# Augmented Probability Simulation for Accelerated Life Test Design

Nick Polson
*Booth School of Business*
*University of Chicago, USA*

Refik Soyer
*Department of Decision Sciences*
*The George Washington University, USA*

# Augmented Probability Simulation for Accelerated Life Test Design

Nicholas G. Polson
*University of Chicago* [*]

Refik Soyer
*George Washington University* [†]

Sunday 16th April, 2017

### Abstract

Designing accelerated life tests presents a number of conceptual and computational challenges. We propose a Bayesian decision-theoretic approach for selecting an optimal stress-testing schedule, and develop an augmented probability simulation approach to obtain the optimal design. The notion of a "dual utility probability density" enables us to invoke the concept of a conjugate utility function. For accelerated life tests, this allows us to construct an augmented probability simulation which simultaneously optimizes and calculates the expected utility. In doing so, we circumvent many of the computational difficulties associated with evaluating pre-posterior expected utilities. To illustrate our methodology, we consider a single-stage accelerated life test design; our approach naturally extends to multiple stage designs. Finally, we conclude with suggestions for further research.

*Keywords: Bayesian optimal design, Life testing, Decision theory, Monte Carlo optimization, Particle methods.*

[*]email: ngp@chicagobooth.edu
[†]email: soyer@gwu.edu

1

# 1 Introduction

Designing accelerated life tests (ALTs) involves testing systems in severe stress environments relative to a standard use conditions environment. Bayesian methods are often attractive for analyzing accelerated life test data; see for example, Soyer (2007). One caveat is the computational challenge of solving for the optimal design of the accelerated stress levels together with the number of items to be tested at each level. The challenge is to calculate a pre-posterior expected utility and to optimize over the set of design choices.

We take a simulation-based approach which simultaneously calculates the pre-posterior expected utility and stochastically finds the optimal design. Our Bayesian decision theoretic set up builds on Chaloner and Larntz (1990), Verdinelli, Polson and Singpurwalla (1993), Polson (1993), and Erkanli and Soyer (2000). Our methodology extends to sequential designs. One theoretically new addition is our use of conjugate utility structures following the work of Lindley (1976), and the use of augmented probability simulation (Muller (1999)) to compute optimal designs. This helps in the construction of both Markov chain Monte Carlo and particle-based simulation algorithms. We tailor a particle-based algorithm to calculate and optimize pre-posterior utilities; see also Muller (1999) and Ekin, Polson and Soyer (2014).

Interest in optimal design of ALTs starts with the original works of Zelen (1959) and Chernoff (1962). There is a considerable literature on Bayesian ALT designs dating back to Martz and Waterman (1978) and DeGroot and Goel (1979). The seminal work on optimal Bayesian designs for linear models is discussed by Chaloner (1984), and this can be used for ALT designs when the underlying lifetime probability model is normal or a lognormal. For example, Chaloner and Larntz (1990) find Bayesian designs for Type I censored tests with uncertainty about whether the underlying life model is lognormal or Weibull when several fractiles of the lifelength distribution at the use stress are of interest. The optimality criterion is proportional to the expected asymptotic variance of the fractiles of interest. Menzefricke (1992) formulates the optimal design of Type II censored ALTs when the lifelength model is lognormal. Verdinelli, Polson and Singpurwalla (1993) identify designs for a complete ALT that maximizes Shannon information. Zhang and Meeker (2006) consider large sample results for Bayesian ALT designs as well as simulation based methods. For alternative life models, such as exponential or Weibull, optimal designs can be obtained by use of either numerical methods such as non-parametric surface estimation; see Muller and Parmigiani (1995) and Erkanli and Soyer (2000), who analyze single stage accelerated life test designs, or by special techniques; see Vopatek (1992) and Soyer and Vopatek (1995).

The remainder of the paper is outlined as follows. Section 2 introduces the ALT design problem and discusses our conjugate utility functions approach. Section 3 presents the augmented probability simulation method. Section 4 discusses the implementation of augmented probability for ALT design. We provide both the-

oretical and applied examples illustrating the augmented probability simulation model and also introduce a particle-based approach to augmented simulation. Finally, section 5 concludes with some suggested directions for future research.

# 2 Bayesian Design of ALTs

The Bayesian determination of an optimal design is based on the paradigm of maximizing pre-posterior expected utility; see for example, Lindley (1985). A comprehensive review of Bayesian experimental design can be found in Verdinelli (1992) and Chaloner and Verdinelli (1995).

## 2.1 The ALT Design Problem

An accelerated test environment is created by increasing the level of one or more of the stress variables such as temperature, voltage, etc. to values which are higher than those at normal operating conditions. In our development we assume that the environment is characterized by a single stress with extensions to multiple stresses being straightforward; see for example, Escobar and Meeker (1995) and Zhang and Meeker (2006).

Let $d$ denote a design, that is, the level of the stress variable characterizing the accelerated test environment. It is possible that tests will be conducted at $K$ accelerated levels of the stress variable which are specified in advance.

The Bayesian ALT design problem requires specification of three components:

1. Utility (loss): reflecting the consequences of selecting a specific accelerated environment $d$

2. Probability model: life distribution at the accelerated stress level(s)

3. Prior distribution: reflecting a priori beliefs about all unknown quantities such as the parameters of the probability model.

Components 2 and 3 together constitute the predictive model in the design problem which is obtained by integrating out the parameters of the probability model using the prior. As pointed out by one of the reviewers, separation of probabiliy and utility (components 1 and 2) is a foundational issue which requires caution. Kadane and Winkler (1988) discuss conditions under which elicitation of probabilities can be separated from utilities and study the implications when such separation is not possible. We refer the interested reader to their paper.

To fix our notation, let $y$ denote an observable and $\theta$ a parameter associated with the life distribution, $p(y|\theta)$. Once a design $d$ is specified and $y$ is observed from the ALT, uncertainty about $\theta$ is revised according to Bayes' rule. Given a

specific form for the utility function $u(y, \theta, d)$, the design problem requires finding $d^* = \underset{d}{argmax}\, u(d)$, where

$$u(d) = E_{y|d}E_{\theta|y,d}[u(y, \theta, d)] = E_{y|\theta,d}E_{\theta|d}[u(y, \theta, d)],$$

is the pre-posterior expected utility. Evaluation of $u(d)$ requires the computation of

$$u(d) = \int u(y, \theta, d)p(y, \theta|d)\, d\theta\, dy. \tag{1}$$

In the context of accelerated life tests, $u(y, \theta, d) = -V(\theta|y, d)$, the negative of the posterior variance, is a common choice. The optimal design then minimizes the pre-posterior variance $E_{y,\theta}\Big(V(\theta|y, d)\Big)$. The negative of the posterior variance is used in many fields such as economics and finance, engineering and machine learning especially in point estimation and control problems where a symmetric loss/utility is appropriate.

An alternative utility function is based on an information-theoretic criterion, such as the negative *entropy*. Here the optimal design is selected via maximization of the expected gain of information from experimentation (Lindley, 1956 and Verdinelli et al.,1993) given by the quantity

$$\underset{d}{argmax}\; E\Big(\int \log[p(\theta|y, d)]p(\theta|y, d)d\theta\Big). \tag{2}$$

This is equivalent to choosing the design that maximizes the expected Kullback-Leibler divergence between posterior and prior distributions, the so called *Lindley's measure*. It is well known that use of Lindleys measure of information about parameter $\theta$ as a utility function yields the posterior distribution of $\theta$ as the optimal decision in the problem of reporting a probability distribution from the space of all distributions; see Bernardo (1979) and Ebrahimi, Soofi and Soyer (2010). Information-theoretic utility functions are used in many other fields including economics and engineering.

### 2.1.1 Time Transformation Function

The objective of an ALT is to make statements of uncertainty about lifetime at the use environment based on data from accelerated environments. Let $d_i$ denote the level of the stress variable at the *ith* accelerated test environment $i = 1, \ldots K$, and $T_i$ denote the lifetime of an item tested at environment $d_i$. Also let $T_u$ denote the lifetime at the use environment $d_u$ such that $d_u < d_i, i = 1, \ldots K$. An important component of ALTs is the *time transformation function* (*acceleration function*) that describes the relationship between the failure characteristic such as the mean or median life time and the applied level of the stress. Typically such relationship is based on physics of failure.

4

For example, if $T_i$ is lognormally distributed with location parameter $\mu_i$, then a *power law* time transformation function can be written in terms of the median lifetime at stress environment $d_i$ as

$$e^{\mu_i} = \alpha d_i^\beta, \tag{3}$$

where $\alpha$ and $\beta$ are unknown parameters. On the other hand, if $T_i$ follows an exponential model then the failure rate or mean life time can replace the median life time in (3). Other commonly used time transformation functions include the Arrhenius and Eyring laws; see Singpurwalla (2006).

Verdinelli et al. (1993) considered a lognormal model with the power law (3) by assuming a utility function based on (negative) posterior variance of loglifetime at the use stress $d_u$. For the case of one-point designs (testing at one stress level) they showed that if $\beta$ is known then it does not matter at what level one tests. On the other hand, if $\alpha$ is known then the optimal design is to test at the highest stress.

Erkanli and Soyer (2000) considered an exponential lifetime model where the failure rate $\lambda_i$ at stress $d_i$ follows a power law relation as in (3). They assumed the posterior variance of the failure rate at use stress as the utility function and obtained optimal designs by adopting the Monte Carlo based approach of Mueller and Parmigiani (1995) where the expectation step is replaced by a scatter plot smoother and the optimization step is replaced by the optimization of the fitted smoother.

There is a plethora of other failure models and different class of utility functions that require a use of Monte Carlo methods to obtain optimal designs. Choice of the class of utility functions may help in reducing dimension in computation of expected utilities. Lindley (1976) introduced a class of *conjugate utility* functions that are helpful in reducing dimension for computing expected utilities.

## 2.2 Conjugate Utility

For most choices of utility functions, there are two computational difficulties. First, we need to be able to marginalize over $(y, \theta)$ to be able to calculate expected utility of a specific design $d$. Second, we need to be able to optimize over $d$ to find $d^* = \underset{d}{argmax}\, u(d)$.

Lindley (1976) proposes a class of probability models and utility functions that are "conjugate" in the sense that the marginal utility $u(d)$ can be calculated in closed form. He only considers the case of an unknown parameter $\theta$. We extend this conjugate conjugate utility structure to incorporate both future data and parameters. We build on Lindley's approach to facilitate the computation of $u(d)$ and show how it can be adapted to marginalize over both $(y, \theta)$.

We describe below the concept of a conjugate utility which proves helpful in both developing an augmented simulation approach, and constructing particle learning algorithms for optimal sequential design. In the simplest case, suppose

5

that our observable $y$ has density conditional on $\theta$ as

$$p(y|\theta) = e^{y\theta}H(y)G(\theta)$$

where $G^{-1}(\theta) = \int e^{x\theta}H(x)dx$ is a suitable normalization constant. Note that the conditional density $p(y|\theta)$ is a member of the exponential family. If the prior distribution of $\theta$ is given by

$$p(\theta|y_0, n_0) = K(y_0, n_0)e^{y_0\theta}G(\theta)^{n_0},$$

for suitable hyperparameters $(y_0, n_0)$, then we can calculate

$$K^{-1}(y_0, n_0) = \int e^{y_0\theta}G(\theta)^{n_0}d\theta.$$

The associated posterior distribution $p(\theta|y_0, n_0, y)$ is conjugate within the same family as the prior distribution.

A natural class of conjugate utility functions is defined by

$$u(d, \theta) = e^{y(d)\theta}G(\theta)^{n(d)}F(d), \tag{4}$$

where $(y(d), n(d), F(d))$ are suitable functions of a decision variable, $d$. Since our augmented probability simulation method requires $u(d, \theta)$ to be a positive valued function, $F(d)$ will be restricted to be positive in (4). As a function of $\theta$, the conjugate utility function (4) can be thought as a unnormalized density function. Thus, it is typically a unimodal function. As pointed out by Lindley, this is helpful in point-estimation problems as well as in other applications such as the inventory control problems. Properties and roles of functions $y(d)$ and $n(d)$ are studied and conditions for a bounded utility functions are discussed in Lindley (1976, Section 3). Diaconis and Ylvisaker (1979) showed that mixtures conjugate priors can be used to approximate any prior. The same reasoning applies to conjugate utility functions, that is, any utility function $u(y, d)$ can be expressed as a mixture of the conjugate utility functions given by (4).

With $N = n + n_0$, $y = \sum_{i=1}^{n} y_i$, we can calculate the expected utility

$$u(d) = \int e^{(y+y(d))\theta}G(\theta)^{N+n(d)}K(N, y)F(d)d\theta \tag{5}$$

$$= \frac{K(N, y)F(d)}{K(N + n(d), y + y(d))}. \tag{6}$$

Availability of the expected utility (6) is attractive from a computational viewpoint. However optimizing this objective function can still be challenging. One of the aims of our paper is to show how this can be achieved using augmented probability simulation and particle methods.

# 3  Augmented Probability Simulation (APS) Model

Evaluation of $u(d)$ via traditional MC techniques involves simulation of $(y, \theta)^{(g)}, g = 1, \ldots, G$, from the joint model

$$p(y, \theta | d) = p(y | \theta, d) p(\theta | d)$$

and estimation of expected utility

$$\widehat{u}(d) = \frac{1}{G} \sum_{g=1}^{G} u((y, \theta)^{(g)}, d). \tag{7}$$

Optimizing $\widehat{u}(d)$ over $d$ can be very inefficient as MC errors in estimating $\widehat{u}(d)$ may overwhelm the optimization effort. Put simply, most of the draws $(y, \theta)^{(g)}$ are from a poor part of the space and provide little information. Performance of the MC approximation of $u(d)$ by (7) worsens with high dimensional integration. Bielza et al. (1999) and Mueller (1999) proposed an alternative strategy by treating the decision variable $d$ as a random variable and recasting the problem as a problem of drawing samples from an augmented probability model.

The augmented probability model is given by

$$\pi(y, \theta, d) \propto u(y, \theta, d) p(y, \theta | d), \tag{8}$$

where distribution of $d$ is generally specified as a uniform distribution over the decision space. The marginal distribution of $d$, that is, $\pi(d)$ is proportional to

$$u(d) = \int u(y, \theta, d) p(y, \theta | d) \, d\theta \, dy. \tag{9}$$

Thus, the optimal design can be obtained by simulating samples from the marginal distribution of $d$ and finding the mode of $\pi(d) \propto u(d)$. A Markov chain Monte Carlo (MCMC) scheme can be used to draw from the augmented distribution $\pi(y, \theta, d)$. We use a Gibbs sampler in our development. This requires simulating from the conditional distribution $\pi(y, \theta | d)$ which is a "tilted" version of the distribution $p(y, \theta | d)$. The tilted conditional distribution is obtained as proportional to (8). The Gibbs sampler requires also simulation from the conditional distribution $\pi(d | y, \theta) \propto u(y, \theta, d)$.

For dealing with higher dimensions of $d$ where it may not be easy to find the mode and with flat expected utility surfaces Mueller (1999) proposed to replace the $\pi(d)$ with a power type transformation on the expected utility. By drawing $J$ samples $(y_j, \theta_j)_{j=1}^{J}$ for each design $d$, we can obtain

$$\pi_J(y^J, \theta^J, d) \propto \prod_{j=1}^{J} u(y_j, \theta_j, d) p(\theta_j, y_j | d), \tag{10}$$

where $\theta^J = (\theta_1, \ldots, \theta_J)$ and $y^J = (y_1, \ldots, y_J)$. In this case the implied marginal distribution $\pi_J(d) \propto u^J(d)$; see Muller et al. (2004). This distribution has a number of useful features to note. First of all, as $J$ gets large $\pi_J(d)$ will concentrate around the mode which gives the optimal decision $d^*$. We can always use MCMC to draw from the distribution $d$ and for parameter learning.

The optimal design problem requires both evaluation and optimization of the expected utility function $u(d)$. Our proposed approach performs these two tasks simultaneously by treating the design variable $d$ as stochastic and simulating $d$, together with $(y, \theta)$ from the augmented probability model (8). In doing so, the approach tilts MC draws to regions of high utility values. This avoids inefficiency of the standard MC approach and reduces MC errors since samples are drawn more frequently from the augmented space with high utility values. Our proposed approach for ALT design using the APS differs from Muller (1999) by our use of conjugate utility functions. This provides conditional conjugacy and allows for use of Gibbs sampler in simulating the design variable. Also, to the best of our knowledge, particle based APS for design has not been considered before.

## 4  APS for ALT Design

To illustrate our proposed methodology, assume that life lengths $T_i$ under stress environment $d_i$ are exponentially distributed with failure rate $\theta d_i$; that is, $T_i | \theta, d_i \sim Exp(\theta d_i)$. If we test $n$ items under $d_i$ then distribution of total time on test $Y_i = \sum_{j=1}^{n} T_{ij}$ will have a gamma distribution with shape parameter $n$ and scale $\theta d_i$, that is,

$$p(y_i | \theta, d_i) = \frac{(\theta d_i)^n}{\Gamma(n)} y_i^{n-1} e^{-\theta d_i y_i}. \tag{11}$$

Here we assume a conjugate gamma prior for $\theta$ with parameters $a$ and $b$. We will also use a conjugate utility function. In this case, the joint distribution $p(y, \theta_i | d_i)$ is a gamma-gamma distribution, and the marginal $p(y_i | d_i)$ is a scaled beta prime (inverted beta) distribution given by

$$p(y_i | d_i) = \frac{\Gamma(a+n)d_i / b}{\Gamma(n)\Gamma(a)} \frac{(y_i d_i / b)^{n-1}}{(1 + y_i d_i / b)^{a+n}}.$$

The objective now is to select the accelerated environment $d_i > d_u$ to learn about the failure rate $\theta d_u$ at the use stress. We will consider a single point design, drop the index $i$, and assume a utility function $u(y, d)$ as

$$u(y, d) = \frac{1}{(d/d_u)^\alpha} e^{-ky}, \tag{12}$$

where $\alpha$ and $k$ are positive constants with $\alpha > 1$.

The utility function (12) is specified to reflect the consequences of choosing a design $d$ and observing a total time on test $y$ from the ALT. For making inferences

about life lengths at $d_u$, it is desirable to test at $d$ closer to $d_u$. However, smaller values of $d$ will imply a lower failure rate and thus, larger total time on test $y$. The utility function (12) reflects the trade-off between testing closer to $d_u$ and testing for a shorter amount of time. We can reflect the trade-off by appropriate choice of constants $\alpha$ and $k$. In (12) $\alpha$ reflects the cost of selecting a design $d$ away from $d_u$ whereas $k$ reflects the cost of testing. Without loss of generality we let $d_u = 1 < d$.

Figure 1 shows a plot of the utility function $u(y, d)$ for values $k = 0.5$ and $\alpha = 1.5$. As we can see from the plot, $u(y, d)$ goes down with larger amount of testing as well as with increasing level of acceleration, that is, utility is decreasing in both $y$ and $d$.
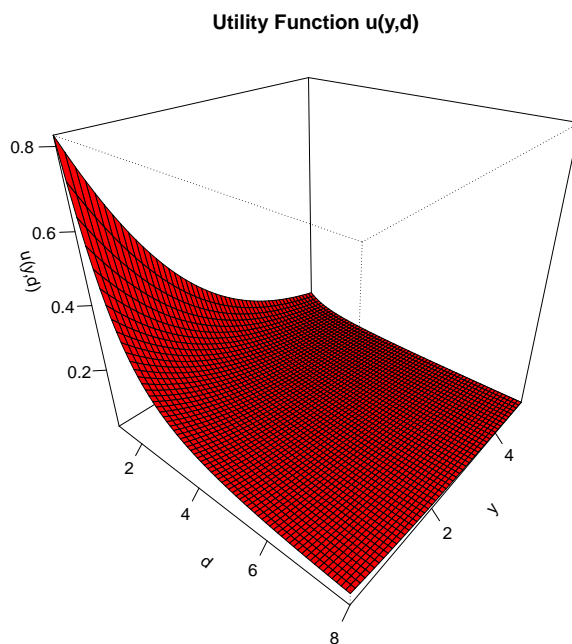


Figure 1: Plot of $u(y, d)$ for $k = 0.5$ and $\alpha = 1.5$.

The utility function $u(y, d)$ is a conjugate utility function in the sense of Lindley (1976). We can obtain

$$u_1(\theta, d) = E_Y[u(y, d)] = \frac{1}{d^\alpha} \left( \frac{\theta d}{\theta d + k} \right)^n. \tag{13}$$

Figure 2 shows plots of the utility function $u_1(\theta, d)$ for two different values of $k$ when $\theta$ and the parameters are fixed as $\theta = 0.125$, $n = 3$, and $\alpha = 1.5$. As previously mentioned, $k$ represents the cost of testing. Figure 2 shows that as $k$ gets larger the utility goes down and the optimal value of $d$ increases. In other words, as the testing cost gets larger it is preferred to use a more accelerated environment to cut down test time.
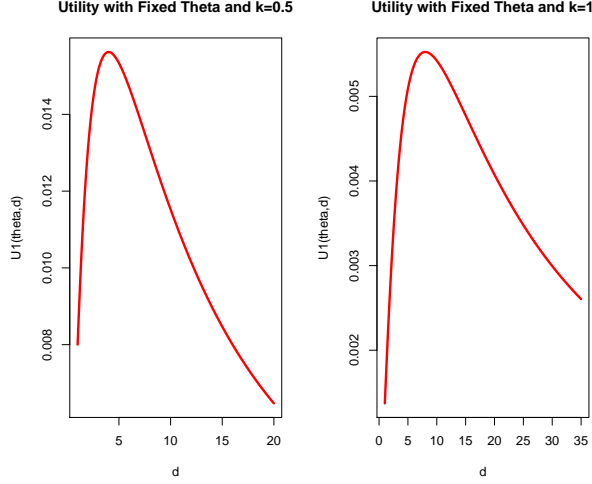
9

Figure 2: Plots of $u_1(\theta, d)$ for $k = 0.5$ and $k = 1$.

The expected utility $u(d)$ can be evaluated as

$$u(d) = E_\theta[u_1(d, \theta)] = \int \frac{1}{d^\alpha} \left( \frac{\theta d}{\theta d + k} \right)^n \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \, d\theta, \tag{14}$$

which can be written as

$$u(d) \propto d^{n-\alpha} \int \frac{\theta^{a+n-1}}{(1 + \theta d/k)^n} e^{-b\theta} \, d\theta. \tag{15}$$

The integrand in (15) is proportional to a *Kummer distribution* [see Armero and Bayarri (1997)] whose density for a nonnegative random variable $X$ is

$$f(x|\phi, \beta, \delta, \gamma) = C \frac{x^{\phi-1} e^{-\beta x}}{(1 + \delta x)^\gamma}. \tag{16}$$

The normalization constant $C$ in (16) is given by

$$C^{-1} = \frac{\Gamma(\phi)}{\delta^\phi} K(\phi, \phi + 1 - \gamma, \beta/\delta)$$

where $K(\phi, \phi + 1 - \gamma, \beta/\delta)$, is a Kummer function (or a confluent hypergeometric function) of second type; see Armero and Bayarri (1997, pp. 248). By setting $\phi = a + n$, $\beta = b$, $\gamma = n$, and $\delta = d/k$ in our case, we can obtain

$$u(d) \propto \frac{K(a+n, \ a+1, \ bk/d)}{d^{\alpha+a}}. \tag{17}$$

Due to conjugacy of the utility function (12), it is possible to obtain $u(d)$ in the analytical form (17). However, maximization of $u(d)$ with respect to $d$ involves evaluation of the Kummer function $K(a+n, \ a+1, \ bk/d)$ and as a result optimization of $u(d)$ is not straightforward. This can pose more computational challenges in the multiple point design problems where $d$ is multi-dimensional.

10

## 4.1 APS Implementation

Using the conjugate utility (12) and the gamma prior for $\theta$, the augmented probability model of (10) can be written as

$$\pi_J(y^J, \theta^J, d) \propto \prod_{j=1}^{J} \frac{1}{d^\alpha} e^{-ky_j} (\theta_j d)^n y_j^{n-1} e^{-\theta_j dy_j} \theta_j^{a-1} e^{-\theta_j b}. \tag{18}$$

We assume that $d$ is uniform over $(1, d_{max})$. Samples can be drawn from the augmented probability model (18) using MCMC. The conjugacy of the utility function enables us to use a Gibbs sampler where all full conditional distributions are available.

For the full conditional of $d$ we have

$$\pi_J(d|y^J, \theta^J) \propto \prod_{j=1}^{J} d^{n-\alpha} e^{-\theta_j dy_j} = d^{J(n-\alpha)} e^{-d\sum_{j=1}^{J} \theta_j y_j} \tag{19}$$

which is a gamma with parameters $J(n-\alpha) + 1$ and $s_J = \sum_{j=1}^{J} \theta_j y_j$.

Full conditional of $\theta_j$'s, for $j = 1, \ldots, J$, is given by

$$\pi_J(\theta_j|y_j, d) \propto \theta_j^{n+a-1} e^{-\theta_j(b+dy_j)} \tag{20}$$

which is a gamma density with parameters $(n + a)$, and $(b + dy_j)$. Finally, for $y_j, j = 1, \ldots, J$, we have

$$\pi_J(y_j|\theta_j, d) \propto y_j^{n-1} e^{-y_j(k+d\theta_j)} \tag{21}$$

a gamma density with parameters $n$ and $(k + d\theta_j)$. By drawing iteratively from the full conditionals (19)-(21), we obtain $g = 1, \ldots, G$ samples $(y_j^{(g)}|\theta_j^{(g)}, d^{(g)}; j = 1, \ldots, J)$ from the augmented probability model (18). The mode of the $d^{(g)}$'s histogram collapses on the optimal design. Implementation of the algorithm requires specification of $J$ which will affect the convergence of the Gibbs sampler. In most cases, values as large as $J = 5$ will be adequate. As suggested by Ekin et al. (2014), for practical purposes, one can increase the value of $J$ until the samples stabilize.

For illustrating our approach, we chose the utility function (12) with $k = 2$ and $\alpha = 1.25$ and $\alpha = 1.75$. Note that choice of $k = 2$ implies a higher cost of testing than what is considered in Figure 2. We assumed a sample size of $n = 2$ and used a gamma prior for $\theta$ with shape $a = 2$ and scale $b = 20$ implying a mean of 0.1 and a standard deviation 0.07. In Figure 3, we present the distribution of $d$ with $J = 5$ for values of $\alpha = 1.25$ and $\alpha = 1.75$. As expected, the higher values of $\alpha$ increase penalty of being away from the use stress and thus the optimal design value shifts to the left.

Figure 4 illustrates that as the value of $J$ is increasing, the distribution of $d$ becomes more peaked around the optimal design $d^*$.
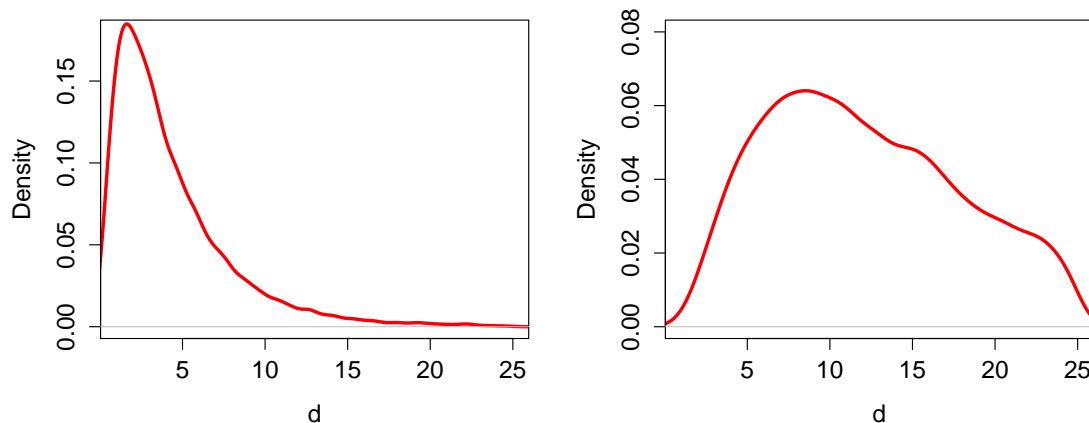
11

Figure 3: Plots of the distribution of $d$ for $\alpha = 1.25$ (right panel) and $\alpha = 1.75$ (left panel).
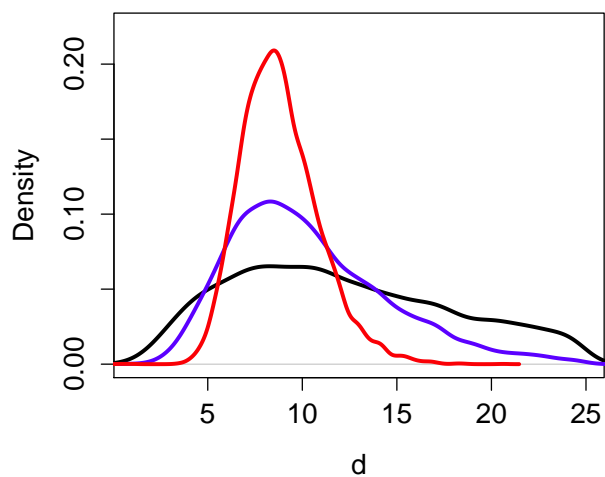


Figure 4: Plots of the distribution of $d$ for values of $J = 5, 15, 50$.

### 4.1.1 Augmented Simulation for Multiple Point Designs

Another useful feature of augmented probability models and conjugate utility functions is that they can be extended to multiple point ALT designs. For example, for a two-point fixed designs where we need to choose $d_1$ and $d_2$ so as to that we maximize the expected value of utility function $u(y_1, y_2, d_1, d_2)$ given by the prod-

12

uct

$$u(y_1, y_2, d_1, d_2) = \frac{1}{(d_1 d_2)^\alpha} e^{-k(y_1 + y_2)}. \tag{22}$$

Here $y_1$ and $y_2$ denote the testing time at stress environments $d_1$ and $d_2$ where we test $n_1$ and $n_2$ items respectively.

Our augmented probability model is given by

$$\pi_J(\theta^J, y^J, d) \propto \left( \prod_{j=1}^{J} u(y_{1j}, y_{2j}, d_1, d_2) p(y_{1j}, y_{2j} | \theta_j, d_1, d_2) p(\theta_j) \right) \tag{23}$$

where $y^J = (y_1^J, y_2^J)$ and $d = (d_1, d_2)$. This leads to a joint distribution

$$\pi_J(\theta^J, y^J, d) \propto \prod_{j=1}^{J} \frac{1}{d_1^\alpha d_2^\alpha} e^{-k(y_{1j}+y_{2j})} (\theta_j d_1)^{n_1} y_{1j}^{n_1-1} e^{-\theta_j d_1 y_{1j}} (\theta_j d_2)^{n_2} y_{2j}^{n_2-1} e^{-\theta_j d_2 y_{2j}} \theta_j^{a-1} e^{-\theta_j b}.$$

To implement our simulation, we still use a Gibbs sampler as all full conditionals are available. Specifically, we have

$$\pi_J(d_i | y_i^J, \theta^J) \propto \prod_{j=1}^{J} d_i^{n_1-\alpha} e^{-\theta_j d_i y_{ij}} = d_i^{J(n_1-\alpha)} e^{-d_i \sum_{j=1}^{J} \theta_j y_{ij}},$$

for $i = 1, 2$. Thus, $\pi_J(d_i | \theta^J, y_i^J)$ is a gamma density with parameters $J(n_1 - \alpha) + 1$ and $\sum_{j=1}^{J} \theta_j y_{ij}$. Furthermore, we obtain

$$\pi_J(\theta_j | y_j, d) \propto \theta_j^{n_1 + n_2 + a - 1} e^{-\theta_j (b + d_1 y_{1j} + d_2 y_{2j})}$$

which is a gamma density with parameters $(n_1 + n_2 + a)$ and $(b + d_1 y_{1j} + d_2 y_{2j})$. For the total time on test variables we have

$$\pi_J(y_{ij} | \theta_j, d_i) \propto y_{ij}^{n_i - 1} e^{-y_{ij}(k + d_i \theta_j)}$$

which is $Gam(n_i, k + d_i \theta_j)$ for $i = 1, 2$.

The main computational difficulty is that we have to simulate from the high dimensional joint distributions (18) and (23). In some cases, MCMC methods can be inefficient due to high correlation induced by the common design variable, $d$. An alternative is to use particle-based methods that can learn static parameters such as $d$. We now construct a tailored particle-based APS approach.

## 4.2 Particle-Based APS

The stochastic simulation problem is to draw $G$ samples $(y_j^{(g)}, \theta_j^{(g)}, d^{(g)})_{g=1}^{G}$, for $j = 1, \ldots, J$, from the joint distribution $\pi_J(y^J, \theta^J, d) \propto \prod_{j=1}^{J} u(y_j, \theta_j, d) p(y_j, \theta_j | d)$ as

13

defined by (18). The basic idea behind the particle simulation is to resample using the utility function and then propagate with the weighted distribution.

Carvalho et al. (2010) provide a framework for particle methods when there are static parameters. Static parameters are handled by using particle methods to draw from the corresponding conditional sufficient statistics. The key is for $j = 1, ..., J$, to construct sufficient statistics so that the conditional posterior simplifies to

$$\pi_j(d|y^j, \theta^j) = \pi_j(d|s_j), \tag{24}$$

where $y^j = (y_1, ..., y_j)$ and $\theta^j = (\theta_1, ..., \theta_j)$. The sufficient statistics, $s_j$'s, also have the property that they satisfy a recursion $s_j = s(s_{j-1}, y_j, \theta_j)$. This can be used to propagate sufficient statistic particles in an efficient manner. Given $Jth$ stage particles $s_J^{(g)}, g = 1, ..., G$, we can Rao-Blackwellize to obtain $\pi_J(d)$

$$\pi_J(d) = \frac{1}{G} \sum_{g=1}^{G} \pi_J(d|s_J^{(g)}). \tag{25}$$

This suggests that using conjugate utility functions is helpful in order to obtain the sufficient statistics.

Given particles $s_j^{(g)}, g = 1, ..., G$, our particle design algorithm proceeds as a four-step procedure:

1. Resample $\bar{s}_j^{(g)}; g = 1, \ldots, G$, using weights proportional to $u(s_j^{(g)})$ where

$$u(s_j^{(g)}) \propto \int_d \int_{\theta_{j+1}} \int_{y_{j+1}} u(y_{j+1}, d) p(y_{j+1}|\theta_{j+1}, d) p(\theta_{j+1}) \pi_j(d|s_j^{(g)}).$$

2. Propagate $(y_{j+1}^{(g)}, \theta_{j+1}^{(g)})$ using $\pi_{j+1}(y_{j+1}^{(g)}, \theta_{j+1}^{(g)}|\bar{s}_j^{(g)})$.

3. Calculate sufficient statistic $s_{j+1}^{(g)} = s(\bar{s}_j^{(g)}, y_{j+1}^{(g)}, \theta_{j+1}^{(g)})$.

4. Update $d$ using $\pi_{j+1}(d|s_{j+1})$.

If $j = J$ estimate $\pi_J(d) = \frac{1}{G} \sum_{g=1}^{G} \pi_J(d|s_J^{(g)})$ and obtain optimal decision

$$d^* = \frac{1}{G} \sum_{g=1}^{N} E(d|s_J^{(g)}).$$

14

### 4.2.1 Example for Particle-Based Design

To illustrate our particle-based approach, we consider the augmented probability model (18) with the conjugate utility function (12).

Step 1 of the particle algorithm requires evaluation of the resampling weights $u(s_j^{(g)})$ which are marginal expected utilities at stage $j$. We can write $u(s_j)$ as

$$u(s_j) \propto \int_d \int_{\theta_{j+1}} u_1(\theta_{j+1}, d) p(\theta_{j+1}) \pi_j(d|s_j). \tag{26}$$

It follows from (19) that $\pi_j(d|s_j)$ is a gamma density with shape parameter $j(n - \alpha) + 1$ and scale parameter $s_j = \sum_t^j \theta_t y_t$. Since the prior for $\theta_{j+1}$ is also gamma with parameters $a$ and $b$, we can write

$$u(s_j) \propto \int_d d^{n-\alpha} \pi_j(d|s_j) \int_{\theta_{j+1}} \frac{\theta_{j+1}^{a+n-1}}{(1 + \theta_{j+1}d/k)^n} e^{-b\theta_{j+1}}$$

and note that $\pi_{j+1}(\theta_{j+1}|d)$ is a Kummer distribution as in (16) with parameters $(a + n, a + 1, bk/d)$. Thus,

$$u(s_j) \propto \int_d \int_{\theta_{j+1}} \pi_{j+1}(\theta_{j+1}|d) \pi_j(d|s_j). \tag{27}$$

and the resampling weights are given by

$$u(s_j) \propto \int_d \frac{K(a + n,\, a + 1,\, bk/d)}{d^{\alpha+a}} \pi_j(d|s_j). \tag{28}$$

The integration in (28) can not be performed analytically, but we can evaluate (28) by noting that $u(s_j) \propto \pi_j(s_j)$. In obtaining $\pi_j(d|s_j)$, the posterior distribution of $d$, $\pi_j(s_j)$ is the marginal likelihood term in Bayes' law. This term can be evaluated numerically since all densities in the numerator and denominator of the Bayes' law are of known forms.

For the propogation of next particles, $(\theta_{j+1}^{(g)}, y_{j+1}^{(g)})$ we need to generate from the conditional posterior, $\pi_{j+1}(y_{j+1}, \theta_{j+1}|\bar{s}_j^{(g)})$, which can be written as

$$\pi_{j+1}(\theta_{j+1}, y_{j+1}|\bar{s}_j^{(g)}) = \int_d \pi_{j+1}(y_{j+1}|\theta_{j+1}, d) \pi_{j+1}(\theta_{j+1}|d) \pi_j(d|\bar{s}_j^{(g)}).$$

Since both $\pi_{j+1}(y_{j+1}|\theta_{j+1}, d)$ and $\pi_j(d|\bar{s}_j^{(g)})$ are gamma densities and $\pi_{j+1}(\theta_{j+1}|d)$ is a Kummer density, given the resampled particles $\bar{s}_j^{(g)}$ we can propogate $(\theta_{j+1}^{(g)}, y_{j+1}^{(g)})$. Thus, starting with a prior set of particles $s_0^{(g)}$ we only need carry over particles of sufficient statistics from one stage to the next in the augmented probability simulation.

Particle methods are particularly useful for sequential designs as they can be updated in an online fashion as opposed to MCMC simulation which has to be a run for the entire system. In the context of Bayesian designs see Amzal et al. (2006), Gramacy and Polson (2011) and by Azadi et al. (2014).

# 5   Discussion

Designing ALT's requires one to maximize expected utility. We propose an augmented probability simulation approach. One advantage of our method over traditional ones is that our proposed approach enables us to sample more frequently from high utility regions of the augmented design and uncertainty space. From a theoretical perspective, we exploit the use of class of conjugate utility functions as given in Lindley (1976). This provides simplification in implementation of the MCMC and particle-based methods. The key property that conjugate utilities and prior distributions possess is that they allow us to obtain conditional sufficient statistics. We show how this provides further computational efficiency for calculating the optimal design. We provide an extension of our approach to multiple point fixed designs. There are many areas for future research, in particular fully sequential settings where particle filtering methods; see see Johannes and Polson (2009), can be used for obtaining optimal designs. Furthermore, alternative sampling methods such as nested sampling can be incorporated into APS as recently proposed by Ekin, Polson and Soyer (2016).

# References

Amzal, B., Bois, F. Y., Parent, E. and Robert, C. P. (2006). Bayesian-optimal design via interacting particle systems, *Jour. of Amer. Stat. Assoc.*, Vol. 101, pp. 773-785.

Armero, C. and Bayarri, M. J. (1997). A Bayesian analysis of a queueing system with unlimited service, *Jrnl. of Statist. Plan. & Inf*, Vol. 58, pp. 241-261.

Azadi, N. A., Fearnhead, P., Ridall, G. and Blok, J. H.(2014). Bayesian sequential experimental design for binary response data with application to electromyographic experiments, *Bayesian Analysis*,Vol. 9, pp. 287-306.

Bernardo, J. M. (1979). Expected information as expected utility, *Annals of Statistics*, Vol. 7, pp. 686690.

Bielza, C., Muller, P. and Insua-Rios, D. (1999). Decision analysis by augmented probability simulation, *Management Science*, Vol. 45, pp. 995-1007.

Carvalho, C., M.S. Johannes, H.F. Lopes and N.G. Polson (2010). Particle learning and smoothing, *Statistical Science*, Vol. 25, pp. 88-106.

Chaloner, K. (1984). Optimal Bayesian experimental designs for linear models, *Ann. of Statist.*, Vol. 12, pp. 283−300.

Chaloner, K. and Larntz, K. (1992). Bayesian design for accelerated life testing, *Jrnl. of Statist. Plan. & Inf*, Vol. 33, pp. 245−259.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review, *Statistical Science*, Vol. 10, pp. 273-304.

Chernoff, H. (1962). Accelerated life designs for estimation. *Technometrics*, Vol. 4, pp.381-408.

DeGroot, M. H. and Goel, P. (1979). Bayesian estimation and optimal designs in partially accelerated life testing. *Naval Res. Logistics Quarterly*, Vol. 26, pp. 223-235.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families, *The Annals of Statistics*, Vol. 7, pp. 269-281

Ebrahimi, N., Soofi, E. S. (2010). On the sample information about parameter and prediction, *Statistical Science*, Vol. 25, pp. 348-367.

Ekin, T., Polson, N. G. and Soyer, R. (2014). Simulation-based two-stage stochastic programming with recourse, *Decision Analysis*, Vol. 11, pp. 250-264.

Ekin, T., Polson, N. G. and Soyer, R. (2016). Augmented nested sampling for stochastic programs with recourse and decision dependent uncertainty, under review.

Erkanli, A. and Soyer, R. (2000). Simulation-based designs for accelerated life tests, *Jrnl. of Statist. Plan. & Inf.*, Vol. 90, pp. 335-348.

Escobar, L. A. and Meeker, W. Q. (1995). Planning accelerated life tests with two or more experimental factors, *Technometrics*, Vol. 37, pp. 411-427.

Gramacy, R. B. and Polson, N. G. (2011). Particle learning of Gaussian process models for sequential design and optimization, *Journal of Computational and Graphical Statistics*, Vol. 20, pp. 102-118.

Johannes, M. and N.G. Polson (2009). Particle filtering, in *Handbook of Financial Time Series*. (Eds. T. G. Anderson, et al.), 1015-1029.

Kadane, J. B. and Winkler, R. L. (1988). Separating probability elicitation from utilities, *Jour. of Amer. Stat. Assoc.*, Vol. 83, pp. 357-362.

Lindley, D. V. (1956). On a measure of information provided by an experiment, *Ann. Math. Statist.*, Vol. 27, pp. 9861005.

Lindley, D.V. (1976). A class of utility functions, *The Annals of Statistics*, Vol. 4, pp. 1-10.

Lindley, D.V. (1985). *Making Decisions*, John Wiley and Sons, London.

Martz, H. F. and Waterman, M. S. (1978). A Bayesian model for determining the optimal test stress for a single test unit, *Technometrics*, Vol. 20, pp. 179-185.

Menzefricke, U. (1991). Designing accelerated life tests when there is Type II censoring. *Commun Statist.-Theory Meth.*, Vol. 21, pp. 2569-2590.

Muller, P. and Parmigiani, G. (1995). Optimal design via curve fitting of Monte Carlo experiments, *Jour. of Amer. Stat. Assoc.*, Vol. 90, pp. 503-510.

Muller, P. (1999). Simulation-based optimal design, *Bayesian Statistics 6*, (J. O. Berger, J. M. Bernardo, A. P. Dawid and A. F. M. Smith, Eds.), pp. 459-474.

Muller, P., Sanso B., and De Iorio, M. (2004). Optimal Bayesian design by inhomogeneous Markov chain simulation, *Jour. of Amer. Stat. Assoc.*, Vol. 99, pp. 788798.

Polson, N. G. (1993). A Bayesian perspective on the design of accelerated life tests, in *Advances in Reliability*, (A. P. Basu, Ed.), North Holland, pp. 321-330.

Singpurwalla, N. D. (2006). *Reliability and Risk: A Bayesian Perspective*, John Wiley and Sons, UK.

Soyer, R. and Vopatek, A. L. (1995). Adaptive Bayesian designs for accelerated life testing, in *Adaptive Designs*, (N. Flournoy and W. F. Rosenberger, Eds.), Institute of Mathematical Statistics, Lecture Notes, Vol. 25, 1995, pp. 263-275.

Soyer, R. (2007). Accelerated Life Tests: Bayesian Design, in *The Encyclopedia of Statistics in Quality and Reliability* (F. Ruggeri, R. S. Kenett and F. W. Faltin, Eds.), Vol. 1, pp. 12-20, John Wiley and Sons, UK.

Verdinelli, I. (1992). Advances in Bayesian experimental design. In *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). Oxford: University Press, pp. 46748.

Verdinelli, I., Polson, N. G., and Singpurwalla, N. D. (1993). Shannon information and Bayesian design for prediction in accelerated life testing, in *Reliability and Decision Making*, pp. 247-256. (R. E. Barlow, C. A. Clariotti and F. Spizzichino, Eds.). Chapman & Hall.

Vopatek, A. L. (1992). Design of Accelerated Life Tests: A Bayesian Approach. *Doctoral Dissertation*, The George Washington University.

Zelen, M. (1959). Factorial experiments in life testing, *Technometrics*, Vol. 1, pp. 269288.

Zhang, Y. and Meeker, W. Q. (2006) Bayesian Methods for Planning Accelerated Life Tests, *Technometrics*, Vol. 48, pp. 49-60.