM at point \( \phi_i \), and \( L(\theta) \) is the log-likelihood function.

By Caratheodory’s theorem (see Fedorov 1972, p. 66 for details), it is known that for any design there is always another design with at most \( \frac{m+1}{2} \) different points in its support such that their information matrices are the same, where \( m \) is the number of unknown parameters to be estimated. This helps the search for an optimal design and makes it realizable in practice.

The FIM for a nonlinear model, as in this case, depends on the explanatory variable, \( \phi \), and upper extremes of the possible outlet diameters of the silo, \( \mathcal{X}_\phi = [\phi_L, \phi_U, \phi_C] \). Assuming \( \phi \) is the probability mass function on the design space. The optimal design problem is concerned with finding a weighted average of the two criterion functions, say \( \Phi \), defined on \( \mathcal{X}_\phi \) instead of \( \Phi(M(\xi, \theta)) \).

In this work, D-optimal designs are computed. D-optimality is the most popular criterion that minimizes the volume of an approximate confidence ellipsoid of the vector of unknown parameters \( \theta \). This volume is inversely proportional to the square root of the determinant of the information matrix. Then, this criterion is defined by

\[
\Phi_D(\xi) = -\log |M(\xi, \theta)|,
\]

where \( |\cdot| \) stands for the determinant.

When the experimenter must pay attention to two goals represented by two convex criterion functions, a compromise criterion can be expressed as a single compound criterion based on a weighted average of the two criterion functions, say \( \Phi_1 \) and \( \Phi_2 \),

\[
\Phi_\lambda(\xi) = (1 - \lambda)\Phi_1(\xi) + \lambda\Phi_2(\xi), \quad \lambda \in [0, 1].
\]

An important result, the General Equivalence Theorem (GET; see e.g., Atkinson, Donev, and Tobias 2007, sec. 9.2 for details), gives a tool for checking whether a design is optimal or not under such optimality criteria. Additionally, it provides a powerful tool for algorithmic computation. Let \( \psi(\phi, \xi, \theta) \) be the Fréchet’ directional derivative in the direction of a one-point design at \( \phi \). This function is frequently called the sensitivity function. The GET states that under some conditions of the criterion function, \( \psi(\phi, \xi, \theta) \) achieves its minimum value, zero, at the support points of the optimal design. In particular, the sensitivity
function for D-optimality is 
\[ \psi(\phi, \xi^*, \theta) = m - \text{tr} \left[ M^{-1}(\xi^*, \theta) I(\phi, \theta) \right]. \]

The \( \Phi \)-efficiency measures how good an arbitrary design is with respect to the optimal design,
\[ \text{Eff}_\Phi(\xi) = \frac{\Phi(\xi^*)}{\Phi(\xi)}. \]

If \( \Phi \) is positive homogeneous, the efficiency has a practical interpretation. For instance, if the efficiency of a particular design is 60%, then 40% of the experiments would be saved from the optimal design to get the same statistical inferences.

As mentioned above, the FIM depends on the nominal value of the vector of unknown parameters \( \theta \), and so does the criterion function. Prior knowledge on the possible values of the vector of unknown parameters of model (1) can be used to deal with this problem. Locally optimal designs are computed considering specific nominal values of the parameters and doing the optimization according to the criterion chosen.

For a Bayesian approach, the prior information on the vector of unknown parameters, \( \theta \), is represented by a prior distribution, \( \pi(\theta) \). Chaloner and Verdinelli (1995) provided a general overview of different Bayesian optimality criteria used in the literature giving a general structure of the problem. Using the approach given by Lindley and Smith (1972, pp. 19–20), the problem is reduced to finding the design that maximizes the expected utility of the experiment outcome. The optimal design problem is then determined by the expected utility of the best decision,
\[ \xi^* = \arg \max_{\xi \in \mathcal{S}} \int_{\mathcal{S}} \max_{d \in \mathcal{D}} \int_{\Theta} U(d, \theta, \xi, s) \pi(\theta|\xi, s) f(s|\xi) d\theta ds, \]
where \( \Theta \) is the parameters space, \( U(d, \theta, \xi, s) \) is the utility function chosen to satisfy certain goal, \( s \) is the observed data from a sample space \( \mathcal{S} \), \( f(s|\xi) \) is a suitable marginal density for the response, and \( d \) is a decision on the model, for example, the estimator of the parameters. Thus, the decision consists of two parts: first, selection of \( \xi \), and then the choice of an optimal estimator of the parameters, \( d \).

Chaloner and Verdinelli (1995, sec. 4.2) stated that in the case of nonlinear models the expected utility needs an approximation. Using Shannon information as utility function, a normal approximation for the estimator, and taking the predictive distribution of the estimator to be the prior distribution lead to the following criterion described by Chaloner and Verdinelli (1995),
\[ \Phi_B(\xi) = E_\theta(\Phi(M(\xi, \theta))). \]  

For a sequential design approach it is necessary to choose an initial design, for example, a locally optimal design for some nominal values of the parameters, perform the experiments, and estimate the value of the vector of unknown parameters using maximum likelihood estimation. Then a new optimal point is chosen using these estimators as nominal values of the parameters and considering the previous design including the new point. Chaudhury and Mykland (1993) provided conditions, first for guaranteeing that the inverse of the information matrix converges to the covariance matrix, then for proving the convergence of the probability distribution of the estimators to the Normal distribution.

The statistical techniques of OED have been widely used. An extensive bibliography can be cited, for example, the works of Fedorov (1972), Pázman (1986), Pukelsheim (1993), or Atkinson et al. (2007).

### 3. D-OPTIMAL DESIGN

In this section, D-optimal designs are computed for model (1). The D-optimal design, \( \xi_{D^*}, \) minimizes the criterion function \( \Phi_D(\xi) \). The information matrix at point \( \phi \) is
\[ I(\phi, \theta) = \begin{pmatrix} \log(\phi_C - \phi) & \frac{1}{\phi C^2} \log(\phi_C - \phi) & \frac{1}{2} \gamma \log(\phi_C - \phi) \\ \frac{1}{\phi C} & \frac{1}{A^2} & \frac{1}{2} \gamma \phi C \\ \frac{1}{\phi C} & \frac{1}{2} \gamma \phi C & \frac{1}{x^2} \end{pmatrix}. \]

For simplicity, the transformation \( x = \phi_C - \phi \) is used and then the information matrix at point \( x \) is
\[ I(x, \theta) = \begin{pmatrix} \log(x^2) & -\frac{\log(x)}{x^2} & -\frac{\gamma \log(x)}{x^2} \\ -\frac{\log(x)}{x} & \frac{1}{A^2} & \gamma \phi C \\ -\frac{\gamma \log(x)}{x^2} & \frac{\gamma \phi C}{x^2} & \frac{1}{2} \gamma \phi C \end{pmatrix}, \]
and the new design space is then \( X' = [\phi_C - \phi_U, \phi_C - \phi_L] = [x_1, x_2] \). Thus, the new design space depends on \( \phi_C \).

It is easy to check that if the support of the D-optimal design equals the number of parameters of the model, which is frequently the case, then the determinant is the product of the weights of the design and a function depending solely on the support points (see Atkinson et al. 2007, No. 8 of sec. 11.1; Pázman 1986, proof of proposition VI.4.). Then the product of the weights is minimized for equal weights no matter the shape of the model considered. The problem reduces to identification of those support points.

Thus, we first consider three-point equal weights candidates,
\[ x = \left\{ \begin{array}{l} x_1 \\ 1/3 \\ x_2 \\ 1/3 \\ x_3 \\ 1/3 \end{array} \right\}, \]
with \( x_1 < x_2 < x_3 \). The FIM for this design is then
\[ M(\xi, \theta) = \frac{1}{3} I(x_1, \theta) + \frac{1}{3} I(x_2, \theta) + \frac{1}{3} I(x_3, \theta) \]
and the determinant,
\[ |M(\xi, \theta)| = \frac{\gamma^2}{27 A^2} \times (x_1x_2x_3) \log(x_1 + x_2(x_1 - x_3))^2 + x_1^2x_2^2x_3^2 \times x_1^2x_2^2x_3^2. \]

It is fortunate that although (1) is a nonlinear model with respect to \( \gamma \), the D-optimal design does not depend on it. From here, the following result is derived.

**Theorem 3.1.** Given the model defined by (1) and the design space, \( X_\phi = [\phi_L, \phi_U] \), the three-point equally weighted design...
maximizing the determinant is
\[
\xi_D^* = \begin{bmatrix} \phi_L & \phi^* & \phi_U \\ 1/3 & 1/3 & 1/3 \end{bmatrix},
\]
where \( \phi^* = \phi_C - \frac{(\phi_C - \phi_L)(\phi_C - \phi_U)\log(\phi_C - \phi_L) - \log(\phi_C - \phi_U)}{\phi_C - \phi_U} \) and \( \phi_L < \phi^* < \phi_U \).

Moreover, if \( \psi(\phi, \xi_D^*, \theta) \geq 0 \) for all \( \phi \), then \( \xi_D^* \) is actually a D-optimal design.

Proof. See Appendix A in the online supplementary materials. \( \square \)

Example. Janda et al. (2008) considered \( X_\phi = [1.53, 5.63] \) as the design space. The estimated parameters are used as nominal values here, \( \theta^{(0)} = (12.7, 1.1 \times 10^{11}, 8.5) \). Then
\[
\xi_D = \begin{bmatrix} 1.53 & 4.17 & 5.63 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.
\]

Figure 2 shows the sensitivity function \( \psi(\phi, \xi_D, \theta) \) proving this is actually the D-optimal design for this example.

Janda et al. (2008) considered a design with eight nonreplicated points. Thus, we put the same weights for all of them as approximate design,
\[
\xi_k = \begin{bmatrix} 1.53 & 2.17 & 2.51 & 3.02 & 3.52 & 4.12 & 4.36 & 4.8 \\ 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \end{bmatrix}.
\]

The D-efficiency of this design is
\[
\text{Eff}_D(\xi_k) = \left( \frac{|M(\xi_k, \theta)|}{|M(\xi_D, \theta)|} \right)^{1/3} = 33.57%.
\]

Therefore, with the D-optimal design, the number of experiments can be reduced by 66.43% to achieve the same D-efficiency as the design \( \xi_k \).

Remark 3.1. According to the Elfving’s graphic method proposed by López-Fidalgo and Rodríguez Díaz (2004), the D-optimal design computed for \( \phi_C \) for the example given by Janda et al. (2008) is
\[
\xi_D^* = \begin{bmatrix} 1.53 & 3.67 & 5.63 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}.
\]

The \( D^* \)-efficiency of (6) is 73.6%, which is reasonably high taking into account the relevant role of \( \phi_C \). It is important to emphasize the importance of the parameter \( \phi_C \) by its physical meaning, but it is also just as important to have a good fit of the model to make predictions.

4. COMPOUND CRITERION

The purpose of this article is to compute the best values of \( \phi \) to run the experiments for estimating the vector of unknown parameters, \( \theta \), for model (1). But, an experiment is usually conditioned in practice by economical, temporary factors or lack of resources. In the example considered here the waiting time between two successive jamming events cannot be too long. As a matter of fact it would be natural to define some cost function increasing with this waiting time, that is, with the diameter of the outlet. A criterion combining D-optimality and a gain function is defined by
\[
\Phi_\lambda(\xi) = (1 - \lambda)\Phi_D(\xi) + \lambda \sum_{i=1}^{k} \xi_i \log(\phi_i),
\]
where \( 0 \leq \lambda \leq 1 \) and \( G(\phi) \) is the gain (opposite to the cost) function associated to \( \phi \). The transformation \( x = \phi_C - \phi \) is used again. We set the following compound criterion function to be maximized,
\[
\Phi_\lambda(\xi) = (1 - \lambda) \log |M(\xi)| + \lambda \sum_{i=1}^{k} \xi_i \log(x_i + 1).
\]

where a gain (instead of loss) function of experimentation for a fixed \( \phi_i \) is considered here,
\[
G(\phi) = \log(\phi_C - \phi + 1).
\]

This provides a compromise scale for combining D-optimality and the cost of experimentation.

Table 1 shows the optimal designs for three different values of \( \lambda \) considering again the example from Janda et al. (2008). For \( \lambda = 0 \), we get the D-optimal design computed in Section 3. For \( \lambda = 1 \), the optimal design for the cost function is supported at the minimum possible value of \( \phi \), where the expected waiting time between consecutive jams is minimum.

The sensitivity function for the compound criterion is
\[
\psi(\phi, \xi^*, \theta) = (1 - \lambda) \left( \right.
\frac{\text{tr}(M^{-1}(\xi^*)I(\phi, \theta)) - m}{\phi_C - \phi}
+ \lambda \left( \log(\phi_C - \phi + 1) - \sum_{i=1}^{m} \xi_i \log(\phi_C - \phi_i + 1) \right) \right) / \phi_C - \phi_1 + 1 \right).
\]

We have checked numerically that the GET is satisfied for \( 0 \leq \lambda \leq 1 \). The solid and dashed lines of Figure 3 show the efficiencies of the compound designs depending on the value

| \( \lambda \) | Design | \text{Eff}_{D-opt} | \text{Eff}_{Cost} |
|--------|--------|-----------------|------------------|
| 0.0    | \begin{bmatrix} 1.53 & 4.17 & 5.63 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} | 1.00 | 0.935 |
| 0.745  | \begin{bmatrix} 1.53 & 4.01 & 5.63 \\ 0.46 & 0.31 & 0.23 \end{bmatrix} | 0.95 | 0.95 |
| 1.0    | \begin{bmatrix} 1.53 \\ 1 \end{bmatrix} | 3.62 \times 10^{-5} | 1.00 |
of \( \lambda \) with respect to the D-optimal design and the design minimizing the cost function, respectively. Equal efficiencies are reached for \( \lambda = 0.745 \). The efficiencies of the designs for any value of \( \lambda \) with respect to the cost function are always greater than 90%. But, the D-efficiency decreases dramatically for values of \( \lambda > 0.745 \).

5. SENSITIVITY ANALYSIS AND ESTIMATION STRATEGIES

In this section, we focus on two issues. First, we check how sensitive is the optimal design to the choice of the nominal value of \( \phi_C \). For that, the efficiency of the optimal design with the nominal value \( \phi_C^{(0)} = 8.5 \) is computed assuming different values, say \( \phi_C^* \), around it. Second, a comparative study of the least-squares estimation and maximum likelihood estimation of the vector of unknown parameters \( \theta \) is performed.

5.1 Sensitivity Analysis for \( \phi_C \)

One of the main goals of this problem is to estimate a critical diameter for the outlet above which jamming is not possible (in practical terms). As has been explained in Section 3, the computation of the D-optimal design depends on the nominal value of \( \phi_C \). The experimenters have estimated the value of the unknown parameter \( \phi_C = 8.5 \) using least squares. This value has been used as a nominal value for computing the optimal design throughout the article.

The sensitivity analysis is performed with respect to the D-optimal design, (6). We consider different values for the critical diameter in a neighborhood of the nominal value, \( \phi_C^{(0)} \in [\phi_C^{(0)} - \varepsilon_1, \phi_C^{(0)} + \varepsilon_2] \), where \( \phi_C^{(0)} - \varepsilon_1 > \phi_U \) and \( \varepsilon_1, \varepsilon_2 > 0 \). For values \( \phi_C^{(0)} - \varepsilon_1 < \phi_U \), the point

\[
\phi^* = \phi_C - \frac{(\phi_C - \phi_U)(\phi_C - \phi_L)(\log(\phi_C - \phi_U) - \log(\phi_C - \phi_L))}{\phi_L - \phi_U}
\]

does not belong to the design space \( X_\theta \). The relative D-efficiency of the D-optimal design obtained with the nominal value of the parameter is defined as

\[
\text{Eff}_D(\xi_{\phi_C^*}) = \left( \frac{M(\xi_{\phi_C^*}, \phi_C^*)}{M(\xi_{\phi_C}, \phi_C)} \right)^{1/m},
\]

where \( M(\xi_{\phi_C}, \phi_C^*) \) is the FIM assuming the value of \( \phi_C^* \) for \( \phi_C \).

A sensitivity analysis may be performed in a similar way for the compound criterion defined in Section 4. In this case, the sensitivity analysis is performed with respect to the optimal design computed for \( \lambda = 0.745 \).

Figure 4 shows the behavior of the efficiency with respect to the value of \( \phi_C^* \) for both cases. The range for the possible true values of the parameter was selected as follows. The estimator of \( \phi_C \) follows asymptotically a Normal distribution with mean 8.5 and standard deviation 4.29 (estimated from the FIM). In general, possible values for \( \phi_C^* \) may be generated from this distribution and the relative efficiency computed if it is computationally too demanding. But in this case we were able to compute a number of efficiencies to complete the plots of Figure 4. We used the distribution to select the range of possible true values not farther than three standard deviations from the mean, that is, \( 8.5 + 3 \times 4.29 = 21.37 \). On the other hand, it must be above the upper limit of the range for \( \phi_C \), that is, 5.63. This gives reasonable limits, (5.63, 21.37), for checking the sensitivity. Thus, Figure 4 (left) shows that for possible true values below 8.5, the D-efficiency of the D-optimal design decays quite fast while for possible true values greater than 8.5, the D-efficiency of the D-optimal design is over 94%. The efficiency of the optimal design for the compound criterion (Figure 4, right) decays slightly for possible true values below 8.5 but faster than for values greater
than 8.5. Therefore, it is better to underestimate the nominal value of the parameter \( \phi_C \) leaving less chances to the left for the possible true values. As shown in Figure 4, the behavior of the sensitivity analysis is very similar for different values for \( \phi_C^{(0)} \).

5.2 Estimation Strategies

In previous sections, optimal designs have been computed to run experiments from different perspectives. These designs are always based on the maximum likelihood estimation from the exponential distribution of the observations. In practice the experimenters have used least squares to estimate the unknown parameters using the methodology developed by Chaloner and Verdinelli (1995) considering prior information about the vector of unknown parameters \( \theta \). This information is represented by a prior distribution, \( \pi(\theta) \). In addition, we compute sequential D-optimal designs using the methodology developed by Chaudhury and Mykland (1993).

6. ALTERNATIVE APPROACHES

So far, locally optimal designs have been computed for the nominal value of the vector of unknown parameters. In this section, we compute Bayesian D-optimal designs following the methodology developed by Chaloner and Verdinelli (1995) considering prior information about the vector of unknown parameters \( \theta \). This information is represented by a prior distribution, \( \pi(\theta) \). In addition, we compute sequential D-optimal designs using the methodology developed by Chaudhury and Mykland (1993).

6.1 Bayesian Designs

Using knowledge of the researchers in the experimental field, it is assumed that \( \phi_U \) is always a positive real number greater than the upper extreme of the design space, \( \chi_U \). In particular, two continuous prior distributions are considered for \( \phi_C \). A Log-normal distribution shifted to the right starting at the upper limit of the design space and a truncated Normal distribution to the left of the upper limit of the design space are used.

Let \( \mu_{LN} = 8.5 \) and \( \sigma_{LN} \) be varying on (0,5]. According to the Bayesian methodology for nonlinear models (Chaloner and Verdinelli 1995), the design

\[
\xi^{*}_B = \left\{ \phi_L, \phi_U \right\}
\]

is the Bayesian D-optimal design in the sense of (3) if it maximizes

\[
E_\theta \left( \log |M(\xi, \theta)| \right) = \int_{\phi_U}^{\infty} \log(|M(\xi, \theta)|) \pi(\phi_C) d\phi_C,
\]

where \( \pi(\phi_C) \) is the probability density function for the Log-normal distribution shifted to the right and the sensitivity function for the GET is

\[
\psi(\phi, \xi^{*}_B, \theta) = \int_{\phi_U}^{\infty} \text{tr} \left[ M^{-1}(\xi^{*}_B, \theta) I(\phi, \theta) \right] \pi(\phi_C) d\phi_C - m.
\]

Now, consider a truncated Normal distribution with parameters \( \mu \) and \( \sigma \) with support \([\phi_U, \infty)\) for the value of \( \phi_C \). Then the

Table 2. Mean squared error of the estimators of unknown parameters using LSE and MLE for the D-optimal design and the design \( \xi_s \)

|       | LSE         | MLE         | LSE         | MLE         |
|-------|-------------|-------------|-------------|-------------|
| \( \xi_s \) | 2.77 \times 10^9 | 4.91 \times 10^9 | 2.69 \times 10^9 | 4.59 \times 10^9 |
| \( \gamma \) | 0.0846      | 0.0144      | 0.0759      | 0.00667     |
| \( \phi_C \) | 0.0872      | 0.00572     | 0.0805      | 0.00137     |

Table 3. Bias of the estimators of unknown parameters using LSE and MLE for the D-optimal design and the design \( \xi_s \)

|       | LSE         | MLE         | LSE         | MLE         |
|-------|-------------|-------------|-------------|-------------|
| \( \xi_s \) | 2.499 \times 10^9 | 2.532 \times 10^9 | 2.285 \times 10^9 | 2.08 \times 10^9 |
| \( \gamma \) | -0.00275 | 0.000658   | -0.0112    | 0.00582     |
| \( \phi_C \) | -0.00387 | 0.00658    | 0.00624    | 0.00177     |
mean and variance of this distribution are

$$
\mu_T = \frac{1}{1 - \Pi(\phi_U)} \int_{\phi_U}^{\infty} \phi_C \pi(\phi_C) d\phi_C,
$$

$$
\sigma_T^2 = \frac{1}{1 - \Pi(\phi_U)} \int_{\phi_U}^{\infty} \phi_C^2 \pi(\phi_C) d\phi_C - \mu_T^2,
$$

(11)

where $\Pi(\cdot)$ is the cumulative distribution function and $\pi(\cdot)$ is the probability density function of the Normal distribution $N(\mu, \sigma)$. The mean is again fixed to $\mu_T = 8.5$ and the standard deviation $\sigma_T$ is allowed to vary on $(0, 5]$.

In this case, the design (10) is the Bayesian D-optimal design if it maximizes

$$
E_{\theta}(\log |M(\xi, \theta)|) = \int_{\phi_U}^{\infty} \frac{\log(M(\xi, \theta)) \pi(\phi_C)}{1 - \Pi(\phi_U)} d\phi_C,
$$

and the sensitivity function for the GET is

$$
\psi(\phi, \xi^*_B, \theta) = \int_{\phi_U}^{\infty} \text{tr} \left[ (M^{-1}(\xi^*_B, \theta) M) \right] \frac{\pi(\phi_C)}{1 - \Pi(\phi_U)} d\phi_C - m.
$$

The GET is satisfied for all $\sigma_{LN}, \sigma_T \in (0, 5]$. Figures 5 and 6 show the middle optimal point, $\phi^*_B$, and the efficiency with respect to the locally D-optimal design (6). While the value of $\sigma_{LN}$ and $\sigma_T$ increases, the D-efficiency decreases, but it is always over 96%, and increases the value of $\phi^*_B$. The D-efficiencies for design (7) with respect to the Bayesian D-optimal designs are always lower than 0.342 for all $\sigma_{LN} \in (0, 5]$ and lower than 0.353 for all $\sigma_T \in (0, 5]$.

6.2 Sequential Designs

For the computation of a sequential design it is necessary to start by choosing an initial design, then perform the experiments, and estimate the value of the vector of unknown parameters using maximum likelihood estimation. Considering the initial design (6) with the data from Janda et al. (2008), $\hat{\theta} = (13.7, 1 \times 10^{10}, 9.5)$. Thus, the new point, $\phi_3$, to be added to the design maximizes

$$
\phi_3 = \arg \max_{\phi \in \lambda_0} \frac{3}{4} M(\xi^*_D, \hat{\theta}) + \frac{1}{4} I(\phi, \hat{\theta}),
$$

and corresponds in this case to the upper extreme of $\lambda_0$. The D-efficiency of the new design with respect to the D-optimal design, (6), is 94.5%. Let consider now the initial design $\xi_8$, then $\hat{\theta} = (11.79, 1 \times 10^{10}, 8.1)$. The new point, $\phi_9$, maximizes

$$
\phi_9 = \arg \max_{\phi \in \lambda_0} \frac{8}{9} M(\xi_8, \hat{\theta}) + \frac{1}{9} I(\phi, \hat{\theta}),
$$

and corresponds to the lower extreme of $\lambda_0$. The D-efficiency of the new design with respect to the D-optimal design, (6), is 34.3%.

The conditions stated by Chaudhury and Mykland (1993) for convergence are satisfied here.

7. DISCUSSION

In this work, a practical physical problem has been considered. The experimenters are interested in the law that governs the process of jam formation during the discharge by gravity of granular material stored in a two-dimensional silo. The experimenters assume a critical outlet diameter such that jamming is known to be impossible above that critical point.

Prior to our work, the process of choosing diameters to run the experiments has been carried out without any rigorous rule. In this article, we propose optimal experimental designs using the D-optimality criterion and a compound criterion with a cost function. Those designs depend on the values of the
parameters. Locally optimal designs have been computed for nominal values of the parameters. In addition, prior knowledge of the possible values of the parameter $\phi_C$ has been considered to search for optimal experimental designs under a Bayesian approach. A sequential design approach to deal with the problem of dependence on parameters has also been proposed and studied. Zhu, Dasgupta, and Huang (2014) proposed a Bayesian sequential algorithm for obtaining D-optimal designs and a geometric approach to compute locally D-optimal designs.

Parameter $\phi_C$ is of special interest for further application since the practitioners are concerned with choosing the right diameter so as to guarantee there will not be a jam within a certain period of time. This is especially important when there is a high cost in breaking a jam, as is the case for the mining industry. After a sensitivity analysis, we can conclude that from the point of view of the experimental design it is better to underestimate the value of the true parameter $\phi_C$.

Another important issue is the correct estimation of the unknown parameters of the model. The experimenters have previously used least-squares estimation although it is known that the data come from an exponential distribution. In this work, we propose the use of maximum likelihood estimation for estimating the unknown parameters. The optimal designs are given for this setup. Simulations show that the variance of the MLE is lower than the variance of the LSE.

Another interesting point to note is that the response may be measured in two different ways. On one hand, the time elapsed between two successive jamming events. This fits the assumed exponential distribution used in the article. An alternative way to measure the response is to count the number of beads that fall between two successive jamming events, which can be approximated by the total weight of material falling between two successive jamming events. Even more than that, just the weight of them is taken as a measurement of this “between two successive jamming events.” Thus, a Poisson distribution may be used instead. At this point a detailed study of the statistical technique of D-optimal experimental design applied to a Poisson model describing the elimination of a radioactive substance in the human body can be found in López-Fidalgo and Villarroel (2007).

The process of jam formation in a two-dimensional silo has been studied widely and several models have been proposed in practice. A comparative study between these models and our model will be considered in future research.

Censoring the time between two successive jamming events may be considered if a particular diameter takes too long to produce a jam. Some limiting time can be fixed to shorten waiting times, and so costs, in the experimentation. This should be done both for estimating the parameters and for finding optimal designs.

SUPPLEMENTARY MATERIALS

Proof: A detailed proof of Theorem 1. (theorem1proof, pdf file)

Simulations: Results of the simulations of Section 5.2. (simulation. pdf file)

ACKNOWLEDGMENTS

The authors were supported by Ministerio de Economía y Competitividad and Fondos FEDER MTM2010–20774-C03-01 and MTM2013–47879-C2-1-P and Junta de Comunidades de Castilla la Mancha PII10–0291-1850. The authors thank the associate editor and the referees for their comments, which have contributed very much to improve the article.

[Received September 2013. Revised April 2015.]

REFERENCES

Atkinson, A. C., Donev, A. N., and Tobias, R. D. (2007), Optimum Experimental Designs, With SAS, New York: Oxford University Press. [270,271]

Chaloner, K., and Verdinelli, I. (1995), “Bayesian Experimental Design: A Review,” Statistical Science, 10, 273–304. [271,274]

Chaudhuri, P., and Mykland, I. A. (1993), “Nonlinear Experiments: Optimal Design and Inference Based on Likelihood,” Journal of the American Statistical Association, 88, 538–546. [271,274,275]

Fedorov, V. V. (1972), “Theory of Optimal Experiments,” Technical report, Department of Statistics, Purdue University. [270,271]

Janda, A., Zuriguel, I., García-Martín, A., Pugnaloni, L. A., and Maza, D. (2008), “Jamming and Critical Outlet Size in the Discharge of a Two-Dimensional Silo,” Europhysics Letters, 84, 44002. [269,272,274,275]

Lindley, D. V., and Smith, A. F. M. (1972), “Bayes Estimates for the Linear Model,” Journal of the Royal Statistical Society, Series B, 34, 1–41. [271]

López-Fidalgo, J., and Rodríguez Díaz, J. M. (2004), “Elfving’s Method for m-Dimensional Models,” Metrika, 59, 235–244. [272]

López-Fidalgo, J., and Villarroel, J. (2007), “Optimal Designs for Radiation Retention With Poisson Correlated Response,” Statistics in Medicine, 26, 1999–2016. [276]

Pázman, A. (1986), Foundations of Optimum Experimental Design, Dordrecht: D. Reidel. [271]

Pukelsheim, F. (1993), Optimal Design of Experiments, New York: Wiley. [271]

Zhu, L., Dasgupta, T., and Huang, Q. (2014), “A D-Optimal Design for Estimation of Parameters of an Exponential-Linear Growth Curve of Nanostructures,” Technometrics, 56, 432–442. [276]