Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2014
August 17-20, 2014, New York, United States

DETC2014-34790

ENHANCED ADAPTIVE CHOICE-BASED CONJOINT ANALYSIS INCORPORATING ENGINEERING KNOWLEDGE

Yi Ren*
Mechanical Engineering
University of Michigan
Ann Arbor, Michigan, 48109
Email: yiren@umich.edu

Panos Y. Papalambros
Mechanical Engineering
University of Michigan
Ann Arbor, Michigan, 48109
Email: pyp@umich.edu

ABSTRACT

Conjoint analysis from marketing has been successfully integrated with engineering analysis in design for market systems. The long questionnaires needed for conjoint analysis in relatively complex design decisions can become cumbersome to the human respondents. This paper presents an adaptive questionnaire generation strategy that uses active learning and allows incorporation of engineering knowledge in order to identify efficiently designs with high probability to be optimal. The strategy is based on viewing optimal design as a group identification problem. A running example demonstrates that a good estimation of consumer preference is not always necessary for finding the optimal design and that conjoint analysis could be configured more effectively for the specific purpose of design optimization. Extending the proposed method beyond a homogeneous preference model and noiseless user responses is also discussed.

1 Introduction

Conjoint analysis [1, 2] has been widely studied and practiced in the marketing and design communities where statistical models of consumer preferences are employed. Research to improve theoretical and practical issues with conjoint analysis is actively pursued. For example, research on heterogeneous consumer preference models based on questionnaire responses has addressed modeling formulations (e.g., nested logit [3], mixed logit [4], mixture model [5], consideration sets [6]), efficiency in constructing the models (e.g., through convex optimization [7] and variational methods [8]), and design of the questionnaires such as adaptive questionnaires [9–11].

Such developments in conjoint analysis are mainly motivated by marketing questions such as how to achieve a good estimation of human preferences within limited resources. In design for market systems, the goal is not the preference model itself that is typically incorporated in a demand model within a profit maximization objective. Rather, the goal is the identification of the optimal design to bring to the market, accounting for business and engineering considerations at the same time. Therefore, high accuracy of the preference model per se may not be important or may be important only at relevant points in the design space, for example, points visited by the optimization algorithm and eventually the optimal design. This approach is similar to the use of local and global models in surrogate-based optimization [12] and in design under uncertainty [13]. This idea of incorporating engineering and marketing models in conjoint analysis was noted in Feit [14], where the author proposed a design of experiments in place of D-optimal design for maximizing the expected profit. That approach was found to be computationally intractable and no simulation or experimental results were reported.

In this paper, we propose an adaptive questionnaire directly for design optimization, motivated by existing theoretical work on active learning. Specifically, we develop a view of conjoint analysis as a group identification problem. We use a well-documented dial-readout scale design from Michalek et al. [15] as a running example to demonstrate that (i) a good estimation

*Address all correspondence to this author.

TABLE 1. Product Attributes and Price Levels from [15]

k	Description	Metric	Units	Levels				
z_1	Weight Capacity	Weight Causing a 360° Dial Turn	lbs	200	250	300	350	400
z_2	Aspect Ratio	Platform Length Divided by Width	-	6/8	7/8	8/8	8/7	8/6
z_3	Platform Area	Platform Length Times Width	in.^{2}	100	110	120	130	140
z_4	Tick Mark Gap	Distance between 1-lb Tick Marks	in.	2/32	3/32	4/32	5/32	6/32
z_5	Number Size	Length of Readout Number	in.	0.75	1.00	1.25	1.50	1.75
p	Price	US Dollars	\$	10	15	20	25	30

of consumer preference is not necessary for finding the optimal design, and (ii) we can improve the effectiveness of questionnaires used for the purpose of design optimization. To simplify the discussion, we assume a homogeneous noiseless consumer preference model and a deterministic engineering model. In the remainder of the paper, Section 2 describes the running example, Sections 3 and 4 present the theory and algorithm for the proposed adaptive questionnaire, which is applied to the running example in Section 5, and Sections 6 and 7 offer discussion and conclusions, respectively.

2 Running Example: Dial-Readout Scale Design

We first introduce the well-documented scale design problem from [15] to show that finding the optimal design does not necessarily require an accurate consumer preference model, and then provide a more theoretical explanation for this argument.

The scale design problem consists of a profit model and an engineering model, as we elaborate below.

2.1 Profit model

The scale has six attributes that affect consumer purchase decisions: weight capacity (z_1), aspect ratio (z_2), platform area (z_3), tick mark gap (z_4), number size (z_5) and price (z_6). Discrete levels of these attributes are listed in Table 1. We denote a design by its binary attribute vector $\mathbf{z} := (z_{1,1}, \ldots, z_{1,5}, \ldots, z_{6,1}, \ldots, z_{6,5})$, where $z_{i,j} = 1$ if the ith attribute is at level j, or otherwise $z_{i,j} = 0$. A homogeneous population-level preference model was derived through a conjoint study in [15] and the resulting partworth vector $\mathbf{w}_0$ is presented in Table 2. While the true preference model is never known, these partworth values will be considered as the "true" partworths throughout the running example for demonstration purposes.

Let p be the price level induced by the binaries $(z_{6,1}, \ldots, z_{6,6})$ associated with the price, and $c_V = \$3$ be the constant manufacturing cost of the scale. The profit is proportional to the following function of design attributes:

$$\text{profit} \propto f(\mathbf{z}, \mathbf{w}) = \frac{p - c_V}{1 + \exp(-\mathbf{w}^T\mathbf{z})}, \tag{1}$$

TABLE 2. Partworth values ($\mathbf{w}_0$), from [15]

z_1		z_2		z_3	
200 lbs.	-0.534	100 in.^{2}	0.015	0.75 in.	-0.744
250 lbs.	0.129	110 in.^{2}	-0.098	1.00 in.	-0.198
300 lbs.	0.228	120 in.^{2}	0.049	1.25 in.	0.235
350 lbs.	0.104	130 in.^{2}	0.047	1.50 in.	0.291
400 lbs.	0.052	140 in.^{2}	-0.033	1.75 in.	0.396
z_4		z_5		p	
0.75	-0.058	2/32 in.	-0.366	\$10	0.719
0.88	0.253	3/32 in.	-0.164	\$15	0.482
1.00	0.278	4/32 in.	0.215	\$20	0.054
1.14	-0.025	5/32 in.	0.194	\$25	-0.368
1.33	-0.467	6/32 in.	0.100	\$30	-0.908

assuming no competing products exist in the market. In the case where competitors exist, the profit will be proportional to

$$\text{profit} \propto \frac{(p - c_V)\mathbf{w}^T\mathbf{z}}{1 + \mathbf{w}^T\mathbf{z} + \sum_{c=1}^{C} \exp(\mathbf{w}^T\mathbf{z}^{(c)})}, \tag{2}$$

where $\mathbf{z}^{(c)}$ for $c = 1, \ldots, C$ are the designs of the C competitors. Since addition of competitors will not affect the algorithmic development, the running example will use Equation (1) for profit calculation.

2.2 Engineering model and feasibility

The engineering analysis detailed in [15] derived mappings of some 14 design variables and 13 design parameters to the five attributes, alongside eight mathematical or geometric constraints to determine design feasibility. The engineering model is used here to identify a set of 2455 feasible design attribute combinations. Specifically, let $\mathbf{z}(\mathbf{x})$ be the mapping from design variables to attributes and $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ be the constraints. For each combination of design attributes $\mathbf{z}$, we solve the problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & ||\mathbf{z} - \mathbf{z}(\mathbf{x})||_2 \\ \text{subject to} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}. \end{aligned} \tag{3}$$

The combination $\mathbf{z}$ will be considered feasible if a feasible solution $\mathbf{x}$ is found and the resultant minimal discrepancy $||\mathbf{z} - \mathbf{z}(\mathbf{x})||_2$ is less than 10^{-3}. Note that unlike [15] where optimal design variables $\mathbf{x}$ are found for maximizing profit, the goal in this study is to identify the optimal combination of design attributes from the feasible set. We will discuss in Subsection 5.4 the extension from optimizing the attribute combination to optimizing design variables.

2.3 Parametric study

Based on the profit and engineering models, and the revealed partworth of the population, one can identify the optimal design, i.e., the most profitable combination of design attributes.

We now investigate how the optimal design changes when the estimated partworths are different from the true ones. We vary each element in the true partworth vector $\mathbf{w}_0$, one at a time, from -2 to 2 at an interval of 0.01, while keeping other partworths unchanged. Following Equation (1), each test leads to a most profitable design. The result is presented in Figure 1, where the rows correspond to the thirty attribute levels, from $z_{1,1}$ at the top to $z_{6,5}$ at the bottom. The true partworth values of $\mathbf{w}_0$ are marked in red. Take the first row as an example: It is composed of a sequence of hypothetical partworth vectors where the partworth value for $z_{1,1}$ goes from -2 to 2. The dark section to the right of the row represents partworth vectors that yield different optimal designs than the true one, while the white section represents those that retain the true optimal design. We also observed that beyond the region $[-2, 2]$ there will be no further flip of colors.

This result shows that the optimal solution is insensitive to the partworth estimates for some of the attribute levels. For example, the partworths for low price levels, namely, $z_{6,1}$ and $z_{6,2}$, do not affect the optimal solution on their own. This is because the potentially larger market share induced by these price levels will not compensate for the lower unit profit, according to the profit model.

The findings from this parametric study indicate that for identifying the optimal design, e.g., with maximum profit, it is not necessary to acquire all the information needed for correctly modeling the preference. In other words, the questionnaire could be tailored for design optimization purposes, by considering the profit and engineering models.

We can further use the illustration in Figure 2 to support the findings: Consider the entire space where the true partworth vector lies in as the square, and each binary question, e.g., "Among design A and design B, which one do you prefer more?", as a cutting hyperplane in that space. The feasible space for the true

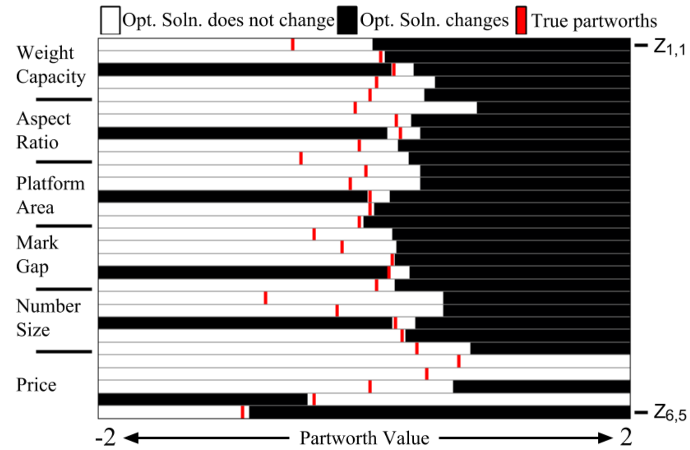

FIGURE 1. Moving one of the partworth values away from the truth may not change the optimal design.

partworth to reside will be reduced by each new question and response from the user. In a traditional marketing study where questionnaires are designed for estimating the true partworth, cuts would be adaptively designed to reduce the remaining space as much as possible. In the case of optimal design, consider the entire space to be segmented by a set of designs, each of which is the optimal in the corresponding segment of the space according to the profit model. The strategy for finding the true optimal design is to cut the space efficiently so that ideally its remainder will consist only of the segment belonging to the true optimal design in a few iterations. Intuitively, this cutting strategy is not necessarily the same as reducing the space as much as possible and may involve less cuts. We will show this theoretically in the algorithm development in Section 3. This explains the motivation for developing a different adaptive questionnaire strategy for design optimization.

3 Adaptive Questionnaire

Motivated by the parametric study from the running example, we propose an adaptive questionnaire design specific for design optimization. The theoretical support of the questionnaire comes from the Group Generalized Binary Search (GGBS) algorithm used for minimizing the number of questions to solve a group identification problem [16]. This algorithm was previously extended to preference elicitation [17] where the design with the highest expected preference was identified. Under the same theory and similar implementation, we now apply the GGBS algorithm to identify the design with the highest expected profit.

The rest of this section introduces the group identification problem and shows that identifying an optimal design through binary questions can be casted as such a problem.

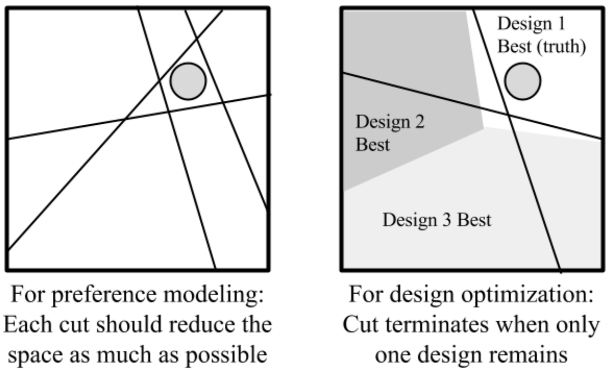

FIGURE 2. A comparison between adaptive questionnaire for preference modeling and for design optimization.The true preference model (the circle) is a point in a preference space (the square). Binary questions are cuts (the lines). With the profit model, the space is segmented, where segments are labeled by the corresponding optimal designs.

Group	Object	Q1: On ground?	Q2: Takes many people?	Q3: Moves vertically?
Ground vehicle	Motorcycle	Yes	No	No
	Bus	Yes	Yes	No
Aircraft	Helicopter	No	No	Yes
	Plane	No	Yes	No

FIGURE 3. An example of the group identification problem. Among the three candidate questions, "Q1" is the best to ask.

3.1 Group identification

Group identification can be explained by the following game with two players: Assume Players A and B are both familiar with and only with a list of objects, say, the one in Figure 3. Player A first picks an object from the list without letting Player B know. Player B then picks yes/no questions to ask (see figure). Based on Player A's responses, B guesses which group the object belongs to. To Player B, the ordering of questions will affect how quickly the correct group can be identified. For the example in Figure 3, the best question to ask is "Q1", since its answer directly determines which group the object is from. To help Player B to find the best strategy for asking questions in such identification problems, Bellala et al. proposed the GGBS algorithm that adaptively chooses questions and minimizes the expected number of questions needed [16][1].

3.2 Questionnaire as group identification

We now show that identifying the optimal design can be considered as a group identification problem. Consider a simple case with three designs A, B and C. Let an "object" be a ranking of the designs according to their profits calculated from Equation (1). For example, let design A be the most profitable and C the least, i.e., $A \succ B \succ C$. A "group" is a set of rankings with the same most profitable design, e.g., the group "design A is the best" contains two objects: $A \succ B \succ C$ and $A \succ C \succ B$. The questionnaire consists of a set of binary questions, i.e., pairwise comparisons. The cumulative binary responses from a questionnaire provide constraints on the feasible space of the partworth vector and induce probabilities of each design being the most profitable one, a process similar to identifying a group. Based on these analogies, we see that finding the most profitable design is equivalent to identifying a group, given that the set of candidate designs is finite. Therefore GGBS can be used to design the questionnaire adaptively.

On the side, the same analogy can be applied to preference modeling: Consider that the true model induces a ranking of all designs. To retrieve this unknown ranking, we may treat all possible rankings as "objects". Therefore a conjoint analysis can be considered as the problem of object identifying, where the Generalized Binary Search algorithm can apply.

4 The GGBS Algorithm

We now elaborate on the technical details of the GGBS algorithm for identifying the optimal design. The majority of the development has been reported in previous work on finding the most preferred design [17] and is restated here for completeness.

4.1 Preliminaries

Some notations and definitions are introduced first.

Let the set of K feasible designs be $\{\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(K)}\}$. A profit ranking θ can be derived for a given partworth vector $\mathbf{w}$ using Equation (1), for example, $\theta = \mathbf{z}^{(1)} \succ \mathbf{z}^{(2)} \succ ... \succ \mathbf{z}^{(K)}$. Let the total number of rankings be M and the ranking set be $\Theta = \{\theta_1, \theta_2, ..., \theta_M\}$. Each ranking is then labeled according to its top-ranked product, e.g., if $\theta_m = \mathbf{z}^{(1)} \succ ...$, then a label $y_m = 1$ is assigned. We denote by $\Theta_k = \{\theta_m \in \Theta : y_m = k\}$ the rankings for which product k is the most preferred.

For each ranking θ_m, we use $\pi_{\theta_m} = \Pr(\theta = \theta_m)$ to represent its probability to be the correct one, and the set

$$\Pi_\theta := (\pi_{\theta_1}, ..., \pi_{\theta_M}) \tag{4}$$

for the set of probabilities of all M rankings (objects), with $\sum_{m=1}^{M} \pi_{\theta_m} = 1$. Similarly, we use $\pi_{\Theta_k} = \sum_{\theta \in \Theta_k} \pi_\theta$ for the probability of product k being the most profitable, and the set

$$\Pi := (\pi_{\Theta_1}, ..., \pi_{\Theta_K}) \tag{5}$$

for the set of probabilities of all K products (groups). Note that Π can be determined jointly by the design set $\{\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(K)}\}$ and the prior distribution of $\mathbf{w}$, $p(\mathbf{w})$, and GGBS requires Π (therefore $p(\mathbf{w})$) as an input. In this study, without any prior knowledge of user preferences, we assume $p(\mathbf{w})$ to be standard multivariate normal. Relaxation on this assumption will be discussed in Subsection 6.1.

[1] Golovin et al. [18] also studied the group identification problem. We defer a comparison of their approach with this study to future work.

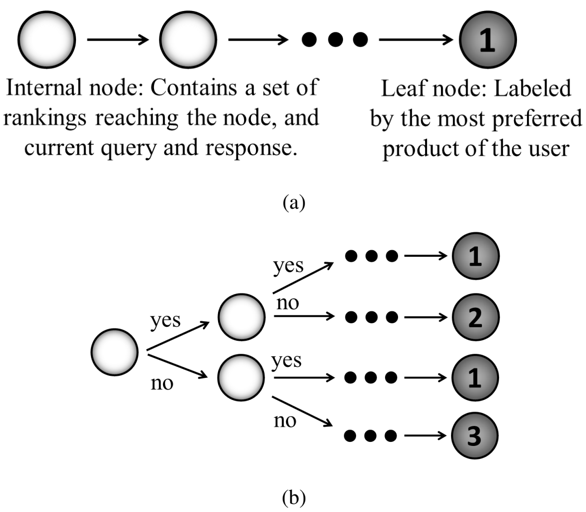

FIGURE 4. (a) An individual interaction as a path; (b) a binary decision tree composed of all possible paths

4.2 Questionnaire as a binary decision tree

Each questionnaire can be regarded as a user-computer interaction, consisting of a sequence of questions or "queries". In this study, a query is a pairwise comparison task generated by the computer and completed by the respondent. The entire interaction can be considered as a "path" where a set of "nodes" are connected by "edges", see Figure 4(a). We call the last node of the path the "leaf" node, which is labeled by the most profitable product. The rest of the nodes are called "internal" ones. Each internal node "a" contains (1) a set of rankings $\Theta^a \subseteq \Theta$ that reaches the node based on previous query responses, and (2) a new query made at this node, the response to which will lead to the next node.

Since each query could result in two responses depending on $\mathbf{w}$, the collection of possible paths forms a binary decision tree, each internal node of which has a query and each leaf node an optimal design corresponding to some realizations of $\mathbf{w}$, see Figure 4(b). We call an arrangement of queries along this binary decision tree a "query strategy."

Note from the figure that the binary decision tree can have multiple paths with leaf nodes labeled by the same design. This is because different partworth vectors could yield the same optimal design, as we demonstrated in Subsection 2.3.

4.3 The algorithm

Given a query strategy and the distribution $p(\mathbf{w})$, we can calculate the expected number of queries of the decision tree, denoted as L; varying the choice of query strategy could change that expectation. While optimizing L over all possible arrangements of queries is shown to be NP complete [19], Bellala et at. [16] proved that for any query strategy, L can be decomposed into a set of additive terms L^a with respect to each internal node "a":

$$L = \sum_{a \in \substack{\text{all internal} \\ \text{nodes}}} L^a + \text{constant}. \tag{6}$$

Thus L can be heuristically minimized by greedily minimizing L^a at each node with respect to the choice of query. This local objective L^a has the following form and requires some explanation:

$$L^a = 1 - H(\rho^a) + \sum_{k=1}^{K} \frac{\pi_{\Theta^a_k}}{\pi_{\Theta^a}} H(\rho^a_k). \tag{7}$$

For a given query (pairwise comparison) with binary response, the current node "a" will lead to the "left" and "right" child nodes denoted as "$l(a)$" and "$r(a)$", respectively. The sets $\Theta^{l(a)}$ and $\Theta^{r(a)} \subseteq \Theta^a$ contain rankings that fall into these two child nodes. The symbol ρ^a in Equation (7) is called the "reduction factor" and is defined as

$$\rho^a = \max\{\pi_{\Theta^{l(a)}}, \pi_{\Theta^{r(a)}}\}/\pi_{\Theta^a}, \tag{8}$$

where $\pi_{\Theta^a} := \sum_{\{i:\theta_i \in \Theta^a\}} \pi_{\theta_i}$ is the total probability mass of rankings at node "a", which then splits into $\pi_{\Theta^{l(a)}}$ and $\pi_{\Theta^{r(a)}}$ for the "left" and "right" child nodes. Similarly, the "group reduction factor" ρ^a_k is defined as

$$\rho^a_k = \max\{\pi_{\Theta_k^{l(a)}}, \pi_{\Theta_k^{r(a)}}\}/\pi_{\Theta^a_k}, \tag{9}$$

where $\pi_{\Theta^a_k} := \sum_{\{i:\theta_i \in \Theta^a_k\}} \pi_{\theta_i}$ is the total probability mass of a group labeled by product k at node "a", which is then separated into $\pi_{\Theta_k^{l(a)}}$ and $\pi_{\Theta_k^{r(a)}}$ for the given query.

Note that both the reduction factor and the group reduction factor are functions of the choice of query. The term $\pi_{\Theta^a_k}/\pi_{\Theta^a}$ represents the conditional probability of product k being the most profitable at node "a". Finally, we denote by $H(\rho) := -\rho \log_2 \rho$ the Shannon entropy of some scalar ρ, and define $0 \log_2 0 = 0$.

To demonstrate how GGBS works, let us revisit the group identification problem from Figure 3. Once Player A picked one object from the four, Player B needs to pick a query. To start, it is reasonable for Player B to believe that the four candidate objects have equal chances to be the correct one, i.e., $\Pi_\theta = (\pi_{\text{Motorcycle}}, \pi_{\text{Bus}}, \pi_{\text{Helicopter}}, \pi_{\text{Plane}}) = (0.25, 0.25, 0.25, 0.25)$ and therefore $\Pi = (\pi_{\Theta_{\text{Ground vehicle}}}, \pi_{\Theta_{\text{Aircraft}}}) = (0.5, 0.5)$ for the two groups. Player B will now calculate Equation (7) for each query and pick the one with minimal L^a. The calculation is performed in Table 3, where the left (right) child node corresponds to the answer "yes" ("no"). From the table, "Q1" leads to the minimum L^a, therefore should be queried. The presented calculation can be applied at each node to determine the best query.

5 Adaptive Questionnaire for the Scale Design Problem

Let us now review the scale design problem presented in Section 2. The dial-readout scale has six design attributes, each

TABLE 3. Query selection for the case in Figure 3

Query	$\pi_{\Theta^{l(a)}}$ $(\pi_{\Theta^{r(a)}})$	ρ^a	$\pi_{\Theta^{l(a)}_{\text{Ground vehicle}}}$ $(\pi_{\Theta^{r(a)}_{\text{Ground vehicle}}})$	$\pi_{\Theta^{l(a)}_{\text{Aircraft}}}$ $(\pi_{\Theta^{r(a)}_{\text{Aircraft}}})$	$\rho^a_{\text{Ground vehicle}}$	ρ^a_{Aircraft}	L^a
Q1	0.5 (0.5)	0.5	1 (0)	0 (1)	1	1	**0.5**
Q2	0.5 (0.5)	0.5	0.5 (0.5)	0.5 (0.5)	0.5	0.5	1
Q3	0.25 (0.75)	0.75	0 (1)	0.5 (0.5)	1	0.5	1.16

with five levels. We assume that the homogeneous user preference can be modeled by a linear utility function and the profit model by Equation (1). Based on engineering constraints, the set of all $K = 2455$ feasible combinations of attribute levels are chosen as the candidate design set. The questionnaire contains a sequence of pairwise comparison questions, each consisting of a pair of designs selected from the candidate set. For each given question, the simulated user will deterministically choose the design with higher utility, according to the true partworth vector $\mathbf{w}_0$.

5.1 Optimal design under full information

We first investigate the extreme case where answers of all $K(K-1)/2 = 3012285$ questions are known. These answers provide a set of half-spaces in $\mathbb{R}^{30}$, the intersection of which defines a narrow feasible partworth space[2]. As previously described, this running example assumes that $\mathbf{w}$ follows a standard multivariate normal distribution. Under full information, the distribution of $\mathbf{w}$ is constrained in the narrow feasible space, and a numerical integration within this space can be performed to calculate the conditional probabilities of each design being the optimal[3]. The result in Table 4 shows that with full information, there are two designs, with the only difference being the price level, that have non-zero probabilities to be the optimum. In addition, the first design from the table is the true optimum under the given models and the true partworth values.

TABLE 4. Optimal design under full information

#	z_1	z_2	z_3	z_4	z_5	p	Prob.
1	300lbs.	120in.^{2}	1.25in.	1.14	$5/32$	\$25	0.37
2	300lbs.	120in.^{2}	1.25in.	1.14	$5/32$	\$30	0.63

[2] In traditional conjoint analysis, it is required that identification constraints are imposed on the feasible space of $\mathbf{w}$. This is because infinite solutions exist that maximize the likelihood under the logit model. However, under hierarchical Bayesian models (or with regularization on $\mathbf{w}$), such constraints are not necessary as a constant shift in the partworths will cause the posterior likelihood of $\mathbf{w}$ to change. Therefore we do not impose identification constraints on $\mathbf{w}$ in this running example.

[3] Under full information, the feasible space is almost a line. Therefore we used a one-dimensional truncated normal distribution for numerical integration.

5.2 GGBS implementation

Here we describe the GGBS implementation for selecting the best query, using the procedure elaborated in Subsection 4.3. In each round, with the accumulated responses from previous questions, we first calculate the conditional probabilities of each design to be the optimal. In order to allow tractable computation in choosing the best query, we pick the 10 designs with the highest conditional probabilities, leading to a set of $10(10-1)/2 = 45$ candidate queries to choose from in each iteration.

The best query is chosen according to Equation (7). To do so, the distributions Π and $\pi_{\Theta^a_k}$ shall be calculated. Under the assumption that $\mathbf{w}$ is standard multivariate normal, each probability mass π_{Θ_k} from the distribution Π can be calculated as

$$\pi_{\Theta_k} = \int_{\mathbf{w}\in\mathbb{R}^{30}} \mathbb{1}\{f(\mathbf{z}^{(k)},\mathbf{w}) > f(\mathbf{z}^{(k')},\mathbf{w}), \forall k' \neq k\} \, p(\mathbf{w})d\mathbf{w}. \tag{10}$$

Here $\mathbf{z}^{(k)}$ is the kth design from the candidate set, $f(\mathbf{z}^{(k)},\mathbf{w})$ is the profit of the kth design given $\mathbf{w}$, $\mathbb{1}\{\text{conditions}\}$ is an indicator function that equals 1 when conditions are satisfied or 0 otherwise. Similarly, the calculation of $\pi_{\Theta^a_k}$ for $k = 1,\ldots,K$ takes the form:

$$\pi_{\Theta^a_k} = \int_{\mathbf{w}\in\mathbb{R}^{30}} \mathbb{1}\{f(\mathbf{z}^{(k)},\mathbf{w}) > f(\mathbf{z}^{(k')},\mathbf{w}), \forall k' \neq k\} p(\mathbf{w}|a)d\mathbf{w},$$

where $p(\mathbf{w}|a)$ is a truncated normal distribution where $\mathbf{w}$s are constrained by all query responses prior to node "a". The numerical integration is performed using a Gibbs sampler following Rodriguez-Yam et al. [20], with a sample size of 10^4 and the first half discarded as burn-in samples.

5.3 Alternative questionnaires

Two alternative query strategies are used for comparison: The "Adaptive for Preference" ("AP") approach has queries adaptively chosen for the purpose of estimating the partworth vector; the "Best Profit" ("BP") approach has each query consisting of the designs with the highest probabilities to be the most profitable. "AP" is used to show the advantage of a questionnaire designed for optimization purpose from one designed for partworth recovery; "BP" is used to show the performance improvement by GGBS from a naïve questionnaire approach for design optimization.

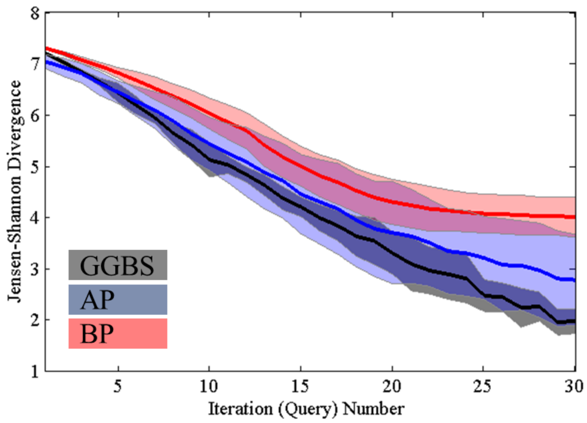

FIGURE 5. Averaged Jensen-Shannon divergence from the target during the questionnaire. Best viewed in color.

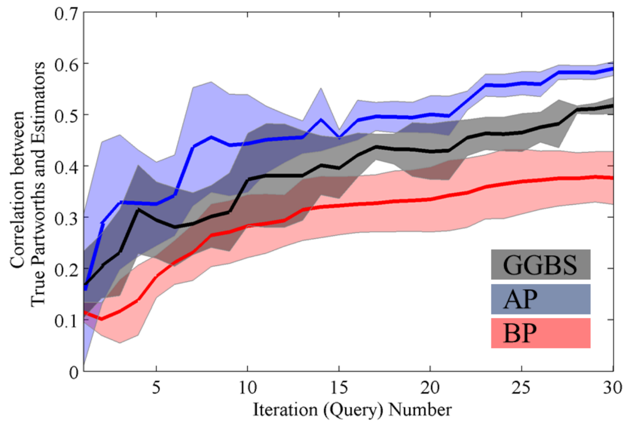

FIGURE 6. Averaged correlation between the true partworths and the estimators during the questionnaire. Best viewed in color.

The "AP" approach implemented here can be considered as a special case of adaptive preference elicitation where each object (ranking of designs) is a group by itself. The best query at each node "a" can be found by minimizing $L^a = 1 - H(\rho^a)$ where the reduction factor ρ^a follows Equation (8), and both Π and $\pi_{\Theta_k^a}$ are calculated based on the linear preference utility model rather than the profit model. This approach is essentially an uncertainty sampling scheme as the query that maximizes the Shannon entropy $H(\rho^a)$ should contain a pair of designs with the smallest difference in their utilities. The method is also similar in concept to Tong et al. [21], Toubia et al. [10] and Jamieson et al. [22]. Similar to the GGBS implementation, "AP" also uses a reduced set of 45 candidate queries in each iteration.

5.4 Simulation results

For all three algorithms, we run the simulated questionnaire for 30 queries (iterations). We use the Jensen-Shannon divergence [23] to track the difference between the conditional probability distribution to the target one in Table 4. The decrease in divergence indicates the convergence from the guess (Π) to the truth. Jensen-Shannon divergence is chosen because the two distributions are sparse, considering that only a small set of designs will retain non-zero probabilities to be optimal conditioned on all the query responses. For each algorithm, twenty independent experiments are conducted and the averaged divergence plots are summarized in Figure 5, along with standard errors. The result shows that the GGBS algorithm performs significantly better than the naive BP approach in convergence and also consistently but marginally better than AP on average. Note that the randomness in these simulation results is solely due to the Gibbs sampler.

Note that the proposed GGBS algorithm is not supposed to be the best choice when the goal of the questionnaire is to efficiently model the preference. For verification, we derive the estimator of $\mathbf{w}_0$ at each iteration using an L2-regularized logistic regression and the LIBLINEAR solver [24][4], and track the change in the correlation between the true partworth vector and the estimator during the questionnaire. The comparison on this correlation from the same simulations is shown in Figure 6, where the performance of AP is superior than the other two, both supporting our argument and indicating the correct implementation of AP.

The above conclusion is warranted only under the current settings in this simulated study. Future scrutiny is required to further generalize the conclusion. For example, more queries beyond the current query limit must be simulated to observe the performance trends of the three algorithms in longer questionnaires. Parametric studies should also be performed to understand the influence of algorithmic parameters such as the sampling size of the Gibbs sampler and the number of candidate queries at each iteration. Other alternative adaptive conjoint analysis algorithms, such as Abernethy et al. [25][5], should be included for benchmarking in addition to AP.

6 Discussion

Here we provide discussion beyond the results presented for the running example to show how the proposed method could be extended to address some of its current drawbacks.

6.1 Relaxing the assumptions on $p(\mathbf{w})$

One significant drawback of the GGBS algorithm is its reliance on the distribution of $\mathbf{w}$ (and hence the probability masses Π) as a known input. Therefore, a good approximation of $p(\mathbf{w})$

[4] In the LIBLINEAR solver context, the penalty on training error is set to $C = 10^6$ since we assumed no random noise in user utility.

[5] When multiple most uncertain queries exist, Abernethy et al. proposes to choose the one that minimizes an approximation of the variance of the estimates.

of the target user population will lead to an efficient query strategy and vice versa. Since $p(\mathbf{w})$ is commonly unknown, it is necessary for the algorithm to refine its approximation of $p(\mathbf{w})$ through a sequence of questionnaires. Indeed, based on accumulated user responses, a flexible model can be efficiently learned, e.g., using convex optimization methods from Chapelle et al. [26] or Evgeniou et al. [7], to replace the initial guess of a standard normal assumption on $p(\mathbf{w})$ used in the running example.

In addition, while GGBS requires $p(\mathbf{w})$ to be known, it does not rely on homogeneous user preferences. Therefore the refined model of $p(\mathbf{w})$ can be directly applied in GGBS while addressing the heterogeneity of preferences. Nonetheless, it should be noted that the algorithm requires $p(\mathbf{w})$ to be fixed during the questionnaire.

6.2 Noisy human choice

The presented algorithm assumes that human utilities have zero noise, which is inconsistent with the profit model in Equation (1) which inherently incorporates choice noise. While GGBS performed well in the zero-noise setting, it does not necessarily outperform the alternative methods when noise exists. In the preference learning context, Jamieson et al. proposed to query the same question multiple times and pick the response based on majority vote [22]. However, the algorithm requires the probability of a false choice from the user to be strictly lower than 0.5. Whether such an algorithm can be applied to robust questionnaire design for design optimization requires further investigation.

6.3 High computational cost

Due to the use of Markov Chain Monte Carlo simulations for numerical integration, the GGBS algorithm has high computation cost, making it impractical for real-time human computation interactions[6]. Several treatments should be looked into: The calculation of conditional probabilities could be shortened by a more efficient Gibbs sampler, such as in Pakman et al. [27], while the samplers can be called in parallel; when both the total numbers of queries and the designs are limited, it is also possible to offload the computational burden during the questionnaire by generating query sequences offline so that the next query can be looked up.

6.4 From optimal attribute combination to optimal design variables

Recall that the scale problem has six design attributes governed by 14 design variables. While the present study aimed at identifying the optimal feasible combination of design attributes, it would be more valuable if the optimal values for the design variables could be found from a continuous design space, by conducting a sequence of questionnaires. However, this requires a conditional probability density function on the design space to be constructed at each iteration. Therefore this goal could only be achieved when the computational cost of sampling method is significantly lowered.

6.5 Questions outside of the candidate set

In the running example, we showed that even with all query responses, the true optimal design cannot be identified with probability 1. In fact, it is not even ranked with the highest probability. Considering that the example only uses queries formed from feasible designs, it would be interesting to see if some infeasible queries, i.e., a design pair containing infeasible designs, could lead to higher L^a values. Therefore further study is needed to investigate whether new queries can be created, rather than picked based on feasible designs, to improve the effectiveness of the questionnaire.

7 Conclusions

We offered an analogy between identifying an optimal design from engineering and profit models and the general problem of identifying the group label of objects, and showed that questionnaires can be directly and adaptively designed for design optimization. We demonstrated using an existing dial-readout scale design problem that the proposed GGBS algorithm is superior to an adaptive conjoint analysis method aimed for preference modeling in converging to the correct conditional probabilities of designs being the optimal. The conclusion from the study requires future investigation in order to be generalized.

8 Acknowledgement

This work has been supported by the National Science Foundation under Grant No. CMMI-1266184. This support is gratefully acknowledged.

REFERENCES

[1] Green, P. E., and Srinivasan, V., 1990. "Conjoint analysis in marketing: New developments with implications for research and practice". *The Journal of Marketing*, pp. 3–19.

[2] Netzer, O., Toubia, O., Bradlow, E., Dahan, E., Evgeniou, T., Feinberg, F., Feit, E., Hui, S., Johnson, J., and Liechty, J., 2008. "Beyond conjoint analysis: Advances in preference measurement". *Marketing Letters,* ***19***(3), pp. 337–354.

[3] Berkovec, J., and Rust, J., 1985. "A nested logit model of automobile holdings for one vehicle households". *Transportation Research Part B: Methodological,* ***19***(4), pp. 275–285.

[6] On a PC with i5 CPU and 8G memory, it takes 10 minutes for comparing among all 45 candidate queries by calling the Gibbs sampler iteratively. Note that this time changes for a different parameter setting of the Gibbs sampler.

[4] McFadden, D., and Train, K., 2000. "Mixed MNL models for discrete response". *Journal of Applied Econometrics,* ***15***(5), pp. 447–470.
[5] Lenk, P. J., and DeSarbo, W. S., 2000. "Bayesian inference for finite mixtures of generalized linear models with random effects". *Psychometrika,* ***65***(1), pp. 93–119.
[6] Hauser, J. R., Toubia, O., Evgeniou, T., Befurt, R., and Dzyabura, D., 2010. "Disjunctions of conjunctions, cognitive simplicity, and consideration sets". *Journal of Marketing Research,* ***47***(3), pp. 485–496.
[7] Evgeniou, T., Pontil, M., and Toubia, O., 2007. "A convex optimization approach to modeling consumer heterogeneity in conjoint estimation". *Marketing Science,* ***26***(6), pp. 805–818.
[8] Jaakkola, T., and Jordan, M., 1997. "A variational approach to bayesian logistic regression models and their extensions". In Sixth International Workshop on Artificial Intelligence and Statistics, Citeseer.
[9] Toubia, O., Hauser, J., and Simester, D., 2004. "Polyhedral methods for adaptive choice-based conjoint analysis". *Journal of Marketing Research,* ***41***(1), pp. 116–131.
[10] Toubia, O., Simester, D., Hauser, J., and Dahan, E., 2003. "Fast polyhedral adaptive conjoint estimation". *Marketing Science,* ***22***(3), pp. 273–303.
[11] Toubia, O., Evgeniou, T., and Hauser, J., 2007. "Optimization-based and machine-learning methods for conjoint analysis: Estimation and question design". *Conjoint Measurement: Methods and Applications*, p. 231.
[12] Haftka, R. T., 1991. "Combining global and local approximations". *AIAA Journal,* ***29***(9), pp. 1523–1525.
[13] Li, J., Mourelatos, Z. P., Kokkolaras, M., Papalambros, P., and Gorsich, D., 2010. "Validating designs through sequential simulation-based optimization". In Proceedings of the ASME, pp. 15–18.
[14] Feit, E. M., 2009. "Essays on integrating product design and marketing research". PhD thesis, The University of Michigan, Ann Arbor.
[15] Michalek, J. J., Feinberg, F. M., and Papalambros, P. Y., 2005. "Linking marketing and engineering product design decisions via analytical target cascading". *Journal of Product Innovation Management,* ***22***(1), pp. 42–62.
[16] Bellala, G., Bhavnani, S., and Scott, C., 2012. "Group-based active query selection for rapid diagnosis in time-critical situations". *IEEE Transactions on Information Theory,* ***58***(1), pp. 459–478.
[17] Ren, Y., Scott, C., and Papalambros, P., 2013. "A scalable preference elicitation algorithm using group generalized binary search". In Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.
[18] Golovin, D., Krause, A., and Ray, D., 2010. "Near-optimal bayesian active learning with noisy observations.". In NIPS, Vol. 10, pp. 766–774.
[19] Hyafil, L., and Rivest, R., 1976. "Constructing optimal binary decision trees is NP-complete". *Information Processing Letters,* ***5***(1), pp. 15–17.
[20] Rodriguez-Yam, G., Davis, R. A., and Scharf, L. L., 2004. "Efficient Gibbs sampling of truncated multivariate normal with application to constrained linear regression". *Unpublished Manuscript*.
[21] Tong, S., and Koller, D., 2002. "Support vector machine active learning with applications to text classification". *The Journal of Machine Learning Research,* ***2***, pp. 45–66.
[22] Jamieson, K., and Nowak, R., 2011. "Active ranking using pairwise comparisons". *arXiv preprint arXiv:1109.3701*.
[23] Lin, J., 1991. "Divergence measures based on the Shannon entropy". *Information Theory, IEEE Transactions on,* ***37***(1), pp. 145–151.
[24] Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., and Lin, C.-J., 2008. "LIBLINEAR: A library for large linear classification". *The Journal of Machine Learning Research,* ***9***, pp. 1871–1874.
[25] Abernethy, J., Evgeniou, T., Toubia, O., and Vert, J.-P., 2008. "Eliciting consumer preferences using robust adaptive choice questionnaires". *Knowledge and Data Engineering, IEEE Transactions on,* ***20***(2), pp. 145–155.
[26] Chapelle, O., and Harchaoui, Z., 2004. "A machine learning approach to conjoint analysis". In Neural Information Processing Systems.
[27] Pakman, A., and Paninski, L., 2013. "Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians". *Journal of Computational and Graphical Statistics*.