JOURNAL OF QUALITY TECHNOLOGY
https://doi.org/10.1080/00224065.2020.1829215

Optimal experimental designs for ordinal models with mixed factors for industrial and healthcare applications

Joshua Lukemire[a], Abhyuday Mandal[b], and Weng Kee Wong[c]

[a]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia; [b]Department of Statistics, University of Georgia, Athens, Georgia; [c]Department of Biostatistics, University of California at Los Angeles, Los Angeles, California

ABSTRACT

Models with ordinal outcomes are an important part of generalized linear models and design issues for them are less studied, especially when the model has discrete and continuous factors. We propose an effective and flexible Particle Swarm Optimization (PSO) algorithm for finding locally *D*-optimal approximate designs for experiments with ordinal outcomes modeled using the cumulative logit link. We apply our technique to obtain a locally *D*-optimal approximate design for an odor removal experiment with both discrete and continuous factors and show that this design is superior to the design obtained by discretizing the continuous factor. Additionally, we find a pseudo-Bayesian *D*-optimal approximate design for this problem and study the performance of both designs under a range of plausible parameter values. We also (i) demonstrate PSO's versatility by finding locally *D*-optimal approximate designs for a manufacturing example with surface defects and multiple continuous factors, and (ii) use PSO to find other optimal designs for estimating percentiles in a dose-response study.

KEYWORDS

approximate design; dose-response; particle swarm optimization; qualitative factors; quantitative factors

1. Introduction

This work concerns experiments conducted to determine the effect of a set of factors on a categorical outcome with J ordered levels. For example, the quality of a manufactured product might be classified as "low," "medium," or "high," and the experimenter may be interested to learn what process settings will increase the probability of having a "high" quality product. Similarly, in a healthcare study, a patient's response after an intervention may be classified as "no improvement," "some improvement," or "significant improvement." The outcomes of such ordinal response experiments are often modeled using cumulative logit models, which is a form of generalized linear model.

Analysis of experimental data using cumulative logit models is well described (Agresti 2003). In particular, the ordinal categories are modeled using a continuous latent variable, Z, and the probability of the response falling into each category depends on the value z of the latent variable through cut-points, with the jth cut-point denoted by θ_j. For a model with J outcome categories, there will be $J-1$ cut-points. For example, Figure 1 displays the latent continuous factor for an experiment with a three-level response: a, b, and c. If $z < \theta_1$, outcome a is observed, if $\theta_1 \leq z < \theta_2$ outcome b is observed, and if $z \geq \theta_2$ outcome c is observed. It is important to note that these cut-points are *not* determined by the experimenter, and thus $\boldsymbol{\Theta} = \{\theta_1, ..., \theta_{J-1}\}^T$ must be treated as a vector of parameters to be estimated.

This modeling strategy has a rich history and is used in a variety of settings including healthcare (Clark et al. 2000; Lall et al. 2002) and industrial experiments (Nair 1986). However, optimal design for experiments with ordered outcomes is not well studied. The design problem is to determine the set of optimal experimental settings that most efficiently estimate the model parameters, or some function of the model parameters. The efficiency of the estimation is measured as a function of the "size" of the Fisher Information matrix. For example, *D*-optimality seeks to maximize the logarithm of the determinant of the Fisher Information matrix, and the resulting *D*-optimal design minimizes the volume of the confidence ellipsoid of the model parameters.

There is much research on addressing design issues for generalized linear models. Atkinson and Woods

CONTACT Abhyuday Mandal amandal@stat.uga.edu Department of Statistics, University of Georgia Franklin College of Arts and Sciences, Athens, GA 30602, USA.

Supplemental data for this article is available online at https://doi.org/10.1080/00224065.2020.1829215

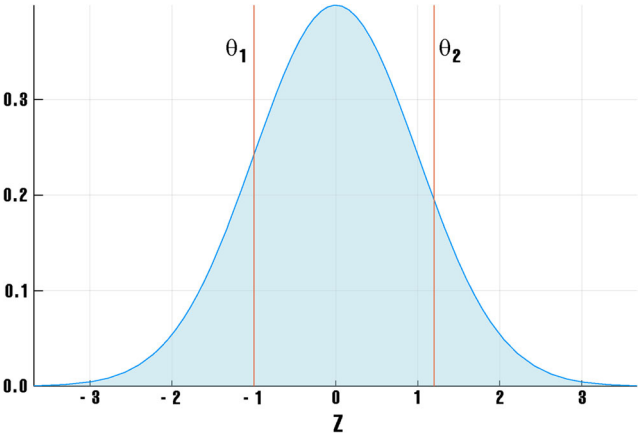

Figure 1. An example of the latent variable and cut-points for a cumulative logit model with a three-level outcome. The *x*-axis shows the values of Z and the *y*-axis is the corresponding density of Z at these values. The cut-points θ_1 and θ_2 are marked with red lines.

(2015) provides an overview. For example, Russell et al. (2009) constructs optimal designs for Poisson regression models. In the binary response case, Yang, Mandal, and Majumdar (2012) describes techniques for obtaining locally D-optimal designs for experiments with two factors, and Yang, Mandal, and Majumdar (2016) extends this technique to experiments with an arbitrary number of discrete factors. Dette (2004) discusses designs for risk estimation that are robust to assumed success rates, and Wu, Wong, and Crespi (2017) discusses optimal design for cluster randomized trials with binary outcomes. Xu et al. (2019) uses differential evolution to find designs for logistic regression models with multiple interacting factors.

Optimal design results for models with a categorical outcome with several levels are more scarce. Zocchi and Atkinson (1999) considered D-optimal designs for experiments with a continuous factor in the model. Perevozskaya, Rosenberger, and Haines (2003) studied optimal design problems for quantile estimation, and is an example for finding an optimal design to estimate a function of the model parameters. Recently, Yang, Tong, and Mandal (2018) developed a technique for cumulative logit models with the restriction that all factors in the model are discrete.

Our main goals are to develop and implement an effective and flexible algorithm for practitioners to generate new D-optimal designs with outcomes modeled using cumulative logit models with both discrete and continuous factors, henceforth referred to as experiments with "mixed factors." For such experiments, a common practice is to either discretize the range of the continuous factors and use methodology for designing experiments with all discrete factors, or to fix the continuous factor at a particular setting for the experiment. The former strategy can be computationally inefficient, as the algorithm will become slow when the number of continuous factors grows. The latter is clearly undesirable as it does not capture information about responses at other settings of the continuous factor or factors.

Table 1. The five factors in the odor removal experiment, with four discrete and one continuous factor.

		Levels	
Type	Factor	−	+
Discrete	Algae	Catfish Algae	Solix Microalgae
Discrete	Scavenger	Activated Carbon	Zeolite
Discrete	Resin	Polyethylene	Polypropylene
Discrete	Compatibilizer	Absent	Present
Continuous	Temperature	5°C to 35°C	

Our work on mixed factor experiments with ordinal response is motivated by a bio-plastics experiment carried out at the University of Georgia (Wang et al. 2017). The investigators were interested in the effects of several factors on whether or not a manufactured bio-plastic emitted an odor. The factors of interest are listed in Table 1. The ordinal outcome had five levels: almost no odor, mild odor, medium odor, strong odor, and serious odor. The original experiment considered only the discrete factors in Table 1 and held the continuous variable temperature fixed at 25 °C without explanation. We revisit this experiment and generate designs that more realistically include all factors of interest in the study.

Our approach to designing experiments with mixed factors is not to discretize the continuous factors. We propose a Particle Swarm Optimization (PSO) algorithm for finding D-optimal approximate designs for experiments with ordinal response modeled using a cumulative logit model. Section 2 provides background and discusses the form of the Fisher information matrix for cumulative logit models. Section 3 describes the PSO algorithm for generating D-optimal approximate designs for experiments with mixed factors. In Section 4 we use PSO and generate a locally D-optimal approximate design for the odor removal study that includes temperature as a continuous variable and investigate its robustness properties to mis-specification in the model parameter associated with the temperature variable. We further use PSO to generate a pseudo-Bayesian approximate design for this problem, and compare the performance of the locally D-optimal and pseudo-Bayesian designs to a range of plausible parameter values. Additionally, Section 5 uses PSO to generate a locally D-optimal approximate design for a surface defects experiment with a large number of continuous factors. We conclude in Section 6 with comments on future directions. In the Supplementary Materials, we demonstrate

another application of PSO to find other types of optimal designs for a dose-response model with ordinal outcomes and show some of the optimal designs published earlier were not really optimal. General Julia codes for finding designs for experiments with ordered categorical outcomes are provided as a Supplementary Material including markdown files that reproduce all analyses from the paper. The Julia language was chosen due to its fast computation speed and ease of use. All reported computation times were obtained using a 2020 iMac (3.2 GHz 6-Core Intel Core i7 processor).

2. Notation and design background

Consider an experiment with response Y having J ordered categories, K factors in the model, p regression parameters, and analyzed using a cumulative logit model (Agresti 2003). Here we have N available experimental units which can be distributed across L support points ($L \leq N$), where L is user-defined. Define $\mathbf{x}_i = \{x_{i,1}, x_{i,2}, \ldots x_{i,K}, x_{i,K+1}, \ldots, x_{i,p}\}^T$ for $i = 1, \ldots, L$, where the first K components represent experimental settings for the K factors (i.e. the support points), and the remaining values for $\mathbf{x}_i$ are determined from the levels of the K factors based on the postulated interactions in the model. The experimental design is the collection of the support points, along with the number of replicates at each support point. If the ith support point has n_i replicates, we have that $\sum_{i=1}^{L} n_i = N$. The design space is the set of all possible combinations of factor settings and given an objective function, the goal is to determine the optimal support points and n_i, $i = 1, \ldots, L$. The resulting design is known as an optimal exact design, or strictly speaking an L-point optimal exact design.

Finding an optimal exact design analytically is very challenging, and numerical searches can be prohibitively computationally demanding, especially when the model has many interacting factors. An alternative option is to find an approximate design (Kiefer 1959), where the optimal proportion (w_i) of experimental units at the support point $\mathbf{x}_i$ is found instead of the optimal number of replicates. If there are L support points, the weights $w_1, \ldots, w_L$ sum to unity. In practice, an approximate design is implemented by taking N_{w_i} observations at the ith support point, and N_{w_i} is obtained by rounding each N_{w_i} to the nearest integer, subject to the constraint that $N_{w_1}, \ldots, N_{w_L}$ sum to N.

The worth of a model-based design is measured by the Fisher information matrix for the parameters $\boldsymbol{\Theta} = \{\boldsymbol{\beta}^T, \boldsymbol{\theta}^T\}^T$, where $\boldsymbol{\theta}$ is the vector of cut-points and $\boldsymbol{\beta}$ is the vector of p coefficients in the linear predictor of the generalized linear model. Here we make the common proportional odds assumption that this set of p coefficients is the same across different outcome categories. Let the linear predictor for the jth outcome at $\mathbf{x}_i$ be $\eta_{i,j} = \theta_j - \mathbf{x}_i^T\boldsymbol{\beta}, i = 1, \ldots, L, j = 1, \ldots, J-1$. In the cumulative logit model, we model the probability that the outcome Y_i falls into category j or below as $P(Y_i \leq j) = g(\eta_{i,j})$, where $g(\cdot)$ is the logistic link function. The inverse link function is $g_{i,j}^{-1} = \exp(\eta_{i,j})/(1 + \exp(\eta_{i,j}))$, and the probability that observation i falls into category j is $\pi_{i,j} = g_{i,j}^{-1} - g_{i,j-1}^{-1}$.

We represent the *ith* point, $i = 1, \ldots, L$ as

$$\underset{(p+J-1)\times(J-1)}{\mathbf{X}_i} = \begin{bmatrix} -x_{i,1} & -x_{i,1} & \ldots & -x_{i,1} \\ -x_{i,2} & -x_{i,2} & \ldots & -x_{i,2} \\ \vdots & \vdots & \vdots & \vdots \\ -x_{i,p} & -x_{i,p} & \ldots & -x_{i,p} \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

This representation is chosen so that $\mathbf{X}_i^T\boldsymbol{\Theta} = \boldsymbol{\eta}_i$, the $J-1$ dimensional vector of linear predictors for

$$\underset{(J-1)\times(J-1)}{\mathbf{W}_i} = \begin{bmatrix} \left(\frac{\partial g_{i,1}^{-1}}{\partial \eta_{i,1}}\right)^2 & \left(\frac{\partial g_{i,1}^{-1}}{\partial \eta_{i,1}}\right)\left(\frac{\partial g_{i,2}^{-1}}{\partial \eta_{i,2}}\right) & 0 & \ldots & 0 \\ \left(\frac{\partial g_{i,2}^{-1}}{\partial \eta_{i,2}}\right)\left(\frac{\partial g_{i,1}^{-1}}{\partial \eta_{i,1}}\right) & \left(\frac{\partial g_{i,2}^{-1}}{\partial \eta_{i,2}}\right)^2 & \left(\frac{\partial g_{i,2}^{-1}}{\partial \eta_{i,2}}\right)\left(\frac{\partial g_{i,3}^{-1}}{\partial \eta_{i,3}}\right) & \ldots & 0 \\ 0 & \left(\frac{\partial g_{i,3}^{-1}}{\partial \eta_{i,3}}\right)\left(\frac{\partial g_{i,2}^{-1}}{\partial \eta_{i,2}}\right) & \left(\frac{\partial g_{i,3}^{-1}}{\partial \eta_{i,3}}\right)^2 & \ldots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \left(\frac{\partial g_{i,J-1}^{-1}}{\partial \eta_{i,J-1}}\right)^2 \end{bmatrix} \quad (2)$$

the ith support point. Next, we define $\mathbf{W}_i$, a tri-diagonal matrix for each support point in Eq. [2].

Using these two definitions, the Fisher Information matrix for a design ξ is

$$\mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\theta}}(\xi) = \sum_{i=1}^{L} w_i \mathbf{X}_i' \mathbf{W}_i \mathbf{X}_i, \tag{3}$$

(Yang, Tong, and Mandal 2018). The Fisher information matrix depends on the parameter vector $\boldsymbol{\Theta}$ through the inverse of the link function. Consequently, the optimal design depends on the parameters we wish to estimate and we have a circular problem: an accurate set of parameters is required to obtain an optimal design that can accurately estimate the set of parameters.

There are several common solutions to this problem, each posing a different set of difficulties. First, in the event that a best guess for the parameters is already known, say, from a pilot experiment, then this best guess can be used to generate the design. This best guess is referred to as the set of nominal values for the parameters, and the resulting design is known as a locally optimal design (Chernoff 1953). Such designs can be quite efficient if the set of nominal values is close to the true parameter values, but robustness studies should be undertaken to examine the sensitivity of the optimal design to the set of nominal values before implementing an approximate design. If the performance of the locally optimal design is poor and there is uncertainty in the parameter values, maximin optimal designs provide alternative design strategies (Chen et al. 2015, 2017; Wu, Wong, and Crespi 2017). These approaches assume there is a plausible set containing all possible values for the unknown, true parameters of interest, and the optimal design seeks to maximize the minimal gain arising from mis-specification of the nominal values from the plausible set.

Alternatively, the experimenter can elicit a prior distribution on the parameter vector and generate a Bayesian D-optimal design. A Bayesian D-optimal design maximizes the logarithm of the determinant of the Fisher information averaged over a user-selected prior distribution for the parameters. This option can be unattractive in practice because the numerical integration involved is frequently costly in terms of computation time. Less computationally intensive methods have been proposed to approximate the expected Fisher information, including work by Gotwalt, Jones, and Steinberg (2009) and Overstall and Woods (2017), and other methods such as compromise designs (Woods et al. 2006) may also be used.

The pilot study by Wang et al. (2017) provides us with a set of nominal values for the odor removal experiment, which we can use to generate a locally D-optimal design. The caveat is that a nominal value for the coefficient of the temperature variable is still required since this variable was set to a fixed value in the earlier study. Accordingly we make a sensible guess for its value and ascertain whether the locally D-optimal design is sensitive to its mis-specification. Section 4.2 shows the efficiency of a locally D-optimal design can be seriously reduced if this guess is dramatically different from the true value.

We compare the performance of one design relative to another using the D-efficiency. For two designs, ξ_1 and ξ_2, the D-efficiency of ξ_1 relative to ξ_2 is

$$\text{eff}_D(\xi_1, \xi_2) = \left(\frac{\det\{\mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\theta}}(\xi_1)\}}{\det\{\mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\theta}}(\xi_2)\}} \right)^{\frac{1}{p+J-1}}. \tag{4}$$

If the D-efficiency of design ξ_1 relative to another design ξ_2 is 1/3, then the design ξ_1 requires three times as many observations to do as well as ξ_2 in terms of estimating the model parameters. If ξ_2 is a locally D-optimal approximate design, then the ratio is the D-efficiency relative to the locally D-optimal approximate design.

Sensitivity functions allow us to evaluate the optimality of an approximate design when the design criterion is a concave or convex function on the space of all approximate designs. Atkinson and Woods (2015) provides examples of use of sensitivity functions for nonlinear models. For our models and D-optimality, it can be shown that the sensitivity function for the design expressed in terms of the $\mathbf{X}_i$'s in (1) is given by

$$tr\left(\mathbf{W}_i \mathbf{X}_i^T \left\{ \sum_{j=1}^{L} w_j \mathbf{X}_j^T \mathbf{W}_j \mathbf{X}_j \right\}^{-1} \mathbf{X}_i \right) - (p + J - 1) \tag{5}$$

(Perevozskaya, Rosenberger, and Haines 2003). If the sensitivity function in (5) is less than or equal to 0 for all possible $\mathbf{X}_i$, with equality at the support points of the design, the design is a locally D-optimal approximate design. Such reasoning is based on the equivalence theorem for a convex functional and is derived by studying the directional derivative of a convex functional, see for example, Fedorov (1972) or Silvey (1980). The major use of the equivalence theorem is to confirm optimality of an approximate design among all possible designs on the user selected compact design space.

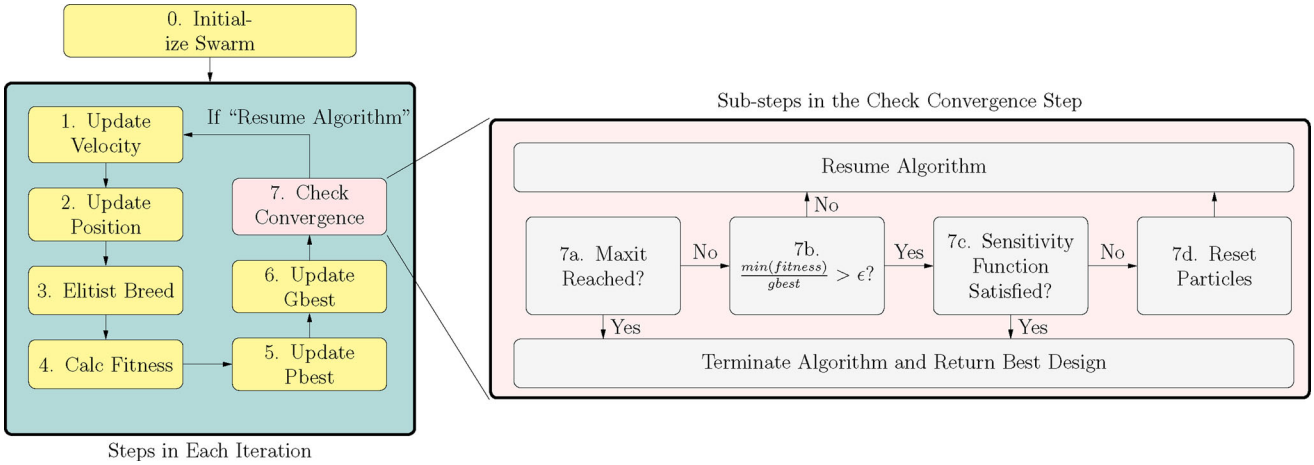

Figure 2. Diagram of the steps in the PSO algorithm. The sub-steps of the "Check Convergence" step are enlarged and presented in the red box.

If the design does not satisfy the equivalence theorem (i.e. there is a $\mathbf{X}_i$ in the design space such that the sensitivity function is positive), the current design is not locally D-optimal. In this case, its proximity to the unknown optimum can be measured via a lower bound on the D-efficiency of the design relative to the theoretical locally D-optimal approximate design. Specifically, let δ be the maximum positive value of the sensitivity function across the design space. A lower bound on the D-efficiency of the design ξ is given by $\exp\{-\delta/(p+J-1)\}$ (Pázman 1986), and we use this lower bound as one of the stopping criteria in our algorithm.

3. Particle swarm optimization

PSO is an optimization approach based on the behavior of a flock of birds searching an area for a food source, where the quality of the food source corresponds to the fitness of the solution (Eberhart and Kennedy 1995). The key idea is that as the flock of birds searches an area, each bird has some idea of where the best food source is based on its own previous experience in the field (cognitive knowledge). This position is known as the *pbest* position and the corresponding fitness is denoted the *pbest*. Additionally, each bird has some idea of where the best food source is based on the behavior of the other birds (social knowledge); if many birds have selected a location then it is likely that there is a good food supply at that location. The overall best position found by any bird is known as the *gbest* position, with its fitness labeled the *gbest*.

In PSO, birds are drawn to both their *pbest* position and the *gbest* position at each iteration, and they "fly" through the search space at a velocity determined by their distance from these two positions. In doing so, the birds get the chance to explore many new solutions between the *pbest* and *gbest* positions, hopefully identifying better solutions. Some of the most appealing features of PSO are that it is virtually assumption free, flexible, and easy to use. For example, the behavior of the swarm is governed by two simple equations: the velocity update and the position update equations. As such, the PSO algorithm and its many variants have been widely applied to many types of optimization problems in engineering and computer science research, and increasingly in many other disciplines.

PSO has become popular in recent years due to its ease of implementation and its ability to find high quality solutions for complicated optimization problems. Variants of PSO have been applied to solve a number of optimal design problems (Wong et al. 2015; Lee et al. 2018; Chen et al. 2015). However, to our knowledge, this is the first work to use a nature-inspired metaheuristic algorithm to find designs for experiments with mixed factors and ordinal outcomes modeled using a cumulative logit link function. We selected PSO to find optimal designs not only because of its speed and frequent successes in generating quality solutions to complication optimization problems, but also for its ability to search a continuous search space without requiring the space to be discritized. The latter capability is especially useful for finding locally D-optimal approximate designs for experiments with mixed factors.

We now describe our PSO algorithm, schematically depicted in Figure 2, for obtaining D-optimal approximate designs for experiments with ordinal response modeled under a cumulative logit model. To ease the exposition, we assume that all discrete factors have two levels as in the odor removal experiment.

Step 0: Initialization: The swarm is initialized with s particles, $\xi_1, \ldots, \xi_s$, where each particle is a candidate design, and s is user-determined. The initialization is random, thus random factor settings and proportions allocated to each support point are generated for each particle, where the starting proportions are normalized to sum to unity. Because each particle is a design, we can calculate the log determinant of the Fisher information matrix for each particle. This value is the starting fitness, as well as the starting *pbest* fitness. The initial *gbest* fitness is the maximum of all *pbest* fitness values. The initial velocity is set to 0 for all positions.

Step 1: Update velocity: The velocity for particle i at position k, $v_{i,k}, k = 1, \ldots, K$ is updated at iteration t according to:

$$v_{i,k}^{(t)} = \phi_d v_{i,k}^{(t-1)} + U_1 \phi_c (\xi_{i,k,pbest} - \xi_{i,k}^{(t-1)}) + U_2 \phi_s (\xi_{k,gbest} - \xi_{i,k}^{(t-1)}), \quad (6)$$

where U_1, U_2 are independent random draws from a continuous Uniform(0, 1) distribution, ϕ_d is the inertia factor, and ϕ_c and ϕ_s are the cognitive and social learning factors, respectively. We take $\phi_c = \phi_s = 2$ as is common in the literature (Eberhart and Kennedy 1995). Using these settings allows the particle's own experience to have equal weight to the global best. The inertia factor is slowly reduced from 1 to 0.4 over the course of the iterations. After the velocity update, any velocity that exceeds the width of a factor's range is set to be equal to that width. For example, if a factor can take values in $[-2, 2]$, and a particle's velocity for a setting of that factor is 5, then the velocity will be trimmed down to 4 (Eberhart and Shi 2000).

Step 2: Update position: The position update is performed differently for discrete factors and continuous factors. If k indexes a continuous factor or a proportion of experimental units assigned to a support point, then the position is updated as:

$$\xi_{i,k}^{(t+1)} = \xi_{i,k}^{(t)} + v_{i,k}^{(t+1)}, \quad (7)$$

and any positions exceeding the factor's boundaries are set to the boundary. For example, in the odor removal experiment, if Eq. [7] results in a temperature value of 40, that value will instead be set to 35 and the velocity will be reduced accordingly. The proportions are normalized to sum to unity, and proportions that are equal to zero correspond to unused support points. This means that PSO can find designs with fewer than L support points.

The discrete factors are updated one support point at a time using Hamming distance as described by Xi et al. (2016). For each support point, a reference point, $\mathbf{x}_r$, is created using crossover from the *pbest* and the *gbest* positions. For a randomly selected crossover point, all settings to the left of the crossover point come from the particle's *pbest* position for that support point. The setting at the crossover point and all factors to its right come from the *gbest* discrete factor settings for that support point. The crossover point is obtained using a randomly drawn index.

After this reference support point is created, we count the number of elements that are different between the particle's position and the reference point. The probability of a factor changing setting (e.g. from -1 to 1) is then taken to be the ratio of the number of different elements to the total number of discrete factors. Under this updating scheme, discrete factor settings that are identical to the reference point do not change at all. On the other hand, if the discrete factor settings of a support point do differ from the reference point, then each discrete factor setting for that support point has equal probability of changing, regardless of whether the individual point was equal to the reference setting.

Step 3: Elitist breeding: Elitist breeding for PSO was first suggested by Yang, Wu, and Min (2015) as a technique for encouraging exploration of the search space by "breeding" current solutions with other, better solutions. It was applied to PSO for finding locally D-optimal designs by Lukemire, Mandal, and Wong (2019) and is used here in a similar manner. In our implementation, for each particle, each support point has probability 1/3 of breeding at each iteration. If a support point is selected for breeding, then it is replaced by a randomly selected support point from that particle's *pbest* position with probability 1/2, or from the *gbest* position with probability 1/2. This operation helps to prevent the swarm from becoming stuck in a local maxima, and it also acts to slow the swarm from converging prematurely. We address this problem via our convergence criteria described below.

Step 4: Calculate fitness: At iteration t, if a locally D-optimal design is sought then a particle's fitness is the logarithm of the determinant of the Fisher information matrix for the design given by its position. Alternatively, if a pseudo-Bayesian design is sought, the fitness is the approximation to the expected logarithm of the determinant of the Fisher information as described in Gotwalt, Jones, and Steinberg (2009).

Steps 5–6: Update bests: For each particle, if the fitness at iteration t is greater than that particle's *pbest* from the last $t-1$ iterations, then the *pbest* fitness is updated to the new fitness and the *pbest* position is changed to the particle's current position. Similarly,

the *gbest* fitness and position at iteration t are taken to be the best of all *pbest*s at iteration t.

Step 7: Check convergence: Eq. [6] implies that as all of the particles approach the same position, their velocities will go to zero. Therefore we can measure convergence of the swarm by detecting when all of the particles have stopped moving. However, because we have implemented the elitist breeding mutation, it is likely that at least a few particles will always have non-zero velocity due to the breeding mechanism replacing some of their support points. This leads us to instead measure convergence by looking at the ratio of the smallest *pbest* fitness to the *gbest* fitness. If this ratio is larger than 0.995 then we say the swam has converged and proceed to verify the local D-optimality of the *gbest* design using the sensitivity function.

Our PSO codes also check the sensitivity function (Eq. [5]) for the *gbest* position using a grid search over the continuous factor settings for each possible combination of discrete factor settings. If the design is locally D-optimal, then the value of the sensitivity function will be 0 at the support points and less than 0 elsewhere in the design space. We calculate the maximum value of the sensitivity function for the *gbest* position and obtain a lower bound of the efficiency relative to the locally D-optimal approximate design as described in Section 2. If this lower bound is greater than ϵ, then we say that the swarm has converged to the locally D-optimal approximate design and terminate the search. Note that when searching for pseudo-Bayesian designs this stopping criteria is not used, and instead the user-specified maximum number of iterations is used.

If the sensitivity function check fails, then the *gbest* results are set aside and the swarm is reset (re-initialized) in hopes that the new positions will allow the swarm to search a new part of the search space. The re-initialized swarm is not "aware" of the previous *gbest* results until it converges, at which point the old *gbest* is reintroduced and the swarm is allowed to converge again, this time with knowledge of the old *gbest*. The swarm terminates when the locally D-optimal design has been found or when the maximum number of iterations has been reached, and the *gbest* design found by the swarm across all re-initializations is returned as the PSO-generated approximate design.

Like all meta-heuristic algorithms, an important consideration is the choice of tuning parameters. Tuning the PSO algorithm requires choosing the number of particles, the number of support points, the maximum number of iterations (maxit), and the lower bound, ϵ, required to terminate the algorithm. In our experience, we have found that generally 10–20 particles works well. For the number of support points, we have found that setting $L = 2 \times (p + J - 1)$ is generally adequate. PSO may generate a design with fewer than L support points by setting some of the weights to 0. We recommend using a maximum of 20,000–50,000 iterations as a starting point for more complicated designs and then increasing this number as needed. For our purposes we generally set $\epsilon = 0.99$, however this is rather strict and can be relaxed if the algorithm fails to find a locally D-optimal approximate design. Recall that this value is a lower bound, and in practice the efficiency of the PSO-generated design relative to the locally D-optimal approximate design might be significantly higher.

4. Odor removal experiment

4.1. Locally D-optimal approximate design

In this section, we obtain a locally D-optimal approximate design for the odor removal experiment that motivated this work. The nominal values for the discrete factors and the cutpoints are taken from Wang et al. (2017) and are displayed in Table 2. As discussed earlier, temperature was not studied in the original experiment, thus we must select a reasonable value. We take the nominal value for the temperature effect to be 0.2, as we expect that a warmer storage temperature will result in a stronger odor in the final product.

We used PSO with 20 particles and the nominal parameter values in Table 2 to search for a locally D-optimal approximate design with up to $L = 20$ support points. We set $\epsilon = 0.99$, meaning that if the generated design has a minimum D-efficiency of 99 percent, we will treat the design as optimal in practice.

Table 2. The nominal values used in obtaining a locally D-optimal approximate design for the odor removal experiment, and the independent priors used for the pseudo-Bayesian D-optimal approximate design. U indicates a uniform distribution.

Parameter	Nominal Value	Prior	Outcome Cutpoint	Nominal Value	Prior
Algae	2.890	U(1.5, 3.0)	θ_1	−4.270	U(-5.0, −4.0)
Scavenger	0.841	U(0.25, 1.75)	θ_2	0.362	U(0.0, 1.0)
Resin	−1.476	U(-2.0, −0.5)	θ_3	3.309	U(3.0, 4.0)
Compatibilizer	−0.024	U(-0.75, 0.75)	θ_4	5.451	U(5.0, 6.0)
Temperature	0.200	U(-1.0, 1.5)			

Alternatively, we can also run the algorithm up to 20,000 iterations as another stopping criterion. Table 3 displays the 13 point PSO-generated locally D-optimal approximate design for estimating all nine parameters in the model. This design was generated in a little under 10 seconds. The determinant of the Fisher information matrix is 1.51×10^{-6}. Figure 3 plots the sensitivity function for each combination of discrete factor settings, and the graph confirms the local D-optimality of the design.

Table 3 illustrates several important aspects of the optimal design with mixed factors. First, the locally D-optimal design can be generated using only 13 points. This is advantageous because in practice it can be quite costly to change design settings. Second, the smallest proportion allocated to a support point is 0.03, or roughly 3 percent. The larger the proportion assigned to each support point, the easier it is to implement the approximate design; otherwise allocation proportions with extremely small values will require large sample sizes to implement. Most importantly, many of the optimal continuous factor settings do not require values at either extreme of the temperature range. These points could easily be missed if the design space was discritized using a coarse grid.

Another popular approach for finding locally D-optimal designs is the coordinate exchange algorithm (Meyer and Nachtsheim 1995; Gotwalt, Jones, and Steinberg 2009), and hence we use it as a baseline for comparing with PSO. This algorithm performs a greedy search, cycling through each factor setting for each design point and finding the optimal setting. Apart from being intuitive and easy to follow, another advantage of this approach is that it is very fast, and thus it can be used to quickly find designs. The downside is that it is quite vulnerable to getting stuck in local minima. To demonstrate, we use the coordinate exchange algorithm to attempt to find a design for the odor removal experiment. We consider designs of size 12–18 points, and for each design size we run the coordinate exchange algorithm using 1000 different starting designs, as recommended in Goos and Jones (2011). The best and average determinants of the Fisher information matrix across the 1000 starting designs are listed in Table 4. From the results, it is clear that the coordinate exchange algorithm is not able to find designs that are as efficient as those from PSO.

Table 3. The 13-point PSO-generated locally D-optimal approximate design for the odor removal experiment. The determinant of the Fisher information matrix for this design is 1.51×10^{-6}.

Algae	Scav.	Resin	Comp.	Temp.	w_i
−1.0	−1.0	1.0	−1.0	5.0	0.06
1.0	1.0	1.0	1.0	14.78	0.08
1.0	−1.0	−1.0	−1.0	5.0	0.12
−1.0	−1.0	−1.0	1.0	35.0	0.12
1.0	1.0	1.0	−1.0	5.08	0.05
−1.0	1.0	−1.0	1.0	5.0	0.10
−1.0	−1.0	1.0	−1.0	30.07	0.10
−1.0	1.0	1.0	1.0	35.0	0.04
−1.0	1.0	1.0	−1.0	35.0	0.07
−1.0	1.0	−1.0	−1.0	5.0	0.06
−1.0	1.0	−1.0	−1.0	30.03	0.03
1.0	−1.0	1.0	1.0	5.0	0.10
−1.0	−1.0	1.0	1.0	5.0	0.07

4.2. Robustness to mis-specification in the effect of storage temperature

The nominal parameter values for the discrete factor settings and cut-points were based on a pilot study. However, we did not have an estimate for the effect of storage temperature from the pilot study (β_5). Consequently, we should be most worried that our nominal value for this parameter will be incorrect. In this subsection, we examine the robustness of our design in Table 3, designated ξ_0, to mis-specification in the effect of storage temperature. We do this by varying the true value for the effect of storage temperature, generating locally D-optimal designs under the true value, and examining the D-efficiency of our design in Table 3 relative the locally D-optimal design for each considered value of β_5 using (5).

We first assume $\beta_5 \in [0, 1]$ and divide the interval into 100 equally spaced points. For each of these values of β_5, we use PSO with 20 particles to search for a locally D-optimal approximate design with up to $L = 20$ support points. We use $\epsilon = 0.99$ or 400,000 maximum iterations as the stopping criterion, whichever comes first. This procedure produces 100 locally D-optimal approximate designs, one for each choice of β_5. We then compare the relative D-efficiency of the PSO-generated locally D-optimal design with nominal value $\beta_5 = 0.2$ relative to its true value which could be any value in $[0, 1]$ using (4). Figure 4 plots the D-efficiencies for varying amounts of mis-specification in the temperature parameter β_5. For a range of roughly ± 100 percent the nominal value, i.e. $(0, 0.4)$, the design is highly optimal. However, if we are very wrong about the magnitude of the effect, then the efficiency drops off steeply to as low as 20 percent.

4.3. Pseudo-Bayesian D-optimal approximate design

Section 4.2 demonstrated that mis-specification of the model parameters, even if it is just for one parameter,

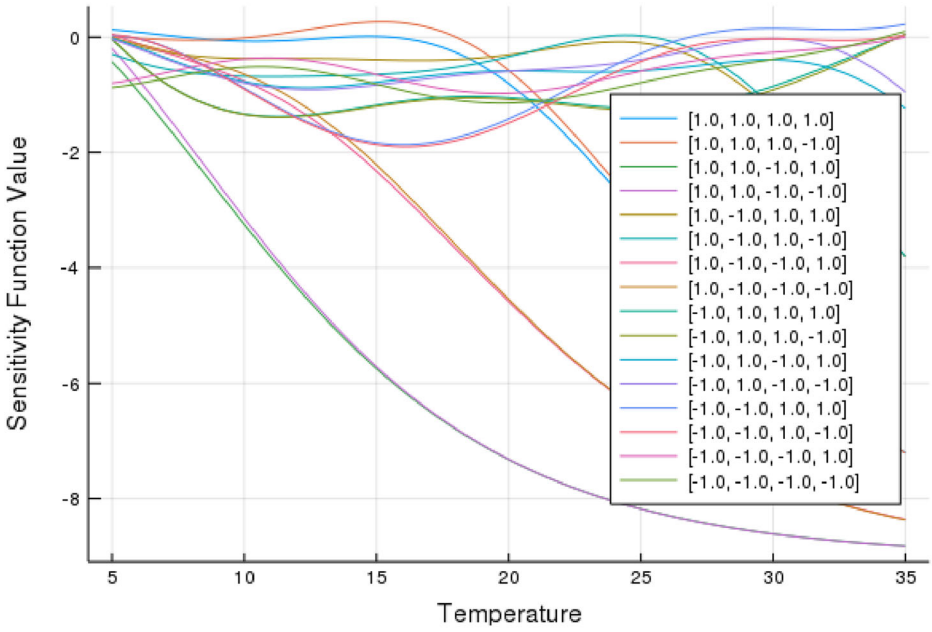

Figure 3. Plots of the sensitivity functions of the PSO-generated locally *D*-optimal approximate design for the odor removal experiment at different configurations of the discrete factors.

can have a noticeable effect on the quality of the design in practice. In general an experimenter will likely not have nominal values for some or all of the parameters, but likely will have varying amounts of information on the true values for each of the parameters. The experimenter then elicits a prior distribution for the model parameters.

Motivated by this problem we next use PSO to find a pseudo-Bayesian design for the odor removal experiment. We follow the technique outlined in Gotwalt, Jones, and Steinberg (2009) to find a pseudo-Bayesian design. Table 2 displays the priors for the model parameters. The priors are uniform and fairly wide, implying that we do not have reliable prior knowledge about the parameter values. We use PSO with 20 particles to search for a pseudo-Bayesian *D*-optimal approximate design with up to $L = 20$ support points. We use 3,000 maximum iterations as the stopping criterion. Table 5 shows the 18 point design found by PSO. Searches for Bayesian designs are more computationally intensive. In this case, the search algorithm required about 11 minutes to find the design.

4.4. Robustness to mis-specification of all model parameters

Next, we conduct a robustness study to compare the locally *D*-optimal approximate design from Table 3 with the pseudo-Bayesian approximate design in Table 5. In particular, we are interested in the relative performance of the locally *D*-optimal approximate design

Table 4. Summary of the results from 1000 runs of the coordinate exchange algorithm. The time reported is the time across all 1000 runs for the corresponding number of design points. For comparison, the determinant of the Fisher information matrix for the PSO-generated design is 1.51×10^{-6}.

Number of Design Points	Best Determinant	Average Determinant	Time (seconds)
12	1.28×10^{-6}	1.11×10^{-6}	28.06
13	1.34×10^{-6}	1.18×10^{-6}	33.38
14	1.42×10^{-6}	1.22×10^{-6}	36.75
15	1.40×10^{-6}	1.25×10^{-6}	42.18
16	1.42×10^{-6}	1.28×10^{-6}	46.57
17	1.45×10^{-6}	1.31×10^{-6}	52.51
18	1.45×10^{-6}	1.33×10^{-6}	58.80

to the pseudo-Bayesian approximate design over a wide range of possible parameter values. To fix ideas, let ξ_l be the locally *D*-optimal approximate design in Table 3 and let ξ_b be the pseudo-Bayesian design in Table 5. Our procedure is as follows. First, we draw a set of parameters from the prior in Table 2, $\boldsymbol{\Theta}^{(m)} = \{\boldsymbol{\beta}^{(m)^T}, \boldsymbol{\theta}^{(m)^T}\}^T, m = 1, \ldots, 1000$. Then, for each draw m, we calculate the relative efficiency of ξ_l to ξ_b at $\boldsymbol{\Theta}^{(m)}$ as

$$\gamma^{(m)} = \left(\frac{\det\{\mathbf{I}_{\boldsymbol{\beta}^{(m)}, \boldsymbol{\theta}^{(m)}}(\xi_l)\}}{\det\{\mathbf{I}_{\boldsymbol{\beta}^{(m)}, \boldsymbol{\theta}^{(m)}}(\xi_b)\}} \right)^{\frac{1}{9}}.$$

Values of $\gamma^{(m)} < 1$ indicate that ξ_b is more efficient than ξ_l when $\boldsymbol{\Theta}^{(m)}$ is the true parameter vector, and $\gamma^{(m)} > 1$ indicates the opposite.

Figure 5 displays a histogram of the $\gamma^{(m)}$ values. From the histogram, we can see that in the majority of cases the pseudo-Bayesian design outperforms the

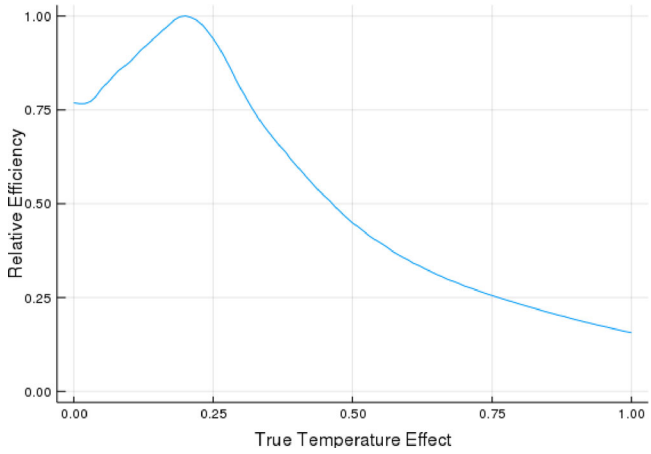

Figure 4. *D*-efficiencies of the PSO-generated locally *D*-optimal approximate design for the odor removal experiment assuming $\beta_5 = 0.2$ relative to the locally *D*-optimal approximate design when the true value for $\beta_5 \in [0, 1]$.

locally *D*-optimal approximate design. However, there are some cases in which the locally *D*-optimal approximate design is able to significantly outperform the pseudo-Bayesian design. In practice, when choosing between two designs, it is important that a researcher examines robustness to parameter mis-specification over a range of plausible values before deciding which design to use.

5. Minimizing surface defects

Wu (2008) describes a study for manufacturing circuits for the optimal parameter settings of a polysilicon deposition process with the six factors in Table 6. The original experiment discretized all of the continuous factors at three levels. The experiment had 5 outcomes, labeled as Category I through Category V. We selected this example from the literature because the odor-removal experiment has only one continuous factor, and thus we want to investigate PSO's utility as the number of continuous factors increases. Finding a locally *D*-optimal design for this experiment is quite difficult, as we need to find the optimal settings for five continuous factors. We used PSO with 20 particles and the nominal parameter values in Table 6 to search for a locally *D*-optimal approximate design with up to $L = 24$ support points. We set 50,000 maximum iterations as the stopping criterion and the set of nominal values we used for the four cutpoints is $\{\theta_1, \theta_2, \theta_3, \theta_4\}^T = \{-1.113, 0.183, 1.518, 2.639\}^T$ based on the experiment in Wu (2008). Table 7 displays the 14-point PSO-generated locally *D*-optimal approximate design with the determinant for its Fisher information matrix. This design was generated in about 25 seconds.

Table 5. The 18-point PSO-generated pseudo-Bayesian approximate design for the odor removal experiment.

Algae	Scav.	Resin	Comp.	Temp.	w_i
1.0	1.0	−1.0	1.0	5.0	0.10
1.0	1.0	1.0	1.0	5.0	0.03
−1.0	1.0	1.0	1.0	11.76	0.04
−1.0	−1.0	1.0	1.0	5.0	0.06
1.0	1.0	1.0	−1.0	5.0	0.10
1.0	−1.0	−1.0	−1.0	34.98	0.02
−1.0	1.0	−1.0	−1.0	5.0	0.05
−1.0	−1.0	−1.0	−1.0	10.47	0.05
−1.0	1.0	−1.0	1.0	5.0	0.08
1.0	−1.0	1.0	1.0	7.68	0.03
−1.0	1.0	1.0	−1.0	28.79	0.02
−1.0	1.0	1.0	−1.0	8.38	0.04
−1.0	−1.0	1.0	−1.0	5.0	0.08
−1.0	−1.0	1.0	1.0	32.37	0.05
−1.0	1.0	−1.0	−1.0	15.13	0.02
1.0	−1.0	−1.0	−1.0	5.0	0.11
1.0	−1.0	1.0	1.0	5.0	0.07
−1.0	−1.0	−1.0	1.0	5.0	0.04

It is not feasible to plot the six-dimensional sensitivity function for this design across the whole design space, and even if we could, it is hard to identify important features from the high dimensional graph. This keeps us from being able to verify visually whether this approximate design is a locally *D*-optimal. Instead, we conduct a simulation study by generating 100 designs using PSO with the same settings as above and examine the distribution of these designs' efficiencies relative to the design reported in Table 7. If the efficiencies are generally close to unity, we can be confident that the design we have found is locally *D*-optimal, or is at least highly efficient. On the other hand, if the relative efficiencies are highly variable, then we know that we are likely doing a poor job exploring the search space and thus we should be concerned about the quality of our design.

Figure 6 displays a histogram of the resulting relative efficiencies. It is clear from the image that all 100 designs we found are nearly as efficient as the design in Table 7. This allows us to be reasonably confident that the design is locally *D*-optimal, or at least highly efficient.

6. Discussion

In this work we proposed a modified PSO algorithm to search for locally *D*-optimal and pseudo-Bayesian approximate design of experiments with mixed factors and an ordinal response. The key advantages of PSO are that (i) codes are widely available and easy to use, (ii) it is generally fast and free of assumptions, and (iii) it is a general purpose optimization algorithm that can be used to solve many optimization problems in science and engineering. Unlike other algorithms, PSO does not require the design space to be discretized and so it can capture more information from

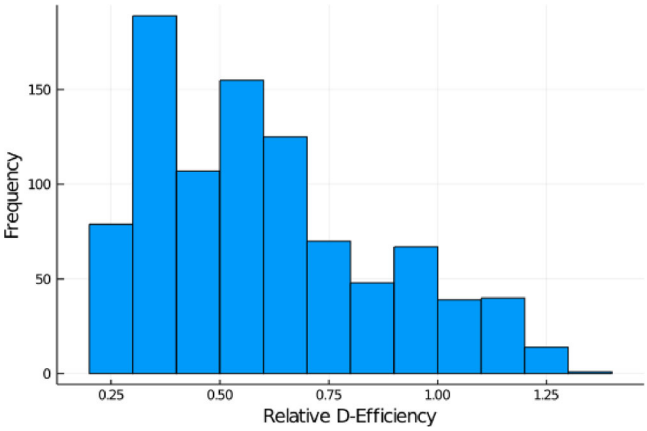

Figure 5. *D*-efficiencies of the PSO-generated locally *D*-optimal approximate design for the odor removal experiment relative to the pseudo-Bayesian design for 1000 draws from the prior vector in Table 2.

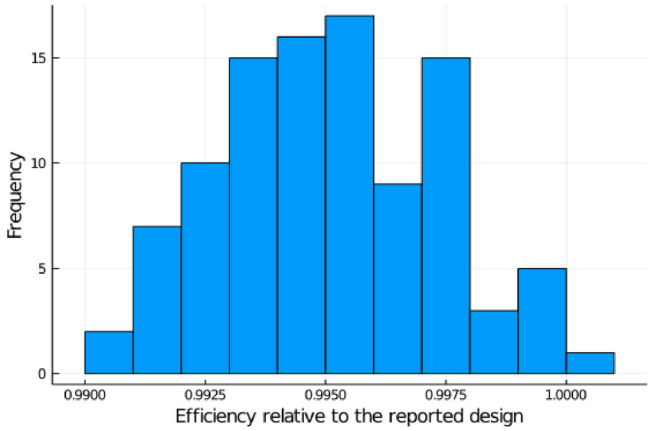

Figure 6. Efficiencies of the 100 PSO-generated designs for the surface defects experiment relative to the design reported in Table 7.

Table 6. The six factors of interest in the surface defects experiment and their nominal parameter values. CM is an abbreviation for "cleaning method." Continuous factor ranges are relative to a base value.

		Levels		Nominal Value
Type	Factor	−	+	
Discrete	Cleaning Method	CM1	CM2	−0.970
Continuous	Temperature	−25 to 25		0.077
Continuous	Pressure	−200 to 200		0.008
Continuous	Nitrogen Flow	−150 to 0		−0.007
Continuous	Silane Flow	−100 to 0		0.007
Continuous	Setting Time	0 to 16		0.056

Table 7. The 14 point PSO-generated locally *D*-optimal approximate design for the surface defects experiment. The determinant of the Fisher information matrix for this design is 6.71×10^9.

Cleaning Method	Temp.	Pressure	Nit. Flow	Sil. Flow	Set. Time	w_i
−1.0	−25.0	199.96	−150.0	−100.0	16.0	0.05
1.0	23.14	196.35	0.0	−100.0	16.0	0.07
1.0	−24.99	199.99	−150.0	0.0	16.0	0.07
−1.0	25.0	−200.0	0.0	0.0	16.0	0.11
−1.0	−25.0	−200.0	−150.0	0.0	16.0	0.09
−1.0	25.0	−199.96	−150.0	−100.0	0.0	0.09
1.0	7.46	199.98	−150.0	0.0	0.0	0.10
1.0	25.0	−199.97	0.0	0.0	0.0	0.09
1.0	25.0	−200.0	−150.0	−100.0	0.0	0.05
1.0	25.0	−200.0	−149.98	−100.0	16.0	0.05
1.0	−24.48	200.0	0.0	−100.0	16.0	0.07
−1.0	−25.0	200.0	0.0	−100.0	0.0	0.09
−1.0	25.0	−199.78	0.0	−100.0	0.0	0.01
−1.0	−25.0	200.0	0.0	−0.01	0.0	0.06

the response surface than algorithms that require the space to be discretized.

For this paper, we focus on applying PSO to find optimal designs for an ordinal response model. It is a flexible algorithm in the sense that if the link function is not the cumulative logit link function, the codes can be easily amended to generate and compare the resulting optimal designs. Likewise, PSO can be directly amended to find optimal designs for models with nominal response categories. Further demonstration of PSO flexibility for finding different types of optimal designs can be found in Wong et al. (2015), where they found *D*-optimal designs for various models in mixture experiments and in Chen et al. (2017), where they found various types of maximin optimal designs for enzyme-kinetic models. We applied this algorithm to re-design a bioplastic odor removal study with nine parameters without fixing the settings of an important continuous factor in the study. We investigated the robustness of this design to mis-specification in the temperature value. We also demonstrated PSO's ability to find a locally *D*-optimal approximate design for a surface defects example with ordinal response, one discrete factor, and five continuous factors, which is a difficult search problem. While the emphasis of this paper was on experiments with mixed factors, the method is also applicable to experiments with only discrete or continuous factors. In the Supplementary Materials, we provide an application of our PSO approach to find other kinds of designs for dose-response experiments, and show that it can produce more efficient designs than those in the literature.

About the authors

Joshua Lukemire is a Ph.D. candidate in Biostatistics at Emory University. He received his bachelor's in Statistics from the University of Georgia, and his master's degree in Biostatistics from Emory University. His research interests include design of experiments and neuroimaging statistics.

Abhyuday Mandal is a Professor and the Undergraduate Coordinator of the University of Georgia, Department of Statistics. He received his bachelor's and master's degrees from Indian Statistical Institute, Kolkata, another master's

degree from University of Michigan and Ph.D. in 2005 from Georgia Institute of Technology. His research interests include design of experiments, optimization techniques and survey sampling. Currently he is an Associate Editor of *Sankhya - Series B, Statistics and Probability Letters* and *Journal of Statistical Theory and Practice*.

Weng Kee Wong is a Professor of Biostatistics at UCLA Fielding School of Public Health. He is a Fellow of the Institute of Mathematical Statistics, a Fellow of the American Statistical Association, a Fellow of the American Association for the Advancement of Science and an elected member of the International Statistical Institute. He received his Ph.D. in Statistics from the University of Minnesota in 1990 and has served as Principal Investigator on multiple R01 grants from NIH. He is currently associate editor for multiple statistics journals and his current research area is in use of metaheuristics to find various types of optimal designs for challenging and high-dimensional problems in the biomedical and industrial areas.

Acknowledgments

The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding

The research of Dr. Wong was partially supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639.

References

Agresti, A. 2003. Logit models for multinomial responses. In *Categorical data analysis*, 2nd ed., 267–313. Hoboken, NJ: Wiley.

Atkinson, A. C., and D. C. Woods. 2015. Designs for generalized linear models. In *Handbook of design and analysis of experiments*, ed. A. Dean, M. Morris, J. Stufken and D. Bingham, 471–514. Boca Raton, FL: Chapman & Hall/CRC.

Chen, P.-Y., R.-B. Chen, H.-C. Tung, and W. K. Wong. 2017. Standardized maximim D-optimal designs for enzyme kinetic inhibition models. *Chemometrics and Intelligent Laboratory Systems: An International Journal Sponsored by the Chemometrics Society* 169:79–86. doi: 10.1016/j.chemolab.2017.08.009.

Chen, R. B., S. P. Chang, W. Wang, H. C. Tung, and W. K. Wong. 2015. Minimax optimal designs via particle swarm optimization methods. *Statistics and Computing* 25 (5): 975–88. doi: 10.1007/s11222-014-9466-0.

Chernoff, H. 1953. Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics* 24 (4): 586–602. doi: 10.1214/aoms/1177728915.

Clark, N., M. Gong, M. Schork, N. Kaciroti, D. Evans, D. Roloff, M. Hurwitz, L. Maiman, and R. Mellins. 2000. Long-term effects of asthma education for physicians on patient satisfaction and use of health services. *The European Respiratory Journal* 16 (1):15–21. doi: 10.1034/j.1399-3003.2000.16a04.x.

Dette, H. 2004. On robust and efficient designs for risk estimation in epidemiological studies. *Scandinavian Journal of Statistics* 31 (3):319–31. doi: 10.1111/j.1467-9469.2004.03_037.x.

Eberhart, R. C., and Y. Shi. 2000. Comparing inertia weights and constriction factors in particle swarm optimization. Proceedings of the 2000 Congress on Evolutionary Computation, vol. 1, IEEE, 84–88.

Eberhart, R., and J. Kennedy. 1995. A new optimizer using particle swarm theory. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995. MHS'95, IEEE, 39–43.

Fedorov, V. V. 1972. *Theory of optimal experiments*. Moscow, Russia: Elsevier.

Goos, P., and B. Jones. 2011. *Optimal design of experiments: A case study approach*. West Sussex, UK: John Wiley & Sons.

Gotwalt, C. M., B. A. Jones, and D. M. Steinberg. 2009. Fast computation of designs robust to parameter uncertainty for nonlinear settings. *Technometrics* 51 (1):88–95. doi: 10.1198/TECH.2009.0009.

Kiefer, J. 1959. Optimum experimental designs. *Journal of the Royal Statistical Society: Series B (Methodological)* 21 (2):272–319. doi: 10.1111/j.2517-6161.1959.tb00338.x.

Lall, R., M. Campbell, S. Walters, K. Morgan, and M. C. Co-Operative. 2002. A review of ordinal regression models applied on health-related quality of life assessments. *Statistical Methods in Medical Research* 11 (1):49–67. doi: 10.1191/0962280202sm271ra.

Lee, J. H., J. Y. Song, D. W. Kim, J. W. Kim, Y. J. Kim, and S. Y. Jung. 2018. Particle swarm optimization algorithm with intelligent particle number control for optimal design of electric machines. *IEEE Transactions on Industrial Electronics* 65 (2):1791–8. doi: 10.1109/TIE.2017.2760838.

Lukemire, J., A. Mandal, and W. K. Wong. 2019. d-QPSO: A quantum-behaved particle swarm technique for finding D-optimal designs with discrete and continuous factors and a binary response. *Technometrics* 61 (1):77–87. doi: 10.1080/00401706.2018.1439405.

Meyer, R. K., and C. J. Nachtsheim. 1995. The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* 37 (1):60–9. doi: 10.1080/00401706.1995.10485889.

Nair, V. N. 1986. Testing in industrial experiments with ordered categorical data. *Technometrics* 28 (4):283–91. doi: 10.2307/1268974.

Overstall, A. M., and D. C. Woods. 2017. Bayesian design of experiments using approximate coordinate exchange. *Technometrics* 59 (4):458–70. doi: 10.1080/00401706.2016.1251495.

Pázman, A. 1986. *Foundations of optimum experimental design*. Netherlands: Springer.

Perevozskaya, I., W. F. Rosenberger, and L. M. Haines. 2003. Optimal design for the proportional odds model. *Canadian Journal of Statistics* 31 (2):225–35. doi: 10.2307/3316068.

Russell, K. G., D. C. Woods, S. Lewis, and J. Eccleston. 2009. D-optimal designs for poisson regression models. *Statistica Sinica* 19 (2):721–30.

Silvey, S. 1980. *Optimal design*. London: Chapman & Hall.

Wang, K., A. Mandal, E. Ayton, R. Hunt, M. Zeller, and S. Sharma. 2017. Modification of protein rich algal-biomass to form bioplastics and odor removal. In *Protein byproducts*, 107–17. London, UK: Elsevier.

Wong, W. K., R. B. Chen, C. C. Huang, and W. Wang. 2015. A modified particle swarm optimization technique for finding optimal designs for mixture models. *PLoS One* 10 (6):e0124720 doi: 10.1371/journal.pone.0124720.

Woods, D. C., S. M. Lewis, J. A. Eccleston, and K. Russell. 2006. Designs for generalized linear models with several variables and model uncertainty. *Technometrics* 48 (2): 284–92. doi: 10.1198/004017005000000571.

Wu, F. C. 2008. Simultaneous optimization of robust design with quantitative and ordinal data. *International Journal of Industrial Engineering: Theory, Applications and Practice* 15 (2):231–8.

Wu, S., W. K. Wong, and C. M. Crespi. 2017. Maximin optimal designs for cluster randomized trials. *Biometrics* 73 (3):916–26. doi: 10.1111/biom.12659.

Xi, M., J. Sun, L. Liu, F. Fan, and X. Wu. 2016. Cancer feature selection and classification using a binary quantum-behaved particle swarm optimization and support vector machine. *Computational and Mathematical Methods in Medicine* 2016:1–9. doi: 10.1155/2016/3572705.

Xu, W., W. K. Wong, K. C. Tan, and J. X. Xu. 2019. Finding high-dimensional *d*-optimal designs for logistic models via differential evolution. *IEEE Access : Practical Innovations, Open Solutions* 7:7133–46. doi: 10.1109/ACCESS.2018.2890593.

Yang, J., A. Mandal, and D. Majumdar. 2012. Optimal designs for two-level factorial experiments with binary response. *Statistica Sinica* 22 (2):885–907. doi: 10.5705/ss.2010.080.

Yang, J., A. Mandal, and D. Majumdar. 2016. Optimal designs for 2^k factorial experiments with binary response. *Statistica Sinica* 26 (1):385–411.

Yang, J., L. Tong, and A. Mandal. 2018. D-optimal designs with ordered categorical data. *Statistica Sinica* 27 (4): 1879–902. doi: 10.5705/ss.202016.0210.

Yang, Z. L., A. Wu, and H. Q. Min. 2015. An improved quantum-behaved particle swarm optimization algorithm with elitist breeding for unconstrained optimization. *Computational Intelligence and Neuroscience* 2015:1–53. doi: 10.1155/2015/326431.

Zocchi, S. S., and A. C. Atkinson. 1999. Optimum experimental designs for multinomial logistic models. *Biometrics* 55 (2):437–44. doi: 10.1111/j.0006-341x.1999.00437.x.