A Bayesian optimal design for accelerated degradation testing based on the Inverse Gaussian process

Xiaoyang Li, Yuqing Hu, Enrico Zio, Rui Kang

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 15, 2017; accepted March 8, 2017, date of publication March 20, 2017, date of current version May 17, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2683533

A Bayesian Optimal Design for Accelerated Degradation Testing Based on the Inverse Gaussian Process

XIAOYANG LI[1], YUQING HU[2], ENRICO ZIO[3,4], (Member, IEEE), AND RUI KANG[1]

[1]School of Reliability and Systems Engineering, Science and Technology on Reliability and Environmental Engineering Laboratory, Beihang University, Beijing 100191, China
[2]School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China
[3]Department of Energy, Politecnico di Milano, Milano 20156, Italy
[4]Chair System Science and the Energy Challenge, Fondation Electricite de France, CentraleSupelec, Universite Paris-Saclay, Paris 92290, France

Corresponding author: X. Li (leexy@buaa.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61603018, Grant 61104182, and Grant 61573043, and in part by the Fundamental Research Funds for the Central Universities under Grant YWF-16-JCTD-A-02-06.

ABSTRACT Accelerated degradation testing (ADT) is commonly used to obtain degradation data of products by exerting loads over usage conditions. Such data can be used for estimating component lifetime and reliability under usage conditions. The design of ADT entails to establish a model of the degradation process and define the test plan to satisfy given criteria under the constraint of limited test resources. Bayesian optimal design is a method of decision theory under uncertainty, which uses historical data and expert information to find the optimal test plan. Different expected utility functions can be selected as objectives. This paper presents a method for Bayesian optimal design of ADT, based on the inverse Gaussian process and considering three objectives for the optimization: relative entropy, quadratic loss function, and Bayesian D-optimality. The Markov chain Monte Carlo and the surface fitting methods are used to obtain the optimal plan. By sensitivity analysis and a proposed efficiency factor, the Bayesian D-optimality is identified as the most robust and appropriate objective for Bayesian optimization of ADT.

INDEX TERMS Accelerated degradation testing, Bayesian optimal design, inverse Gaussian process, Markov chain Monte Carlo (MCMC), surface fitting.

I. INTRODUCTION

Accelerated testing is an effective method to access the reliability and lifetime of products in a short time, especially for high reliability and long life products [1]. The optimal design of accelerated life/degradation testing (ALT/ADT) aims to find a cost-effective testing plan, which trades off different objectives and constraints. Traditional optimal design is based on an acceleration life/degradation model with specified parameter values. The optimal test plan designed is, then, generally referred to as the local optimal solution. Differences between the specified parameter values and the true values cause the optimal test plan to be sub-optimal. Based on available historical data and expert information, a prior distribution can be assigned to account for parameter uncertainty, and then, the optimal test plan can be obtained by averaging over the parameter space and sample space.

Recently, the Bayesian optimal design method has been applied in ALT design. Zhang and Meeker [2] studied the Bayesian ALT design based on censored data from a log-location-scale distribution. Xu and Tang [3] proposed a Bayesian optimal design with the objective of expected relative entropy between the posterior and the prior distributions of the parameters. ALT can shorten testing time, but for some products with extremely high reliability and long lifetime, even Bayesian ALT methods are not suitable, as few physical failures can be observed even in ALT. On the other hand, failure of a product is associated with the degradation of its characteristics. Degradation of a product accumulates over time and causes failure when it exceeds a failure threshold. This naturally provides a relationship between degradation and life of a product [4], [5]. For this reason, ADT has been developed to avoid the aforementioned problem in ALT. Limited published literature is available on Bayesian ADT optimal design, compared to ALT. Models of stochastic degradation processes are used in ADT, different from the models used in ALT; consequently, the optimal methods proposed in ALT cannot be directly used for ADT. Shi and Meeker [6] presented a Bayesian method for accelerated destructive degradation tests (ADDT), under a class of nonlinear degradation models with one accelerating variable. The optimal objective was to maximize the precision of a specified failure-time distribution quantile under usage conditions. A large-sample approximation for the posterior distribution was made to provide a useful simplification to the planning criterion. The general equivalence theorem was applied to verify the global optimality of the numerically optimized test plans. Li *et al.* [7] presented a Bayesian methodology for designing step stress accelerated degradation testing (SSADT) with the objective of relative entropy. It is assumed that the degradation model follows a drift Brownian motion and the acceleration model follows Arrhenius equation, and the corresponding parameters follow normal and Gamma prior distributions. The Markov Chain Monte Carlo (MCMC) and surface fitting method are chosen to solve for optimality. Liu and Tang [8] proposed a Bayesian design method for ADT, with physically based statistical models. The hot-carrier-induced degradation of modern MOS field is considered as case study of transistor degradation. A single-path power-law statistical degradation model with nonlinear stresslife relationships is developed. Based on this model, the optimal objective is to minimize the expected pre-posterior variance of the quantile life at the use condition.

Bayesian design is based on the theory of making optimal decisions under uncertainty. The aim is to find an optimal test plan by maximizing the expected utility of the test outcome. The optimal plan is targeted to specific objectives. There are different mathematical expressions for the expected utility function, such as relative entropy, quadratic loss function and Bayesian alphabetic optimal. Relative entropy is one of the most widely used objectives [3], [7], [9], providing a measurement of the distance between prior and posterior distributions. From the Shannon information point of view, relative entropy can also represent the increased information provided by testing. There is also some published literature on Bayesian optimal design with the quadratic loss function as the objective [2], [6], [8], [10], which minimizes the asymptotic variance of the maximum likelihood estimator (MLE) of a specified quantile lifetime at the usage condition. Bayesian alphabetic optimality considers the precision of model parameters, minimizing the asymptotic variance of the MLE of a function of the model parameters. The most common alphabetic optimization is D-optimality [11], [12] which minimizes the determinant of the covariance matrix of the model parameter estimates [13]. Various optimal objectives have been used to design ALT/ADT plans, and different objectives are expressed by different mathematical functions. However, limited published literature is concerned about the difference and applicability of the objectives. In this paper, the ADT optimal design with the objectives of relative entropy, quadratic loss function and D-optimality are investigated with respect to the difference and applicability of these objectives.

The other important element of ADT Bayesian optimal design is the degradation model. There are two popular models for ADT data, which are the Wiener process and the Gamma process. Although the Wiener process and the Gamma process have received much attention in degradation data analysis, they cannot handle all degradation problems. For example, Wang and Xu [14] found that neither models fit the GaAs laser degradation data well ([15, Example 13.5]), and built a degradation model based on inverse Gaussian (IG) process that fits the GaAs degradation data better. Ye and Chen [16] systematically investigated the IG process and showed its advantages as degradation model. Ye *et al.* [17] provided Bayesian optimal design of constant stress ADT considering the stochastic IG process with the objective of quadratic loss function. Wang *et al.* [18] provided an optimal SSADT plan for the IG degradation process. Under the constraint of total experimental budget, design variables were optimized by minimizing the asymptotic variance of the estimated p-quantile of the lifetime distribution of the product. The sensitivity and stability of the SSADT plan were studied to verify that the optimal test plan is quite robust for a moderate departure from the values of the parameters.

The purpose of this paper lies in two aspects: firstly, it is to investigate the Bayesian planning method for SSADT using the IG process, with the objectives of relative entropy, quadratic loss function and D-optimality; secondly, it is to present the different advantages of the three objectives by comparison. The remaining paper is organized as follows. In Section 2, the framework of Bayesian optimal design for SSADT based on the IG process is presented. In Section 3, the Bayesian optimal criterion is presented with the optimal objectives of relative entropy, quadratic loss function, Bayesian D-optimality, respectively, and the optimal model is constructed with cost constraint. The Markov Chain Monte Carlo (MCMC) and surface fitting method are utilized to obtain the optimal plan. In Section 4, applications are presented and conclusions are drawn based on a number of comparisons.

II. IG PROCESS IN ADT

We assume that the degradation path of a product satisfies the IG process. Let $Y(t)$, $t \geq 0$ be the degradation path of a product. The product fails when its degradation path $Y(t)$ reaches a predefined threshold level Y_D and the associated first-passage-time is denoted by T_D.

A. THE IG PROCESS

If a degradation process has the following three properties, we say that it is an IG process,

(1) $Y(0) = 0$ with probability one;

(2) $Y(t)$ has independent increments, i.e., $Y(t_2) - Y(t_1)$ and $Y(t_4) - Y(t_3)$ are independent, for $0 \leq t_1 < t_2 \leq t_3 < t_4$;

(3) Each increment follows an IG distribution, i.e., $\Delta Y(t) \sim \mathcal{IG}(\mu\Delta\Lambda(t), \lambda\Delta\Lambda(t)^2)$, where $\mu > 0$, $\lambda > 0$, $\Delta\Lambda = \Lambda(t) - \Lambda(s)$, $\Lambda(t)$ is a given, monotone increasing function of time t with $\Lambda(0) = 0$.

For any $x > 0$, the probability density function (PDF) of $\mathcal{IG}(u, v)$, $u > 0$, $v > 0$, with mean u and variance u^3/v is defined by,

$$f_{\mathcal{IG}}(x; u, v) = \sqrt{\frac{v}{2\pi x^3}} \cdot \exp\left[-\frac{v(x-u)^2}{2u^2 x}\right]. \tag{1}$$

Then, the degradation process can be characterized by $Y(t) \sim \mathcal{IG}(\mu\Lambda(t), \lambda\Lambda(t)^2)$. The mean and variance of $Y(t)$ are $\mu\Lambda(t)$ and $\mu^3\Lambda(t)/\lambda$, respectively. Substituting $u = \mu\Lambda(t)$ and $v = \lambda\Lambda^2(t)$ into (1) yields the PDF of $\mathcal{IG}(\mu\Lambda(t), \lambda\Lambda^2(t))$ as,

$$f_{\mathcal{IG}}(x; \mu, \lambda) = \sqrt{\frac{\lambda(\Lambda(t))^2}{2\pi x^3}} \cdot \exp\left[-\frac{\lambda(x-\mu\Lambda(t))^2}{2\mu^2 x}\right] \tag{2}$$

The parameter μ could be assumed as the degradation rate of a product and it is a function of the accelerated stress S, i.e. it is an acceleration model, which is denoted as $\mu(S)$ and could be written as follows,

$$\mu(S) = \exp[a + b\varphi(S)] \tag{3}$$

where, the parameters a and b need to be estimated from ADT. For convenience, the stress level can be standardized using a normalization scheme. In this paper, the linear normalization method is applied. Let S_0 and S_H be the usage stress level and the highest stress level that can be used in the test, respectively. Then, $\varphi(S)$ is a standardized function of S and expressed as,

$$\varphi(S) = \left(\xi(S_j) - \xi(S_0)\right)/(\xi(S_H) - \xi(S_0)) \tag{4}$$

where $\xi(S)$ represents a pregiven function of S. For example, if the accelerated stress is temperature, $\xi(S) = 1/S$, and if the accelerated stress is electricity, $\xi(S) = \ln(S)$.

The parameter λ has no physical meaning and is a constant, i.e. if there are K accelerated stress levels in an ADT, then $\lambda_1 = \lambda_2 = \cdots = \lambda_K$. Since μ and λ are not dependent on time, the process is a homogeneous IG process or simple IG process. Generally, there are three different shapes of performance degradation trend, linear, convex and concave. Hence, it is appropriate to assume $\Lambda(t) = t^\beta$, $\beta > 0$ [10], because when $0 < \beta < 1$, the trend is conve; when $\beta = 1$, the trend is linear; and when $\beta > 1$, then trend is concave. As the path of the IG process is strictly increasing, the cumulative distribution function (CDF) of T_D can be expressed as,

$$F_{Y_D}(t) = P(Y(t) \geq Y_D) = \Phi\left[\sqrt{\frac{\lambda}{Y_D}}(t^\beta - \frac{Y_D}{\mu})\right] - \exp\left(\frac{2\lambda t^\beta}{\mu}\right) \cdot \Phi\left[-\sqrt{\frac{\lambda}{Y_D}}\left(t^\beta + \frac{Y_D}{\mu}\right)\right]. \tag{5}$$

According to [16], when both $\mu\Lambda(t)$ and t are large, $Y(t)$ is approximately normally distributed with mean $\mu\Lambda(t)$ and variance $\mu^3\Lambda(t)/\lambda$. Therefore, the CDF of Y_D can be approximated as,

$$F_{Y_D}(t) = 1 - \Phi\left[\frac{Y_D - \exp(a + b\varphi(S))t^\beta}{\sqrt{\exp(a + b\varphi(S))^3 t^\beta/\lambda}}\right] \tag{6}$$

The p-quantile lifetime of Y_D, based on this approximation is,

$$t_p = \Lambda^{-1}\left[\frac{\mu}{4\lambda}\left(z_p + \sqrt{(z_p)^2 + 4Y_D\lambda/\mu^2}\right)^2\right] \tag{7}$$

where z_p is the standard normal p-quantile, and $\Lambda^{-1}(.)$ is the inverse function of $\Lambda(.)$.

For the sake of simplicity, we assume that the four parameters (i.e., a, b, λ and β) in (6) are independent from each other and compose the parameter vector $\boldsymbol{\theta} = (a, b, \lambda, \beta)$.

In practice, it is appropriate to consider $\boldsymbol{\theta}$ a vector of random variables. In this study, when the degradation increment x follows an IG distribution, λ and β should be positive. Therefore, common positive distributions (i.e. Gamma, Lognormal and Beta distributions) can be used to describe λ and β. Normal, Logistic and Gumbel distributions can be used to describe a and b. The selection of the distributions is discussed in section IV.

B. ADT SETTINGS AND BAYESIAN INFERENCE

We assume that there are n samples for ADT. The K accelerated stress levels are numerically ordered as $S_0 < S_{\min} \leq S_1 < S_2 < \cdots < S_K \leq S_{\max} \leq S_H$. Let $S_{\min}$ and $S_{\max}$ be the lowest and highest stress levels which will be used in ADT. There are m_l degradation measurements on the l^{th} stress level, and the cumulative measurement times of the whole test is M and $M = \sum_{l=1}^{K} m_l$.

Let τ be the non-overlapped interval of degradation measurement constant during SSADT; then, the test duration t_l on the l^{th} stress level is $t_l = \tau m_l$, and the total test duration is $T_0 = \tau * M$.

Let $Y(t_{ilj})$ be the measurement result of the j^{th} measurement of the i^{th} item on the l^{th} stress level at time t_{ilj} $(i = 1, 2, \cdots, n, l = 1, 2, \cdots, K, j = 1, 2, \cdots, m_l)$. The degradation increment is $x_{ilj} = Y(t_{il(j+1)}) - Y(t_{ilj})$ and follows (2). Then, based on (2), the likelihood function is (8), as shown at the top of the next page, and, the posterior distribution $p(\boldsymbol{\theta}|x)$ of $\boldsymbol{\theta}$ is,

$$p(\boldsymbol{\theta}|x) = \frac{p(x|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_\Theta p(x|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{9}$$

According to Equation (7) and (9), the Bayesian posterior p-quantile lifetime of Y_D is expressed as follows,

$$\begin{aligned} t(p, \boldsymbol{\theta}|x) &= t(p|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}|x) \\ &= \Lambda^{-1}\left[\frac{\mu}{4\lambda}\left(z_p + \sqrt{(z_p)^2 + 4Y_D\lambda/\mu^2}\right)^2\right] \cdot p(\boldsymbol{\theta}|x) \end{aligned} \tag{10}$$

III. BAYESIAN OPTIMAL MODEL

The optimal design can be obtained by maximizing the expected utility of the experiment [19]. If a design plan $\boldsymbol{\eta}$ is chosen from the possible plan set $\boldsymbol{D}$, then the sample data x will be collected, and a decision rule d from the decision rule set H is selected with the given $\boldsymbol{\eta}$ and observed x.

$$p(x|\boldsymbol{\theta}) = \prod_{l=1}^{K}\prod_{i=1}^{n}\prod_{j=1}^{m_l}\sqrt{\frac{\lambda\left[\left(m_{il(j+1)}\tau\right)^{\beta}-\left(m_{ilj}\tau\right)^{\beta}\right]^2}{2\pi x_{ilj}^3}}\cdot\exp\left\{-\frac{\lambda\left[x_{ilj}-\exp\left(a+b\varphi(S_l)\right)\right]\left[\left(m_{il(j+1)}\tau\right)^{\beta}-\left(\tau m_{ilj}\right)^{\beta}\right]^2}{2\left[\exp\left(a+b\varphi(S_l)\right)\right]^2 x_{ilj}}\right\} \tag{8}$$

The utility function can be denoted as $U(d, \boldsymbol{\eta}, x, \boldsymbol{\theta})$; then for any design $\boldsymbol{\eta}$, the expected utility of the best decision is expressed as:

$$E(\boldsymbol{\eta}) = \int_{\Omega}\max_{d\in H}\int_{\Theta} U(d, \boldsymbol{\eta}, x, \boldsymbol{\theta})p(\boldsymbol{\theta}\,|x\,, \boldsymbol{\eta})p(x\,|\boldsymbol{\eta}\,)d\boldsymbol{\theta}dx \tag{11}$$

where $p(x|\boldsymbol{\eta})$ denotes the likelihood function under the given $\boldsymbol{\eta}$ and $p(\boldsymbol{\theta}|x, \boldsymbol{\eta})$ denotes the posterior distribution of $\boldsymbol{\theta}$ under the given $\boldsymbol{\eta}$ and observed x. The pre-posterior expected utility $E(\boldsymbol{\eta})$ of the best decision rule is taken as the expectation in the parameter space Θ to account for the uncertainty of the unknown $\boldsymbol{\theta}$ and in the sample space Ω. The Bayesian best plan $\boldsymbol{\eta}^*$can be obtained by maximizing

$$E(\boldsymbol{\eta}^*) = \max_{\boldsymbol{\eta}\in D}\int_{\Omega}\max_{d\in H}\int_{\Theta} U(d, \boldsymbol{\eta}, x, \boldsymbol{\theta})p(\boldsymbol{\theta}\,|x\,, \boldsymbol{\eta})p(x\,|\boldsymbol{\eta}\,)d\boldsymbol{\theta}dx \tag{12}$$

In the case of SSADT, the sample size n, the total number of degradation measurements M, the specified accelerated stress levels $\boldsymbol{S}$, $\boldsymbol{S} = (S_1, S_2, \cdots, S_K)$, and the specified number of degradation measurements on each accelerated stress level $\boldsymbol{m}$, $\boldsymbol{m} = (m_1, m_2, \cdots, m_K)$, are the key elements of a design plan $\boldsymbol{\eta}$, rewritten as $\boldsymbol{\eta}(n, M, \boldsymbol{S}, \boldsymbol{m})$.

To some extent, the utility embodies the effectiveness obtained and the cost spent in an ADT. Moreover, the budget of an ADT is always given in advance and the test should be conducted within this budget. With the utility as the objective and the cost as the constraint of a test, an optimal problem is formulated and the optimal plan should be obtained by a trade-off between them. Therefore, in this optimal problem, n, M, $\boldsymbol{S}$ and $\boldsymbol{m}$ are the decision variables.

A. OPTIMAL OBJECTIVES

1) RELATIVE ENTROPY

In the Bayesian theory, the relative entropy is expressed as showed in [7],

$$RE(\boldsymbol{\eta}) = \int\int p(x|\boldsymbol{\theta}, \boldsymbol{\eta})\log\left(p(x|\boldsymbol{\theta}, \boldsymbol{\eta})\right)\mathrm{d}x - \int p(x)\log p(x)\mathrm{d}x \tag{13}$$

where $p(x)$ is the marginal likelihood function and $p(x|\boldsymbol{\theta}, \boldsymbol{\eta})$ denotes the likelihood function with known parameter vector $\boldsymbol{\theta}$; then, (13) can be expressed as,

$$RE(\boldsymbol{\eta}) = E_x E_{\boldsymbol{\theta}}\log\left(p(x\,|\boldsymbol{\theta}, \boldsymbol{\eta}\,)\right) - E_x\log p(x). \tag{14}$$

Since $p(x|\boldsymbol{\theta}, \boldsymbol{\eta})$ denotes the likelihood function, it is easy to calculate $E_x E_{\boldsymbol{\theta}} log\,[p(x|\boldsymbol{\theta}, \boldsymbol{\eta})]$ using Monte Carlo simulation. However, it is a significant challenge to calculate the marginal likelihood function $p(x)$. The Markov Chain Monte Carlo (MCMC) sampling method is one of the solutions to this problem and can be implemented, for example, using a software like WinBUGS [20]. The harmonic mean estimator introduced by Newton and Raftery [21] will be used to estimate $p(x)$, expressed as,

$$p\,(x) \approx \left\{\frac{1}{N}\sum_{i=1}^{N}\left[p(x, \boldsymbol{\theta}|\boldsymbol{\eta})\right]^{-1}\right\}^{-1} \tag{15}$$

To some extent, the relative entropy is explained as the information gain from a test. Therefore, as an objective, the optimal plan should be obtained by maximizing (14) and written as max $RE(\boldsymbol{\eta})$.

2) QUADRATIC LOSS FUNCTION

Based on (10), the posterior variance of $t(p, \boldsymbol{\theta}|x, \boldsymbol{\eta})$ is the quadratic loss of p-quantile lifetime on the usage condition of Y_D. Since we need to calculate this quadratic loss, denoted as $Var(t(p, \boldsymbol{\theta}|x, \boldsymbol{\eta}))$ before ADT data are collected, its corresponding expectation should be with respect to Îÿ$\boldsymbol{\theta}$ and x. Consequently, we can get the pre-posterior variance as follows.

$$Q(\boldsymbol{\eta}) = E_x E_{\boldsymbol{\theta}}[Var(t(p, \boldsymbol{\theta}|x, \boldsymbol{\eta}))]. \tag{16}$$

The pre-posterior variance is interpreted as the quadratic loss; therefore, the optimal plan should be obtained by minimizing (16) or maximizing $-Q(\boldsymbol{\eta})$, written as max $-Q(\boldsymbol{\eta})$.

3) BAYESIAN D-OPTIMALITY

Chaloner and Larntz [22] proposed the Bayesian D-optimality criterion, written as follows,

$$\Phi(\boldsymbol{\eta}) = E_{\boldsymbol{\theta}}[log(det(I(\boldsymbol{\eta}, \boldsymbol{\theta})))] = \int \log(det(I(\boldsymbol{\eta}, \boldsymbol{\theta})))p(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{17}$$

where the symbol det denotes the determinant of the matrix, and $I(\boldsymbol{\eta}, \boldsymbol{\theta})$ denotes the Bayesian information matrix expressed as follows,

$$I(\boldsymbol{\eta}, \boldsymbol{\theta}) = \begin{pmatrix} E\left(-\frac{\partial^2 L}{\partial a^2}\right) & E\left(-\frac{\partial^2 L}{\partial a\partial b}\right) & E\left(-\frac{\partial^2 L}{\partial a\partial \lambda}\right) & E\left(-\frac{\partial^2 L}{\partial a\partial \beta}\right) \\ & E\left(-\frac{\partial^2 L}{\partial b^2}\right) & E\left(-\frac{\partial^2 L}{\partial b\partial \lambda}\right) & E\left(-\frac{\partial^2 L}{\partial b\partial \beta}\right) \\ & & E\left(-\frac{\partial^2 L}{\partial \lambda^2}\right) & E\left(-\frac{\partial^2 L}{\partial \lambda\partial \beta}\right) \\ \text{symmetrical} & & & E\left(-\frac{\partial^2 L}{\partial \beta^2}\right) \end{pmatrix} \tag{18}$$

where, the symbol *symmetrical* denotes the symmetric matrix, and the matrix $I(\boldsymbol{\eta}, \boldsymbol{\theta})$ must be non-singular. The expressions for all the elements of $I(\boldsymbol{\eta}, \boldsymbol{\theta})$ are given in Appendix A. Then, the Bayesian D-optimality is to maximize (17), which can be expressed as max $\Phi(\boldsymbol{\eta})$.

B. CONSTRAINTS

1) COST CONSTRAINTS

The test cost is mainly related to the test item and operation. The cost related to the test item is expressed as the product of the sample unit price and sample size. The cost related to test operation mainly includes the cost of the resources consumed in the test, such as test labor, power resource, etc. For simplicity, the operation cost is expressed as the product of operation unit price and total test duration. Therefore, the total test cost can be written as follows,

$$n \cdot C_1 + M \cdot \tau \cdot C_2 \leq C_0 \tag{19}$$

where C_1 denotes the sample unit price, C_2 denotes the operation unit price, C_0 denotes the given budget of an ADT.

2) OTHER CONSTRAINTS

In reality, besides the test budget, there might be other constraints to the decision variables given in advance, such as the value range, the relationship between them, etc. From the point of view of practice, we present the following requirements:

a)Sample size n

Generally, we require $n \geq 3$; but the sample size cannot be infinite, there should be a specified limit $n_{\max}$ for the sample size.

b)The number of accelerated stress levels K.

In order to ensure the feasibility and accuracy of the extrapolation in stress dimension, generally, the value of K could be chosen in the range of values $3 \sim 6$.

c) Number of degradation measurements under each accelerated stress levels $\boldsymbol{m} = (m_1, m_2, \cdots, m_K)$.

Since the degradation rate in higher stress levels is greater than in lower stress levels, in order to guarantee that enough useful degradation information can be obtained in all stress levels, more degradation measurements should be assigned in lower stress levels than in higher stress levels, i.e. $m_1 > m_2 >$ ȩ $> m_K (l = 1, 2, \cdots, K)$.

C. OPTIMAL MODEL

Based on the analysis in sections III-A and III-B, the optimal model for Bayesian SSADT design can be expressed as,

$$\begin{cases} \max & E(\boldsymbol{\eta}) \\ s.t. & n \cdot C_1 + M \cdot \tau \cdot C_2 \leq C_0 \\ & 3 \leq n \leq n_{\max} \\ & S_{\min} < S_1 < S_2 < \ldots < S_K \leq S_{\max} \\ & m_1 \geq m_2 \geq \ldots \geq m_K > 0, \quad M = \sum_{l=1}^{K} m_l \end{cases} \tag{20}$$

where $E(\boldsymbol{\eta})$ can be substituted to (14), (16) and (17). By solving the optimal model (20), we can obtain an optimal plan $\boldsymbol{\eta}^*(n^*, M^*, \boldsymbol{S}^*, \boldsymbol{m}^*)$.

D. OPTIMIZATION PROCEDURE

According to the model (20), the settings for the design space $\boldsymbol{D}$ are defined as follows:

1)The subspace $\boldsymbol{A}_{nM}$ is determined by the constraint (19), which accounts for the sample size and the total number of measurements. We assume that there are R_n choices of sample size and R_M choices of total measurements, and then there are $R_{nM} = R_n * R_M$ choices of $\boldsymbol{A}_{nM}$, i.e., $\boldsymbol{A}_{nM} = \{(\boldsymbol{n}_M)\}$, $\boldsymbol{n}_M = ((n, M)_1, (n, M)_2, \cdots, (n, M)_{R_{nM}})$.

2)The subspace $\boldsymbol{A_m}$ describes the allocation of performance measurement times of each accelerated stress level, i.e., $\boldsymbol{A_m} = \{\boldsymbol{m}\}$, $\boldsymbol{m} = (m_1, m_2,$ ȩ$, m_K)$. We assume that there are $R_{\boldsymbol{m}}$ choices of the combination of measurement times on each accelerated stress level, i.e., $\boldsymbol{A_m} = \left\{\boldsymbol{m}_1, \boldsymbol{m}_2, \cdots, \boldsymbol{m}_{R_m}\right\}$.

3)The subspace $\boldsymbol{A_S}$ describes the specified value of each accelerated stress level, i.e., $\boldsymbol{A_S} = \{\boldsymbol{S}\}$, $\boldsymbol{S} = (S_1, S_2,$ ȩ$, S_K)$. Through equation (4), we make a simplification for the determination of the stress level, that is, the interval between $\xi(S_l)$ and $\xi(S_{l+1})$ is constant. We assume that there are $R_{\boldsymbol{S}}$ choices of the combination of accelerated stress levels, i.e., $\boldsymbol{A_S} = \left\{\boldsymbol{S}_1, \boldsymbol{S}_2, \cdots, \boldsymbol{S}_{R_S}\right\}$.

The subspaces mentioned above consist of $\boldsymbol{D} = \boldsymbol{A}_{nM} \times \boldsymbol{A_m} \times \boldsymbol{A_S}$, as shown in Fig. 1. Therefore, there is $R = R_n \times R_M \times R_{\boldsymbol{m}} \times R_{\boldsymbol{S}}$ choices of design in $\boldsymbol{D}$.

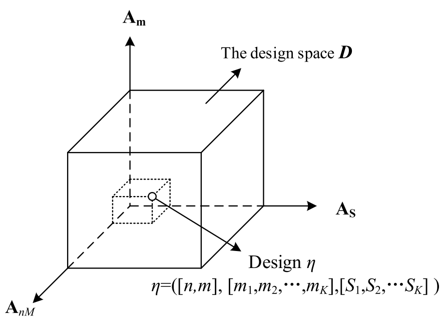


FIGURE 1. Design space D.

In order to alleviate the computational burden, limited choices of $\boldsymbol{\eta}$ in $\boldsymbol{D}$ are firstly considered and the corresponding $E(\boldsymbol{\eta})$ is calculated. Then, the surface fitting method is utilized to surf the R calculated values of $E(\boldsymbol{\eta}_r)$, $r = 1, 2, \cdots, R$, to find out the optimal plan $\boldsymbol{\eta}^*$ by maximizing $E(\boldsymbol{\eta})$. The procedure for solving the model (20) is shown as Algorithm 1:

IV. CASE STUDY

A. NUMERICAL CASE

The stress relaxation data of [23, Example 8.7] are employed in this case study. The data refers to accelerated degradation of electrical connectors. The stress relaxation is the stress loss of a component under constant stress loading. An electrical connector would fail due to excessive stress relaxation.

Algorithm 1 The Procedure for Solving the Model(20)

for each $\boldsymbol{\eta}_r \subseteq \boldsymbol{D}$ %$\boldsymbol{D}$ is determined according to the sections C and D
  Sample parameter $\boldsymbol{\theta}_r$ from their corresponding prior distributions
  for each $\boldsymbol{\eta}_r$
    Generate degradation data x_{rq} from the sampling distribution (2) for R_1 times ($q = 1, 2, \cdots, R_1$).
    Estimate the posterior distributions $p(\boldsymbol{\eta}_{rq}|x_{rq}, \boldsymbol{\eta}_r)$ by the MCMC algorithm
    $q = q + 1$;
  endfor
if the optimal objective is relative entropy (14)
  Calculate $E_x E_{\boldsymbol{\theta}} p(x|\boldsymbol{\theta}, \boldsymbol{\eta})$ based on $\boldsymbol{\theta}_r$ and x_{rq};
  Draw $\boldsymbol{\theta}_{rqs}$ ($s = 1, \cdots, R_2$) from their corresponding posterior distributions $p(\boldsymbol{\eta}_{rq}|x_{rq}, \boldsymbol{\eta}_r)$
  and calculate $Ex(log(p(x)))$ based on (15);
  Calculate the relative entropy $E(\boldsymbol{\eta}_r)$ based on (14);
else if the optimal objective is the quadratic loss function (16)
  Draw $\boldsymbol{\theta}_{rqs}$ ($s = 1, \cdots, R_2$) from the corresponding posterior distributions $p(\boldsymbol{\eta}_{rq}|x_{rq}, \boldsymbol{\eta}_r)$.
  Calculate $Var(t(p, \theta_{rq}|x_{rq}, \eta_r)) = \frac{1}{R_2}\sum_{s=1}^{R_2}\left[\frac{1}{R_2}\sum_{s=1}^{R_2} t(p|\theta_{rqs}) - t(p|\theta_{rqs})\right]^2$
  Calculate $Var(t(p, \theta_r|x_r, \eta_r)) = \frac{1}{R_1}\sum_{q=1}^{R_1} Var(t(p, \theta_{rq}|x_{rq}, \eta_r))$
else if the optimal objective is the Bayesian D-optimality (17)
  Calculate the elements of equation (18) based on the drawn $\boldsymbol{\theta}_r$ and x_{rq} according to appendix A;
  Calculate the value of the derivative of the information matrix;
endif
r=r+1;
endfor
Fit the surface based on the data pair $(\boldsymbol{\eta}_r, E(\boldsymbol{\eta}_r))$
The optimal plan for SSADT is obtained by taking the maximum of this surface.

TABLE 1. Estimated values of the model parameters.

Estimated parameters	$\hat{a}$	$\hat{b}$	$\hat{\lambda}$	$\hat{\beta}$
Mean	-1.8966	1.7379	0.6337	0.4493
Variance	0.1903	0.1738	0.1968	0.0178

In general, an electrical connector is defined as failed when the stress relaxation is over 30% i.e. Y_D = 30. The accelerated degradation data were obtained under the conditions of $S_1 = 60°C$, $S_2 = 85°C$ and $S_3 = 100°C$, respectively. These collected ADT data and the corresponding measurement times are shown in Table 11 and Table 12 in Appendix B.

In [17], it is proved that the collected stress relaxation data follow the IG process, in which the stress function $\xi(S)$ could be rewritten as $1/S$. By using the maximum likelihood estimation method and square root of the diagonal of the Fisher matrix, the mean and the variance of the model parameters can be obtained, as reported in Table 1.

Based on the parameter estimations in Table 1, the prior distribution can be determined. As the mean and variance have been estimated already, it is obvious that when the distribution form is selected, the hyper-parameters are determined. In this section, parameters a and b follow normal distributions and parameters λ and β follow Gamma distributions; then, the prior distribution of $\boldsymbol{\theta}$ can be determined as reported in Table 2.

Reference [20] has pointed out that the normal stress level S_0 is $40°C$; therefore, $S_{\min}$ and $S_{\max}$ are set to $50°C$ and $100°C$, respectively. Without loss of generality, let K be equal to 3 and $n_{\max}$ be equal to 5. The testing cost (C_1, C_2, C_0) is set to be $(2, 0.02, 30) \times 10^2$ dollars, and the measurement interval τ is 10h. With the aforementioned settings, the optimization model (20) could be rewritten as follows,

$$\begin{cases} \max & E(\boldsymbol{\eta}) \\ s.t. & n \cdot 200 + M \cdot \tau \cdot 20 \leq 3000 \\ & 3 \leq n \leq n_{\max} \\ & S_{\min} < S_1 < S_2 < \ldots < S_K \leq S_{\max} \\ & m_1 \geq m_2 \geq \ldots \geq m_K > 0, \quad M = \sum_{l=1}^{K} m_l \end{cases} \tag{21}$$

Since n has only three choices, i.e. 3, 4 and 5 in the constraints of model (20), the corresponding M could be determined based on (19), i.e. $nC_1 + M\tau C_2 \leq C_0$, and their possible values are reported in Table 3.

In order to further reduce the number of design choices in $\boldsymbol{D}$, the following simplification is used. First, let S_1 = [50 55 60 65 70 75 80] $°C$ and $S_3 = S_{\max} = 100°C$; then, by using the interval between $\xi(S_l)$ and $\xi(S_{l+1})$ constant, S_2 = [73 76 78 81

TABLE 2. Prior distributions of the model parameters.

Parameter	a	b	λ	β
Prior distribution	$Normal$ $(-1.90, 0.19)$	$Normal$ $(1.74, 0.17)$	$Gamma$ $(2.04, 0.31)$	$Gamma$ $(11.34, 0.04)$

TABLE 3. Evaluation of test variables *n* and *M*.

n	3	4	5
M	120	110	100

84 87 89] °C. Similarly, let m_1 = [40 50 60 70], m_3 = [40 30 20 10] and $m_2 = 1/2(m_1 + m_3)$; then, m_2 = [40 40 40 40]. In this way, only the optimization of S_1 and m_1 is needed, because $\boldsymbol{S}$ and $\boldsymbol{m}$ will be determined as long as S_1 and m_1 are chosen. Consequently, $R_{nM} = 3$, $R_{\boldsymbol{S}} = 7$ and $R_{\boldsymbol{m}} = 4$; then, there are $R = R_{nM} \times R_{\boldsymbol{S}} \times R_{\boldsymbol{m}} = 84$ choices of the design plan in $\boldsymbol{D}$. With the above simplification, the decision variables in the plan $\boldsymbol{\eta}$ include (n, M, S_1, m_1).

When the different objectives, as formulated by (14), (16) and (17) are selected, the corresponding optimization procedure proposed in section III-D can be applied. In this study, the locally weighted linear regression algorithm is used to smooth the data $(E(\boldsymbol{\eta}), S_1, m_1)$ for every combination of values n and M of Table 3. The optimal results are shown in Table 4; the surface fitting results for $n = 3$ and $M = 120$ are shown in Fig. 2.

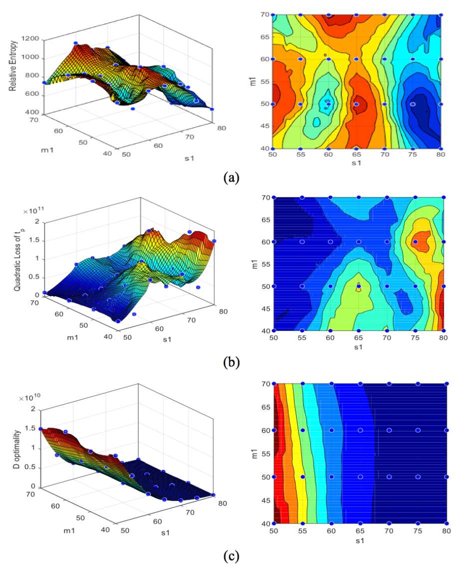


FIGURE 2. Optimal results for *n* = 3, *m* = 120 under the objectives of (a) relative entropy, (b) quadratic loss function and (c) D-optimality.

From Table 4, it is obvious that $E(\boldsymbol{\eta})$ increases with n increasing but decreases with M. If we define the data amount as $z = n \times M$; then, when $n = 3$, $M = 120$, $z = 360$; when $n = 4$, $M = 110$ $z = 440$; when $n = 5$, $M = 100$, $z = 500$: more data, more information and, therefore, the optimal plans obtained under $n = 5$, $M = 100$ would be the best ones with all objectives.

B. SENSITIVITY ANALYSIS

In practice, it can be different to select the prior distributions of the parameters. Hence, a sensitivity analysis is carried out to study the robustness to different distribution forms of the parameters. Without loss of generality, we assumed that the sensitivities of the optimal plans under $n = 3$ and $M = 120$ are of our interest.

Different prior distributions for a, b, λ and β are selected. The details to the selection of the prior distributions are explained as below.

- The first selection of prior distributions is shown in Table 2 and its corresponding optimal results will be regarded as the baseline of the following comparisons;
- The second selection takes Lognormal distributions for λ and β, since these two parameters should be positive; the other two prior distributions for a and b are the same;
- The third selection takes Logistic distributions for a and b, while the prior distributions of λ and β are the same as the first selection;
- The fourth selection considers non-informative priors for all four parameters; so, the uniform distribution is used and the corresponding upper and lower bounds are obtained by $\pm 3\sigma$ from the mean.
- For the first three selections, we keep the mean and the variance of the prior distribution of each parameter the same as in the first selection.

The four selections of prior distributions are given in Table 5.

The optimal plans and the corresponding values of objectives obtained with the different optimizations (shown in Table 6, Table 7, Table 8, respectively) are reported in Table 5.

Fig. 3 represents the test plan with different prior distributions under the same optimization objective. Plans 1, 2, 3 and 4 represent the plans designed with distributions selection I, II, III, IV, respectively.

It is obvious that different optimal objectives result in different plans, as different prior distributions also do. Some discussions can be given as follows:

(a) From the perspective of robustness of optimal design, the optimization with D-optimality is the most stable one, since all four optimal plans concentrate, whereas the results of the other two optimizations scatter around the whole design space (see Fig 3). Furthermore, from Table 6 to Table 8, the values of D-optimality slightly change with different prior

TABLE 4. The estimated values of the model parameters.

n,M	Optimization objective	$S_1(°C), S_2(°C), S_3(°C)$	m_1, m_2, m_3	$E(\boldsymbol{\eta})$
3,120	Relative entropy	60,78, 100	70, 40, 10	1.0825e+03
	Quadratic loss	50,73,100	70, 40, 10	1.1957e+10
	D-optimality	60,78, 100	60, 40, 20	1.7648e+10
4, 110	Relative entropy	55, 76, 100	70, 25, 15	1.2413e+03
	Quadratic loss	50,73,100	70, 25, 15	4.0202e+10
	D-optimality	50,73,100	70, 25, 15	4.0312e+10
5, 100	Relative entropy	55, 76, 100	60, 25, 15	1.7237e+03
	Quadratic loss	55, 76, 100	60, 25, 15	4.9688e+09
	D-optimality	50, 73, 100	50, 30, 20	8.2072e+10

TABLE 5. Different prior distrubutions of the model papameters.

Parameter	a	b	λ	β
Distribution I	$Normal$ $(-1.90, 0.19)$	$Normal$ $(1.74, 0.17)$	$Gamma$ $(2.04, 0.31)$	$Gamma$ $(11.34, 0.04)$
Distribution II	$Normal$ $(-1.90, 0.19)$	$Normal$ $(1.74, 0.17)$	$Lognormal$ $(-0.66, 0.40)$	$Lognormal$ $(-0.84, 0.08)$
Distribution III	$Logistic$ $(-1.90, 0.24)$	$Logistic$ $(1.74, 0.23)$	$Gamma$ $(2.04, 0.31)$	$Lognormal$ $(11.34, 0.04)$
Distribution IV	$Uniform$ $(-3.20, 0.58)$	$Uniform$ $(0.49, 2.99)$	$Uniform$ $(0.10, 1.96)$	$Uniform$ $(0.05, 0.85)$

TABLE 6. Optimal SSADT plans with relative entropy as objective.

Plan	Prior distribution	Test plan $(S_1(°C), S_2(°C), S_3(°C); m_1, m_2, m_3)$	$RE(\boldsymbol{\eta})$
1	Distribution I	(60,78, 100; 70,40,10)	1.0825e+03
2	Distribution II	(70, 84, 100; 70,40,10)	2.2092e+03
3	Distribution III	(60, 78, 100; 50,40,30)	7.4474e+03
4	Distribution IV	(75, 87, 100; 60,40,20)	1.7771e+05

TABLE 7. Optimal SSADT plans with quadryic loss as objective.

Plan	Prior distribution	Test plan $(S_1(°C), S_2(°C), S_3(°C); m_1, m_2, m_3)$	$Q(\boldsymbol{\eta})$
1	Distribution I	(50,73,100; 70,40,10)	1.1957e+10
2	Distribution II	(65, 81,100; 50,40,30)	5.9048e+09
3	Distribution III	(50,73,100; 70,40,10)	2.0270e+09
4	Distribution IV	(50,73,100; 50,40,30)	3.6442e+103

TABLE 8. Optimal SSADT plans with D-optimality as objective.

Plan	Prior distribution	Test plan $(S_1(°C), S_2(°C), S_3(°C); m_1, m_2, m_3)$	$\Phi(\boldsymbol{\eta})$
1	Distribution I	(50,73,100; 60,40,20)	1.7648e+10
2	Distribution II	(50,73,100; 70,40,10)	1.8692e+10
3	Distribution III	(55,76,100; 60,40,20)	5.9970e+10
4	Distribution IV	(50,73,100; 70,40,10)	4.8879e+10

distributions, whereas the other two change a lot. In other words, Bayesian D-optimality is not so sensitive to the prior distribution selection.

(b) From the perspective of engineering practice, the wider the range of accelerated stress levels, the more comprehensive information collected and the higher the

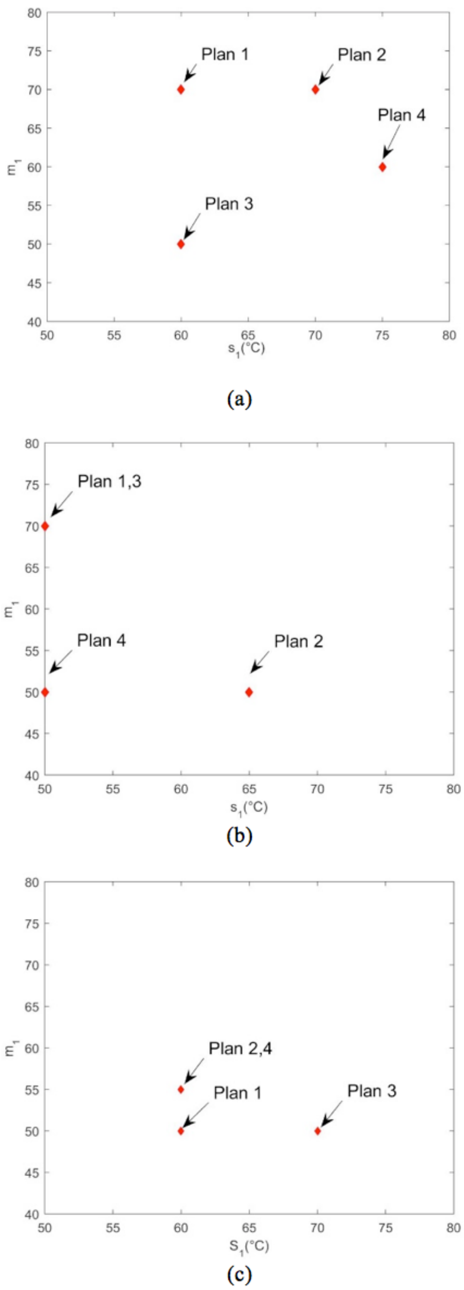


FIGURE 3. Optimal results of different prior distributions under the objectives of (a) relative entropy, (b) quadratic loss function and (c) D-optimality.

prediction precision; the closer the lowest accelerated stress level to normal stress level, the less extrapolation of the stress dimension and the more credible the prediction. Based on this analysis, the optimal plans obtained from relative entropy are not so credible, since the range of accelerated stress levels is narrow and the lowest level is far away from the normal level.

(c) From the perspective of the three optimal objectives, when the prior distribution is selected as the non-informative uniform distribution, the relative entropys value and the quadratic losss value of their corresponding optimal plans, respectively, are both quite large, since the former objective focuses on the information gain collected from the test, while the latter describes the variance of the estimated p-quantile lifetime.

Furthermore, to some extent, although the D-optimality describes the variance, it focuses on the precision of the estimated parameters; contrarily, the results of the quadratic loss are rather large, because the variance has been enlarged when the estimated parameters are used to predict the p-quantile lifetime. This is possibly the reason that the values of the D-optimality are quite stable no matter the prior distribution.

C. COMPARISON ANALYSIS OF OPTIMAL OBJECTIVES

From the sensitive analysis above, we can know which optimization is more robust. But when the optimal plans obtained by the proposed methods are used for some specific populations which have their own true values of the model parameters, how efficient the optimal plans are? To evaluate this, we propose an efficiency factor, which is the ratio between the value of the optimal objective under the corresponding optimal design plan, denoted as $E(\boldsymbol{\eta}^*)$. And the value of the optimal objective calculated under the corresponding optimal design plan based on the true values of the model parameters, denoted as $E_T(\boldsymbol{\eta}^*)$. The efficiency factor $\psi_{E(\boldsymbol{\eta}^*)}$ is expressed as,

$$\psi_{E(\boldsymbol{\eta}^*)} = \begin{cases} E_T\left(\boldsymbol{\eta}^*\right)/E\left(\boldsymbol{\eta}^*\right), & \text{for RE and D-optimality} \\ E\left(\boldsymbol{\eta}^*\right)/E_T\left(\boldsymbol{\eta}^*\right), & \text{for quadratic loss} \end{cases} \tag{22}$$

TABLE 9. True values of model parameters.

Parameters	A_T	B_T	λ_T	B_T
$\boldsymbol{\theta}_{T1}$	-1.5	1.5	0.5	0.5
$\boldsymbol{\theta}_{T2}$	-1.8966	1.7379	0.6337	0.4493
$\boldsymbol{\theta}_{T3}$	-1.46	2.15	1.08	0.58
$\boldsymbol{\theta}_{T4}$	-2.33	1.32	0.19	0.32

From (22), the more efficient the plans, the higher $\psi_{E(\boldsymbol{\eta}^*)}$. Different true values are selected for the parameters in model (6) (see Table 9). In Table 9, the subscript "T"means "true value", $\boldsymbol{\theta}_{T1}$ is randomly selected, $\boldsymbol{\theta}_{T2}$ is the mean shown in Table 1, while $\boldsymbol{\theta}_{T3}$ and $\boldsymbol{\theta}_{T4}$ are the results obtained by using $\boldsymbol{\theta}_{T2}$ plus and minus the corresponding variance from Table 1, respectively.

When the plans under the condition of $n = 3$ and $m = 12$ shown in Table 4 are chosen as $\boldsymbol{\eta}^*$, we use the Algorithm 1 mentioned in section III-D to calculate $E(\boldsymbol{\eta}^*)$, where $\boldsymbol{\eta}_r$ is selected as $\boldsymbol{\eta}^*$ and the parameters in Table 9 is used as $\boldsymbol{\theta}_r$. By following Algorithm 1, the results of $E_T(\boldsymbol{\eta}^*)$ can be obtained as reported in Table 10.

From Table 10, generally speaking, the efficiency of relative entropy is the highest, then is D-optimal. As for the efficiency of quadratic loss, it is too unstable to draw any conclusion.

TABLE 10. Efficienty factor of different optimization objectives.

Objective	$S_1(°C)$, $S_2(°C)$, $S_3(°C)$	m_1, m_2, m_3	$E(\boldsymbol{\eta}^*)$	$\boldsymbol{\theta}$	$E_a(\boldsymbol{\eta}^*)$	$\psi_{E(\boldsymbol{\eta}^*)}$
Relative entropy	60,78, 100	70, 40, 10	1082.5	$\boldsymbol{\theta}_{T1}$	794.61	73.40%
				$\boldsymbol{\theta}_{T2}$	979.54	90.49%
				$\boldsymbol{\theta}_{T3}$	802.64	74.14%
				$\boldsymbol{\theta}_{T4}$	2175.7	200.99%
Quadratic loss	50,73, 100	70, 40, 10	1.1957e+10	$\boldsymbol{\theta}_{T1}$	2.0303e+8	5889.28%
				$\boldsymbol{\theta}_{T2}$	1.1111e+10	107.61%
				$\boldsymbol{\theta}_{T3}$	9.5795e+5	1248186.23%
				$\boldsymbol{\theta}_{T4}$	1.5022e+18	0.00%
Relative entropy	60,78, 100	70, 40, 10	1.7648e+10	$\boldsymbol{\theta}_{T1}$	1.0111e+10	57.29%
				$\boldsymbol{\theta}_{T2}$	1.6398e+10	92.91%
				$\boldsymbol{\theta}_{T3}$	1.0189e+10	57.73%
				$\boldsymbol{\theta}_{T4}$	2.4020e+10	136.11%

TABLE 11. Stress relaxation degradation data of electrical connectors under different accelerated stress levels.

T	ID	Stress loss										
$60°C$	1	2.12	2.7	3.52	4.25	5.55	6.12	6.75	7.22	7.68	8.46	9.46
	2	2.29	3.24	4.16	4.86	5.74	6.85	*	7.4	8.14	9.25	10.55
	3	2.4	3.61	4.35	5.09	5.5	7.03	8.24	8.81	9.629	10.27	11.11
	4	2.31	3.48	5.51	6.2	7.31	7.96	8.57	9.07	10.46	11.48	12.31
	5	3.14	4.33	5.92	7.22	8.14	9.07	9.44	10.09	11.2	12.77	13.51
	6	3.59	5.55	5.92	7.68	8.61	10.37	11.11	12.22	13.51	14.16	15
$85°C$	7	2.77	4.62	5.83	6.66	8.05	10.61	11.2	11.98	13.33	15.64	-
	8	3.88	4.37	6.29	7.77	9.16	9.9	10.37	12.77	14.72	16.8	-
	9	3.18	4.53	6.94	8.14	8.79	10.09	11.11	14.72	16.47	18.66	-
	10	3.61	4.37	6.29	7.87	9.35	11.48	12.4	13.7	15.37	18.51	-
	11	3.42	4.25	7.31	8.61	10.18	12.03	13.7	15.27	17.22	19.25	-
	12	5.27	5.92	8.05	9.81	12.4	13.24	15.83	17.59	20.09	23.51	-
$100°C$	13	4.25	5.18	8.33	9.53	11.48	13.14	15.55	16.94	18.05	19.44	-
	14	4.81	6.16	7.68	9.25	10.37	12.4	15	16.2	18.24	20.09	-
	15	5.09	7.03	8.33	10.37	12.22	14.35	16.11	18.7	19.72	21.66	-
	16	4.81	7.5	9.16	10.55	13.51	15.55	16.57	19.07	20.27	22.4	-
	17	5.64	6.57	8.61	12.5	14.44	16.57	18.7	21.2	22.59	24.07	-
	18	4.72	8.14	10.18	12.4	15.09	17.22	19.16	21.57	24.35	26.2	-

V. CONCLUSION AND FUTURE WORK

A Bayesian SSADT optimal design method based on IG Process is proposed. The objectives of the proposed optimal design methodology are relative entropy, quadratic loss function and D-optimality, and the testing cost is regarded as the constraint. MCMC method is adopted in the simulation and the surface fitting technique is utilized for the optimal design.

By sensitivity analysis of the prior distribution, D-optimality has been identified as the most robust design. From the perspective of engineering practice, its associated optimal plans are more reasonable than the ones from the relative entropy and the quadratic loss. An efficiency factor is proposed to quantify how close the designed plan is to the right optimal plan. The efficiency factor is the ratio between the value of the optimal objective under the corresponding optimal design plan and that calculated under the corresponding optimal design plan based on the true values of the model parameters. In the numerical case study, the optimization with relative entropy as objective is the most efficient, whereas quadratic loss is the worst one.

Combining sensitive analysis and efficiency factor, we can conclude that D-optimality is a good choice as the optimal objective in Bayesian design, even though its efficiency is slightly lower than relative entropy.

Bayesian decision theory has been proved as an effective method to help us to obtain the optimal plans for ADT, hence, it is high attractive and potential to apply the other Bayesian method for ADT. For example, naive Bayes classification [24]–[26] can be utilized to carry out fault diagnosis based on ADT data.

TABLE 12. Inspection time under different stress levels.

T	Performance inspection time										
65°C	108	241	534	839	1074	1350	1637	1890	2178	2513	2810
85°C	46	108	212	408	632	764	1011	1333	1517	2586	-
100°C	46	108	212	344	446	626	729	972	1005	1218	-

APPENDIX A

In order to simplify the expression of the equations in (18), we assume that $\Lambda_{ilj} = t_{il(j+1)} - t_{ilj}) = (m_{il(j+1)}\tau) - (m_{ilj}\tau)$, and then,

$$\frac{\partial \Lambda_{ilj}}{\partial \beta} = \ln\left(m_{il(j+1)}\tau\right)\left(m_{il(j+1)})\tau\right)^{\beta} - \ln\left(m_{ilj}\tau\right)\left(m_{ilj}\tau\right)^{\beta}$$

$$\frac{\partial^2 \Lambda_{ilj}}{\partial^2 \beta} = \ln^2\left(m_{il(j+1)}\tau\right)\left(m_{il(j+1)})\tau\right)^{\beta} - \ln^2\left(m_{ilj}\tau\right)\left(m_{ilj}\tau\right)^{\beta}$$

The elements of $I(\boldsymbol{\eta}, \boldsymbol{\theta})$ in (18) are derived as follows,

$$E\left(-\frac{\partial^2 L}{\partial a^2}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{2\lambda x_{ilj}}{\exp^2\left[a+b\xi\left(S_l\right)\right]} - \frac{\lambda \Lambda_{ilj}}{\exp\left[a+b\xi\left(S_l\right)\right]}\right)$$

$$E\left(-\frac{\partial^2 L}{\partial a \partial b}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{2\lambda x_{ilj}}{\exp^2\left[a+b\xi\left(S_l\right)\right]} - \frac{\lambda \Lambda_{ilj}}{\exp\left[a+b\xi\left(S_l\right)\right]}\right) \bullet \xi\left(S_l\right)$$

$$E\left(-\frac{\partial^2 L}{\partial a \partial \lambda}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{\Lambda_{ilj}}{\exp\left[a+b\xi\left(S_l\right)\right]} - \frac{x_{ilj}}{\exp^2\left[a+b\xi\left(S_l\right)\right]}\right)$$

$$E\left(-\frac{\partial^2 L}{\partial a \partial \beta}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{\lambda \frac{\partial \Lambda_{ilj}}{\partial \beta}}{\exp\left[a+b\xi\left(S_l\right)\right]}\right)$$

$$E\left(-\frac{\partial^2 L}{\partial b^2}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{2\lambda x_{ilj}}{\exp^2\left[a+b\xi\left(S_l\right)\right]} - \frac{\lambda \Lambda_{ilj}}{\exp\left[a+b\xi\left(S_l\right)\right]}\right) \bullet \xi^2\left(S_l\right)$$

$$E\left(-\frac{\partial^2 L}{\partial b \partial \lambda}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{\Lambda_{ilj}}{\exp\left[a+b\xi\left(S_l\right)\right]} - \frac{x_{ilj}}{\exp^2\left[a+b\xi\left(S_l\right)\right]}\right) \bullet \xi\left(S_l\right)$$

$$E\left(-\frac{\partial^2 L}{\partial b \partial \beta}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{\lambda \frac{\partial \Lambda_{ilj}}{\partial \beta}}{\exp\left[a+b\xi\left(S_l\right)\right]}\right) \bullet \xi\left(S_l\right)$$

$$E\left(-\frac{\partial^2 L}{\partial \lambda^2}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\left(\frac{1}{2\lambda^2}\right)$$

$$E\left(-\frac{\partial^2 L}{\partial \lambda \partial \beta}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\frac{\Lambda_{ilj}\left(\frac{\partial \Lambda_{ilj}}{\partial \beta}\right)}{x_{ilj}} - \frac{\frac{\partial \Lambda_{ilj}}{\partial \beta}}{\exp\left[a+b\xi\left(S_l\right)\right]}$$

$$E\left(-\frac{\partial^2 L}{\partial \beta^2}\right) = \sum_{l=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m_l-1}\frac{\left(\frac{\partial \Lambda_{ilj}}{\partial \beta}\right)^2}{\Lambda_{ilj}{}^2} - \left(\frac{1}{\Lambda_{ilj}} + \frac{\lambda}{\exp\left[a+b\xi\left(S_l\right)\right]}\right) \times \frac{\partial^2 \Lambda_{ilj}}{\partial^2 \beta} + \frac{\lambda}{x_{ilj}}\left[\left(\frac{\partial \Lambda_{ilj}}{\partial \beta}\right)^2 + \Lambda_{ilj}\left(\frac{\partial^2 \Lambda_{ilj}}{\partial^2 \beta}\right)\right]$$

APPENDIX B

The stress relaxation data and the measurement times are tabulated in Table 11 and Table 12.

REFERENCES

[1] P. Lall, P. Sakalaukus, and L. Davis, "Reliability and failure modes of solid-state lighting electrical drivers subjected to accelerated aging," *IEEE Access*, vol. 3, pp. 531–542, 2015.

[2] Y. Zhang and W. Q. Meeker, "Bayesian methods for planning accelerated life tests," *Technometrics*, vol. 48, no. 1, pp. 49–60, 2006.

[3] A. Xu and Y. Tang, "A Bayesian method for planning accelerated life testing," *IEEE Trans. Rel.*, vol. 64, no. 4, pp. 1383–1392, Dec. 2015.

[4] X. Wang, "Wiener processes with random effects for degradation data," *J. Multivariate Anal*, vol. 101, no. 2, pp. 340–351, 2010.

[5] Z. S. Ye, L. C. Tang, and H. Y. Xu, "A distribution-based systems reliability model under extreme shocks and natural degradation," *IEEE Trans. Rel.*, vol. 60, no. 1, pp. 246–256, Mar. 2011.

[6] Y. Shi and W. Q. Meeker, "Bayesian methods for accelerated destructive degradation test planning," *IEEE Trans. Rel.*, vol. 61, no. 1, pp. 245–253, Mar. 2012.

[7] X. Li, M. Rezvanizaniani, Z. Ge, M. Abuali, and J. Lee, "Bayesian optimal design of step stress accelerated degradation testing," *J. Syst. Eng. Electron.*, vol. 26, no. 3, pp. 502–513, 2015.

[8] X. Liu and L. C. Tang, "A Bayesian optimal design for accelerated degradation tests," *Quality Reliab. Eng. Int.*, vol. 26, no. 8, pp. 863–875, 2010.

[9] M. Hamada, H. Martz, C. Reese, and A. Wilson, "Finding near-optimal Bayesian experimental designs via genetic algorithms," *Amer. Statist.*, vol. 55, no. 3, pp. 175–181, 2001.

[10] W. Peng, Y. Liu, Y.-F. Li, S.-P. Zhu, and H.-Z. Huang, "A Bayesian optimal design for degradation tests based on the inverse Gaussian process," *J. Mech. Sci. Technol.*, vol. 28, no. 10, pp. 3937–3946, 2014.

[11] S. Roy and C. Mukhopadhyay, "Bayesian D-optimal accelerated life test plans for series systems with competing exponential causes of failure," *J. Appl. Statist.*, vol. 43, no. 8, pp. 1477–1493, 2016.

[12] L. A. Escobar and W. Q. Meeker, "Planning accelerated life tests with two or more experimental factors," *Technometrics*, vol. 37, no. 4, pp. 411–427, 1995.

[13] W. B. Nelson, "An updated bibliography of accelerated test plans," in *Proc. 61th Rel. Maintainability Symp.*, Jan. 2015, pp. 1–6.

[14] X. Wang and D. Xu, "An inverse Gaussian process model for degradation data," *Technometrics*, vol. 52, no. 2, pp. 188–197, 2010.

[15] A. Escobar and Q. Meeker, *Statistical Methods for Reliability Data*. Hoboken, NJ, USA: Wiley, 1998.

[16] Z.-S. Ye and N. Chen, "The inverse Gaussian process as a degradation model," *Technometrics*, vol. 56, no. 3, pp. 302–311, 2014.

[17] Z.-S. Ye, L.-P. Chen, L. C. Tang, and M. Xie, "Accelerated degradation test planning using the inverse Gaussian process," *IEEE Trans. Rel.*, vol. 63, no. 3, pp. 750–763, Sep. 2014.

[18] H. Wang, G.-J. Wang, and F.-J. Duan, "Planning of step-stress accelerated degradation test based on the inverse Gaussian process," *Rel. Eng. Syst. Safety*, vol. 154, pp. 97–105, Oct. 2016.

[19] D. V. Lindley and D. Lindley, *Bayesian Statistics: A Review*. Philadelphia, PA, USA: SIAM, 1972.

[20] I. Ntzoufras, "Bayesian model and variable evaluation," in *Proc. Bayesian Modeling Using WinBUGS*, 2009, pp. 389–433.

[21] M. A. Newton and A. E. Raftery, "Approximate Bayesian inference with the weighted likelihood bootstrap," *J. Roy. Statist. Soc. Ser. B Methods*, vol. 56, no. 1, pp. 3–48, 1994.

[22] K. Chaloner and K. Larntz, "Optimal Bayesian design applied to logistic regression experiments," *J. Statist. Planning Inference*, vol. 21, no. 2, pp. 191–208, 1989.

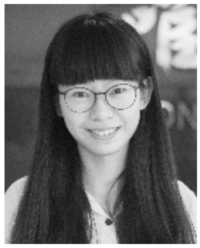
[23] G. Yang, *Life Cycle Reliability Engineering*. Hoboken, NJ, USA: Wiley, 2007.

[24] J. Wu, S. Pan, X. Zhu, P. Zhang, and C. Zhang, "SODE: Self-Adaptive One-Dependence Estimators for classification," *Pattern Recognit.*, vol. 51, pp. 358–377, Mar. 2016.

[25] N. A. Zaidi, J. Cerquides, M. J. Carman, and G. I. Webb, "Alleviating naive Bayes attribute independence assumption by attribute weighting," *J. Mach. Learn Res.*, vol. 14, no. 1, pp. 1947–1988, 2013.

[26] J. Wu, S. Pan, X. Zhu, Z. Cai, P. Zhang, and C. Zhang, "Self-adaptive attribute weighting for naive Bayes classification," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1487–1502, 2015.

XIAOYANG LI received the Ph.D. degree in aerospace systems engineering from Beihang University, Beijing, China, in 2007. She is currently an Associate Professor with the School of Reliability and Systems Engineering, Beihang University. Her research interests include life modeling and optimal test design for accelerated testing, and quantification method of uncertainty.



YUQING HU received the B.S. degree in quality and reliability from Beihang University (BUAA) in 2015. She is currently pursuing the master's degree with the Science and Technology on Reliability and Environmental Engineering Laboratory, BUAA. Her research interests focus on life prediction, accelerated test modeling, and testing design.

ENRICO ZIO (M'06) received the B.Sc. degree in nuclear engineering from the Politecnico di Milano, Milano, Italy, in 1991, the M.Sc. degree in mechanical engineering from the University of California, Los Angeles, CA, USA, in 1995, and the Ph.D. degree in nuclear engineering from the Politecnico di Milano, in 1995. He is currently the Director of the Chair in Complex Systems, and the Energetic Challenge of Ecole Centrale Paris and Supelec, a Full Professor and a Rector's delegate for the Alumni Association, a past Director of the Graduate School with the Politecnico di Milano, and an Adjunct Professor with the University of Stavanger. He has authored or co-authored over five international books and 170 papers in international journals. His research interests include analysis of the reliability, safety, and security of complex systems under stationary and dynamic conditions, particularly by Monte Carlo simulation methods, and development of soft computing techniques for safety, reliability and maintenance applications, system monitoring, and fault diagnosis and prognosis. He is also the Chairman of the European Safety and Reliability Association, a member of the Korean Nuclear society and China Prognostics and Health Management Society, and a past Chairman of the Italian Chapter of the IEEE Reliability Society.

RUI KANG received the master's degree in reliability engineering from Beihang University, Beijing, China, in 1990. He is currently an Eminent Professor with the School of Reliability and Systems Engineering, Beihang University. His research interests include reliability testing and environment engineering of product.