

Article

# The Algorithm of Gu and Eisenstat and D-Optimal Design of Experiments

Alistair Forbes 

National Physical Laboratory, Teddington TW11 0LW, UK; alistair.forbes@npl.co.uk

**Abstract:** This paper addresses the following problem: given  $m$  potential observations to determine  $n$  parameters,  $m > n$ , what is the best choice of  $n$  observations. The problem can be formulated as finding the  $n \times n$  submatrix of the complete  $m \times n$  observation matrix that has maximum determinant. An algorithm by Gu and Eisenstat for determining a strongly rank-revealing QR factorisation of a matrix can be adapted to address this latter formulation. The algorithm starts with an initial selection of  $n$  rows of the observation matrix and then performs a sequence of row interchanges, with the determinant of the current submatrix strictly increasing at each step until no further improvement can be made. The algorithm implements rank-one updating strategies, which leads to a compact and efficient algorithm. The algorithm does not necessarily determine the global optimum but provides a practical approach to designing an effective measurement strategy. In this paper, we describe how the Gu–Eisenstat algorithm can be adapted to address the problem of optimal experimental design and used with the QR algorithm with column pivoting to provide effective designs. We also describe implementations of sequential algorithms to add further measurements that optimise the information gain at each step. We illustrate performance on several metrology examples.

**Keywords:** design of experiment; D-optimality; metrology; QR factorisation; uncertainty evaluation



**Citation:** Forbes, A. The Algorithm of Gu and Eisenstat and D-Optimal Design of Experiments. *Algorithms* **2024**, *17*, 193. <https://doi.org/10.3390/a17050193>

Academic Editors: Dunhui Xiao and Shuai Li

Received: 27 March 2024

Revised: 25 April 2024

Accepted: 30 April 2024

Published: 2 May 2024



**Copyright:** © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

While machine learning paradigms are based on access to large sets of data to extract the information of interest, many practical problems in science and technology arise in a context in which data are sparse and potentially expensive to generate. Given limited experimental resources, it is important that these resources are used effectively. The design of experiment (DoE) is an important area of study in applied mathematics and statistics. Many DoE studies relate to problems in which a response variable of interest, e.g., the yield of a crop, is possibly influenced by a number of control or stimulus variables, e.g., crop variety, soil type, irrigation regime, fertiliser treatment, etc. The form of the response function is usually unknown and the DoE problem is to design a number of experiments with the control variables set to different values, in order to estimate the response function from a set of candidate response functions and consequently choose control values that optimise the response; see, e.g., [1–8]. The DoE problem has been addressed in applications across many scientific disciplines, including the placement of sensors in a electrical power grid [9,10] and the measurement of thermal properties of materials [11,12], for example.

In this paper, we look at a particular aspect of the design problem: given a known response model, how do we choose a minimal number of measurements, taking into account the associated uncertainties, in order to estimate all the parameters of the system under study as accurately as possible. Thus, the problem can be posed as: given  $m$  potential observations to determine  $n$  parameters,  $m > n$ , what is the best choice of  $n$  observations. In the context of (linear) least squares estimation, a version of the problem can be formulated as finding the  $n \times n$  submatrix of the complete  $m \times n$  observation matrix that has maximum determinant. The problem can be posed as an example of the following general optimisation problem. Suppose  $A = \{a_i, i = 1, \dots, m\}$  is a set with  $m$  elements. Represent the set  $P(A)$

of all subsets of  $A$  as the set of characteristic functions  $\chi : A \rightarrow \{0, 1\}$ : if  $B \subset A$ , then the corresponding function  $\chi_B \in P(A)$  is such that  $\chi_B(a_i) = 1$  if  $a_i \in B$  and is 0 otherwise. Let  $D : P(A) \rightarrow \mathbb{R}$  be a real valued function defined on  $P(A)$ . The optimisation problem is

$$\min_{\chi \in P(A)} D(\chi) \quad \text{subject to} \quad \sum_{i=1}^m \chi(a_i) = n. \tag{1}$$

In theory, the global solution to this DoE problem can be found by a brute force assessment of all choices of subsets, a total of  $m! / (n!(m - n)!)$  possibilities. For all but small  $m$  and  $n$ , this approach will be computationally infeasible, even if the functional  $D$  is cheap to evaluate. Two other related approaches are commonly used to partially solve this type of combinatorial optimisation problem.

The first approach, with a long history, see e.g., [13–15], is to use a sequential approach that at each stage constructs a new candidate subset by removing one element from the subset, replacing it by a new element that leads to a strict improvement in  $D$ . The computational cost per step for such an algorithm will be given by the cost evaluating  $D$  and the cost of searching for a suitable exchange from the  $n(m - n)$  possibilities. Ideally, the computational cost of an exchange should be no worse than  $O(m)$  or  $O(m \log m)$  as a function of  $m$ . The second approach is based on convex relaxation, replacing a combinatorial problem by a convex optimisation problem [16]. Let  $0 \leq w_i = w(a_i) \leq 1$ , where we think of  $w_i$  as specifying the degree to which the  $i$ th element is in the solution optimal subset. If  $W(A)$  is the set of functions  $A \rightarrow [0, 1]$ , then the optimisation problem is defined by

$$\min_{w \in W(A)} D(w) \quad \text{subject to} \quad \sum_{i=1}^m w(a_i) = n. \tag{2}$$

In this approach, the combinatorial constraint  $\chi(a_i) \in \{0, 1\}$  is relaxed to the convex constraint  $0 \leq w(a_i) \leq 1$ . Often, the optimal  $w$  is in a fact a characteristic function associated with an  $n$ -element subset. The optimisation algorithms for such problems are generally iterative and often also have a sequential element, such as in sequential quadratic programming [17] and semidefinite programming [18]. Convex relaxation approaches to DoE problems were considered in [4,9,10,16], for example.

In this paper, we show how two algorithms, both following a sequential approach and developed for numerical linear algebra applications, can be repurposed to address the design problem of choosing  $n$  observations from  $m$  possible observations.

The remainder of this paper is organised as follows. In Section 2, we give an overview of linear least squares estimation, while in Section 3, we describe commonly used aggregate measures of uncertainty and the corresponding criteria used in optimal experimental design. In Section 4, we describe two algorithms that address the DoE problem under consideration, as well as sequential algorithms for choosing additional observations in a step-wise optimal way. Example applications of the algorithms to design problems arising in metrology [19–22] are given in Section 5. Our concluding remarks are given in Section 6.

## 2. Least Squares Problems

For  $m \times n$  observation matrix  $C$ ,  $m \times 1$  vector of measurements  $y$ , the least squares estimate  $\alpha$  is the solution of

$$\min_{\alpha} (y - C\alpha)^T (y - C\alpha).$$

The QR factorisation of  $C$  expresses  $C = QR$  as a product of an  $m \times m$  orthogonal matrix  $Q$  and an upper triangular matrix  $R$ . If

$$C = [Q_1 \ Q_2] \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} = Q_1 R_1,$$

where  $Q_1$  is the submatrix constructed from the first  $n$  columns of  $Q$  and  $Q_2$  that is constructed from the remaining  $m - n$  columns, then

$$\begin{aligned} \mathbf{a} &= (\mathbf{C}^\top \mathbf{C})^{-1} \mathbf{C}^\top \mathbf{y} = \mathbf{R}_1^{-1} \mathbf{Q}_1^\top \mathbf{y}, \\ \hat{\mathbf{y}} &= \mathbf{C} \mathbf{a} = \mathbf{C} (\mathbf{C}^\top \mathbf{C})^{-1} \mathbf{C}^\top \mathbf{y} = \mathbf{Q}_1 \mathbf{Q}_1^\top \mathbf{y}, \\ \mathbf{r} &= \mathbf{y} - \mathbf{C} \mathbf{a} = (\mathbf{I} - \mathbf{Q}_1 \mathbf{Q}_1^\top) \mathbf{y} = \mathbf{Q}_2 \mathbf{Q}_2^\top \mathbf{y}, \end{aligned}$$

are the least squares solution, model approximant, and associated residual vector. The QR factorisation enables  $\mathbf{a}$  to be given by the solution of  $\mathbf{R}_1 \mathbf{a} = \mathbf{Q}_1^\top \mathbf{y}$  and this solution can be constructed using back substitution, solving for  $a_n$  first, then  $a_{n-1}$ , etc., exploiting the upper-triangular form of  $\mathbf{R}_1$ . The QR factorisation also avoids the numerical loss of accuracy that can occur in forming  $\mathbf{C}^\top \mathbf{C}$  and its inverse in the case in which  $\mathbf{C}$  is ill-conditioned [23] (Chapter 5). The computational cost of evaluation the least-squares solution in  $O(mn^2)$ .

The  $n \times m$  matrix  $\mathbf{S} = \mathbf{R}_1^{-1} \mathbf{Q}_1^\top$  is the sensitivity matrix of the solution parameters  $\mathbf{a}$  with respect to the data  $\mathbf{y}$  with  $s_{ji} = \partial a_j / \partial y_i$ . If the variance matrix associated with  $\mathbf{y}$  is  $\mathbf{V}_y$ , then the variance matrix associated with the solution  $\mathbf{a}$  is

$$\mathbf{V}_a = \mathbf{S} \mathbf{V}_y \mathbf{S}^\top.$$

For the case  $\mathbf{V}_y = \mathbf{I}$ ,

$$\mathbf{V}_a = \mathbf{S} \mathbf{S}^\top = (\mathbf{R}_1^\top \mathbf{R}_1)^{-1} = (\mathbf{C}^\top \mathbf{C})^{-1}.$$

For this case, the trace of the variance matrix is given by

$$\text{Tr}(\mathbf{V}_a) = \sum_j \sum_i s_{ji}^2,$$

the variance associated with  $a_j$  is  $u^2(a_j) = \sum_i s_{ji}^2$  and  $\sum_j s_{ji}^2$  is the contribution of the uncertainty associated with the  $i$ th observation to the trace of  $\mathbf{V}_a$ .

The least-squares solution  $\mathbf{a}$  is the maximum likelihood estimate of  $\boldsymbol{\alpha}$  for the model  $\mathbf{y} \in \mathcal{N}(\mathbf{C}\boldsymbol{\alpha}, \mathbf{I})$ . If  $\mathbf{y}$  is associated with the variance matrix  $\mathbf{V}_y$ , i.e.,  $\mathbf{y} \in \mathcal{N}(\mathbf{C}\boldsymbol{\alpha}, \mathbf{V}_y)$ , and  $\mathbf{V}_y$  is full rank and has Cholesky decomposition  $\mathbf{V}_y = \mathbf{L}\mathbf{L}^\top$ , where  $\mathbf{L}$  is lower triangular [23], then the maximum likelihood estimate  $\mathbf{a}$  of  $\boldsymbol{\alpha}$  solves the Gauss–Markov problem

$$\min_{\boldsymbol{\alpha}} (\mathbf{y} - \mathbf{C}\boldsymbol{\alpha})^\top \mathbf{V}_y^{-1} (\mathbf{y} - \mathbf{C}\boldsymbol{\alpha}). \tag{3}$$

If

$$\tilde{\mathbf{C}} = \mathbf{L}^{-1} \mathbf{C}, \quad \tilde{\mathbf{y}} = \mathbf{L}^{-1} \mathbf{y},$$

then the solution of (3) is the least-squares solution of  $\tilde{\mathbf{C}}\boldsymbol{\alpha} \approx \tilde{\mathbf{y}}$ :

$$\mathbf{a} = (\tilde{\mathbf{C}}^\top \tilde{\mathbf{C}})^{-1} \tilde{\mathbf{C}}^\top \tilde{\mathbf{y}}.$$

In particular, if  $\mathbf{V}_y$  is a diagonal matrix with  $u_i^2 > 0$  in the  $i$ th diagonal element, then  $\tilde{\mathbf{C}}$  is constructed from the weighted rows of  $\mathbf{C}$ , and similarly for  $\tilde{\mathbf{y}}$ :

$$\tilde{c}_i = w_i c_i, \quad \tilde{y}_i = w_i y_i, \quad w_i = 1/u_i.$$

In general, the cost of evaluating the Cholesky factorisation of  $\mathbf{V}_y$  is  $O(m^3)$ , which could be problematic for large  $m$ . Often,  $\mathbf{V}_y$  has a structure that enables the factorisation to be performed in  $O(m)$ , as in the simple but common case of a diagonal variance matrix.

### 3. Aggregate Measures of Uncertainty

A central issue in design of experiment analysis is to maximise some measure of the information gain subject to constraints on resources. In a metrology context, in particular, it is natural to express the information gain in terms of uncertainties associated with the fitted parameters; i.e., related to the  $n \times n$  variance matrix  $V_{\mathbf{a}}$ . Given an aggregate measure of uncertainty derived from  $V_{\mathbf{a}}$ , one experiment will be judged to be more informative than another if it is associated with a lower aggregate uncertainty. There are a number of aggregate measures that are commonly used. In this paper, we are concerned with two:  $A$ -optimality and  $D$ -optimality.

A full rank variance matrix  $V$  is a symmetric positive definite matrix and has an eigenvalue decomposition with eigenvalues  $\lambda_j, j = 1, \dots, n$ . The  $A$ -measure is the trace of  $\text{Tr}(V)$  of  $V$ , the sum of the diagonal elements of  $V$ , as well as the sum of the eigenvalues of  $V$ :

$$\text{Tr}(V) = \sum_{j=1}^n \lambda_j.$$

We note that  $\text{Tr}(\sigma^2 V) = \sigma^2 \text{Tr}(V)$ . The  $D$ -measure is the determinant  $|V|$  of  $V$ , and also the product of the eigenvalues of  $V$ :

$$|V| = \prod_{j=1}^n \lambda_j.$$

We also evaluate

$$\bar{d}(V) = |V|^{1/n} = \left( \prod_{j=1}^n \lambda_j \right)^{1/n}, \tag{4}$$

the geometric mean of the eigenvalues, noting that  $\bar{d}(\sigma^2 V) = \sigma^2 \bar{d}(V)$ .

Arguments can be made for and against using each of these two measures. Given  $V$ , the  $A$ -measure can be evaluated in  $O(n)$  steps, while the  $D$ -measure will usually involve  $O(n^3)$  steps, using the Cholesky factorisation of the variance matrix to evaluate its determinant. The  $A$ -measure has the apparent advantage of being most easily interpreted as an aggregate uncertainty, being the sum of the variances  $\sum_j u^2(a_j)$  associated with each of the parameter estimates  $a_j$  or, more precisely, the numerical values of these estimates. From a dimensional analysis point of view [24], the  $A$ -measure can only be meaningful if all the parameters in  $\mathbf{a}$  have the same dimension, length, mass, etc. Otherwise, the trace involves adding quantities of different dimensions. The  $D$ -measure is consistent with a dimensional analysis.

The  $A$ -measure can also be regarded as the variance associated with the sum  $\sum_j a_j$  of parameter estimates, *assuming* these estimates as statistically independent. In general, they will not be independent, and one disadvantage of the  $A$ -measure is that it takes no account of correlation, being defined solely in terms of the diagonal elements of  $V$ . The  $D$ -measure does take into account correlation. Consider the  $2 \times 2$  variance matrix

$$V = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad \text{Tr}(V) = 2, \quad |V| = 1 - \rho^2,$$

associated with parameters  $a_1$  and  $a_2$ . The  $A$ -measure is independent of the correlation coefficient  $\rho$  while the  $D$ -measure depends directly on  $\rho^2$ . If the result of an experiment is to change  $\rho$  from 0 to near 1, there is a gain of information, since  $a_1 - a_2$  is now much more accurately estimated. The  $D$ -measure reflects this information gain, the  $A$ -measure does not.

The  $D$ -measure is invariant to re-parametrizations of the model, while the  $A$ -measure is not. Suppose  $V_{\mathbf{a}} = (C^T C)^{-1}$  and let  $\mathbf{b} = B\mathbf{a}$ , where  $B$  is  $n \times n$  and full rank, so that  $\mathbf{b}$  can

be regarded as an alternative parametrization of the problem. For the same experimental design, the variance matrix associated with the estimates  $\mathbf{b}$  is given by

$$V_{\mathbf{b}} = BV_{\mathbf{a}}B^{\top}.$$

We have

$$|V_{\mathbf{b}}| = |BV_{\mathbf{a}}B^{\top}| = |B|^2|V_{\mathbf{a}}|,$$

so that a strategy that is optimal for the D-measure for parametrization  $\mathbf{a}$  will also be optimal for parametrization  $\mathbf{b}$ . For the A-measure,

$$\text{Tr}(V_{\mathbf{b}}) = \text{Tr}(BV_{\mathbf{a}}B^{\top}) = \text{Tr}(B^{\top}BV_{\mathbf{a}}) = \text{Tr}(V_{\mathbf{a}}B^{\top}B),$$

and, in general, there is no constant  $K$  such that  $\text{Tr}(BV_{\mathbf{a}}B^{\top}) = K\text{Tr}(V_{\mathbf{a}})$ , unless  $B^{\top}B$  is a multiple of the identity matrix. In a similar vein, the D-measure is invariant with respect to changes in unit, the A-measure is not.

The eigenvalues  $\lambda_j$  of  $V_{\mathbf{a}}$  are related to the size of the confidence ellipses centred about the parameter estimates  $\mathbf{a}$ . In fact, for a given confidence level, the lengths of the semi-axes are proportional to  $\sqrt{\lambda_j}$ . Hence, the D-measure is related to the square of the volume of the confidence ellipse, while the A-measure is related to the sum of the squares of the semi-axis lengths.

#### 4. Choosing a D-Optimal Subset of $n$ Measurements

Let  $C$  be an  $m \times n$  observation matrix,  $m > n$ , representing  $m$  measurements that could potentially be made. If we are constrained only to use  $n$  measurements, which  $n$  should we choose, in order to minimise an aggregate measure of the uncertainties associated with the fitted parameters? It is assumed that there is at least one  $n \times n$  submatrix constructed from the rows of  $C$  that is full rank, allowing all the parameters  $\mathbf{a}$  to be estimated. In this section, we consider two algorithms described below in Sections 4.1 and 4.2, for addressing the issue of subset selection with a view to providing a D-optimal solution. Both algorithms were designed to address other problems but can be repurposed for design of experiment applications.

Let  $C_k$  be a full rank  $n \times n$  submatrix of  $C$  defined by a subset of the rows of  $C$  indexed by row indices  $I_k = \{i_1, \dots, i_n\}$ , and set  $V_k = H_k^{-1}$ , where  $H = C_k^{\top}C_k$ . Then,

$$|V_k| = |H_k|^{-1} = |C_k|^{-2},$$

so that choosing  $I_k$  to minimise  $|V_k|$  amounts to choosing  $I_k$  to maximise  $|C_k|$ . For any square  $n \times n$  matrix  $A$  with rows  $\mathbf{a}_i$ ,  $|A|$  is the volume of the parallelepiped spanned by the vectors  $\mathbf{a}_i$  in  $\mathbb{R}^n$ . This volume will depend on the Euclidean lengths  $\|\mathbf{a}_i\|_2$  of the vectors and also their mutual independence. For vectors of equal length,  $|A|$  is maximised if the vectors  $\mathbf{a}_i$  are orthogonal to each other (and can therefore be rotated or reflected to align with the coordinate axes). In terms of the observation matrix  $C$ , the lengths of the row vectors  $\mathbf{c}_i$  reflect the accuracy associated with the measurement  $y_i$ —more accurate measurements will be accorded a greater weight and therefore larger row vector  $\mathbf{c}_i$  in terms of Euclidean length. The orthogonality of the row vectors relates to how independent the sets of information represented by the measurements are from each other. Consider

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = C \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}, \quad C = C(\theta) = \begin{bmatrix} 1 & 0 \\ \cos \theta & \sin \theta \end{bmatrix} \quad \epsilon_1, \epsilon_2 \in \mathcal{N}(0, 1).$$

Both measurements  $y_1$  and  $y_2$  are associated with the equal information gain, since  $\|\mathbf{c}_1\| = \|\mathbf{c}_2\| = 1$ . The data point  $y_1$  provides information about  $a_1$  alone, while the data point  $y_2$  potentially provides information about both  $a_1$  and  $a_2$ . For  $\theta \approx 0$ , the observation equation involving  $y_2$  only serves to provide more information about  $a_1$ , and  $a_2$  remains poorly

estimated with  $|C(\theta)| = |\sin \theta| \approx 0$ . For  $\theta \approx \pi/2$ , the information provided by  $y_2$  provides significant new information about  $a_2$  and  $|C(\theta)| = |\sin \theta| \approx 1$ . In this way, amongst a set of measurements of comparable accuracy, we look for subsets that are mutually maximally independent. Both of the algorithms described below involve trying to select a set of measurements that are sufficiently independent.

4.1. Subset Selection Using the QR Factorisation with Column Pivoting

The QR factorisation of an  $m \times n$  matrix  $C = QR$  has the following geometrical interpretation. The column vectors  $c_j$  of  $C$  define an  $n$  dimensional subspace  $\mathcal{C}$  of  $\mathbb{R}^m$ . The orthogonal matrix  $Q$  defines an axis system for  $\mathbb{R}^m$ , such that the  $n$  columns of  $Q_1$  define an axis system for  $\mathcal{C}$  and the  $m - n$  columns of  $Q_2$  define an axis system for the space of vectors  $\mathcal{C}^\perp$  orthogonal to  $\mathcal{C}$ . The columns for  $Q_1$  are constructed so that  $q_1$  is aligned with  $c_1$ , and so that there is an  $r_{11}$  such that  $c_1 = r_{11}q_1$ . The vector  $q_2$  is chosen to lie in the plane defined by  $c_1$  and  $c_2$ , and so there are scalars  $r_{12}$  and  $r_{22}$  such that  $c_2 = r_{12}q_1 + r_{22}q_2$ , etc. This gives the factorisation

$$[c_1 \cdots c_n] = [q_1 \cdots q_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix},$$

i.e., in matrix notation  $C = Q_1R_1$ .

An equivalent geometrical interpretation is that the orthogonal matrix  $Q^\top$ , a combination of rotations and reflections, transforms  $C$  so that its first column is aligned with the first axis in  $\mathbb{R}^m$ , the second column lies in the plane defined by the first and second axes in  $\mathbb{R}^m$ , the third column lies in the three-dimensional space defined by the first three axes in  $\mathbb{R}^m$ , and so on:  $Q^\top C = R$ .

For the standard QR algorithm, the columns of  $C$  are processed in the order they appear in the matrix  $C$ . For the QR algorithm with column pivoting [23], the order of the columns is changed to improve the numerical properties of the algorithm. The column of largest Euclidean length is aligned with the first axis. Of the remaining columns, the column that is furthest from the space defined by the first axis is then rotated to the plane defined by the first and second axes. At the  $k + 1$ th stage, it is the column furthest from the subspace defined by the first  $k$  axes directions that is chosen to be transformed to lie in the space defined by the first  $k + 1$  axes directions. The output of the algorithm includes the specification of a permutation matrix  $P$  that permutes the columns of  $C$ , along with an orthogonal  $Q$  and upper triangular  $R$  such that  $CP = QR$ .

The first algorithm considered here is based on a subset selection algorithm described in [23] (Section 5.5). The issue addressed by the subset selection algorithm is that of trying to explain a measured response vector  $y$  in terms of basis vectors, usually representing covariates  $x$  and functions of covariates, stored in an  $m \times p$  matrix  $C$ . It is assumed that  $y$  lies close to a  $n$  dimensional subspace defined by a subset of the columns of  $C$ . The problem is to estimate  $n$  and choose  $n$  columns of  $C$  that best explain the response, using the QR with column pivoting to select columns that are optimally mutually independent. This information can then be used to predict other, possibly future, responses based on knowledge of the variates represented by the selected columns. The algorithm uses the singular value decomposition (SVD) to estimate  $n$  and the QR algorithm with column pivoting to select an appropriate set of  $n$  columns.

Our design of experiment problem is different. The observation matrix  $C$  of all potential measurements is assumed to be full rank, so estimating the rank of  $C$  is not required. More importantly, our selection problem is to select from the rows of  $C$ , not the columns. However, the selection of a set of vectors that are maximally mutually independent is directly relevant. Suppose  $C$  has QR factorisation  $C = Q_1R_1$ , where  $Q_1$  is  $m \times n$  and  $R_1$  in  $n \times n$  and upper-triangular and full rank. The matrix  $R_1$  can be used

to re-parametrize the problem in terms of  $\mathbf{b} = R_1^{-1}\mathbf{a}$ . The observation associated with parameters  $\mathbf{b}$  is  $Q_1$ . Since D-optimal designs are independent of parametrization, it is sufficient to work with  $Q_1$  as follows. We use column pivoting to find the QR factorisation of  $Q_1^\top$  so that  $Q_1^\top P = UT$ , where  $U$  is an  $n \times n$  orthogonal matrix and  $T$  is upper-triangular. The first  $n$  columns selected by the QR decomposition of  $Q_1^\top$  specified by  $P$  also specify the  $n$  rows of  $C$  that are a good choice for the D-optimal design. This algorithm is summarised below (Algorithm 1). Its computational complexity is  $O(mn^2)$ .

---

**Algorithm 1:** SSQR: subset selection algorithm using QR with column pivoting to determine a favourable subset of an observation matrix in terms of the D-measure

---

**Data:** An  $m \times n$  observation matrix  $C$ .

**Result:** Index set  $I \subset \{i = 1, \dots, m\}$  where the first  $n$  elements specify  $C_n$  that represent a good selection of the rows of  $C$  in terms of maximising the determinant.

**Steps:**

Determine the QR factorization of  $C = Q_1R_1$ , where  $Q_1$  is an  $m \times n$  orthogonal matrix.

Use column pivoting to find the QR factorisation of  $Q_1^\top$  so that  $Q_1^\top P = UT$ , where  $U$  is an  $n \times n$  orthogonal matrix and  $T$  is upper-triangular.

Assign the index  $I$  by assigning  $I(k) = i_k$  where  $i_k$  specifies the row index of the  $k$ th column of  $P$  such that  $P(i_k, k) = 1$ .

---

The SSQR algorithm is not targeted directly at determining a D-optimal design but is an approach that, at least heuristically, has elements that makes it likely to produce a design that is good in terms of the D-measure. The algorithm is an example of a sequential ‘greedy algorithm’, in that at each step, a choice is made that maximises some measure, irrespective of what previous steps have been made, and takes no account of what choices could arise in a future step. Once a row has been selected, there is no opportunity for the row to be deselected at a future iteration. The algorithm is largely independent of the ordering of the rows of the observation matrix  $C$  and, other than for reasons of rounding error effects and symmetries in the observation matrix, will arrive at the same solution for different row orderings. As has been mentioned, there is no guarantee that this solution is globally optimal. Example calculations using the algorithm are discussed in Section 5.

#### 4.2. Exchange Approach Based on the Gu and Eisenstat Algorithm

The second algorithm described in this paper is directly targeted at determining a D-optimal design. It is based on the iterative algorithm of Gu and Eisenstat [25] for determining a strongly rank-revealing QR decomposition. Below, we show how it can be adapted to provide a reduction in the D-measure at each iteration through exchanging rows and to stop when no reduction is possible. It starts with a selection  $C_0$  of  $n$  rows of the  $m \times n$  observation matrix  $C$ , where  $C_0$  is already full rank; e.g., that provided by the SSQR subset selection algorithm of Section 4.1.

We recall that minimising the D-measure is equivalent to maximising the absolute value of the determinant of the upper-triangular factor of the observation matrix. The idea of the algorithm is as follows. Partition  $C^\top$  as  $C^\top = [ A \ B ]$ , where  $A$  is  $n \times n$  and  $B$  is  $n \times (m - n)$ . We assume that  $A = C_0^\top$  is full rank. The Gu–Eisenstat (GE) algorithm uses the following result:

**Proposition 1.** For  $n \times n$  full rank matrix  $A$  and  $n \times p$  matrix  $B$ , let  $\tilde{A}$  be the matrix constructed by replacing the  $i$ th column of  $A$ ,  $1 \leq i \leq n$ , by the  $j$ th column of  $B$ ,  $1 \leq j \leq p$ . Then,

$$|\tilde{A}| = F_{i,j}|A|$$

where  $F = A^{-1}B$ .

**Proof.** For  $n$ -vectors  $u$  and  $v$ ,

$$|A + uv^T| = |A|(1 + v^T A^{-1}u); \tag{5}$$

see Appendix A.2. The result follows for  $v = e_i$ , the  $i$ th column of the identity matrix and  $u = b_j - a_i$ , noting that  $A^{-1}a_i = e_i$ .  $\square$

This suggests the following approach, assuming that the first  $n$  rows of  $C$  are full rank (Algorithm 2).

---

**Algorithm 2:** Basic steps to determine a  $D$ -optimal subset of an observation matrix

---

**Data:** An  $m \times n$  observation matrix  $C$  such that the first  $n$  rows are full rank, and tolerance factor  $f > 1$ .

**Result:** Index set  $I \subset \{i = 1, \dots, m\}$  where the first  $n$  elements specify  $C_n$  such that no row interchange will increase the determinant of  $C_n$  by more than  $f$ .

**Steps:**

Partition  $C^T = [ A \ B ]$  and form  $F = A^{-1}B$ .

Initialise  $m \times 1$  index set  $I$  with  $I(i) = i$ .

**while**  $\max_{i,j} |F_{i,j}| > f$  **do**

1. Choose  $(i, j)$  such that  $|F_{i,j}|$  is maximised.
2. Row interchange: set  $A := A + (b_j - a_i)e_i^T$ ,  $B := B - (b_j - a_i)e_j^T$ .
3. Update  $I$ , swapping the  $i$ th and  $(n + j)$ th elements.
4. Form  $F = A^{-1}B$ .

**end**

---

The main computational element of this algorithm is forming  $F = A^{-1}B$  at each iteration. Forming  $F$  explicitly in this way involves  $O(n^3)$  steps to evaluate  $A^{-1}$  and a further  $O(mn)$  steps to evaluate the matrix product. Let  $Q$  be orthogonal such  $\tilde{A} = Q^T A$  is upper-triangular and set

$$Q^T [ A \ B ] = [ \tilde{A} \ \tilde{B} ]. \tag{6}$$

Then,  $F = A^{-1}B = \tilde{A}^{-1}\tilde{B}$ , and the latter can be computed efficiently using back-substitution. Thus, step 4 in Algorithm 2 can be replaced by

4.1 Determine orthogonal  $Q$  such that  $Q^T A$  is upper-triangular.

4.2 Set

$$[ A \ B ] := Q^T [ A \ B ], \quad F = A^{-1}B.$$

Since  $|Q^T A| = |A|$ , up to sign, the modified algorithm still determines the optimal index set. As described in the Gu and Eisenstat paper, these computations can be made much more efficient by exploiting the fact that at each iteration we are performing a rank-one modification to update an upper triangular matrix in step 2 in Algorithm 2 above. Suppose we have determined to replace  $i$ th column of the upper-triangular matrix  $A$  by the  $j$ th column of  $B$ . First, we move the  $i$ th column of  $A$  to the  $n$ th column, as shown schematically in (7).

$$\begin{array}{cccccccc}
 * & * & i & * & * & * & * & * & * & * & i \\
 & * & i & * & * & * & * & * & * & * & i \\
 & & i & * & * & * & * & * & * & * & i \\
 & & & * & * & * & * & * & * & * & \\
 & & & & * & * & * & * & * & * & \\
 & & & & & * & * & * & * & * & \\
 & & & & & & * & * & * & * & \\
 & & & & & & & * & * & * & \\
 & & & & & & & & * & * & \\
 & & & & & & & & & * & \\
 & & & & & & & & & & *
 \end{array} \longrightarrow \tag{7}$$

This step can be written as

$$\begin{bmatrix} A_1 & \mathbf{a}_i & A_2 \end{bmatrix} \mapsto \begin{bmatrix} A_1 & A_2 & \mathbf{a}_i \end{bmatrix} = A\Pi_{i,n}$$

where  $\Pi_{i,n}$  is the appropriate permutation matrix. The right-hand matrix is upper-triangular except for the sub-diagonal elements stored in  $A_2$ . Let  $Q$  be the orthogonal matrix where  $\tilde{A} = Q^\top A\Pi_{i,n}$  is upper-triangular and set  $\tilde{B} = Q^\top B$ . Note that  $Q^\top$  only acts on rows and columns  $i$  to  $n$  of  $A\Pi_{i,n}$  and rows  $i$  to  $n$  of  $B$ . We have that

$$\tilde{F} = \tilde{A}^{-1}\tilde{B} = \Pi_{i,n}^\top F.$$

We can now replace the  $n$ th column  $\tilde{\mathbf{a}}_n$  of  $\tilde{A}$  by the  $j$ th column  $\tilde{\mathbf{b}}_j$  of  $\tilde{B}$ , forming matrices

$$\tilde{A} = \tilde{A} + \mathbf{u}\mathbf{e}_n^\top, \quad \tilde{B} = \tilde{B} - \mathbf{u}\mathbf{e}_j^\top, \quad \mathbf{u} = \tilde{\mathbf{b}}_j - \tilde{\mathbf{a}}_n.$$

Note that the matrix  $\tilde{A}$  on the left remains upper-triangular. For the next iteration it remains to calculate  $\tilde{F} = \tilde{A}^{-1}\tilde{B}$ . Using the Sherman–Morrison formula [26] (Appendix A.1), we have

$$\tilde{A}^{-1} = (\tilde{A} + \mathbf{u}\mathbf{e}_n^\top)^{-1} = \tilde{A}^{-1} - \frac{\tilde{\mathbf{u}}\tilde{\mathbf{v}}^\top}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}}, \quad \tilde{\mathbf{u}} = \tilde{A}^{-1}\mathbf{u}, \quad \tilde{\mathbf{v}} = \tilde{A}^{-\top}\mathbf{e}_n,$$

so that

$$\begin{aligned} \tilde{A}^{-1}\tilde{B} &= \tilde{A}^{-1}\tilde{B} - \tilde{A}^{-1}\mathbf{u}\mathbf{e}_j^\top - \frac{\tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \tilde{B})}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}} + \frac{\tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \mathbf{u})\mathbf{e}_j^\top}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}}, \\ &= \tilde{F} + \tilde{\mathbf{w}}\mathbf{e}_j^\top, \end{aligned} \tag{8}$$

where

$$\mathbf{w} = \left( \frac{\tilde{\mathbf{v}}^\top \mathbf{u}}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}} - 1 \right) \mathbf{e}_j - \frac{\tilde{B}^\top \tilde{\mathbf{v}}}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}}.$$

Thus,  $\tilde{F}$  is a rank-one update of  $\tilde{F}$ . The expressions above can be considerably simplified. We note that

$$\tilde{\mathbf{u}} = \tilde{F}(:,j) - \mathbf{e}_n, \quad \frac{\tilde{\mathbf{v}}^\top \mathbf{u}}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}} - 1 = -1/\tilde{F}(n,j), \quad \frac{\tilde{B}^\top \tilde{\mathbf{v}}}{1 + \mathbf{e}_n^\top \tilde{\mathbf{u}}} = -\tilde{F}(n,:)^\top / \tilde{F}(n,j),$$

so that

$$\mathbf{w} = -\frac{1}{\tilde{F}(n,j)}(\tilde{F}(n,:)^\top + \mathbf{e}_j).$$

Recalling that  $\tilde{F} = \Pi_{i,n}^\top A^{-1}B$ , we see that from (8),  $\tilde{A}^{-1}\tilde{B}$  can be constructed as a rank-one update of the previously calculated  $F = A^{-1}B$ . We also note that  $\tilde{F}(n,j) = F(i,j) > 1$  (since the update is only invoked to increase  $|A|$ ) so that the rank-one update of  $F$  is numerically stable. The formation of  $\tilde{A}$  and  $\tilde{A}$  requires  $O(n^2)$  steps and the formation of  $\tilde{F}$  requires  $O(mn)$  steps. These calculations lead to a computationally efficient and numerically stable version (Algorithm 3) of Algorithm 2:

**Algorithm 3:** GE: efficient algorithm to determine a  $D$ -optimal subset of an observation matrix

**Data:** An  $m \times n$  observation matrix  $C$  such that the first  $n$  rows are full rank, and tolerance factor  $f > 1$ .

**Result:** Index set  $I \subset \{i = 1, \dots, m\}$  where the first  $n$  elements specify  $C_n = C(I(1:n), :)$  such that no row interchange will increase the determinant of  $C_n$  by more than  $f$ ; matrices  $[A \ B]$  storing the upper-triangular factor of  $C(I, :)^T$ .

**Steps:**

Partition  $C^T = [A \ B]$  and form  $F = A^{-1}B$ .

Initialise  $m \times 1$  index set  $I$  with  $I(i) = i$ .

**while**  $\max_{i,j} |F_{i,j}| > f$  **do**

1. Choose  $(i, j)$  such that  $|F_{i,j}|$  is maximised. and store  $\mathbf{a}_i$  and  $\mathbf{b}_j$ .
2. Apply permutation matrix  $\Pi_{i,n}$  to move the  $i$ th column of  $A$  to the  $n$ th column  $A := A\Pi_{i,n}$ .
3. Apply  $\Pi_{i,n}^T$  to  $F$ :  $F := \Pi_{i,n}F$ .
4. Determine  $(n - i + 1) \times (n - i + 1)$  orthogonal matrix  $Q$  such that  $Q^T A(i:n, i:n-1)$  is upper-triangular and update:

$$A(i:n, i:n) := Q^T A(i:n, i:n), \quad B(i:n, :) := Q^T B(i:n, :).$$

5. Set  $\mathbf{a}_n = A(:, n)$ ,  $\mathbf{b}_j = B(:, j)$ ,  $\tilde{\mathbf{u}} = F(:, j) - \mathbf{e}_n$  and

$$\mathbf{w} = -\frac{1}{F_{n,j}}(F(n, :)^T + \mathbf{e}_j).$$

6. Interchange columns and update  $I$ .  $A(:, n) = \mathbf{b}_j$ ,  $B(:, j) = \mathbf{a}_j$ . Set  $i_n = I(n)$ ,  $i_j = I(n + j)$  and  $I(n) = i_j$ ,  $I(n + j) = i_n$ .
7. Update  $F$ :  $F := F + \tilde{\mathbf{u}}\mathbf{w}^T$ .

**end**

The algorithm is a sequential descent algorithm, with each exchange producing a decrease in the determinant of the variance matrix by at least a factor of  $1/f^2 < 1$ , ensuring that it stops in a finite number of steps. However, there is no guarantee that the algorithm will determine the globally  $D$ -optimal subset, only a locally  $D$ -optimal solution. A different ordering of the rows of the observation matrix is quite likely to produce a different solution. Consider the following matrix  $C(a)$ :

$$C(a) = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & a \\ 0.50 & 0.50 & 0.50 & 0.50 \\ 0.17 & -0.83 & 0.17 & 0.50 \\ 0.17 & 0.17 & -0.83 & 0.50 \\ -0.83 & 0.17 & 0.17 & 0.50 \end{bmatrix}. \tag{9}$$

The bottom four rows form an orthogonal matrix  $Q$  with  $|Q| = 1$ . However, for  $a > 0.5$ , there is no row interchange that will allow the determinant of the first four rows to be increased, and the GE algorithm stops with the first four rows with determinant  $a$ , sub-optimal if  $a < 1.0$ . For  $0.5 < a < 1.0$ , the SSQR algorithm, by contrast, selects the bottom four rows, in this case, the globally optimal solution.

### 4.3. Combined SSQR-GE Algorithm

An effective approach in practice is to combine the SSQR and GE algorithms, using the SSQR algorithm to find an initial selection of columns, and to use the GE interchange algorithm to improve on, if possible, this selection. This combined algorithm has the advantages of being largely independent of the row ordering and will deliver a local minimum based on the SSQR algorithm starting point. Again, there is no guarantee that the solution will be a global optimum, but the examples discussed in Section 5 show that the combined algorithm works well.

### 4.4. Sequential Approach for Choosing a Set of Additional Observations

The algorithms described in Sections 4.1 and 4.2 are designed to determine a minimal set of observations that produce D-optimal estimates of the model parameters  $\mathbf{a}$ . We now consider the case in which estimates of  $\mathbf{a}$  have already been established. Suppose the (full rank) variance matrix  $V = LL^\top = (R^\top R)^{-1}$  represents the current knowledge about parameters  $\mathbf{a}$  and up to  $m$  (new) measurements are potentially available, modelled according to

$$y_i \sim \mathcal{N}(c_i^\top \mathbf{a}, 1), \quad i = 1, \dots, m.$$

(As before, we assume that, for each  $i$ ,  $y_i$  and  $c_i$  have been weighted so that the uncertainty associated with  $y_i$  is 1.) Which measurements should we make in order to best improve our knowledge about  $\mathbf{a}$ ? We suppose that resources exist to make  $p$  more observations. An exhaustive search of all combinations is only feasible for small  $p$  and modest  $m$ . The sequential algorithms below add the observations one at a time, with the choice at each stage designed to maximise the information. These are further examples of so-called ‘greedy algorithms’.

Let  $C$  be the  $m \times n$  observation matrix whose  $i$ th row is  $c_i^\top$  and  $V_i$  the variance matrix associated with  $\mathbf{a}$  given that the  $i$ th measurement (and only the  $i$ th measurement) is made. Then,

$$V_i = (R^\top R + c_i c_i^\top)^{-1} = V - \frac{V c_i (V c_i)^\top}{1 + c_i^\top V c_i},$$

using the Sherman–Morrison formula [26] (Appendix A.1). Let

$$\mathbf{f}_i = V c_i, \quad g_i^2 = c_i^\top \mathbf{f}_i = c_i^\top V c_i. \tag{10}$$

Then,

$$V_i = V - \frac{\mathbf{f}_i \mathbf{f}_i^\top}{1 + g_i^2} = V - \mathbf{u}_i \mathbf{u}_i^\top, \quad \mathbf{u}_i = \frac{1}{\sqrt{1 + g_i^2}} \mathbf{f}_i. \tag{11}$$

so that  $V_i$  is a symmetric rank-one update of  $V$ .

The eigenvalues  $\lambda_{j,i}$  of  $V_i$  can be related to the eigenvalues  $\lambda_j$  of  $V$  through the following result [23] (Section 8.1), [27] (p. 94):

**Proposition 2.** Suppose  $W = V + \rho \mathbf{u} \mathbf{u}^\top$ , where  $V$  is an  $n \times n$  symmetric matrix,  $\mathbf{u}$  is an  $n$ -vector with  $\mathbf{u}^\top \mathbf{u} = 1$  and  $\rho$  is a constant. Suppose  $V$  and  $W$  have eigenvalues

$$\lambda_1(V) \geq \lambda_1(W) \geq \lambda_2(V) \geq \dots \geq \lambda_n(V),$$

and

$$\lambda_1(W) \geq \lambda_1(W) \geq \lambda_2(W) \geq \dots \geq \lambda_n(W).$$

If  $\rho \geq 0$ , then

$$\lambda_i(W) \in [\lambda_i(V), \lambda_{i-1}(V)], \quad i = 2, \dots, n,$$

while if  $\rho \leq 0$ , then

$$\lambda_i(W) \in [\lambda_{i+1}(V), \lambda_i(V)], \quad i = 1, \dots, n - 1.$$

In either case, there exist  $n$  constants  $\gamma_i, \gamma_i \geq 0$ , with  $\gamma_1 + \dots + \gamma_n = 1$ , such that

$$\lambda_i(W) = \lambda_i(V) + \gamma_i \rho.$$

#### 4.4.1. Sequential D-Optimality Algorithm

We use the relationship (5) to write

$$\begin{aligned} |V_i| &= |(R^\top R + \mathbf{c}_i \mathbf{c}_i^\top)^{-1}|, \\ &= |R^\top R|^{-1} \left(1 + \mathbf{c}_i^\top (R^\top R)^{-1} \mathbf{c}_i\right)^{-1}, \\ &= |V| \left(1 + \mathbf{c}_i^\top V \mathbf{c}_i\right)^{-1}, \\ &= \frac{1}{1 + g_i^2} |V|, \quad g_i^2 = \mathbf{c}_i^\top V \mathbf{c}_i. \end{aligned}$$

A sequential algorithm can therefore be designed which at each stage chooses the  $i$  that maximises  $g_i^2$ . The algorithm can be run in two modes: the first in which the algorithm can choose any row at most one time, the second in which there is no limit in the number of times any row can be selected. Having a row repeated  $k$  times is equivalent to having a single more accurate observation with its uncertainty reduced by a factor of  $1/\sqrt{k}$ . In the second mode, it is often the case that additional measurements correspond to repeat measurements chosen from the D-optimal set of  $n$  points. In the algorithm below (Algorithm 4), the mode is controlled by the flag  $i_R$ :

---

**Algorithm 4:** Sequential approach to optimising the D-measure in choosing  $p$  observations from a possible  $m > p$  observations

---

**Data:**

- $C$ , an  $m \times n$  observation matrix  $C, m > n$ ,
- $V$ , an  $n \times n$  full rank variance matrix  $V$  encoding the current state of knowledge of the parameters,
- $p$ , an integer  $1 \leq p < m$  specifying the maximum number of observations to be made, and
- flag  $i_R = 0$  or  $1$ . If  $i_R = 1$ , rows of  $C$  can be selected more than once.

**Result:** A  $p \times 1$  index set  $I \subset \{i = 1, \dots, m\}$ , updated variance matrix  $V$  representing the variance in the parameters as a result of using observations  $C_p = C(I, :)$ , the  $p$  rows of  $C$  specified by  $I$ , and  $p \times 1$   $t$  where  $0 < t_q < 1$ , the reduction in the determinant of  $V$  at each iteration:  $|V_{q+1}| := t_q |V_q|$ . The rows are chosen to minimise  $t_q$  at each iteration.

Set  $F = VC^\top$  and  $n \times m$  matrix  $G$  with  $G_{j,i} = F_{j,i} C_{j,i}$  with  $g_i = \|G(:, i)\|, i = 1, \dots, m$ .

**for**  $q = 1$  **to**  $p$  **do**

1. Evaluate  $\tau_i = 1/(1 + g_i^2), i = 1, \dots, m$ .
2. If  $q > 1$  and  $i_R = 0$ , set  $\tau_i = 2, i = I(1), \dots, I(q - 1)$ .
3. Find  $k$  such that  $\tau_k \leq \tau_i, i = 1, \dots, m$ .
4. Set  $I(q) = k, t_q = \tau_k, \mathbf{u} = F(:, k)/\sqrt{1 + g_k^2}$ , and  $\mathbf{w} = C\mathbf{u}_k$ .
5. Update  $V := V - \mathbf{u}\mathbf{u}^\top, F = F - \mathbf{u}\mathbf{w}^\top$ , and  $g_i = [(g_i - w_i)(g_i + w_i)]^{1/2}, i = 1, \dots, m$ .

**end**

---

For each  $q$ , the only operations required are matrix–vector multiplications and similar, requiring at most  $O(mn)$  steps. This compares with  $O(mn^3)$  steps for an algorithm not exploiting rank-one updates.

#### 4.4.2. Expected Information Gain Calculations

The quantities  $t_q$  calculated at each stage give the reduction in determinant of the variance matrix at the  $q$ th step. The following argument can be used to provide an expected value for  $t_q$ , under the assumption that the rows of  $C$  each provide approximately the same amount of information. Suppose  $C$  consists of rows with a 1 in one column and zeros elsewhere. A D-optimal design would select rows so that the ones are distributed equally amongst the columns. If a total of  $q$  rows are selected, the variance matrix  $V_q$  is such that

$$V_q \approx \frac{n}{q}I, \quad |V_q| \approx \left(\frac{n}{q}\right)^n.$$

From this argument, the expected value of  $t_q, q > n$ , is such that

$$t_q \approx \tilde{t}(q) = \left(\frac{q-1}{q}\right)^n. \tag{12}$$

In terms of the geometric mean  $\bar{d}$  of the eigenvalues of the variance matrices, we have

$$\bar{d}(V_q) = \frac{q-1}{q}\bar{d}(V_{q-1}).$$

For  $q = n + 1$ , we have

$$\frac{1}{2} \geq \left(\frac{n}{n+1}\right)^n \geq \frac{1}{e} \approx 0.37,$$

where the left hand value corresponds to the case  $n = 1$ , and  $1/e$  is the limiting value as  $n \rightarrow \infty$ .

#### 4.4.3. Sequential A-Optimality Algorithm

We can also construct a counterpart of Algorithm 4 for A-optimality. For any  $v$ ,  $\text{Tr}(vv^\top) = v^\top v = \|v\|^2$ . With  $f_i$  and  $g_i$  as in (10), if

$$\tau_i^2 = f_i^2 / (1 + g_i^2), \quad f_i^2 = f_i^\top f_i,$$

then

$$\text{Tr}(V_i) = \text{Tr}(V) - \tau_i^2.$$

This last expression calculates the change in the A-measure for a given choice of one additional measurement. A sequential algorithm can therefore be designed which at each stage chooses the  $i$  that maximises  $\tau_i^2$ . Each stage requires the calculation of the quantities  $f_i^2$  and  $g_i^2, i = 1, \dots, m$ . If  $V$  is the current variance matrix, then the  $f_i^2$  are given by the sum of the squares of row elements of  $F = VC^\top$ . Suppose measurement  $k$  maximises  $\tau_i^2$ . Then, the update of  $F$  is given by

$$F = (V - u_k u_k^\top)C^\top = F - u_k w^\top, \quad w = C u_k,$$

and the updated of  $g_i^2$  is given by  $g_i^2 - w_i^2$ , with the latter evaluated as  $(g_i - w_i)(g_i + w_i)$  for better numerical accuracy. These steps are summarised in (Algorithm 5):

---

**Algorithm 5:** Sequential approach to optimising the A-measure in choosing  $p$  observations from a possible  $m > p$  observations

---

**Data:**

- $C$ , an  $m \times n$  observation matrix  $C$ ,  $m > n$ ,
- $V$ , an  $n \times n$  full rank variance matrix  $V$  encoding the current state of knowledge of the parameters,
- $p$ , an integer  $1 \leq p < m$  specifying the maximum number of observations to be made, and
- flag  $i_R = 0$  or  $1$ . If  $i_R = 1$ , rows of  $C$  can be selected more than once.

**Result:** A  $p \times 1$  index set  $I \subset \{i = 1, \dots, m\}$ , updated variance matrix  $V$  representing the variance in the parameters as a result of using observations  $C_p = C(I, :)$ , the  $p$  rows of  $C$  specified by  $I$ , and  $p \times 1$   $t$  where  $t_q = \tau_q^2$ , the reduction in the trace of  $V$  at the  $q$ th iteration. The rows are chosen to maximise  $\tau_i^2$  at each iteration.

Set  $F = VC^T$ ,  $f_i = \|F(:, i)\|$ ,  $n \times m$  matrix  $G$  with  $G_{j,i} = F_{j,i}C_{j,i}$  with  $g_i = \|G(:, i)\|$ ,  $i = 1, \dots, m$ .

**for**  $q = 1$  **to**  $p$  **do**

1. Evaluate  $\tau_i^2 = f_i^2 / (1 + g_i^2)$ ,  $i = 1, \dots, m$ .
2. If  $q > 1$  and  $i_R = 0$ , set  $\tau_i^2 = 0$ ,  $i = I(1), \dots, I(q - 1)$ .
3. Find  $k$  such that  $\tau_k^2 \geq \tau_i^2$ ,  $i = 1, \dots, m$ .
4. Set  $I(q) = k$ ,  $t_q = \tau_k^2$ ,  $u = F(:, k) / \sqrt{1 + g_k^2}$  and  $w = Cu$ .
5. Update  $V := V - uu^T$ ,  $F = F - uw^T$ ,  $f_i = \|F(:, i)\|$  and  $g_i = [(g_i - w_i)(g_i + w_i)]^{1/2}$ ,  $i = 1, \dots, m$ .

**end**

---

As for Algorithm 4, the computational requirement for each step is  $O(mn)$ .

## 5. Numerical Examples

### 5.1. Polynomial Calibration Curves

Many instrument response functions are modelled as polynomial curves, most commonly, of order 2 (degree 1), i.e., modelling a linear response, but higher orders arise, e.g., in the calibration of platinum resistance thermometers [28], or in the calibration of a stage motion [29] where polynomials of order 10 are involved. (Design problems for problems for polynomial models augmented by Gaussian process models are considered in [4].) Suppose a polynomial response function of order  $n$  is defined on the interval  $[-1, 1]$ , and that it is desired to calibrate the response function using  $n$  calibration points  $x_i \in [-1, 1]$ . Which choice of  $n$  points is D-optimal? For a linear (quadratic) response  $n = 2$  ( $n = 3$ ), it is intuitively clear that the end points  $x = \pm 1$  (and  $x = 0$ ) are optimal. For higher orders, the optimal choice is not obvious. In fact, the D-optimal set [30] for order  $n$  is given by solutions of

$$(1 - x^2)\dot{L}_n(x) = 0, \tag{13}$$

where  $\dot{L}_n(x)$  is the derivative of Legendre polynomial of order  $n$  (degree  $n - 1$ ). These solutions can be approximated by the so-called ‘arcsine’ points

$$x_i = \cos\left(\pi \frac{n - 1 - i}{n - 1}\right), \quad i = 0, 1, \dots, n - 1. \tag{14}$$

Table 1 gives the results of applying the SSQR algorithm and the combined SSQR-GE algorithm to estimate the D-optimal calibration points for the cases  $n = 4, 5, \dots, 11$  choosing from  $m = 2001$  possible points equally spaced in the interval  $[-1, 1]$  in steps of 0.001. For these calculations, we have used Chebyshev polynomial basis functions [31] for numerical

stability (but the optimal choice of points is independent of the choice of basis functions.) The arcsine estimates are also given in the table, along with the D-optimal points derived from (13); see Appendix B. The results show that, in terms of closeness to the D-optimal points, the SSQR algorithm improves on the arcsine estimates, while the GE algorithm improves the SSQR solution and is accurate to three decimals in almost all cases (there is one exception for the case  $n = 11$ ). Table 2 gives the computed geometric mean measure  $\bar{d}(V_a)$  given in (4) for evenly spaced calibration points, the arcsine solutions, the SSQR solutions, and the SSQR-GE solutions given in Table 1. Also shown is the number of GE exchanges undertaken to improve the SSQR solutions. The table reflects the improvements provided by the SSQR and SSQR-GE algorithms. The number of GE exchanges is small compared to  $m = 2001$ . The evenly spaced calibration points are markedly worse than the other solutions for larger  $n$ .

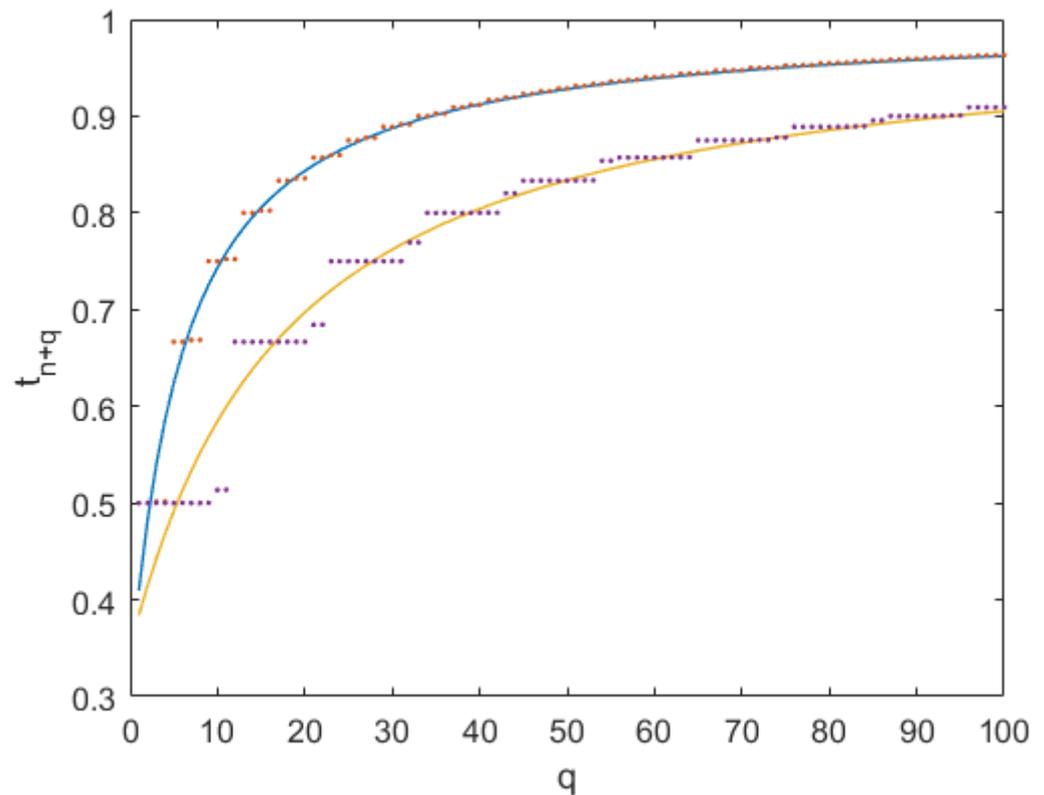
**Table 1.** D-optimal calibration points for polynomial calibration curves of order  $n = 4, \dots, 11$  defined in the interval  $[-1, 1]$ . For each set of four columns, the first column, labelled  $x_0$ , gives the arcsine estimates of the calibration points in (14); the second column, labelled  $x_{QR}$ , gives the optimal the calibration points determined by the SSQR algorithm; the third column, labelled  $x_{GE}$ , gives the estimates determined by the combined SSQR-GE algorithm; while the fourth column, labelled  $x^*$ , gives the optimal the calibration points derived from (13).

$x_0$	$x_{QR}$	$x_{GE}$	$x^*$	$x_0$	$x_{QR}$	$x_{GE}$	$x^*$
$n = 4$				$n = 5$			
-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000
-0.500	-0.488	-0.447	-0.447	-0.707	-0.669	-0.655	-0.655
0.500	0.437	0.447	0.447	0.000	0.006	0.000	0.000
1.000	1.000	1.000	1.000	0.707	0.686	0.655	0.655
				1.000	1.000	1.000	0.000
$n = 6$				$n = 7$			
-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000
-0.809	-0.786	-0.765	-0.765	-0.866	-0.845	-0.830	-0.830
-0.309	-0.286	-0.285	-0.285	-0.500	-0.484	-0.469	-0.469
0.309	0.308	0.285	0.285	0.000	0.002	0.000	0.000
0.809	0.779	0.765	0.765	0.500	0.493	0.469	0.469
1.000	1.000	1.000	1.000	0.866	0.841	0.830	0.830
				1.000	1.000	1.000	0.000
$n = 8$				$n = 9$			
-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000
-0.901	-0.880	-0.872	-0.872	-0.924	-0.908	-0.900	-0.900
-0.623	-0.613	-0.592	-0.592	-0.707	-0.692	-0.677	-0.677
-0.223	-0.211	-0.209	-0.209	-0.383	-0.383	-0.363	-0.363
0.223	0.225	0.210	0.209	0.000	-0.002	0.000	0.000
0.623	0.608	0.592	0.592	0.383	0.376	0.363	0.363
0.901	0.882	0.872	0.872	0.707	0.695	0.677	0.677
1.000	1.000	1.000	1.000	0.924	0.906	0.900	0.900
				1.000	1.000	1.000	0.000
$n = 10$				$n = 11$			
-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000
-0.940	-0.925	-0.920	-0.920	-0.951	-0.938	-0.934	-0.934
-0.766	-0.753	-0.739	-0.739	-0.809	-0.796	-0.784	-0.784
-0.500	-0.493	-0.478	-0.478	-0.588	-0.580	-0.565	-0.565
-0.174	-0.177	-0.165	-0.165	-0.309	-0.311	-0.296	-0.296
0.174	0.168	0.165	0.165	0.000	-0.001	0.000	0.000
0.500	0.497	0.478	0.478	0.309	0.307	0.296	0.296
0.766	0.751	0.739	0.739	0.588	0.582	0.566	0.565
0.940	0.925	0.920	0.920	0.809	0.795	0.785	0.784
1.000	1.000	1.000	1.000	0.951	0.939	0.934	0.934
				1.000	1.000	1.000	0.000

**Table 2.** For  $n = 4, 5, \dots, 11$ , rows two to five give the computed geometric mean measure  $\bar{d}(V_a)$  given in (4) for evenly spaced calibration points, the arcsine solutions, the SSQR solutions, and the SSQR-GE solutions given in Table 1. The final row gives the number of GE exchanges undertaken to improve the SSQR solutions.

$n$	4	5	6	7	8	9	10	11
$\bar{d}_0$	0.4871	0.4152	0.3748	0.3511	0.3379	0.3316	0.3304	0.3332
$\bar{d}_{AS}$	0.4714	0.3789	0.3175	0.2734	0.2403	0.2143	0.1935	0.1763
$\bar{d}_{QR}$	0.4682	0.3746	0.3130	0.2691	0.2362	0.2107	0.1901	0.1733
$\bar{d}_{GE}$	0.4673	0.3735	0.3119	0.2682	0.2354	0.2099	0.1894	0.1726
$n_{GE}$	3	6	7	9	12	17	19	21

Figure 1 graphs the values of  $t_q, q = 1, \dots, 100$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, without repeat measurements, for the cases  $n = 4$ , upper curve, and  $n = 11$ , lower curve, starting with the SSQR-GE solution. The values of  $t_q$  are batched with a significant jump after every  $n$  additional measurement. Without the repeat measurement, the algorithm tends to select available points closest to the D-optimal  $n$  points. If the algorithm is run with repeat measurements allowed, the algorithm selects the same set of D-optimal  $n$  points again and again. This behaviour is reflected in the graphs of  $t_q$  for this case; see Figure 2.



**Figure 1.** Polynomial calibration points: values of  $t_q, q = 1, \dots, 100$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, for the cases  $n = 4$ , upper curve, and  $n = 11$ , lower curve.

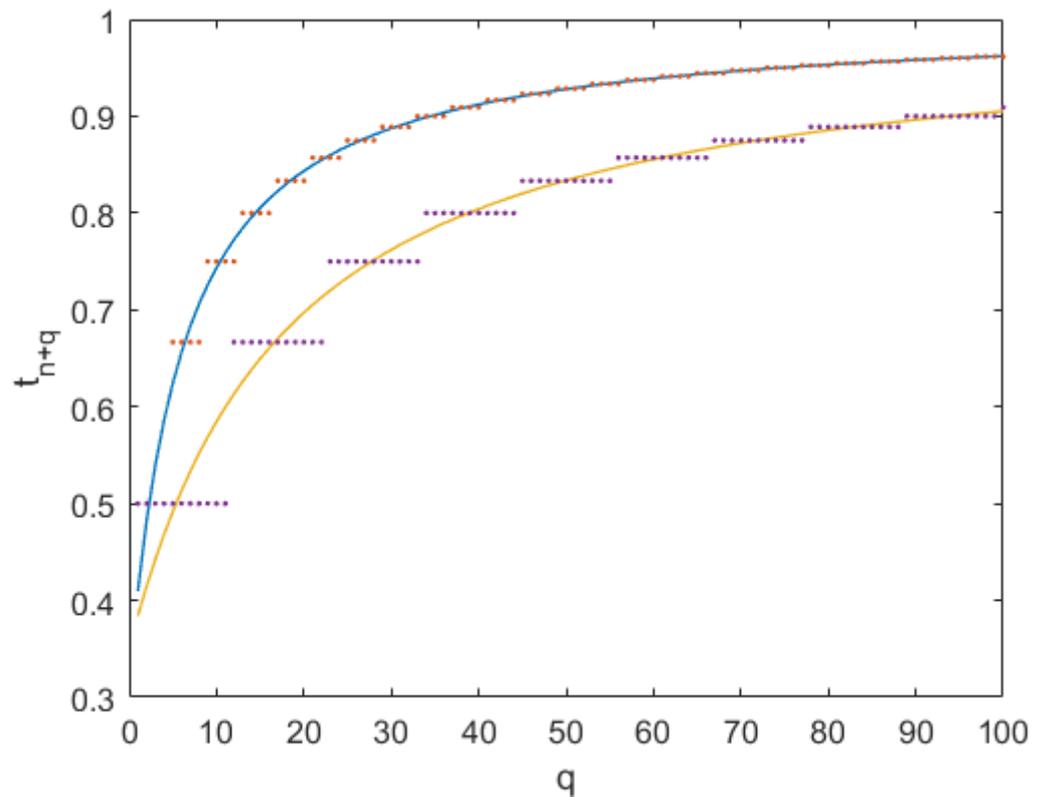


Figure 2. As Figure 1, but for the case where repeated measurements are permitted.

5.2. Tensor Product of Polynomials

Our second example arose from a problem relating to the estimation of heat distribution on an plate [32]. The application involved modelling the heat distribution as a Gaussian process model [33] involving tensor product polynomials, along with spatially correlated effects. A sub-problem is related to choosing a set of calibration points from a  $10 \times 14$  grid of possible sensor locations in order to determine the tensor product mean function. The tensor product model is of the form

$$f(x, y, \mathbf{a}) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} a_{ij} f_i(x) g_j(y),$$

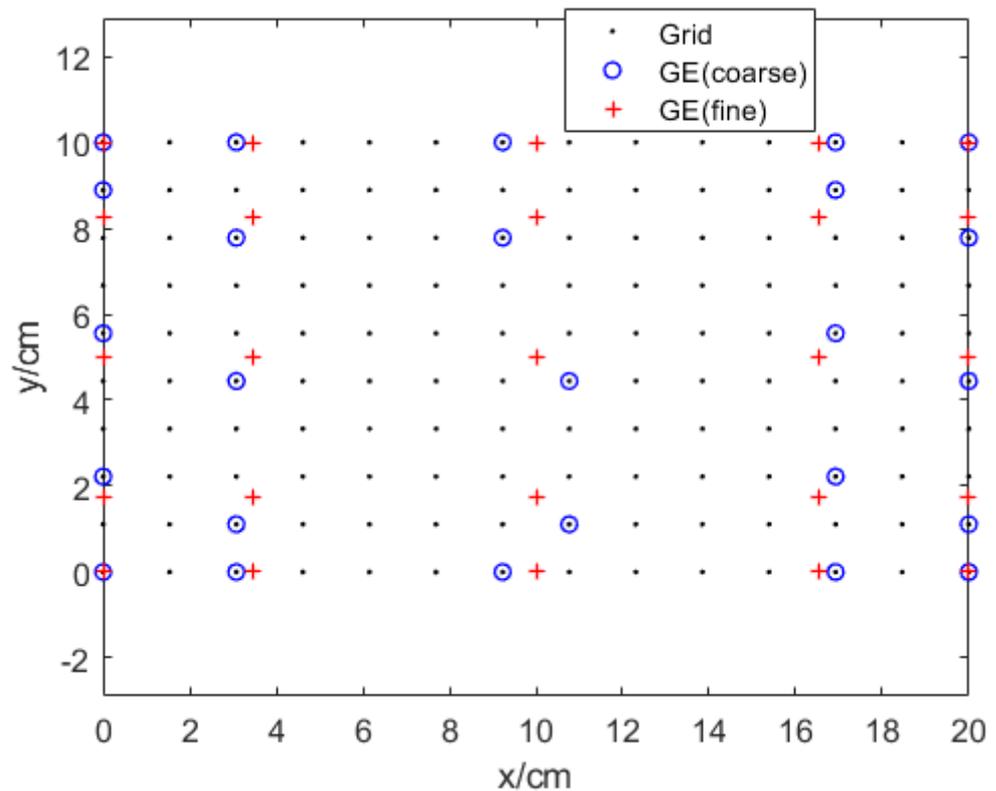
where  $f_i(x)$  and  $g_j(y)$  are Chebyshev polynomial basis functions. Figure 3 shows the solution calibration points for  $n_x = n_y = 5$  constructed from polynomial basis functions up to degree 4. Two solutions are shown: the first in which the calibration points are selected from a  $131 \times 91$  finer grid of 11,921 points, the second in which the calibration points are selected from a  $14 \times 10$  grid of 140 points. The solutions found for the fine grid coincide with a grid  $5 \times 5$  grid of points  $(x_k^*, y_l^*)$ , where the  $x_k^*$ ,  $k = 1, \dots, 5$  are the D-optimal solutions derived from the solution of (13) on the interval  $[0, 20]$  and the  $y_l^*$  are the D-optimal solutions on the interval  $[0, 10]$ . For this arrangement, the observation matrix  $C$  associated with the polynomial fit is the tensor product  $C = C_x \otimes C_y$ . For square, full rank matrices  $A$  and  $B$  of size  $n_A \times n_A$  and  $n_B \times n_B$

$$|A \otimes B| = |A|^{n_B} |B|^{n_A}.$$

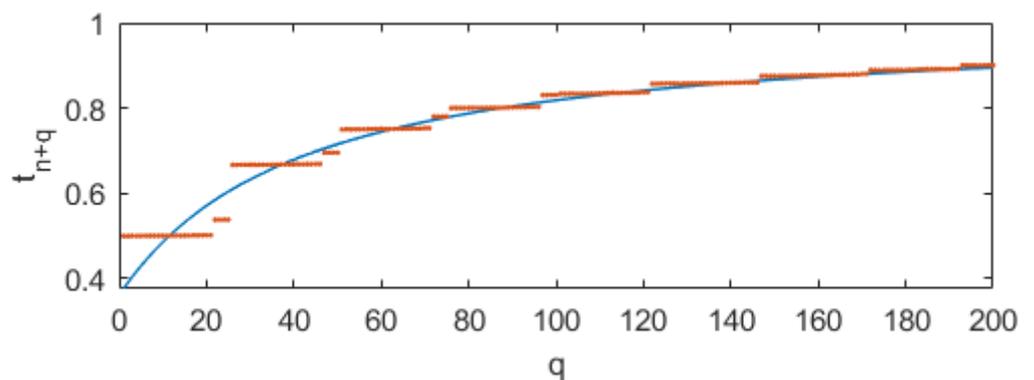
Hence, it is consistent that the D-optimal points for the tensor product are constructed from the D-optimal points for the  $x$ - and  $y$ -basis functions. The solution points for the coarse grid are seen to approximate the solution for the fine grid, given the constraints imposed by the coarse grid. The coarse grid solution points do not represent a regular grid in this case.

For the case of the fine grid, the GE algorithm took 16 iterations to reach a local optimum, starting with the SSQR solution. For the coarse grid, the GE algorithm confirmed that the SSQR solution is locally optimal.

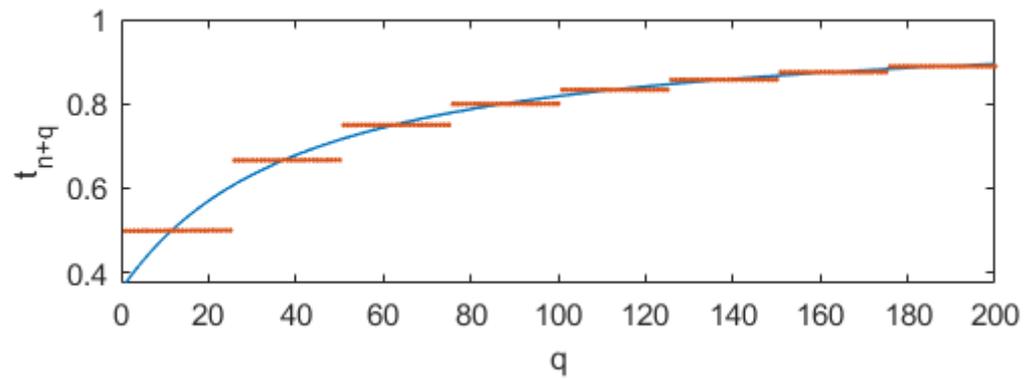
Figures 4 and 5 plot the values of  $t_q, q = 1, \dots, 200$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, for the fine and coarse grids, respectively, with and without repeat measurements. For a coarse grid and no repeat measurements, the additional calibration points that are permitted consistently provide less information compared to the expected information gain given by  $\tilde{t}$  in (12). For a coarse grid of just 140 points, later additional measurements have to be selected from locations that are less informative in terms of improving the D-measure. With repeat measurements on the coarse grid, the actual information gain is in line with the expected information gain.



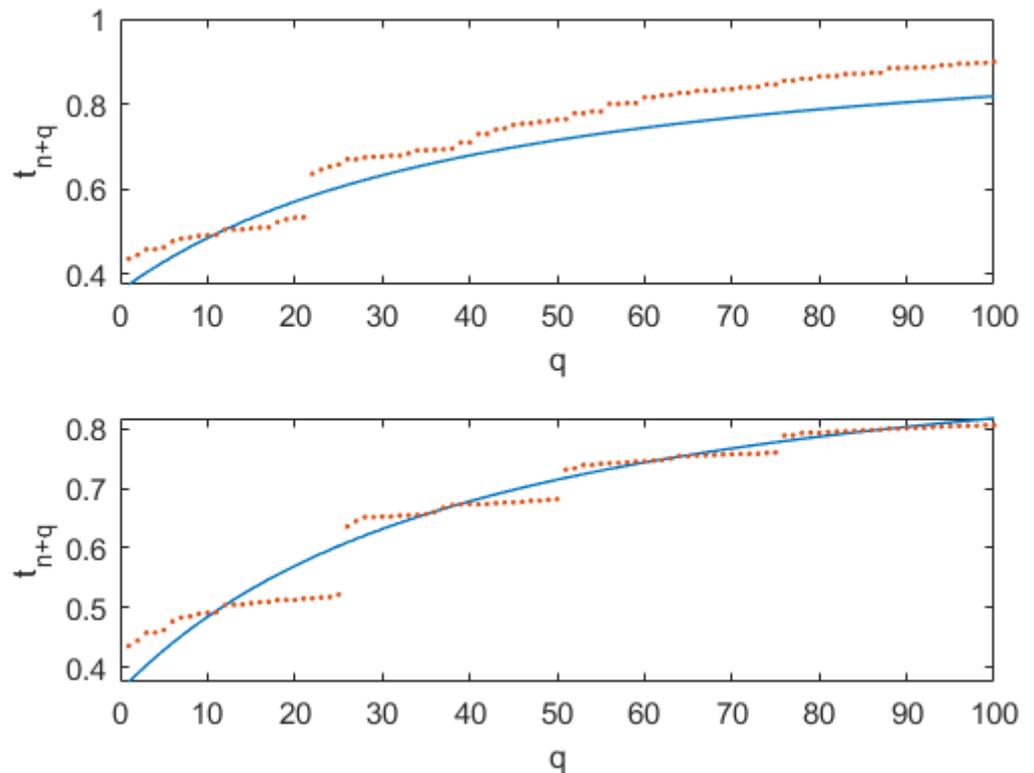
**Figure 3.** Tensor product polynomial calibration points found by the SSQR-GE algorithm. The points marked with a '+' are those selected from a fine grid, those marked 'o' from a coarse grid.



**Figure 4.** Cont.



**Figure 4.** Tensor product polynomial calibration points: values of  $t_q, q = 1, \dots, 100$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, for the fine grid, with no repeat measurements, upper curve, and with repeat measurements, lower curve.



**Figure 5.** Tensor product polynomial calibration points: values of  $t_q, q = 1, \dots, 100$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, for the coarse grid, with no repeat measurements, upper curve, and with repeat measurements, lower curve.

### 5.3. Coordinate Metrology

In coordinate metrology, the size and shape of a geometric surface  $S$  is determined from the measurement of the coordinates  $x_i$  of points lying on the surface. If the geometric surface  $S = S(\mathbf{a})$  is parametrized by  $\mathbf{a} = (a_1, \dots, a_n)^T$  and  $d(\mathbf{x}, \mathbf{a})$  is a measure of the distance from a point  $\mathbf{x}$  to  $S(\mathbf{a})$ , then for measurement points  $\mathbf{x}_{1:m} = \{\mathbf{x}_i, i = 1, \dots, m\}$ ,  $m \geq n$ , estimates of  $\mathbf{a}$  are determined by solving the least squares optimisation problem

$$\min_{\mathbf{a}} = \sum_{i=1}^m d^2(\mathbf{x}_i, \mathbf{a}).$$

If the variance matrix associated with the measured points coordinates is  $\sigma^2 I$ , then the variance matrix  $V_{\mathbf{a}}$  associated with the fitted parameters is approximated by

$$V_{\mathbf{a}} = \sigma^2 (J^T J)^{-1}, \quad J_{ij} = \frac{\partial d}{\partial a_j}(x_i, \mathbf{a}),$$

involving the Jacobian matrix  $J$ . Thus, it is possible to determine optimal measurement strategies, i.e., where to measure on the surface  $\mathcal{S}$ , in terms of the determinant of the Jacobian matrix  $J$ . Often, the case of choosing a minimal set of points with  $m = n$  is of interest. For simple geometries, an optimal set of points is straightforward to arrive at, e.g., three points equally spaced around a circle, but for more complex geometries, the choice of an optimal set may not be obvious, as considered below.

Figure 6 shows a reference artefact designed and calibrated by the Czech Metrology Institute (CMI) [34,35]. The form of the geometric surface is a hyperbolic paraboloid given by the equation

$$z = 24 + (x/8 - 1)(y/8 - 1). \tag{15}$$

A paraboloid with its axis approximately parallel to the  $z$ -axis is parametrized in terms of two rotation angles  $\alpha$  and  $\beta$  and six further parameters  $\mathbf{b}$  such that

$$\hat{z} = b_1 \hat{x}^2 + b_2 \hat{y}^2 + b_3 \hat{x} \hat{y} + b_4 \hat{x} + b_5 \hat{y} + b_6, \quad \hat{\mathbf{x}} = R(\alpha, \beta) \mathbf{x},$$

where  $R(\alpha, \beta) = R_x(\alpha)R_y(\beta)$  is the rotation matrix constructed from the plane rotation matrices

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \quad R_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}.$$

For the surface (15), nominally  $\mathbf{b}^T = (0, 0, 1/64, -1/8, -1/8, 25)$ . Figures 7–9 show the distribution of the  $x$ - and  $y$ -coordinates for three sets  $X_k, k = 1, 2, 3$ , of eight calibration points. The first two are given by ‘expert judgement’; that is, possible choices for a measurement strategy based on engineering practice. The first expert choice,  $X_1$ , Figure 7, represents an approximately uniform distribution, a first guess at a good distribution. This distribution of points is sufficient to determine all the parameters, and the associated  $8 \times 8$  Jacobian matrix is full rank. (If the  $x$ - and  $y$ -coordinates of the data points are rotated through 45 degrees about the  $z$ -axis, the associated Jacobian matrix is rank deficient by three.) The second set  $X_2$  came about through experimenting with moving the inner four points in dataset  $X_1$  closer or further from the origin of the  $xy$ -plane. It turned out that moving the four points further from the origin was better, resulting in dataset  $X_2$ , Figure 8. At first glance, the distribution of points in  $X_2$  does not look a good choice. The third dataset is that obtained using the SSQR-GE algorithm. The GE algorithm performed 11 interchanges starting from the SSQR solution. The optimal choice has seven points along the boundary in the  $xy$ -plane and one point in the interior. Any interior point not too far from the origin would also result in a good selection. In terms of the aggregate measure  $\bar{d}_k$  defined in (4) associated with the three datasets, we have  $\bar{d}_1 = 6.0\bar{d}_3$  and  $\bar{d}_2 = 2.2\bar{d}_3$ .

Table 3 gives the square roots  $v_i$  of the diagonal elements of  $(J_k^T J_k)^{-1}$ , where  $J_k$  is the Jacobian matrix associated with the  $k$ th dataset. If the variance matrix associated with the measurements is  $\sigma^2 I$ , then  $v_i = u(a_j)/\sigma$ , where  $u(a_j)$  is the estimate of the standard uncertainty associated with the  $j$ th parameter [36,37]. In terms of standard uncertainties, the D-optimal dataset  $X_3$  is far better for estimating the surface parameters than the other two datasets.

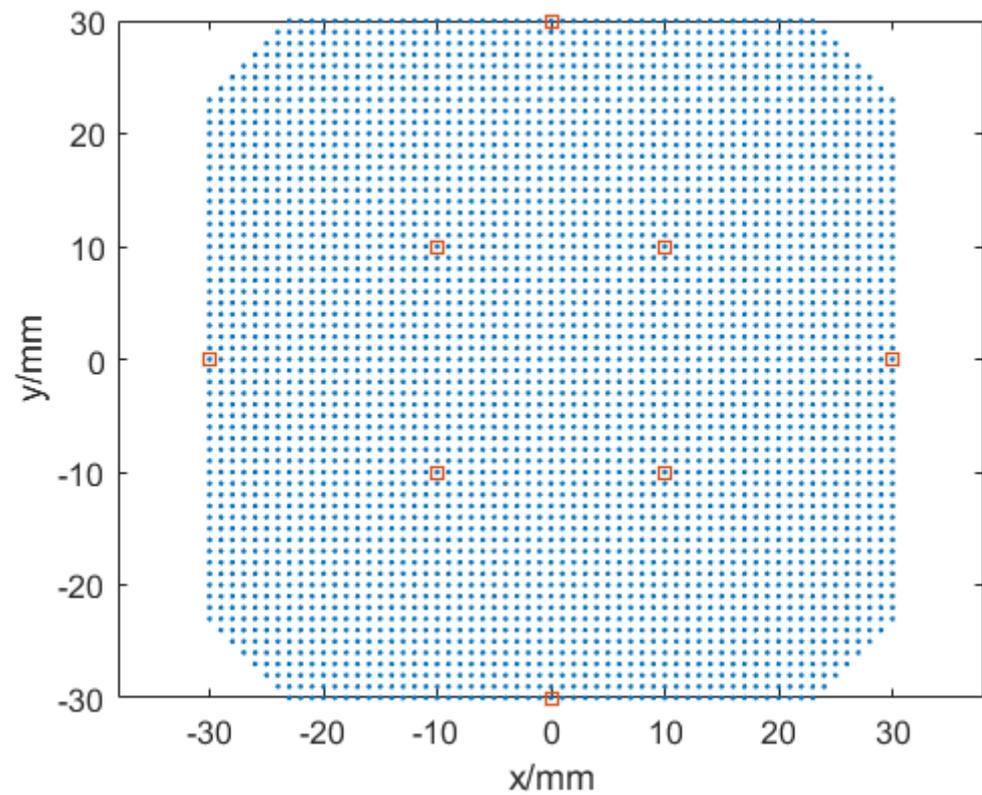
Figure 10 plots the values of  $t_q, q = 1, \dots, 100$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, with and without repeat measurements. For higher values of  $q$ , there is good agreement between  $t_q$  and  $\tilde{t}_q$ .

**Table 3.** Standard uncertainty factors  $u(a_j)/\sigma$  associated with the hyperbolic paraboloid parameters  $a = (\alpha, \beta, b^T)^T$  for three eight-point datasets  $X_k, k = 1, 2, 3$ .

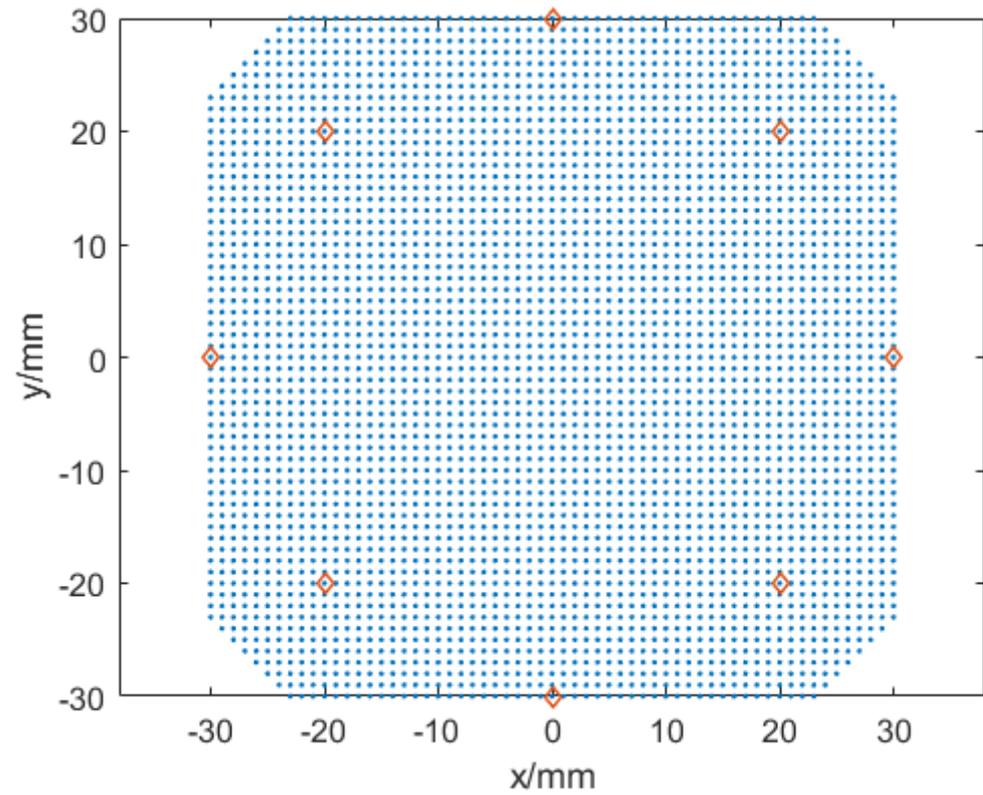
$u(a_j)/\sigma$	$\alpha$	$\beta$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$
$X_1$	2.3867	2.3867	0.0049	0.0049	0.0139	2.6219	2.6219	10.5328
$X_2$	0.7259	0.7259	0.0042	0.0042	0.0035	0.4999	0.4999	2.8188
$X_3$	0.2049	0.2413	0.0012	0.0014	0.0014	0.2883	0.2520	1.4690



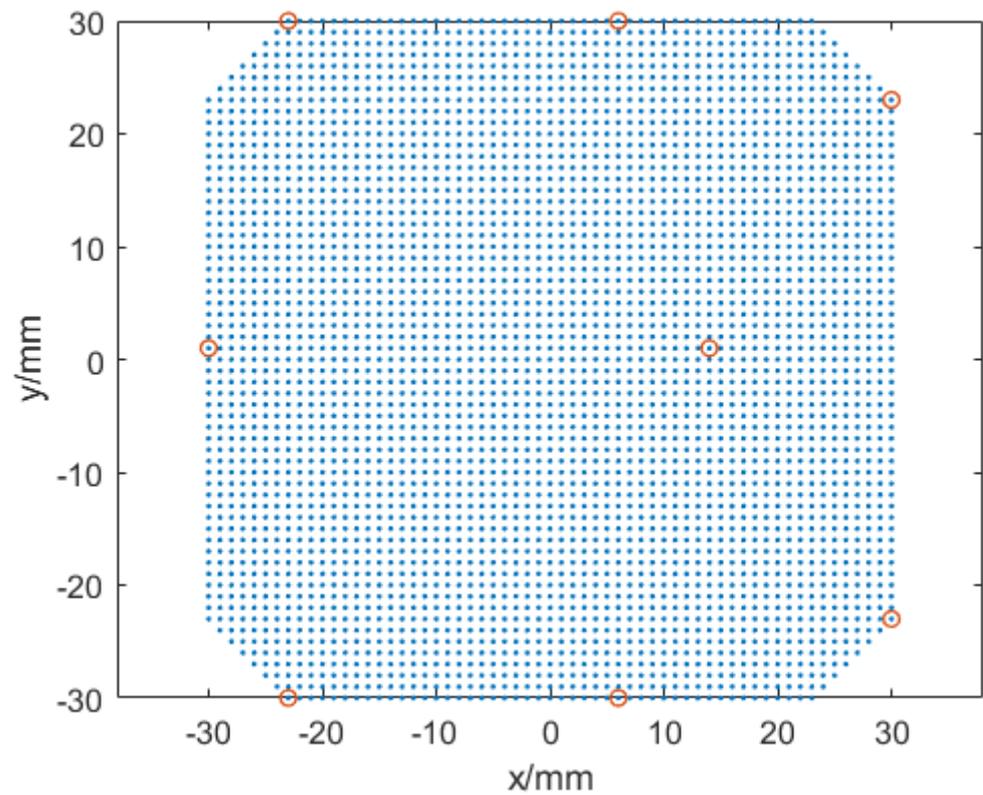
**Figure 6.** The CMI hyperbolic paraboloid reference artefact.



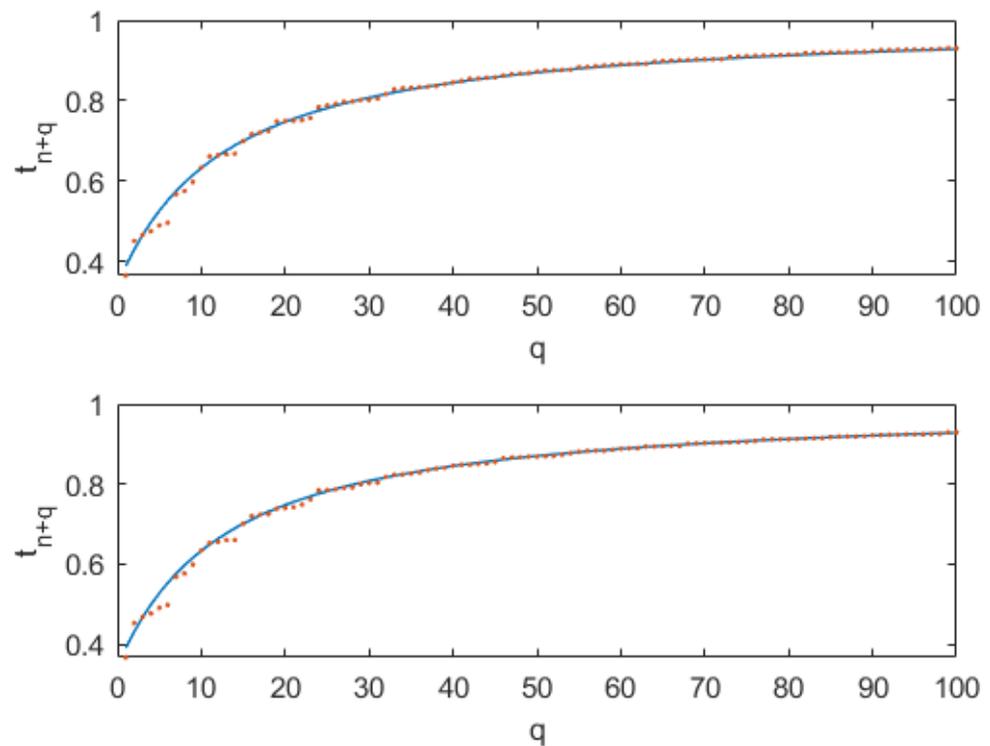
**Figure 7.** Hyperbolic paraboloid calibration points: ‘expert’ guess at a good set of calibration points.



**Figure 8.** Hyperbolic paraboloid calibration points: second 'expert' guess at a good set of calibration points, a modification of  $X_1$ , Figure 7.



**Figure 9.** Hyperbolic paraboloid calibration points: D-optimal points determined by the SSQR-GE algorithm.



**Figure 10.** Hyperbolic paraboloid calibration points: values of  $t_q, q = 1, \dots, 100$ , and its analytical approximation  $\tilde{t}(q)$  given by (12), determined by the sequential D-optimal Algorithm 4, with no repeat measurements, upper curve, and with repeat measurements, lower curve.

#### 5.4. Calibration of a Network of Standards Using a Comparator

In this section, we consider the calibration of a number of standards, starting with one calibrated standard and using a comparator to calibrate the other standards relative to the calibrated standard. An example application is in mass calibration using a mass balance [38–40]. Given a calibrated standard  $A_1$ , nominally of 1 kg, and two uncalibrated masses  $A_2$  and  $A_3$  each of nominal mass 0.5 kg,  $A_2$  and  $A_3$  can be calibrated using two measurements, the first comparing  $A_1$  with the combined mass of  $A_2$  and  $A_3$ , and the second comparing  $A_2$  and  $A_3$ . A second example from length metrology relates to the calibration of gauge blocks [41], where it is possible to align two gauge blocks end to end, a process called ‘wringing’, to define a combined length. This discussion is relevant to the calibration of any set of artefacts defining an extensive quantity where the artefacts can be grouped to define a quantity that is (nominally) the sum of the individual quantities. We use the term ‘network’ to reflect the fact that the estimates of all the quantities are statistically correlated and an experiment that updates the information about one artefact will also update information about the other artefacts.

Let  $\mathbf{a} = (a_1, \dots, a_n)^\top$  represent the values of the quantities associated with the artefacts where the nominal values  $\mathbf{a}^\# = (a_1^\#, \dots, a_n^\#)^\top$  of the quantities are such that the artefacts can be grouped into sets of two groups, with each group being associated with the same nominal value. Thus, there are disjoint subsets  $L_i, R_i \subset I = \{1, 2, \dots, n\}, L_i \cap R_i = \emptyset, i = 1, \dots, m$ , such that

$$\sum_{q \in L_i} a_q \approx \sum_{q \in R_i} a_q.$$

An example considered further below involves a set of artefacts having nominal values 1.00, 0.50, 0.50, 0.20, 0.20, 0.10, 0.10, 0.05, and 0.05. The fact that the two subsets are associated

with the same nominal value allows a comparator measurement to be made, with the associated observation equation

$$0 \approx y_i = \sum_{q \in L_i} a_q - \sum_{q \in R_i} a_q + \epsilon_i, \quad \epsilon_i \in \mathcal{N}(0, \sigma_i^2),$$

where  $\epsilon_i$  represents a random effect associated with the comparator measurement. The design problem is to define measurement strategies encoded by  $L_i$  and  $R_i, i = 1, \dots, n$ , that provide the most accurate calibration of the standards. For the example considered here, there are almost 400 viable experiments.

Tables 4 and 5 show five designs for the calibration of the  $n = 9$  standards. A ‘1’ in column  $j$  indicates that  $j \in L_i$ , while a ‘−1’ indicates that  $j \in R_i$ . Each array of  $9 \times 9$  elements is in fact the unweighted observation matrix  $C$  associated with the particular design. The first row in each design represents the observation associated with the calibration of the first standard using an absolute measurement system. This first observation is the only source of information about the first standard and must be included in the optimal solution. The uncertainties associated with the remaining standards scale with the uncertainty associated with the first observation.

The first design, given in Table 4, is one defined by ‘expert judgement’, and represents a typical set of experiments that would be performed in practice. (In fact, for this set of standards, it is not entirely obvious how to choose a design that will calibrate all of the standards.) The next design, indicated by the first nine rows in Table 5, is that determined by the SSQR-GE algorithm with  $\sigma_1 = \sigma_C = 1.0$  and  $\sigma_i = \sigma_R = 0.5, i = 2, \dots, n$ . The most striking and, at first sight, surprising feature of this second design is that almost all the artefacts are involved in experiments 2 to  $n$ . The reason for this is that, in some sense, the comparator uncertainties represented by  $\sigma_i = \sigma_R$  are partitioned amongst the artefacts involved, and more artefacts being included in an experiment means that each artefact is assigned a smaller uncertainty.

The two designs considered so far are based on the assumption that all the comparator measurements are associated with the same uncertainty  $\sigma_R$ , irrespective of the number of artefacts and the nominal values of the artefacts. A more plausible assignment of uncertainties would take into account both the number and nominal values. For a mass balance, it would be usual for the associated uncertainty to have a component that varied in proportion to the mass. Similarly, a length measuring device will usually have one or more influencing factors, e.g., refractive index effects for the case of a laser interferometric comparator [41], that vary in proportion to length. The use of multiple artefacts in each experiment will also likely introduce effects that will add to the uncertainties.

For each experiment, let  $n_i$  be the total number of artefacts involved and let

$$v_i = \sum_{q \in L_i \cup R_i} a_q^\#$$

be a measure of the total value of the quantities involved in the  $i$ th experiment. Given  $\sigma_R, \sigma_V$  and  $\sigma_N$ , we can assign  $\sigma_i$  according to

$$\sigma_i^2 = \sigma_R^2 + n_{i,2} \sigma_N^2 + v_i^2 \sigma_V^2, \quad n_{i,2} = \max(n_i - 2, 0), \quad i = 2, \dots, n. \tag{16}$$

Table 5 presents the optimal designs calculated using the SSQR-GE algorithm with  $\sigma_i$  calculated as in (16) for the four different values of  $\sigma_R, \sigma_N$  and  $\sigma_V$  shown in Table 6. The third design in Table 5 is the result penalising experiments with a large number of artefacts, while the fourth design in the table is the result of penalising experiments with a larger value of  $v_i$ .

Table 6 gives the uncertainties  $u(a_i)$  associated with  $a_i$  for four different assignments of  $\sigma_i, i = 2, \dots, n = 9$  for the designs in Tables 4 and 5. Also shown in the table is the aggregate measure  $\bar{d} = |V_{\mathbf{a}}|^{1/n}$  of uncertainty, the total number  $\sum_i n_i$  of artefacts and a measure  $\sum_i v_i$  of the total nominal values involved in each set of experiments. The last three rows give the values of  $\sigma_R, \sigma_N$  and  $\sigma_V$  used to calculate  $\sigma_i, i = 2, \dots, n$ . In all cases, the SSQR-GE algorithm leads to significant improvements over the expert judgement design.



**Table 6.** Uncertainties associated with  $a_i$  for four different assignments of  $\sigma_i, i = 2, \dots, n = 9$ . The columns labelled *E* give the uncertainties associated with ‘expert judgement’ design given in Table 4, while the columns labelled ‘O’ give those associated with the optimal designs, given, in Table 5, calculated by the SSQR-GE algorithm. Also shown in the table are the aggregate measure  $\bar{d} = |V_{\mathbf{a}}|^{1/n}$  of uncertainty, the total number  $\sum_i n_i$  of artefacts, and a measure  $\sum_i v_i$  of the total nominal values involved in each set of experiments. The last three rows give the values of  $\sigma_R, \sigma_N$ , and  $\sigma_V$  used to calculate  $\sigma_i, i = 2, \dots, n$ .

$a_i^{\#}$	E	O1	E	O2	E	O3	E	O4
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.50	0.61	0.56	0.66	0.64	0.69	0.69	1.04	1.04
0.50	0.61	0.55	0.66	0.64	0.69	0.69	1.04	1.04
0.20	0.39	0.31	0.43	0.40	0.60	0.58	0.50	0.57
0.20	0.49	0.29	0.52	0.39	0.61	0.58	0.54	0.60
0.10	0.57	0.26	0.61	0.32	0.90	0.44	0.57	0.36
0.10	0.91	0.27	1.03	0.36	1.64	0.45	1.34	0.34
0.05	0.35	0.20	0.36	0.28	0.40	0.48	0.29	0.27
0.05	0.35	0.20	0.36	0.28	0.40	0.48	0.29	0.27
$\bar{d}$	0.17	0.06	0.21	0.12	0.21	0.13	0.21	0.15
$\sum v_i$	7.0	17.4	7.0	11.0	7.0	6.3	7.0	6.3
$\sum n_i$	24	62	24	48	24	22	24	22
$\sigma_R$	0.50	0.50	0.50	0.50	0.20	0.20	0.20	0.20
$\sigma_N$	0.00	0.00	0.20	0.20	0.80	0.80	0.20	0.20
$\sigma_V$	0.00	0.00	0.20	0.20	0.20	0.20	0.80	0.80

### 6. Concluding Remarks

In this paper, we have looked at the experimental design problem: given  $m$  potential observations to determine  $n$  parameters,  $m > n$ , what is the best choice of  $n$  observations. In the context of least squares estimation, the problem was formulated as finding the  $n \times n$  submatrix of the complete  $m \times n$  observation matrix that has maximum determinant, corresponding to the D-optimality criterion. We described two algorithms, the SSQR and GE algorithms, to address this problem. Both were adapted from numerical linear algebra algorithms associated with the QR factorisation of a matrix. We also described algorithms for updating estimates of the problem parameters that are sequentially optimal with the respect to the D-optimality and A-optimality criteria. We illustrated the behaviour of the algorithms on a number of applications drawn from the field of metrology. These algorithms are straightforward to implement. The SSQR algorithm, in particular, is a minor adaption of the standard QR factorisation algorithm with pivoting. The algorithms enable computationally efficient implementations, exploiting rank-one matrix updating techniques. The algorithms are descent-type algorithms that will converge to a local minimum that is not necessarily a global minimum. In the examples considered, the algorithms determined effective experimental designs, often with much better performance than ‘hand-crafted’ designs based on expert judgement.

**Funding:** This work was supported by the UK’s National Measurement System programme for Data Science.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article

**Acknowledgments:** The author is grateful to the anonymous referees for their helpful comments, and to the Czech Metrology Institute for permission to use their photograph of the hyperbolic paraboloid, Figure 6.

**Conflicts of Interest:** The author declares no conflict of interest.

### Appendix A. Numerical Linear Algebra

#### Appendix A.1. Sherman-Morrison Formula

See [23]. For  $n \times n$  full rank matrix  $A$  and  $n$ -vectors  $u$  and  $v$

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u}.$$

If  $\tilde{u} = A^{-1}u$  and  $\tilde{v} = A^{-T}v$ , then

$$(A + uv^T)^{-1} = A^{-1} - \frac{\tilde{u}\tilde{v}^T}{1 + v^T \tilde{u}} = A^{-1} - \frac{\tilde{u}\tilde{v}^T}{1 + \tilde{v}^T u}.$$

#### Appendix A.2. Determinant of a Rank One Update of a Matrix

See, e.g., [42,43]. For  $A$  and square matrix and  $n$ -vectors  $u$  and  $v$

$$|A + uv^T| = |A|(1 + v^T A^{-1}u).$$

Writing

$$A + uv^T = A(I + (A^{-1}u)v^T),$$

it is sufficient to consider the case  $A = I$ :  $|I + uv^T| = 1 + v^T u$ . From the matrix equation

$$\begin{bmatrix} I & \mathbf{0} \\ v^T & 1 \end{bmatrix} \begin{bmatrix} I + uv^T & u \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ -v^T & 1 \end{bmatrix} = \begin{bmatrix} I & u \\ \mathbf{0} & 1 + v^T u \end{bmatrix}$$

it is seen that the determinant of the matrix on the right-hand side is equal to the determinant of the matrix in the middle of the left-hand side, establishing the result for the case  $A = I$ .

### Appendix B. Legendre Polynomials

See, e.g., [44]. Note: in this section, the parameter vector  $a = (a_0, \dots, a_n)^T$  is an  $n + 1$  vector indices starting from 0, not 1.

The D-optimal points for calibration experiments involving a polynomial response function of degree  $n$  are defined in terms of solutions of the polynomial equation

$$F(x) = 0, \quad F(x) = (1 - x^2)\dot{L}_{n-1}(x) = 0,$$

where  $\dot{L}_{n-1}(x)$  is the derivative of Legendre polynomial of degree  $n - 1$ . Given an estimate  $x$  of a solution, an updated solution is given by Newton's method [17,45] according to

$$x := x - F(x)/\dot{F}(x), \quad \dot{F}(x) = -2x\dot{L}_{n-1} + (1 - x^2)\ddot{L}_{n-1}(x)$$

where  $\ddot{L}$  denotes the second derivative of  $L(x)$  with respect to  $x$ . Starting estimates for the  $n + 1$  solutions are given by the arcsine points:

$$x_i = \cos\left(\pi \frac{n-i}{n}\right), \quad i = 0, 1, \dots, n.$$

Legendre polynomials  $L_j(x)$  are orthogonal polynomial basis functions defined on the interval  $[-1, 1]$  and can be evaluated using the three-term recurrence relationship, starting with  $L_0(x) = 1$ ,  $L_1(x) = x$ , and for  $j \geq 2$ ,

$$jL_j(x) = (2j - 1)xL_{j-1}(x) - (j - 1)L_{j-2}(x).$$

If  $f(x)$  is a degree  $n$  polynomial given by a sum of Legendre polynomial basis functions

$$f(x, \mathbf{a}) = \sum_{j=0}^n a_j L_j(x),$$

then  $\dot{f}(x)$ , the derivative of  $f$  with respect to  $x$ , is a degree  $n - 1$  polynomial and can be written as

$$\dot{f}(x, \dot{\mathbf{a}}) = \sum_{j=0}^{n-1} \dot{a}_j L_j(x),$$

where  $\dot{\mathbf{a}} = \dot{A}(a_1, \dots, a_n)^\top$  is defined in terms of the latter  $n$  elements of the  $n + 1$  vector  $\mathbf{a} = (a_0, a_1, \dots, a_n)^\top$ , and  $\dot{A}$  is the  $n \times n$  matrix constructed from diagonals formed from the sequence  $1, 3, 5, \dots, 2n - 1$ . The matrix  $\dot{A}$  for the case  $n = 10$  is

$$\dot{A} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 3 & 0 & 3 & 0 & 3 & 0 & 3 \\ 0 & 0 & 5 & 0 & 5 & 0 & 5 & 0 & 5 & 0 \\ 0 & 0 & 0 & 7 & 0 & 7 & 0 & 7 & 0 & 7 \\ 0 & 0 & 0 & 0 & 9 & 0 & 9 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 11 & 0 & 11 & 0 & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 & 13 & 0 & 13 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 & 0 & 15 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 17 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 19 \end{bmatrix}$$

Similarly, the second derivative  $\ddot{f}(x)$  can be written as

$$\ddot{f}(x, \ddot{\mathbf{a}}) = \sum_{j=0}^{n-2} \ddot{a}_j L_j(x),$$

where  $\ddot{\mathbf{a}} = \ddot{A}(a_2, \dots, a_n)^\top$  is defined in terms of the latter  $n - 1$  elements of  $\mathbf{a} = (a_0, a_1, \dots, a_n)^\top$ , and  $\ddot{A}$  is the  $(n - 1) \times (n - 1)$  matrix with

$$\ddot{A} = \dot{A}(1 : n - 1, 1 : n - 1) \dot{A}(2 : n, 2 : n).$$

## References

1. Atkinson, A.C.; Donev, A.N.; Tobias, R.D. *Optimum Experimental Designs, with SAS*; Oxford University Press: Oxford, UK, 2007.
2. Box, G.E.P.; Hunter, W.G.; Hunter, J.S. *Statistics for Experimenters : Design, Innovation and Discovery*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2005.
3. Chaloner, K. Optimal Bayesian experimental design for linear models. *Ann. Stat.* **1984**, *12*, 283–300. [\[CrossRef\]](#)
4. Forbes, A.B.; Minh, H.D. Design of linear calibration experiments. *Measurement* **2013**, *46*, 3730–3736. [\[CrossRef\]](#)
5. Goos, P. *The Optimal Design of Blocked and Split-Plot Experiments*; Springer: New York, NY, USA, 2002.
6. Goos, P.; Jones, B. *Optimal Design of Experiments: A Case Study Approach*; John Wiley & Sons: New York, NY, USA, 2011.
7. Jones, B.; Nachtsheim, C.J. Effective Model Selection for Definitive Screening Designs. *Technometrics* **2017**, *59*, 319–329. [\[CrossRef\]](#)
8. Montgomery, D.C. *Design and Analysis of Experiments*, 8th ed.; John Wiley & Sons: New York, NY, USA, 2013.
9. Chretien, S.; Clarkson, P. A fast algorithm for the semi-definite relaxation of the state estimation problem in power grids. *J. Ind. Manag. Optim.* **2020**, *16*, 431–443. [\[CrossRef\]](#)
10. Kekatos, V.; Giannakis, G.B. A convex relaxation approach to optimal placement of phasor measurement units. In Proceedings of the 2011 4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), San Juan, PR, USA, 13–16 December 2011; pp. 145–148.
11. Berger, J.; Busser, T.; Dutykh, D.; Nathan Mendes, N. An efficient method to estimate sorption isotherm curve coefficients. *Inverse Probl. Sci. Eng.* **2019**, *27*, 735–772. [\[CrossRef\]](#)
12. Berger, J.; Dutykh, D.; Mendes, N. On the optimal experiment design for heat and moisture parameter estimation. *Exp. Therm. Fluid Sci.* **2017**, *81*, 109–122. doi:10.1016/j.expthermflusci.2016.10.008. [\[CrossRef\]](#)
13. Chernoff, H. *Sequential Analysis and Optimal Design*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1972. [\[CrossRef\]](#)

14. Meyer, R.K.; Nachtsheim, C.J. The Coordinate Exchange Algorithm for Constructing Exact Optimal Designs. *Technometrics* **1995**, *37*, 60–69. [[CrossRef](#)]
15. Wald, A. Sequential Tests of Statistical Hypotheses. *Ann. Math. Stat.* **1945**, *16*, 117–186. [[CrossRef](#)]
16. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
17. Gill, P.E.; Murray, W.; Wright, M.H. *Practical Optimization*; Academic Press: London, UK, 1981.
18. Vandenberghe, L.; Boyd, S. Semidefinite Programming. *SIAM Rev.* **1996**, *38*, 49–95. [[CrossRef](#)]
19. Barker, R.M.; Cox, M.G.; Forbes, A.B.; Harris, P.M. *Software Support for Metrology Best Practice Guide No. 4: Modelling Discrete Data and Experimental Data Analysis*. Technical Report DEM-ES 018; National Physical Laboratory: Teddington, UK, 2007.
20. BIPM. *The International System of Units (SI Brochure (EN))*, 9th ed.; BIPM: Sèvres, France, 2019.
21. BIPM; IEC; IFCC; ILAC; ISO; IUPAC; IUPAP; OIML. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*; JCGM 100:2008; Joint Committee for Guides in Metrology: Sèvres, France, 2008.
22. BIPM; IEC; IFCC; ILAC; ISO; IUPAC; IUPAP; OIML. *Evaluation of Measurement Data—Supplement 2 to the “Guide to the Expression of Uncertainty in Measurement”—Extension to Any Number of Output Quantities*; JCGM 102:2011; Joint Committee for Guides in Metrology: Sèvres, France, 2011.
23. Golub, G.; Van Loan, C. *Matrix Computations*, 4th ed.; Johns Hopkins University Press: Baltimore, MD, USA, 2013.
24. Hart, G.W. *Multidimensional Analysis: Algebras and Systems for Science and Engineering*; Springer: Berlin, Germany, 1995.
25. Gu, M.; Eisenstat, S.C. Efficient Algorithms for Computing a Strong Rank-Revealing QR Factorization. *SIAM J. Sci. Comput.* **1996**, *17*, 848–869. [[CrossRef](#)]
26. Sherman, J.; Morrison, W.J. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Ann. Math. Stat.* **1950**, *21*, 124–127. [[CrossRef](#)]
27. Wilkinson, J.H. *The Algebraic Eigenvalue Problem*; Oxford University Press, Inc.: New York, NY, USA, 1988.
28. Preston-Thomas, H. The International Temperature Scale of 1990 (ITS-90). *Metrologia* **1990**, *27*, 3–10. [[CrossRef](#)]
29. Bartlett, G.; Forbes, A.; Heaps, E.; Raby, A.C.; Yacoot, A. Spatial positioning correction for multi-axis nanopositioning stages. In Proceedings of the ASPE Convention and Expo, Indianapolis, IN, USA, 16–21 September 2022.
30. Pukelsheim, F. *Optimal Design of Experiments*; SIAM: Philadelphia, PA, USA, 2006; Reproduction of 1993 book published by John Wiley and Sons, New York.
31. Handscombe, D.C.; Mason, J.C. *Chebyshev Polynomials*; Chapman & Hall/CRC Press: London, UK, 2003.
32. Forbes, A.B.; Jagan, K.; Dunlevy, J.; Sousa, J.A. Optimization of sensor distribution using Gaussian processes. *Meas. Sens.* **2021**, *18*, 100128. [[CrossRef](#)]
33. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.
34. Linkeová, I.; Zelený, V. Application of ruled surfaces in freeform and gear metrology. *Acta Polytech.* **2021**, *61*, 99–109. [[CrossRef](#)]
35. Zelený, V.; Linkeová, I.; Skalnik, P. Calibration of freeform standard. In Proceedings of the 15th International Conference of the European Society for Precision Engineering and Nanotechnology, EUSPEN 2015, Leuven, Belgium, 1–5 June 2015; pp. 147–148.
36. Forbes, A.B. Parameter estimation based on least squares methods. In *Data Modeling for Metrology and Testing in Measurement Science*; Pavese, F., Forbes, A.B., Eds.; Birkhäuser-Boston: New York, NY, USA, 2009; pp. 147–176.
37. Forbes, A.B. Sensitivity analysis for Gaussian associated features. *Appl. Sci.* **2022**, *12*, 2808. [[CrossRef](#)]
38. Grabe, M. Note on the Application of the Method of Least Squares. *Metrologia* **1978**, *14*, 143. [[CrossRef](#)]
39. Hotelling, H. Some Improvements in Weighing and Other Experimental Techniques. *Ann. Math. Stat.* **1944**, *15*, 297–306. [[CrossRef](#)]
40. Nielsen, L. Evaluation of mass measurements in accordance with the GUM. *Metrologia* **2014**, *51*, S183. [[CrossRef](#)]
41. Lewis, A.J.; Hughes, E.B.; Aldred, P.J.E. Long term study of gauge block interferometer performance and gauge blocks stability. *Metrologia* **2010**, *47*, 473–486. [[CrossRef](#)]
42. Ding, J.; Zhou, A. Eigenvalues of rank-one updated matrices with some applications. *Appl. Math. Lett.* **2007**, *20*, 1223–1226. [[CrossRef](#)]
43. Hogben, L. (Ed.) *Handbook of Linear Algebra*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2007.
44. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions*; Dover: New York, NY, USA, 1964.
45. Fletcher, R. *Practical Methods of Optimization*, 2nd ed.; John Wiley and Sons: Chichester, UK, 1987.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.