

Optimal Setting of Test Conditions and Allocation of Test Units for Accelerated Degradation Tests with Two Stress Variables

Guanqi Fang, Rong Pan, Senior Member, IEEE, and John Stufken

Abstract—Conducting accelerated degradation tests is an effective way to assess reliability of highly reliable products. In the existing literature, most works deal with planning ADT with a single stress variable; however, the situation of more than one stress variable is commonly seen in engineering practice. To fill the gap, in this paper, we provide an analytical approach to address the design issue when two stress variables are in presence. By using a linear mixed-effects model to describe the accelerated degradation process, we demonstrate that the design problem can be solved by, first, finding the optimal setting of test conditions and allocation of test units for a “single-variable” case, and then the initial solution is transformed to the test plan for the case of two stress variables. The transformation is done by maintaining the same value of the asymptotic variance of the estimated p -th quantile lifetime, along with the consideration of reducing the asymptotic variance of model parameters estimation. We also discuss how to find compromise plans that satisfy practical demands. Finally, the proposed framework is illustrated using a real-world example.

Index Terms—Accelerated Degradation Tests (ADTs), C-optimality, D-optimality, Fisher information, linear mixed-effects model, optimal design, test planning

I. INTRODUCTION

A. Background

Life testing is a common engineering tool for assessing reliability of industrial products. However, this method, which needs sample time-to-failure data, is not efficient and, oftentimes, too costly for highly reliable items. Instead, by collecting degradation measurements of an item’s performance characteristic (PC), accelerated degradation tests (ADTs) can be utilized to provide richer information than the traditional life tests (Weaver et al., 2013). Under an ADT, a pre-specified PC of an item, such as the wear resistance of a particular metal alloy (Meeker and Escobar, 1998), is measured repeatedly at some fixed intervals until test termination. This type of experiment is usually conducted on a limited number of test units. Furthermore, to acquire the degradation data more efficiently, engineers usually expose these units to a harsh environmental condition, such as high temperature, humidity, and use rate, etc. Eventually, by fitting a statistical model to

the data, product lifetime under the normal use condition can be predicted. In recent years, a lot of works have been done in the area of ADT data analysis, such as (Fang et al., 2020), (Si et al., 2018), (Pan and Crispin, 2011), (Fang et al., 2018), and (Pan et al., 2016), etc.

To design an ADT, test planners need to make decisions on the total number of test units, measurement time schedules, test stress levels, as well as the proportion of test units to be allocated to each level (Boulanger and Escobar, 1994). These decisions have an effect on the precision of lifetime prediction and they may be constrained by test budget. In this paper, we will investigate the method of designing an ADT plan in terms of statistical efficiency so as to find the optimal setting of test conditions and allocation of test units based on a broad class of linear mixed-effects models. An industry example is used to illustrate our proposed methods.

Three major contributions have been made in this paper. First, we provide a methodology for finding the optimal setting of test conditions and allocation of test units for ADTs with two stress variables. The plans generated by our method are more statistically efficient for estimating product lifetime under the use condition. Secondly, we incorporate practical constraints into the planning process and investigate the robustness of our proposed test plans. Thirdly, we demonstrate a rigorous proof of the optimal allocation of test units as being conjectured by Schwabe et al. (2014). This concludes the theoretical exposition of the numerical result appeared in the rejoinder by Weaver and Meeker (2014b).

The rest of the paper is organized as follows: Section I-B introduces a motivating example of ADT provided by an international standard. Then, a literature review about existing works related to designing an ADT is made in Section I-C. Section II introduces a modeling structure for ADT data and provides the derivation of lifetime distribution. Section III gives planning criteria and planning scope of our study, followed by an optimization scheme we utilize. Section IV provides relevant propositions and theorems we derive to help design an ADT with two stress variables. Section V further develops an in-depth study on the optimal allocation. The results by applying the methodology to the motivating example are given in Section VI. Section VII provides a Monte Carlo simulation study and sensitivity analysis. Lastly, we make conclusions and discussion in Section VIII. All detailed derivations and proofs are provided in Appendix.

Guanqi Fang (e-mail: gfang5@asu.edu) is with School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou, Zhejiang 310018, China.

Rong Pan (corresponding author, e-mail: rong.pan@asu.edu) are with School of Computing, Informatics and Decision Systems Engineering, Arizona State University, 699 S Mill Ave, Tempe, AZ 85287, USA.

John Stufken (e-mail: jstufken@uncg.edu) is with University of North Carolina at Greensboro, Greensboro, NC 27412, USA.

B. A Motivating Example

Nowadays, people use many highly reliable electronic components, such as hard disk drives (HDDs) and digital video discs (DVDs). These optical media are continuously subject to degradation processes during the course of “read” and “written”, while ISO 10995 (2011) is such an international standard providing guidelines for assessing their reliability and predicting archival lifetime. In the standard’s documentation, an ADT dataset for DVDs is described, where the error rate (i.e., the max summed over 8 consecutive error correction blocks, or Max PI Sum8 (CD Associates, 1998)) is monitored and treated as the PC. Two stress variables – temperature and relative humidity (RH) – are elevated to four stress levels, while different measurement schedules are created for each level. In each test cell, either 20 or 30 units are assigned. The failure is defined as the error rate exceeding 280, and the 5th quantile lifetime under use condition (25°C and 50%RH) is to be predicted. Figure 1 presents the degradation paths of the PC of these test units measured in log scale. Table I demonstrates the original design information provided by the standard. In this paper, we will discuss how to find optimum test plans using this example.

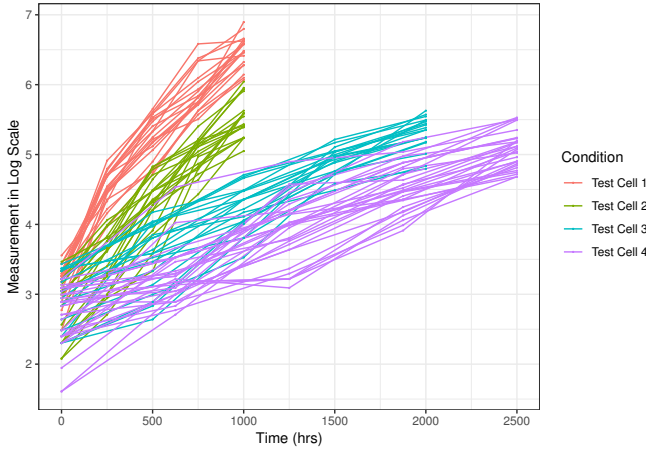


Fig. 1: Degradation Paths of Optical Media.

TABLE I: Original Design Information.

Test Cell	Temperature (°C)	RH (%)	Number of Units	Measuring Time Interval (hrs)	Termination Time (hrs)
1	85	85	20	250	1000
2	85	70	20	250	1000
3	65	85	20	500	2000
4	70	75	30	625	2500

C. Related Work

Numerous studies have investigated the experimental design problems associated with ADTs. As early as 1994, Boulanger and Escobar (1994) provided a complete methodology for designing ADTs in terms of the four decisions mentioned in the background section. In addition, Weaver et al. (2013) gave a systematic study on degradation test planning based on linear mixed-effects models. Following by Weaver and Meeker (2014a), the study of ADT planning was proposed to minimize the asymptotic variance of the estimated quantile lifetime. Furthermore, Kim and Bae (2013) provided a planning procedure

to determine the total sample size and the inspection schedule considering the cost constraint. Ye et al. (2019) discussed a test unit allocation strategy for planning ADTs. All of these studies are developed on the basis of general path models, which are essentially under the framework of (non)linear mixed-effects models. But there are also many papers discussing ADT planning that are based on stochastic process models, such as (Tsai et al., 2012), (Wang et al., 2017), (Zhao et al., 2018a), and (Zhao et al., 2018b), etc.

However, most existing literature only consider a single stress variable, which in fact is not commonly used in practice. For instance, the motivating example provided by ISO 10995 utilizes two variables – temperature and RH. In this paper, we will develop the method of finding an ADT plan to optimize the setting of test conditions and the allocation of test units when two stress variables are present.

II. DEGRADATION MODEL AND LIFETIME DISTRIBUTION

A. ADT Data Modeling

In an ADT, the monitored PC for each test unit is measured periodically at pre-specified intervals under a certain level of environmental condition until termination. Thus, after the test, the data in terms of observations, measurement times and test stress levels are obtained. To explain such rich information contained in the dataset, we need to build a statistical model that is able to 1) establish a relationship between the measurements and actual degradation levels in the presence of measurement errors; 2) incorporate the effect brought by environmental stresses on product degradation; and 3) provide flexibility to account for the existing unit-to-unit variability. Assume there are n test units and m_i measurements for each unit, let $y_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, m_i$ denote the measurement on unit i at time point j . A monotone increasing function, $\tau_{ij} = t_{ij}^\gamma$, is used to transform the time scale with a pre-known parameter γ so as to linearize the degradation paths. Then, a general regression model for the degradation process can be represented by the following function:

$$y_{ij} = D(\tau_{ij}; \boldsymbol{\beta}, \mathbf{b}_i) + \varepsilon_{ij},$$

where $D(\tau_{ij}; \boldsymbol{\beta}, \mathbf{b}_i)$ is the actual degradation level for unit i at (possibly) transformed time τ_{ij} . The measurement errors ε_{ij} ’s are identically and independently distributed (i.i.d.) normal variables with mean 0 and variance σ^2 . $\boldsymbol{\beta} = (\beta_0, \beta_1, \eta_1, \eta_2)'$ denotes the vector of fixed-effects parameters, where β_0 and β_1 are the initial degradation level and degradation rate, respectively. η_1 and η_2 are parameters to capture the acceleration effect brought by stress variables. $\mathbf{b}_i$ represents the vector of the random effects of i^{th} unit. Here, $\{\mathbf{b}_i\}$ and $\{\varepsilon_{ij}\}$ are assumed to be mutually independent, which means that the intrinsic unit-level variability is independent of measurement errors produced by external instruments. Furthermore, we assume that $D(\tau_{ij}; \boldsymbol{\beta}, \mathbf{b}_i)$ is decomposed as below:

$$D(\tau_{ij}; \boldsymbol{\beta}, \mathbf{b}_i) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})\tau_{ij} + (\eta_1 x_{1i} + \eta_2 x_{2i}) \tau_{ij}, \quad (1)$$

where x_{1i} and x_{2i} are the standardized values of stress variables. For instance, for the motivating example, $x_{1i} =$

$\frac{1/T_U - 1/T_i}{1/T_U - 1/T_H}$ and $x_{2i} = \frac{\ln RH_i - \ln RH_U}{\ln RH_H - \ln RH_U}$, where T_i and RH_i are temperature in degrees Kelvin and RH in percentage placed on unit i , respectively. T_H and RH_H are the highest controllable temperature and RH for test chambers. T_U and RH_U are temperature and RH under the normal use condition. Here, the Eyring relation is utilized to characterize the effect of stress variables; but there are many other options depending on the type of stress variables (Zhao et al., 2018b). The standardization makes the design space of stress variables become a unit square in the first quadrant (Seo and Pan, 2017). In addition, we assume that $\mathbf{b}_i = (b_{0i}, b_{1i})'$ is a bivariate random variable, which describes the varying initial degradation value before test and the random degradation rate during test, respectively. Specifically, $\mathbf{b}_i$ is subject to a bivariate normal distribution shown as

$$\mathbf{b}_i = (b_{0i}, b_{1i})' \sim BVN(\mathbf{0}, \mathbf{V}),$$

where the mean is a bivariate vector of zeros and

$$\mathbf{V} = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}$$

is the variance-covariance matrix. And we assume that the elements in $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ are mutually independent. This implies that the test units are independent of each other.

Thus, if we denote $\beta_{0i} = \beta_0 + b_{0i}$ and $\beta_{1i} = \beta_1 + b_{1i}$, then the full model becomes

$$y_{ij} = \beta_{0i} + \beta_{1i}\tau_{ij} + (\eta_1 x_{1i} + \eta_2 x_{2i})\tau_{ij} + \varepsilon_{ij}, \quad (2)$$

and let $\boldsymbol{\theta} = (\beta_0, \beta_1, \eta_1, \eta_2, \sigma_0, \sigma_1, \rho, \sigma)^'$ be the vector of parameters in this model.

B. Lifetime Distribution

For a degradation process, a failure also called “soft” failure occurs when the degradation level passes a pre-specified failure threshold, D_f . Without loss of generality, we assume the degradation path is trending up. Thus, if letting T be the lifetime of a single unit, its probability distribution satisfies the following:

$$\begin{aligned} F_T(t) &= P(T \leq t) \\ &= P((\beta_0 + b_0) + (\beta_1 + b_1)\tau + (\eta_1 x_1 + \eta_2 x_2)\tau \geq D_f) \\ &= P(b_0 + b_1\tau \geq D_f - \beta_0 - \beta_1\tau - (\eta_1 x_1 + \eta_2 x_2)\tau) \\ &= 1 - \Phi(\kappa), \end{aligned} \quad (3)$$

where

$$\kappa = \frac{D_f - \beta_0 - \beta_1\tau - (\eta_1 x_1 + \eta_2 x_2)\tau}{\sqrt{\sigma_0^2 + \tau^2\sigma_1^2 + 2\tau\rho\sigma_0\sigma_1}},$$

and Φ is the cumulative distribution function (cdf) of the standard normal distribution.

Note that for simplicity, we simplify the notation by removing i on b_{0i} , b_{1i} , x_{1i} , and x_{2i} . Also, we remain the notation τ instead of t , where $\tau = t'$. The resulted normal cdf comes from the conclusion $b_0 + b_1\tau \sim N(0, \sigma_0^2 + \tau^2\sigma_1^2 + 2\tau\rho\sigma_0\sigma_1)$.

III. ADT PLANNING AND OPTIMIZATION SCHEME

A. Planning Criteria

When engineers create an ADT plan, it is expected that this plan should be able to provide sufficiently precise estimates of model parameters as well as good evaluation of a specific product population characteristic when applying extrapolation to the use condition (Boulanger and Escobar, 1994). Therefore, we adopt two common planning criteria as shown in the following:

- **D-optimality:** A good design would result in relatively small asymptotic variance of maximum likelihood estimates (MLEs) of model parameters, for which D-optimality is one of such design criteria. Under this criterion, the overall uncertainty in estimation is minimized; and it is equivalent to maximizing the determinant of Fisher information matrix. Thus, our objective can be formulated as

$$\min -\det[\mathbf{I}(\boldsymbol{\theta})],$$

where $\mathbf{I}(\boldsymbol{\theta})$ is the total Fisher information matrix, of which the derivation is given in Appendix A.

- **C-optimality:** Since the ultimate goal of an ADT is to predict the p^{th} quantile lifetime $\hat{t}_p$ under the use condition, minimizing the variance of its estimate would be a natural choice. In general, this criterion is called C-optimality, which seeks to minimize the asymptotic variance of a predetermined linear combination of model parameters (Dean et al., 2015). The asymptotic variance of $\hat{t}_p$ can be approximated by the delta method, which gives the following objective:

$$\min A\text{Var}(\hat{t}_p) = \mathbf{c}'[\mathbf{I}(\boldsymbol{\theta})^{-1}]\mathbf{c},$$

where $\mathbf{c} = \left(\frac{\partial t_p}{\partial \beta_0}, \dots, \frac{\partial t_p}{\partial \sigma}\right)'$. The derivation of t_p and $\mathbf{c}$ is provided in Appendix B and C, respectively.

B. Planning Scope

In order to satisfy the criteria above, care must be taken when designing an ADT. As illustrated by Boulanger and Escobar (1994), practitioners need to make four different decisions: 1) choose stress levels for each environmental variable, 2) decide the proportion of test units to be assigned to each stress level, 3) select measurement schedule, as well as 4) determine the total number of test units. In practice, the number of test units is closely related to the total budget of the experiment. Inspection on test units can also be expensive. As a result, the measurement schedule is often pre-specified by practitioners according to cost constraints. In this paper, we mainly focus on the statistical efficiency of two-stress test plans, thus the first two design concerns are addressed in our study. It is assumed that the number of total test units is pre-fixed. And we adopt an equally-spaced measurement schedule, which is convenient and commonly used in engineering applications. The reader who is interested in the test unit and measurement schedule planning is referred to references – (Lim, 2015), (Tsai et al., 2016) and (Limon et al., 2019).

Suppose there are L stress levels in total created by the two stress variables. Denote a test plan

$$\mathcal{P} = \begin{pmatrix} \xi_1 & \cdots & \xi_l & \cdots & \xi_L \\ \pi_1 & \cdots & \pi_l & \cdots & \pi_L \end{pmatrix},$$

where $\xi_l = (x_1^l, x_2^l)'$ and π_l are the values of the stress variables for the l^{th} stress level and the proportion of test units allocated to that level, respectively. In such case, the total Fisher information matrix under the test plan $\mathcal{P}$, $\mathcal{I}(\theta; \mathcal{P})$ can be expressed as

$$\mathcal{I}(\theta; \mathcal{P}) = n \sum_{l=1}^L \pi_l \mathcal{I}_l(\theta; \xi_l),$$

where $\mathcal{I}_l(\theta; \xi_l)$ is the Fisher information matrix for a single unit at the l^{th} stress level, $\sum_{l=1}^L \pi_l = 1$, and n is the total number of test units.

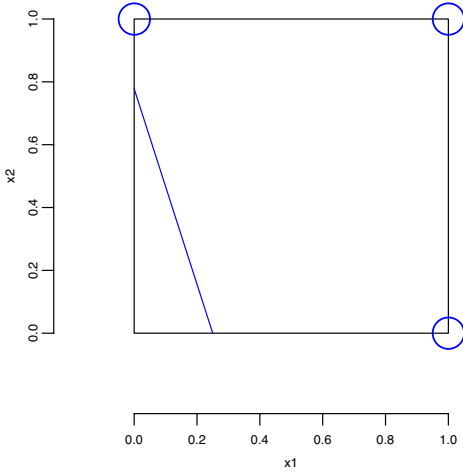


Fig. 2: A Sample Design Space.

In addition, practitioners would like to complete the ADT in a time period they can tolerate; this is also the motivation for conducting accelerated tests. Thus, to reduce the time consumed on experiments as much as possible, a minimum percentage of failure is set to be $a\%$. We expect that this amount of test units would fail at the lowest stress level at the end of test period. This is equivalent to restricting possible test conditions to a region $\mathbb{D}$ bounded by a minimum-stress level line ℓ defined as

$$E[D(\tau'; \boldsymbol{\beta}, \mathbf{b}_i)] = \beta_0 + \beta_1 \tau' + (\eta_1 x_1 + \eta_2 x_2) \tau' \geq a\% D_f,$$

where τ' denotes the transformed test termination time. This provides a constraint $(\eta_1 x_1 + \eta_2 x_2 = a^*)$ on the lowest stress level, where $a^* = \frac{a\% D_f - \beta_0}{\tau'}$.

Under these assumptions above, the initial design space defined by the unit square in the first quadrant shrinks to the region $\mathbb{D}$. Figure 2 shows a sample design space, where the blue solid line represents the minimum-stress level line ℓ and the region on its right is the design space $\mathbb{D}$. The centers of the three blue solid circles indicate a possible choice of combinations of stress levels and the areas of the circles are proportional to the allocation to each level, where an equal amount of test units is assigned in this example.

C. Optimization Scheme

Based on the aforementioned planning criteria and scope, we can formulate corresponding optimization problems shown as below:

$$\begin{aligned} \min \quad & -\det \left[n \sum_{l=1}^L \pi_l \mathcal{I}_l(\theta; \xi_l) \right] && \text{for D-optimality} \\ & \mathbf{c}' \left(n \sum_{l=1}^L \pi_l \mathcal{I}_l(\theta; \xi_l) \right)^{-1} \mathbf{c} && \text{for C-optimality} \\ \text{s.t.} \quad & \sum_{l=1}^L \pi_l = 1 \\ & n\pi_l \in \mathbb{Z}_{>0} \quad \forall l = 1, 2, \dots, L \\ & \xi_l \in \mathbb{D} \quad \forall l = 1, 2, \dots, L \end{aligned}$$

Note that this objective function can be further simplified (see Appendix D) to ease calculation. In this paper, we choose the grid search method to find the optimal solution. In addition, we will look for a design that treats C-optimality as the major criterion while considers D-optimality as a secondary criterion.

IV. OPTIMUM AND COMPROMISE TEST PLANS

As stated in Section I-C, most existing research works about ADT planning mainly focus on the situation in which there is a single stress variable. Adding an additional stress variable would lead to extra decision variables, thus making the optimization problems become harder. As presented in Figure 2, the design space is the region $\mathbb{D}$ when two stress variables are incorporated into an ADT. To ensure the nonsingularity of the Fisher information matrix, it is expected that units are tested at three or more noncollinear stress levels; and we call this plan a nondegenerate plan since all the parameters are estimable. For instance, the plan demonstrated by the three blue solid circles is such a one. As a comparison, a test plan with only two or less stress levels is called a degenerate plan, which only allows estimation of partial parameters.

Inspired by Escobar and Meeker's work (1995) about accelerated life tests, in this section, we provide a method to reduce the two-variable planning problem to a degenerate scenario, which can be treated as a "single-variable" case in a similar way. Then we split the result to achieve a C-optimality nondegenerate plan with the consideration of D-optimality as a supplement in the meantime.

A. The Optimum Plan under a Single-variable Case

Before looking at test plans with two stress variables, we first revisit the scenario with a single variable only. The existing work provided by Weaver and Meeker (2014a) mainly utilizes numerical optimization methods to find the optimum test plan and it takes advantage of the general equivalence theorem to verify global optimality. In the following proposition, we provide a theoretical proof to show that the optimal setting of test conditions always lies at the lowest and highest stress levels for the single-variable case.

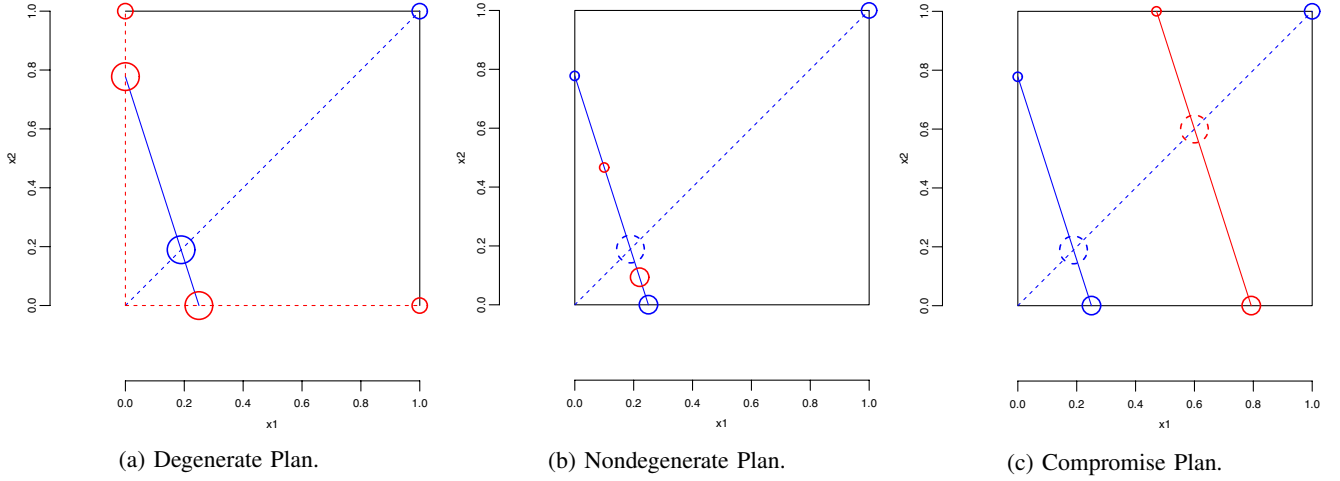


Fig. 3: Various Test Plans.

Proposition 1: For an ADT with a single stress variable, consider a test plan

$$\mathcal{P}_s = \begin{pmatrix} x_H & x_L \\ \pi_H & \pi_L \end{pmatrix},$$

where the two stress levels satisfy $x_L^* \leq x_L < x_H \leq 1$ and x_L^* is the minimum allowable stress level. Then, given any fixed value of π_H (π_L) $\in (0, 1)$, $\text{AVar}(\hat{t}_p)$ achieves its minimum when $x_L = x_L^*$ and $x_H = 1$.

Proof of this proposition is given in Appendix E.

B. The Optimum Degenerate Plan

Then, we consider two-level degenerate plans for the two-variable case. To allow extrapolating to the use condition, the two stress levels must be on a fixed line ℓ_d with slope s that passes through the origin. Figure 3a indicates possible degenerate plans, in which s ranges from 0 to ∞ . The two extreme scenarios (i.e. represented in red dashed lines) are essentially single-variable cases. The scenario represented in the blue dashed line with slope 1 presents the case where $(\beta_0, \beta_1, \eta_1 + \eta_2, \sigma_0, \sigma_1, \rho, \sigma)'$ is the complete set of estimable parameters. Thus, the degenerate plan can be seen as a “single-variable” case such that the acceleration function $\eta_1 x_1 + \eta_2 x_2$ diminishes to $(\eta_1 + s\eta_2)x_1$, or equivalently $(\eta_1/s + \eta_2)x_2$. Following Proposition 1, it can be easily seen that, when the lower stress level ξ_L is the intersection of ℓ_d and the minimum-stress level line ℓ and the higher stress level ξ_H is the intersection of ℓ_d and the boundary of the design space $\mathbb{D}$, $\text{AVar}(\hat{t}_p)$ achieves its minimum given any fixed value of π_H (π_L) and s . Moreover, in the following corollary, we show that among the infinite number of choices for $s, 0 \leq s < \infty$, $\text{AVar}(\hat{t}_p)$ achieves its minimum when $s = 1$.

Corollary 1: For an ADT with two stress variables, consider a two-level degenerate test plan

$$\mathcal{P}_d = \begin{pmatrix} \xi_H & \xi_L \\ \pi_H & \pi_L \end{pmatrix},$$

where the two stress levels are on a line ℓ_d with slope s that passes through the origin. The lower stress level ξ_L is

the intersection of ℓ_d and the minimum-stress level line ℓ . The higher stress level ξ_H is the intersection of ℓ_d and the boundary of the design space $\mathbb{D}$. Then, given any fixed value of π_H (π_L) $\in (0, 1)$, $\text{AVar}(\hat{t}_p)$ achieves its minimum when $s = 1$.

Proof of this corollary is given in Appendix F. Then, the following corollary follows directly from Corollary 1.

Corollary 2: Among the infinite number of choices for $s, 0 \leq s < \infty$, the optimum two-level degenerate plan to minimize $\text{AVar}(\hat{t}_p)$ is unique with test conditions placed at the two stress levels $\xi_H = (1, 1)'$ and $\xi_L = (\frac{a^*}{\eta_1 + \eta_2}, \frac{a^*}{\eta_1 + \eta_2})'$, i.e. when $s = 1$. π_H (π_L) is chosen such that the test plan minimizes $\text{AVar}(\hat{t}_p)$.

Note that the optimum allocation can be found using the grid search method mentioned in Section III-C.

C. The Optimum Nondegenerate Plan

Although a degenerate plan is not directly useful in practice, it does provide a means for finding nondegenerate plans. As explained later, we can split a degenerate plan to obtain a corresponding nondegenerate plan while maintaining the same value of $\text{AVar}(\hat{t}_p)$ achieved. Meanwhile, with the consideration of minimizing the asymptotic variance of parameters estimation, the split should be done in a proper way as shown in the following theorem.

Theorem 1: Consider a two-level degenerate test plan

$$\mathcal{P}_d = \begin{pmatrix} \xi_H & \xi_L \\ \pi_H & \pi_L \end{pmatrix},$$

where the lower stress level ξ_L is the intersection of the minimum-stress level line ℓ and a fixed line ℓ_d with slope s that passes through the origin. The higher stress level ξ_H is the intersection of ℓ_d and the boundary of the design space $\mathbb{D}$. Further define a corresponding nondegenerate split plan

$$\mathcal{P}_n = \begin{pmatrix} \xi_H & \xi_{L1} & \xi_{L2} \\ \pi_H & \pi_{L1} & \pi_{L2} \end{pmatrix},$$

where the two lower stress levels ξ_{L1} and ξ_{L2} are on ℓ and $\pi_L = \pi_{L1} + \pi_{L2}$. Let $\text{AVar}^d(\hat{t}_p)$ and $\text{AVar}^n(\hat{t}_p)$ be the asymptotic variance of $\hat{t}_p$ for the degenerate and nondegenerate plan, respectively. Then, a) $\text{AVar}^n(\hat{t}_p) \geq \text{AVar}^d(\hat{t}_p)$, where the equality

holds when $\pi_{L1}\xi_{L1} + \pi_{L2}\xi_{L2} = \pi_L\xi_L$; b) $\det(I(\theta; \mathcal{P}_n))$ is maximized when the split is done as wide as possible (i.e. until the two lower levels reach the boundaries of $\mathbb{D}$) while maintaining the value of $\text{AVar}^n(\hat{t}_p)$.

Proof of this theorem is given in Appendix G. Note that this theorem holds for any value of s , $0 \leq s < \infty$; but the optimum nondegenerate plan is resulted in when $s = 1$. Figure 3b provides a graphical explanation for finding the optimum nondegenerate plan. Initially, the optimum degenerate plan (i.e. represented in the blue solid circle and the blue dashed circle on the blue dashed line with $s = 1$) is found using Corollary 2. Then, to maintain the same value of $\text{AVar}^d(\hat{t}_p)$ achieved, the units allocated to the lower stress level should be split along the minimum-stress level line according to the proportions specified by Theorem 1. In fact, there exist an infinite number of optimum nondegenerate plans, where the results represented in the two red solid circles and the two blue solid circles are two possible splits. However, to minimize the asymptotic variance of parameters estimation, the plan with three blue solid circles located on the boundaries of $\mathbb{D}$ should be selected.

D. Compromise Plans

Although the optimum plan provides good estimation precision, it may suffer from model uncertainty when the assumed acceleration model deviates significantly from the truth (Meeker and Escobar, 2014). For instance, if there exists an interaction effect between the two stress variables, the current optimum plan is not able to capture this effect since it has only three non-collinear factor levels, which is insufficient for estimating the extra parameter associated with the interaction term. To provide test plans robustness to such model deviation, compromise plans are often created. Similar to the single-variable case, one could allocate a fixed proportion of units (say 20%) at the middle point between an optimized ξ_H and ξ_L . One could also split the units along a line that is parallel with the minimum-stress level line to create a 5-level compromise plan. The red dashed and solid circles in Figure 3c present such idea.

V. AN IN-DEPTH STUDY ON OPTIMAL ALLOCATION

As implied by Proposition 1, once the design space $\mathbb{D}$ is specified, the optimal setting of test conditions can be arranged at the lowest and highest stress levels. In terms of the determination of allocation, Schwabe et al. (2014), in their discussion paper on (Weaver and Meeker, 2014a), proposed a closed-form solution to find the optimal allocation, of which formulas are associated with test conditions only. They suggested that the optimization problem for C-optimality is equivalent to the optimal extrapolation design problem proposed by Klefer and Wolfowitz (1964). Later, in the rejoinder, Weaver and Meeker (2014b) responded to this comment by conducting a numerical study, which verified this claim. Here, we provide a complete theoretical exposition of this conjecture.

Theorem 2: For an ADT with a single stress variable, consider a test plan

$$\mathcal{P}_s = \begin{pmatrix} x_H & x_L \\ \pi_H & \pi_L \end{pmatrix},$$

where the test conditions have been standardized and x_L^* is the minimum allowable stress level. Then, given any fixed value of $x_U < x_L^* \leq x_L < x_H \leq 1$, $\text{AVar}(\hat{t}_p)$ under x_U achieves its minimum when

$$\pi_H = \frac{x_L - x_U}{x_H + x_L - 2x_U} \text{ and } \pi_L = \frac{x_H - x_U}{x_H + x_L - 2x_U}.$$

As a special case of the use condition, i.e. $x_U = 0$,

$$\pi_H = \frac{x_L}{x_H + x_L} \text{ and } \pi_L = \frac{x_H}{x_H + x_L}.$$

Proof of Theorem 2 is given in Appendix H. Then, based on this theorem, it is easy to extend it to the case of two stress variables.

Corollary 3: Following Corollary 2, Theorem 1, and Theorem 2, the optimum two-level degenerate plan to minimize $\text{AVar}(\hat{t}_p)$ under the use condition is unique with solution

$$\mathcal{P}_d^* = \begin{pmatrix} 1 & \frac{a^*}{\eta_1 + \eta_2} \\ 1 & \frac{a^*}{\eta_1 + \eta_2} \\ \frac{a^*}{\eta_1 + \eta_2 + a^*} & \frac{\eta_1 + \eta_2}{\eta_1 + \eta_2 + a^*} \end{pmatrix}.$$

Then, the corresponding optimum nondegenerate split plan is given by

$$\mathcal{P}_n^* = \begin{pmatrix} (1, 1)' & \xi_{L1} & \xi_{L2} \\ \frac{a^*}{\eta_1 + \eta_2 + a^*} & \pi_{L1} & \pi_{L2} \end{pmatrix},$$

where the two lower stress levels, ξ_{L1} and ξ_{L2} , are on the minimum-stress level line ℓ and satisfy $\pi_{L1}\xi_{L1} + \pi_{L2}\xi_{L2} = \pi_L\xi_L$ and $\pi_{L1} + \pi_{L2} = \frac{\eta_1 + \eta_2}{\eta_1 + \eta_2 + a^*}$.

Thus far, one can see that we have already found the closed-form solution to the two-stress variable plan with the optimal setting of test conditions and allocation of test units. Although the optimal allocation of test units is found analytically by this theorem, in practice we still recommend to apply the grid search method to find the exact allocation due to the integer constraint and also to accurately assess the estimation precision.

VI. REVISITING THE MOTIVATING EXAMPLE

In this section, we demonstrate the proposed methods to plan an ADT with two stress variables by revisiting the motivating example.

A. Planning Information

Suppose that the goal is to develop a test plan to evaluate the 5th quantile lifetime of the optical media under the use condition. As specified by Model (2), we assume the degradation process is described by the following model:

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})\tau_{ij} + (\eta_1 x_{1i} + \eta_2 x_{2i})\tau_{ij} + \varepsilon_{ij},$$

where $y_{ij} = \ln(\text{Measurement})$ and $\tau_{ij} = t_{ij}^\gamma$. The planning information is as follows:

- Like the original experiment, 90 units are planned to be tested and 5 repeated measurements for each unit will be taken. For illustrative purposes, we assume measurement is taken every 250 hours until 1,000 hours for each unit.

- The failure threshold is $D_f = \ln(280)$. Our objective under C-optimality is to minimize $AVar(\hat{t}_{0.05})$, i.e. the 5th quantile lifetime under the use condition.
- 85°C (i.e. $x_1 = 1$) and 85% (i.e. $x_2 = 1$) are the maximum allowable testing temperature and RH, respectively. The normal use condition is defined as temperature 25°C and RH 50%, i.e. $x_1 = 0$ and $x_2 = 0$.
- Since the test plan depends on the unknown model parameters, we utilize the following pilot planning values $\theta^* = (2.663, 0.001, 0.056, 0.018, 0.707, 0.002, -0.100, 0.247)'$ and $\gamma = 0.7$. Figure 4 indicates the approximately linearized degradation paths after power transformation.
- We assume that a minimum 80.77% of the failure threshold needs to be achieved by the termination time, which implies that a minimum-stress level line $0.056x_1 + 0.018x_2 = 0.014$ is present.

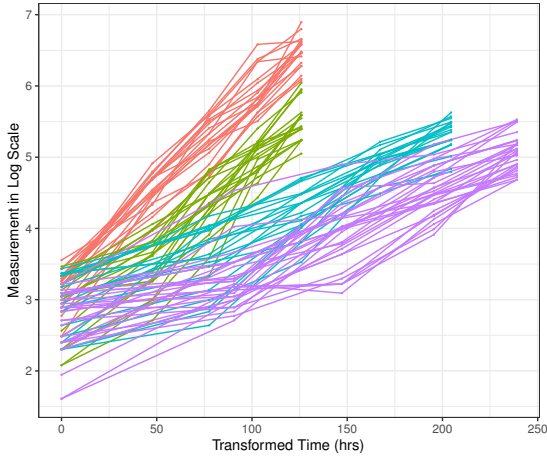


Fig. 4: Approximately Linearized Degradation Paths of Optical Media.

B. Degenerate Plans

Before seeking the optimal design with both stress variables, we first study the degenerate plans. They also contain two single-variable cases: a) temperature is the only varying variable with RH setting at 50%, and b) RH is the only varying variable with temperature setting at 25°C.

Table II demonstrates the degenerate test plan with $s = 1$ and the two single-variable cases with optimal allocations. We also provide a plot indicating the value of $AVar^d(\hat{t}_p)$ versus the slope s in Figure 5. It turns out the degenerate plan with $s = 1$ generates the minimum $AVar^d(\hat{t}_p)$. As already shown in Corollaries 1 and 2, this result comes with no surprise. Also the result of the optimal allocation matches the conclusion of Theorem 2.

C. Nondegenerate Plans

Next, we generate nondegenerate plans according to Theorem 1. Table III presents two nondegenerate plans after two possible splits. As shown in this table, the two plans have approximately the same value of $AVar^n(\hat{t}_p)$, which is also approximately equal to $AVar^d(\hat{t}_p)$ of $\mathcal{P}_d^*$ in Table II; but the

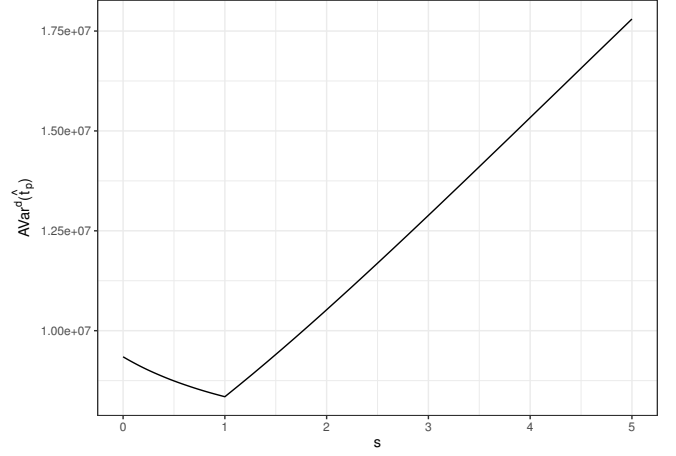


Fig. 5: $AVar^d(\hat{t}_p)$ of Degenerate Plans versus Slope s .

TABLE II: Degenerate Plans.

Type	Stress Levels & Allocation	$AVar^d(\hat{t}_p)$
Single-variable Case a)	$\mathcal{P}_s = \begin{pmatrix} 1 & 0.2500 \\ 18/90 & 72/90 \end{pmatrix}$	9.3462×10^6
Degenerate with $s = 1$	$\mathcal{P}_d^* = \begin{pmatrix} 1 & 0.1892 \\ 14/90 & 76/90 \end{pmatrix}$	8.3469×10^6
Single-variable Case b)	$\mathcal{P}_s = \begin{pmatrix} 1 & 0.7778 \\ 39/90 & 51/90 \end{pmatrix}$	1.0701×10^8

best choice is $\mathcal{P}_n^*$ with a lower value of $-\det(I(\theta; \mathcal{P}_n))$ since it spreads out units to the boundaries the most. The values of $AVar^n(\hat{t}_p)$ approximately match the conclusion of Theorem 1 with a little variation due to the integer constraint of allocation. Also the results of the optimal allocation match the conclusion of Corollary 3.

TABLE III: Nondegenerate Plans.

Type	Stress Levels & Allocation	$AVar^n(\hat{t}_p)$	$-\det(I(\theta; \mathcal{P}_n))$
Nondegenerate	$\mathcal{P}_n = \begin{pmatrix} 1 & 0.1000 & 0.2000 \\ 1 & 0.4670 & 0.1560 \\ 14/90 & 8/90 & 68/90 \end{pmatrix}$	8.3485×10^6	-1.0924×10^{35}
Nondegenerate	$\mathcal{P}_n^* = \begin{pmatrix} 1 & 0 & 0.2500 \\ 1 & 0.7778 & 0 \\ 14/90 & 18/90 & 58/90 \end{pmatrix}$	8.3475×10^6	-1.3113×10^{36}

In addition, we provide some illustrative examples by setting test conditions at the extreme points of the design space. To provide better interpretation of the results, we report the relative efficiency (RE) in percentage for C-optimality. RE is defined as $RE = AVar(\hat{t}_p; \mathcal{P}_n^*) / AVar(\hat{t}_p; \mathcal{P}_n)$, where $AVar(\hat{t}_p; \mathcal{P}_n^*)$ is the value of $AVar^n(\hat{t}_p)$ under the optimum nondegenerate plan and $AVar(\hat{t}_p; \mathcal{P}_n)$ is the value $AVar^n(\hat{t}_p)$ for any other ordinary plan. From Table IV, it can be seen that the best plan is still plan $\mathcal{P}_n^*$, where the test conditions are set at the highest level and the other two lower levels on the minimum-stress level line. For plans 1 and 5 in Table IV, the two lower stress levels and the use condition are on a line. In such cases, having the two lower levels is similar to the case of a test with a single stress variable. As a result, a single unit

TABLE IV: Illustrative Examples for Nondegenerate Plans.

Plan	Stress Levels & Allocation	$AVar^n(\hat{t}_p)$	RE
1	$\mathcal{P}_n = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0.7778 \\ 1/90 & 39/90 & 50/90 \end{pmatrix}$	1.0815×10^8	7.72%
2	$\mathcal{P}_n = \begin{pmatrix} 1 & 0 & 0.2500 \\ 1 & 1 & 0 \\ 15/90 & 15/90 & 60/90 \end{pmatrix}$	8.5043×10^6	98.16%
3	$\mathcal{P}_n = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 30/90 & 30/90 & 30/90 \end{pmatrix}$	1.9272×10^7	43.31%
4	$\mathcal{P}_n = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0.7778 & 0 \\ 38/90 & 35/90 & 27/90 \end{pmatrix}$	1.5335×10^7	54.43%
5	$\mathcal{P}_n = \begin{pmatrix} 1 & 0.2500 & 1 \\ 1 & 0 & 0 \\ 1/90 & 71/90 & 18/90 \end{pmatrix}$	9.3961×10^6	88.84%

allocated to the highest level is enough in order to make all parameters estimable.

D. Compromise Plans

Finally, a compromise plan, $\mathcal{P}_c$ is created by fixing a 20% of the units (i.e. 18 units) at the middle point $-\xi_M = (0.6, 0.6)'$. Let $AVar^c(\hat{t}_p)$ be the asymptotic variance of $\hat{t}_p$ for $\mathcal{P}_c$. Undoubtedly, this 4-level compromise plan would offer a larger value of $AVar^c(\hat{t}_p)$ than the $AVar^n(\hat{t}_p)$ value achieved by $\mathcal{P}_n^*$. Again, given the existing conclusion of Theorem 1, the units allocated at the middle point can be split into two portions; and the split should be done until the two resulted test conditions reach the boundaries of $\mathbb{D}$. As a consequence, this 5-level compromise plan maintains the same value of $AVar^c(\hat{t}_p)$, but obtains a lower $-\det(I(\theta; \mathcal{P}_c))$. Table V gives the results. Comparing it with Table IV, one can also see that some estimation precision is lost by introducing additional stress levels to the original optimum plan.

TABLE V: Compromise Plans.

Plan Type	Stress Levels & Allocation	$AVar^c(\hat{t}_p)$	$-\det(I(\theta; \mathcal{P}_c))$
4-level	$\mathcal{P}_c = \begin{pmatrix} 1 & 0 & 0.2500 & 0.6000 \\ 1 & 0.7778 & 0 & 0.6000 \\ 11/90 & 15/90 & 46/90 & 18/90 \end{pmatrix}$	9.2048×10^6	-1.0159×10^{36}
5-level	$\mathcal{P}_c = \begin{pmatrix} 1 & 0 & 0.2500 & 0.7929 & 0.4714 \\ 1 & 0.7778 & 0 & 0 & 1 \\ 11/90 & 15/90 & 46/90 & 7/90 & 11/90 \end{pmatrix}$	9.2048×10^6	-1.6513×10^{36}

VII. SIMULATION STUDY AND SENSITIVITY ANALYSIS

A. Simulation Study

To compare different test plans derived in this paper and to assess the adequacy of large-sample (i.e. asymptotic) approximations, we conduct a Monte Carlo simulation study using the assumed model and the planning information specified in Section VI-A. On each simulation replication, data are randomly generated by either the optimum nondegenerate plan

or the 5-level compromise plan, as presented in Table III and V, respectively. Then, model parameters are estimated by MLE and they are used to calculate the 5-th quantile lifetime. A total of 1,000 replications are simulated and analyzed for each test plan. We evaluate results using various sample sizes (i.e. the number of test units), $n = 45, 65, 90, 120, 150$, and 180.

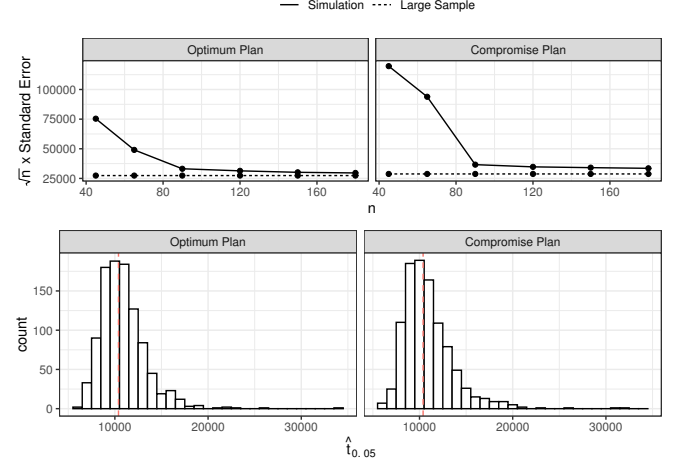


Fig. 6: Results of Monte Carlo Simulation Study.

The upper panel of Figure 6 illustrates the scaled prediction variance of the 5-th quantile, from both large-sample approach and simulation approach, versus sample size. Notice that as sample size increases, the discrepancy between these two approaches becomes smaller for each test plan. The lower panel of Figure 6 provides the histograms of $\hat{t}_{0.05}$ for both plans under the use condition with $n = 180$. It indicates that the distributions for both plans cover the actual 5-th quantile lifetime, which is presented as a red dashed vertical line. In addition, both panels show the increased variability in the compromise plan.

In summary, when sample size is relatively large, e.g. $n \geq 90$, the simulation result agrees with the analytically derived large-sample approximation. The compromise plan tends to have larger standard error than the optimum plan, which matches our previous conclusion. For a small sample size, we recommend using simulation to estimate the variability of predicted quantile value.

B. Sensitivity Analysis

From Theorem 2, it becomes clear that once a design space $\mathbb{D}$ is defined, the optimal allocation of test units can be specified regardless of model parameters' value. However, as implied by Corollary 3, the design decisions made on both test condition and test unit allocation are relevant to both the fixed-effects parameters $-\beta = (\beta_0, \beta_1, \eta_1, \eta_2)'$ and the failure threshold D_f . These terms have an effect on the optimum plan since varying their values does change the design space. Thus, to study the robustness of the optimum plan, we carry out a sensitivity analysis of the optimum plan by increasing or decreasing each parameter's value in β and D_f by 10%. As a convention, we use +1 and -1 to represent increase and decrease in these factors and 0 stands for no change.

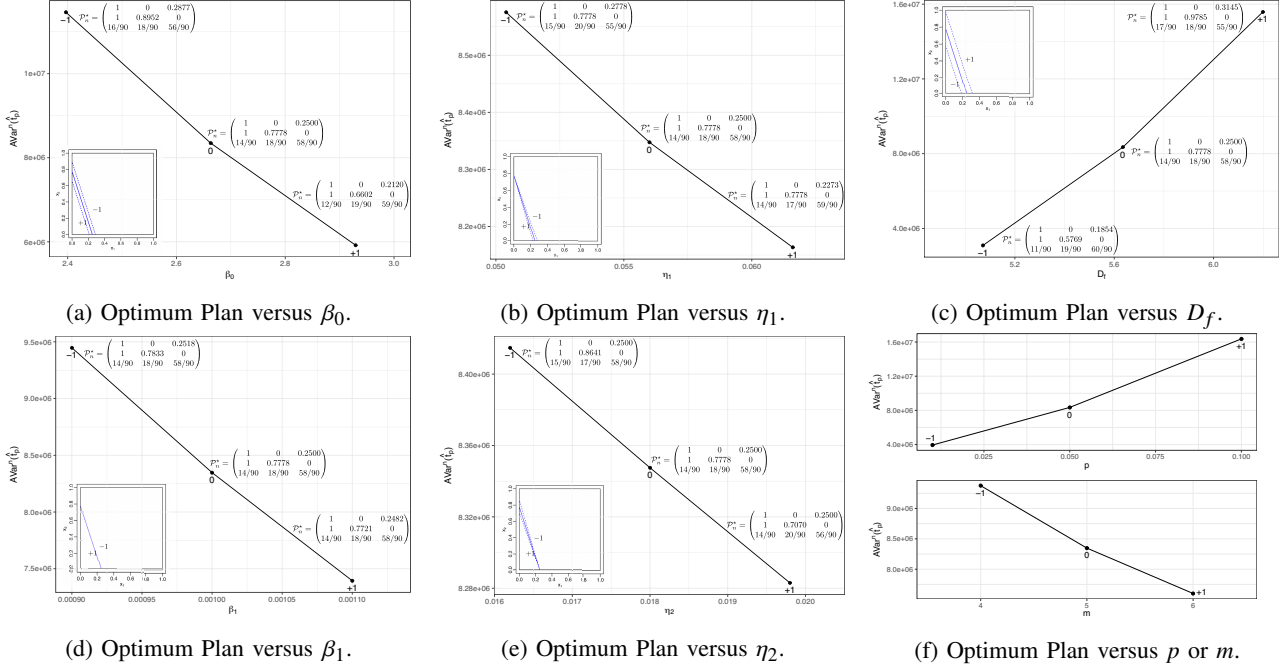


Fig. 7: Results of Sensitivity Analysis.

Results are shown in Figure 7, in which each plot depicts $AVar^n(\hat{t}_p)$ versus the change of a certain factor, with $\mathcal{P}_n^*$ of the corresponding optimum nondegenerate plan showing aside of each point and the resulted design space showing in the corner. In general, the graphs show that the change of each factor makes the minimum-stress level line get either closer or farther to the use condition. As a consequence, $AVar^n(\hat{t}_p)$ becomes larger as the design space becomes smaller. This is because more extrapolation from the higher stress level to the use stress level is resulted in. Particularly, the optimum plan seems to be less sensitive to β_1 due to the small change of design space. We also evaluate the corresponding optimum plans when the value of p (lifetime percentile) or m (number of measurements) varies. The result is shown in Figure 7f. It turns out that $AVar^n(\hat{t}_p)$ increases with p . This is because more uncertainty is involved in inferring a higher percentile lifetime. Furthermore, as expected, $AVar^n(\hat{t}_p)$ decreases with m . The reason behind this is that given more frequent inspection schedule, more information about the degradation process is produced. As a result, the uncertainty in estimating the model variance component (i.e. δ) is reduced. But for both cases, the optimum plan stays unchanged. This result does not come with surprise since p and m are not involved in the specification of optimum plan, as stated in Corollary 3. In addition, notice that random-effects parameters, $\delta = (\sigma_0, \sigma_1, \rho, \sigma)'$, do not affect the optimum plan at all. It is due to the fact that changing test plans would not change the information for estimating these random-effects parameters, as implied by the M_i matrix presented in Equation (5) in Appendix A.

VIII. CONCLUDING REMARKS

Planning an ADT usually involves multiple decisions to make. In the existing literature, the effort has been focused

more on designing the test with a single stress variable alone. As shown in this paper, we extend the design procedure to addressing the case of two stress variables. This procedure is accomplished by a proposed analytical approach. It initially finds the optimal setting of stress levels and allocation of test units for a degenerate scenario and then splits the result to form the optimum nondegenerate plan. The split is done by maintaining the same value of the asymptotic variance of the estimated p -th quantile lifetime achieved with the consideration of the asymptotic variance of parameters estimation meanwhile. This complete methodology – including the modeling framework, the rigorous proofs, and the design procedure – has been demonstrated using the illustrative example.

Notwithstanding that the following issues are beyond the scope of current study, they are definitely worth of a further investigation:

- 1) It is noted that an optimum test plan is designed for a specific planning criterion, so it may happen that the optimum plan under one criterion performs poorly under another criterion. In this paper, we have investigated a design that weighs the asymptotic variance of quantile lifetime under the use condition as the major criterion with the asymptotic variance of parameters estimation as a secondary criterion. This study can be extended to cover other criteria such as V-optimality, which minimizes the average prediction variance over a specific set of design points, or any cost-based design criteria.
- 2) The test plans we developed are based on the assumption that the two stress variables are independent to each other. If there exists an interaction effect, compromise plans can provide some robustness to such a deviation from model assumption. But future research is still needed to characterize the optimum plan under such situation.

- 3) Addressing the optimal design for other types of ADT models including nonlinear mixed-effects models and stochastic process models is much desired. For instance, Tung and Tseng (2019) recently proposed the optimum plan for gamma process-based degradation model with two stress variables under C-optimality. Methodologies handling other types of models such as inverse Gaussian process-based model are still lack of investigation.
- 4) As designing test plans relies on the knowledge of degradation model and model parameters, creating robust designs to model and parameter uncertainties is desired. Bayesian design methodology, as described by Chaloner and Verdinelli (1995), is clearly a future research direction. This issue has recently been discussed by Zhao et al. (2018a) and Weaver and Meeker (2019) too.
- 5) In this paper, the assumption of an equally-spaced measurement schedule is made due to its convenience and commonality in engineering practice. But an unequally-spaced measurement plan may provide better performance in terms of generating precise estimates of percentile lifetime. Typically, one should make more frequent measurements when the degradation process is close to passing its failure threshold. The intuitive reason for this is that more information about the close-to-failure phase of the degradation process is obtained, which results in less uncertainty in extrapolation. But the optimal schedule depends much on the type of the underlying degradation model and can be very sensitive to model choice. Clearly, this problem desires much research effort. For some discussions on it, please refer to the works provided by Boulanger and Escobar (1994).

REFERENCES

- Boulanger, M. and Escobar, L. A. (1994). Experimental design for a class of accelerated degradation tests. *Technometrics*, 36(3):260–272.
- CD Associates, I. (1998). *The DVD1000P Analyzer Manual*. Irvine California, USA.
- Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, pages 273–304.
- Dean, A., Morris, M., Stufken, J., and Bingham, D. (2015). *Handbook of design and analysis of experiments*, volume 7. CRC Press.
- Escobar, L. A. and Meeker, W. Q. (1995). Planning accelerated life tests with two or more experimental factors. *Technometrics*, 37(4):411–427.
- Fang, G., Pan, R., and Hong, Y. (2020). Copula-based reliability analysis of degrading systems with dependent failures. *Reliability Engineering & System Safety*, 193:106618.
- Fang, G., Rigdon, S. E., and Pan, R. (2018). Predicting lifetime by degradation tests: A case study of ISO 10995. *Quality and Reliability Engineering International*, 34(6):1228–1237.
- Harville, D. A. (1998). *Matrix algebra from a statistician's perspective*. Taylor & Francis Group.
- Jennrich, R. I. and Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, 42(4):805–820.
- Kim, S.-J. and Bae, S. J. (2013). Cost-effective degradation test plan for a nonlinear random-coefficients model. *Reliability Engineering & System Safety*, 110:68–79.
- Klefer, J. and Wolfowitz, J. (1964). Optimum extrapolation and interpolation designs, i. *Annals of the Institute of Statistical Mathematics*, 16(1):79–108.
- Lim, H. (2015). Optimum accelerated degradation tests for the Gamma degradation process case under the constraint of total cost. *Entropy*, 17(5):2556–2572.
- Limon, S., Rezaei, E., and Yadav, O. P. (2019). Designing an accelerated degradation test plan considering the Gamma degradation process with multi-stress factors and interaction effects. *Quality Technology & Quantitative Management*, pages 1–17.
- Meeker, W. Q. and Escobar, L. A. (1998). *Statistical methods for reliability data*. Wiley, New York.
- Meeker, W. Q. and Escobar, L. A. (2014). *Statistical methods for reliability data*. John Wiley & Sons.
- Organization, I. S. (2011). Information technology – digitally recorded media for information interchange and storage – test method for the estimation of the archival lifetime of optical media.
- Pan, R. and Crispin, T. (2011). A hierarchical modeling approach to accelerated degradation testing data analysis: A case study. *Quality and Reliability Engineering International*, 27(2):229–237.
- Pan, Z., Sun, Q., and Feng, J. (2016). Reliability modeling of systems with two dependent degrading components based on Gamma processes. *Communications in Statistics-Theory and Methods*, 45(7):1923–1938.
- Rao, C. R., Rao, C. R., Statistiker, M., Rao, C. R., and

- Rao, C. R. (1973). *Linear statistical inference and its applications*, volume 2. Wiley New York.
- Rencher, A. C. and Schaalje, G. B. (2008). *Linear models in statistics*. John Wiley & Sons.
- Schwabe, R., Prus, M., and Graßhoff, U. (2014). Discussion of 'methods for planning repeated measures accelerated degradation tests' by brian p. weaver and william q. meeker. *Applied Stochastic Models in Business and Industry*, 30(6):677–679.
- Seo, K. and Pan, R. (2017). Data analysis of step-stress accelerated life tests with heterogeneous group effects. *IIEE Transactions*, 49(9):885–898.
- Si, W., Yang, Q., Wu, X., and Chen, Y. (2018). Reliability analysis considering dynamic material local deformation. *Journal of Quality Technology*, 50(2):183–197.
- Tsai, C.-C., Tseng, S.-T., and Balakrishnan, N. (2012). Optimal design for degradation tests based on Gamma processes with random effects. *IEEE Transactions on Reliability*, 61(2):604–613.
- Tsai, T., Sung, W., Lio, Y. L., Chang, S. I., and Lu, J. (2016). Optimal two-variable accelerated degradation test plan for Gamma degradation processes. *IEEE Transactions on Reliability*, 65(1):459–468.
- Tung, H.-P. and Tseng, S.-T. (2019). Planning gamma accelerated degradation tests with two accelerating variables. *Naval Research Logistics (NRL)*, 66(5):439–447.
- Wang, H., Zhao, Y., Ma, X., and Wang, H. (2017). Optimal design of constant-stress accelerated degradation tests using the m-optimality criterion. *Reliability Engineering & System Safety*, 164:45–54.
- Weaver, B. P. and Meeker, W. Q. (2014a). Methods for planning repeated measures accelerated degradation tests. *Applied Stochastic Models in Business and Industry*, 30(6):658–671.
- Weaver, B. P. and Meeker, W. Q. (2014b). Rejoinder: Methods for planning repeated measures accelerated degradation tests. *Applied Stochastic Models in Business and Industry*, 30(6):686–690.
- Weaver, B. P. and Meeker, W. Q. (2019). Bayesian methods for planning accelerated repeated measures degradation tests. *Technometrics*, pages 1–10.
- Weaver, B. P., Meeker, W. Q., Escobar, L. A., and Wendelberger, J. (2013). Methods for planning repeated measures degradation studies. *Technometrics*, 55(2):122–134.
- Ye, Z.-S., Hu, Q., and Yu, D. (2019). Strategic allocation of test units in an accelerated degradation test plan. *Journal of Quality Technology*, 51(1):64–80.
- Zhao, X., Pan, R., and Xie, M. (2018a). Bayesian planning of step-stress accelerated degradation tests under various optimality criteria. *Applied Stochastic Models in Business and Industry*.
- Zhao, X., Xu, J., and Liu, B. (2018b). Accelerated degradation tests planning with competing failure modes. *IEEE Transactions on Reliability*, 67(1):142–155.

APPENDIX

A. Derivation of $\mathbf{I}(\theta)$

Recall that the full model (2) is given by

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})\tau_{ij} + (\eta_1 x_{1i} + \eta_2 x_{2i})\tau_{ij} + \varepsilon_{ij},$$

Then, if we let $\mathbf{Y}_i = (y_{i1}, y_{i2}, \dots, y_{im_i})'$ denote the measurements on unit i , it is equivalent to

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i, \quad (4)$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \eta_1, \eta_2)'$, $\mathbf{b}_i = (b_0, b_1)' \sim N(\mathbf{0}, \mathbf{V})$, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{im_i})' \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_i)$ with $\mathbf{I}_i$ being a $m_i \times m_i$ identity matrix. Furthermore, $\mathbf{X}_i$ and $\mathbf{Z}_i$ are design matrices given by

$$\mathbf{X}_i = \begin{pmatrix} 1 & \tau_{i1} & x_{1i}\tau_{i1} & x_{2i}\tau_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \tau_{im_i} & x_{1i}\tau_{im_i} & x_{2i}\tau_{im_i} \end{pmatrix}, \quad \mathbf{Z}_i = \begin{pmatrix} 1 & \tau_{i1} \\ \vdots & \vdots \\ 1 & \tau_{im_i} \end{pmatrix}.$$

In such case, $\mathbf{Y}_i \sim MVN(\mathbf{X}_i \boldsymbol{\beta}, \boldsymbol{\Sigma}_i)$, where

$$\boldsymbol{\Sigma}_i = \text{Var}(\mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i) = \mathbf{Z}_i \mathbf{V} \mathbf{Z}_i' + \sigma^2 \mathbf{I}_i.$$

So the log-likelihood function for unit i is

$$\ell_i = -\frac{m_i}{2} \ln(2\pi) - \frac{1}{2} \ln(|\boldsymbol{\Sigma}_i|) - \frac{1}{2} (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta})' \boldsymbol{\Sigma}_i^{-1} (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}).$$

Then, the total log-likelihood function is $\ell = \sum_{i=1}^n \ell_i$.

Suggested by Weaver et al. (2013), we make use of the existing conclusion about the Hessian matrix and its expectation of a linear mixed-effects model provided by Jennrich and Schluchter (1986). If letting $\boldsymbol{\delta} = (\sigma_0, \sigma_1, \rho, \sigma)'$ denote the model variance component parameters, it can be shown that the Hessian matrix, $\mathbf{H}_i$, for unit i is given by

$$\mathbf{H}_i = \begin{pmatrix} \mathbf{H}_{\beta\beta,i} & \mathbf{H}_{\beta\delta,i} \\ \mathbf{H}_{\delta\beta,i} & \mathbf{H}_{\delta\delta,i} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 \ell_i}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}} & \frac{\partial^2 \ell_i}{\partial \boldsymbol{\beta} \partial \boldsymbol{\delta}} \\ \frac{\partial^2 \ell_i}{\partial \boldsymbol{\delta} \partial \boldsymbol{\beta}} & \frac{\partial^2 \ell_i}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}} \end{pmatrix}.$$

Then, the Fisher information is given by the negative of the expectation of the Hessian matrix.

$$\mathbf{I}_i(\theta) = -E[\mathbf{H}_i] = \begin{pmatrix} \mathbf{X}_i' \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_i \end{pmatrix}, \quad (5)$$

where $\mathbf{M}_i$ is a 4×4 symmetric matrix with elements

$$M_{rs,i} = \frac{1}{2} \text{tr}(\boldsymbol{\Sigma}_i^{-1} \dot{\boldsymbol{\Sigma}}_{ir} \boldsymbol{\Sigma}_i^{-1} \dot{\boldsymbol{\Sigma}}_{is}), \quad r, s = 1, \dots, 4,$$

and

$$\dot{\boldsymbol{\Sigma}}_{ir} = \frac{\partial \boldsymbol{\Sigma}_i}{\partial \delta_r}, \quad r = 1, \dots, 4.$$

It can be further shown that

$$\begin{aligned} \dot{\boldsymbol{\Sigma}}_{i1} &= \frac{\partial \boldsymbol{\Sigma}_i}{\partial \sigma_0} = \mathbf{Z}_i \begin{pmatrix} 2\sigma_0 & \rho\sigma_1 \\ \rho\sigma_1 & 0 \end{pmatrix} \mathbf{Z}_i'; \quad \dot{\boldsymbol{\Sigma}}_{i2} = \frac{\partial \boldsymbol{\Sigma}_i}{\partial \sigma_1} = \mathbf{Z}_i \begin{pmatrix} 0 & \rho\sigma_0 \\ \rho\sigma_0 & 2\sigma_1 \end{pmatrix} \mathbf{Z}_i'; \\ \dot{\boldsymbol{\Sigma}}_{i3} &= \frac{\partial \boldsymbol{\Sigma}_i}{\partial \rho} = \mathbf{Z}_i \begin{pmatrix} 0 & \sigma_0\sigma_1 \\ \sigma_0\sigma_1 & 0 \end{pmatrix} \mathbf{Z}_i'; \quad \dot{\boldsymbol{\Sigma}}_{i4} = \frac{\partial \boldsymbol{\Sigma}_i}{\partial \sigma} = 2\sigma \mathbf{I}_i. \end{aligned}$$

Thus, the total Fisher information is $\mathbf{I}(\theta) = \sum_{i=1}^n \mathbf{I}_i(\theta)$.

B. Derivation of t_p

According to Equation (3), the p -th quantile lifetime t_p under condition x_1 and x_2 satisfies the following:

$$F_T(t_p) = 1 - \Phi\left(\frac{D_f - \beta_0 - \beta_1\tau_p - (\eta_1x_1 + \eta_2x_2)\tau_p}{\sqrt{\sigma_0^2 + \tau_p^2\sigma_1^2 + 2\tau_p\rho\sigma_0\sigma_1}}\right) = p$$

Then,

$$\begin{aligned} \frac{(D_f - \beta_0 - (\beta_1 + \eta_1x_1 + \eta_2x_2)\tau_p)^2}{\sigma_0^2 + \tau_p^2\sigma_1^2 + 2\tau_p\rho\sigma_0\sigma_1} &= [\Phi^{-1}(1-p)]^2 \\ \frac{(h - r\tau_p)^2}{\sigma_0^2 + \tau_p^2\sigma_1^2 + 2\tau_p\rho\sigma_0\sigma_1} &= k \\ (k\sigma_1^2 - r^2)\tau_p^2 + 2(k\rho\sigma_0\sigma_1 + hr)\tau_p + (k\sigma_0^2 - h^2) &= 0, \end{aligned}$$

where $h = D_f - \beta_0$, $r = \beta_1 + \eta_1x_1 + \eta_2x_2$, and $k = [\Phi^{-1}(1-p)]^2$.

Further denote $a = (k\sigma_1^2 - r^2)$, $b = 2(k\rho\sigma_0\sigma_1 + hr)$, and $c = (k\sigma_0^2 - h^2)$. The equation above is equivalent to

$$a\tau_p^2 + b\tau_p + c = 0$$

with solutions

$$\begin{aligned} \tau_p &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-(k\rho\sigma_0\sigma_1 + hr) \pm \sqrt{(k\rho\sigma_0\sigma_1 + hr)^2 - (k\sigma_1^2 - r^2)(k\sigma_0^2 - h^2)}}{k\sigma_1^2 - r^2} \\ &= \frac{-(k\rho\sigma_0\sigma_1 + hr) \pm \sqrt{k^2\rho^2\sigma_0^2\sigma_1^2 + k\sigma_0^2r^2 + k\sigma_1^2h^2 - k^2\sigma_0^2\sigma_1^2 + 2k\rho\sigma_0\sigma_1hr}}{k\sigma_1^2 - r^2}, \end{aligned}$$

then $t_p = \tau_p^{1/\gamma}$.

Note that “ $\pm$ ” becomes “+” and “-” if $0 < p < 0.5$ and $0.5 < p < 1$, respectively. When $p = 0.5$ (i.e. $k = 0$, the case of median lifetime), the equation above reduces to

$$\tau_p = \frac{h}{r} = \frac{D_f - \beta_0}{\beta_1 + \eta_1x_1 + \eta_2x_2}.$$

C. Derivation of $\mathbf{c}$

Recall that $\mathbf{c} = \left(\frac{\partial t_p}{\partial \beta_0}, \dots, \frac{\partial t_p}{\partial \sigma}\right)'$. Let $v = \sqrt{k^2\rho^2\sigma_0^2\sigma_1^2 + k\sigma_0^2r^2 + k\sigma_1^2(\beta_0 - D_f)^2 - k^2\sigma_0^2\sigma_1^2 - 2k\rho\sigma_0\sigma_1r(\beta_0 - D_f)}$, then we will have the followings:

$$\begin{aligned} \frac{\partial t_p}{\partial \beta_0} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \beta_0} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{1}{r^2 - k\sigma_1^2} \left(-r \pm \frac{kh\sigma_1^2 + k\rho\sigma_0\sigma_1}{v} \right) \right\}. \\ \frac{\partial t_p}{\partial \beta_1} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \beta_1} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{2r}{(r^2 - k\sigma_1^2)^2} (-k\rho\sigma_0\sigma_1 - hr \pm v) - \frac{1}{r^2 - k\sigma_1^2} \left(-h \pm \frac{kr\sigma_0^2 + k\rho h\sigma_0\sigma_1}{v} \right) \right\}. \\ \frac{\partial t_p}{\partial \eta_1} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \eta_1} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{2r}{(r^2 - k\sigma_1^2)^2} (-k\rho\sigma_0\sigma_1 - hr \pm v) - \frac{1}{r^2 - k\sigma_1^2} \left(-h \pm \frac{kr\sigma_0^2 + k\rho h\sigma_0\sigma_1}{v} \right) \right\} x_1. \\ \frac{\partial t_p}{\partial \eta_2} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \eta_2} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{2r}{(r^2 - k\sigma_1^2)^2} (-k\rho\sigma_0\sigma_1 - hr \pm v) - \frac{1}{r^2 - k\sigma_1^2} \left(-h \pm \frac{kr\sigma_0^2 + k\rho h\sigma_0\sigma_1}{v} \right) \right\} x_2. \\ \frac{\partial t_p}{\partial \sigma_0} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \sigma_0} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{1}{r^2 - k\sigma_1^2} \left(k\rho\sigma_1 \pm \frac{k^2\sigma_0\sigma_1^2 - kr^2\sigma_0 - k^2\rho^2\sigma_0^2\sigma_1^2 - k\rho h\sigma_1}{v} \right) \right\}. \\ \frac{\partial t_p}{\partial \sigma_1} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \sigma_1} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{1}{r^2 - k\sigma_1^2} \left(k\rho\sigma_0 \pm \frac{k^2\sigma_0^2\sigma_1 - kh^2\sigma_1 - k^2\rho^2\sigma_0^2\sigma_1 - k\rho h\sigma_0}{v} \right) + \frac{2k\sigma_1}{(r^2 - k\sigma_1^2)^2} [k\rho\sigma_0\sigma_1 + hr \pm (-v)] \right\}. \\ \frac{\partial t_p}{\partial \rho} &= \frac{\partial t_p}{\partial \tau_p} \frac{\partial \tau_p}{\partial \rho} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left\{ \frac{1}{r^2 - k\sigma_1^2} \left(k\sigma_0\sigma_1 \pm \frac{-k^2\rho\sigma_0^2\sigma_1^2 - k\rho h\sigma_0\sigma_1}{v} \right) \right\}. \\ \frac{\partial t_p}{\partial \sigma} &= 0. \end{aligned}$$

Again, “ $\pm$ ” becomes “+” and “-” if $0 < p < 0.5$ and $0.5 < p < 1$, respectively. Particularly, when $p = 0.5$, the equations above reduce to

$$\begin{aligned} \frac{\partial t_p}{\partial \beta_0} &= \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left(-\frac{1}{r} \right); \quad \frac{\partial t_p}{\partial \beta_1} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left(-\frac{h}{r^2} \right); \\ \frac{\partial t_p}{\partial \eta_1} &= \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left(-\frac{h}{r^2} \right) x_1; \quad \frac{\partial t_p}{\partial \eta_2} = \frac{1}{\gamma} \tau_p^{(1/\gamma)-1} \left(-\frac{h}{r^2} \right) x_2; \\ \frac{\partial t_p}{\partial \sigma_0} &= \frac{\partial t_p}{\partial \sigma_1} = \frac{\partial t_p}{\partial \rho} = \frac{\partial t_p}{\partial \sigma} = 0. \end{aligned}$$

D. Simplification of the Objective Functions

According to Section III-B, for a given plan $\mathcal{P}$, the total Fisher information matrix is given by

$$\mathcal{I}(\boldsymbol{\theta}; \mathcal{P}) = n \sum_{l=1}^L \pi_l \mathcal{I}_l(\boldsymbol{\theta}; \boldsymbol{\xi}_l),$$

and the existing conclusion in Appendix A shows that

$$\mathcal{I}_l(\boldsymbol{\theta}; \boldsymbol{\xi}_l) = \begin{pmatrix} \mathbf{X}_l' \boldsymbol{\Sigma}_l^{-1} \mathbf{X}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_l \end{pmatrix},$$

where $\mathbf{X}_l$, $\boldsymbol{\Sigma}_l$, and $\mathbf{M}_l$ are the corresponding elements in Equation (5) for a single unit tested at the l^{th} stress level. Then, we further denote $\mathcal{I}(\boldsymbol{\theta}; \mathcal{P})$ by

$$\begin{pmatrix} n \sum_{l=1}^L \pi_l \mathbf{X}_l' \boldsymbol{\Sigma}_l^{-1} \mathbf{X}_l & \mathbf{0} \\ \mathbf{0} & n \sum_{l=1}^L \pi_l \mathbf{M}_l \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{\mathcal{P}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}} \end{pmatrix}.$$

Therefore, the objective functions in Section III-C can be simplified as

- D-optimality:

$$\begin{aligned} -\det \left[\sum_{l=1}^L n_l \mathcal{I}_l(\boldsymbol{\theta}; \boldsymbol{\xi}_l) \right] &= -\det \begin{pmatrix} \mathbf{R}_{\mathcal{P}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}} \end{pmatrix} \\ &= -\det(\mathbf{R}_{\mathcal{P}}) \det(\mathbf{M}_{\mathcal{P}}); \end{aligned}$$

- C-optimality:

$$\begin{aligned} \mathbf{c}' \left(\sum_{l=1}^L n_l \mathcal{I}_l(\boldsymbol{\theta}; \boldsymbol{\xi}_l) \right)^{-1} \mathbf{c} &= \mathbf{c}' \begin{pmatrix} \mathbf{R}_{\mathcal{P}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}} \end{pmatrix}^{-1} \mathbf{c} \\ &= \mathbf{c}' \begin{pmatrix} \mathbf{R}_{\mathcal{P}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}}^{-1} \end{pmatrix} \mathbf{c}. \end{aligned}$$

E. Proof of Proposition 1

In a single-variable case, a test plan is given by

$$\mathcal{P}_s = \begin{pmatrix} x_H & x_L \\ \pi_H & \pi_L \end{pmatrix},$$

and the ADT model becomes

$$y_{ij} = \beta_{0i} + \beta_{1i}\tau_{ij} + \eta x_i\tau_{ij} + \varepsilon_{ij},$$

where the derivation of $\mathcal{I}(\boldsymbol{\theta})$, t_p , and $\mathbf{c}$ follows similarly to Appendices A, B, and C, respectively.

By Appendix D, we have

$$\begin{aligned} \text{AVar}(\hat{t}_p) &= \mathbf{c}' \begin{pmatrix} \mathbf{R}_{\mathcal{P}_s}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}_s}^{-1} \end{pmatrix} \mathbf{c} \\ &= \mathbf{c}' \begin{pmatrix} (n\pi_H \mathbf{X}_H' \boldsymbol{\Sigma}_H^{-1} \mathbf{X}_H + n\pi_L \mathbf{X}_L' \boldsymbol{\Sigma}_L^{-1} \mathbf{X}_L)^{-1} & \mathbf{0} \\ \mathbf{0} & (n\mathbf{M}_i)^{-1} \end{pmatrix} \mathbf{c}, \end{aligned}$$

where X_i^H , X_i^L , and $\mathbf{c}$ can be partitioned in the form

$$X_i^H = \begin{pmatrix} 1 & \tau_{i1} & x_H \tau_{i1} \\ \vdots & \vdots & \vdots \\ 1 & \tau_{imi} & x_H \tau_{imi} \end{pmatrix} \stackrel{\text{def}}{=} (\mathbf{T} \mid x_H \mathbf{t}), X_i^L = \begin{pmatrix} 1 & \tau_{i1} & x_L \tau_{i1} \\ \vdots & \vdots & \vdots \\ 1 & \tau_{imi} & x_L \tau_{imi} \end{pmatrix} \stackrel{\text{def}}{=} (\mathbf{T} \mid x_L \mathbf{t}),$$

$$\text{and } \mathbf{c} = \left(\frac{\partial t_p}{\partial \beta_0}, \dots, \frac{\partial t_p}{\partial \eta} \mid \frac{\partial t_p}{\partial \sigma_0}, \dots, \frac{\partial t_p}{\partial \sigma} \right)' \stackrel{\text{def}}{=} (\mathbf{c}'_1, \mathbf{c}'_2)'$$

① Thus, the proof is **equivalent** to demonstrating that $\mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1$ is a decreasing function of x_H and an increasing function of x_L , $\forall x_H, x_L \in (0, 1)$ since $\mathbf{M}_i$ doesn't involve x_H and x_L .

Note that

$$\begin{aligned} \mathbf{R}_{\mathcal{P}_s} &= \begin{pmatrix} n\pi_H \mathbf{T}' \Sigma_i^{-1} \mathbf{T} & n\pi_H x_H \mathbf{T}' \Sigma_i^{-1} \mathbf{t} \\ n\pi_H x_H \mathbf{t}' \Sigma_i^{-1} \mathbf{T} & n\pi_H x_H^2 \mathbf{t}' \Sigma_i^{-1} \mathbf{t} \end{pmatrix} + \begin{pmatrix} n\pi_L \mathbf{T}' \Sigma_i^{-1} \mathbf{T} & n\pi_L x_L \mathbf{T}' \Sigma_i^{-1} \mathbf{t} \\ n\pi_L x_L \mathbf{t}' \Sigma_i^{-1} \mathbf{T} & n\pi_L x_L^2 \mathbf{t}' \Sigma_i^{-1} \mathbf{t} \end{pmatrix} \\ &= n \begin{pmatrix} \mathbf{T}' \Sigma_i^{-1} \mathbf{T} & (\pi_H x_H + \pi_L x_L) \mathbf{T}' \Sigma_i^{-1} \mathbf{t} \\ (\pi_H x_H + \pi_L x_L) \mathbf{t}' \Sigma_i^{-1} \mathbf{T} & (\pi_H x_H^2 + \pi_L x_L^2) \mathbf{t}' \Sigma_i^{-1} \mathbf{t} \end{pmatrix} \\ &\stackrel{\text{def}}{=} n \begin{pmatrix} \mathbf{R}_{11} & (\pi_H x_H + \pi_L x_L) \mathbf{r}_{12} \\ (\pi_H x_H + \pi_L x_L) \mathbf{r}'_{12} & (\pi_H x_H^2 + \pi_L x_L^2) r_{22} \end{pmatrix}, \end{aligned}$$

where $\mathbf{R}_{11}$ is a 2×2 constant matrix, $\mathbf{r}_{12}$ is a 2×1 constant vector, and r_{22} is a scalar.

Then, according to the Equation (2.51) on page 23 of (Rencher and Schaalje, 2008), we have

$$\begin{aligned} \mathbf{R}_{\mathcal{P}_s}^{-1} &= \frac{1}{n} \begin{pmatrix} \mathbf{R}_{11} & (\pi_H x_H + \pi_L x_L) \mathbf{r}_{12} \\ (\pi_H x_H + \pi_L x_L) \mathbf{r}'_{12} & (\pi_H x_H^2 + \pi_L x_L^2) r_{22} \end{pmatrix}^{-1} \\ &= \frac{1}{nf} \begin{pmatrix} f \mathbf{R}_{11}^{-1} + (\pi_H x_H + \pi_L x_L)^2 \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} & -(\pi_H x_H + \pi_L x_L) \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \\ -(\pi_H x_H + \pi_L x_L) \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} & 1 \end{pmatrix}, \end{aligned}$$

where $f = (\pi_H x_H^2 + \pi_L x_L^2) r_{22} - (\pi_H x_H + \pi_L x_L)^2 \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \mathbf{r}_{12}$.

Note that $\frac{\partial t_p}{\partial \eta} = 0$ under the use condition, we further partition $\mathbf{c}_1$ in the form $\mathbf{c}_1 = \left(\frac{\partial t_p}{\partial \beta_0}, \frac{\partial t_p}{\partial \beta_1}, 0 \right)' \stackrel{\text{def}}{=} (\mathbf{c}'_{11}, 0)'$.

Then,

$$\begin{aligned} \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1 &= \frac{1}{nf} (\mathbf{c}'_{11}, 0)' \begin{pmatrix} f \mathbf{R}_{11}^{-1} + (\pi_H x_H + \pi_L x_L)^2 \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} & -(\pi_H x_H + \pi_L x_L) \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \\ -(\pi_H x_H + \pi_L x_L) \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{c}_{11} \\ 0 \end{pmatrix} \\ &= \frac{1}{nf} \mathbf{c}'_{11} \left[f \mathbf{R}_{11}^{-1} + (\pi_H x_H + \pi_L x_L)^2 \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \right] \mathbf{c}_{11} \\ &= \frac{1}{n} \mathbf{c}'_{11} \left[\mathbf{R}_{11}^{-1} + \frac{(\pi_H x_H + \pi_L x_L)^2}{f} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \right] \mathbf{c}_{11}. \end{aligned}$$

② Thus, the proof is further **equivalent** to demonstrating that $\frac{\partial \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1}{\partial x_H} < 0$ and $\frac{\partial \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1}{\partial x_L} > 0$, $\forall x_H, x_L \in (0, 1)$.

Note that

$$\begin{aligned} \frac{\partial \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1}{\partial x_H} &= \frac{1}{n} \mathbf{c}'_{11} \frac{\partial \left[\mathbf{R}_{11}^{-1} + \frac{(\pi_H x_H + \pi_L x_L)^2}{f} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \right]}{\partial x_H} \mathbf{c}_{11} \\ &= \frac{1}{n} \mathbf{c}'_{11} \left[\frac{\partial \left(\frac{(\pi_H x_H + \pi_L x_L)^2}{f} \right)}{\partial x_H} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \right] \mathbf{c}_{11}. \end{aligned}$$

Similarly,

$$\frac{\partial \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1}{\partial x_L} = \frac{1}{n} \mathbf{c}'_{11} \left[\frac{\partial \left(\frac{(\pi_H x_H + \pi_L x_L)^2}{f} \right)}{\partial x_L} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \right] \mathbf{c}_{11},$$

and we let

$$\begin{aligned} \frac{(\pi_H x_H + \pi_L x_L)^2}{f} &= \frac{(\pi_H x_H + \pi_L x_L)^2}{(\pi_H x_H^2 + \pi_L x_L^2) r_{22} - (\pi_H x_H + \pi_L x_L)^2 \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \mathbf{r}_{12}} \\ &= \frac{(\pi_H x_H^2 + \pi_L x_L^2) r_{22}}{(\pi_H x_H + \pi_L x_L)^2} - \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \\ &\stackrel{\text{def}}{=} \frac{1}{g}. \end{aligned}$$

③ Thus, the proof is further **equivalent** to demonstrating that $\frac{\partial \frac{1}{g}}{\partial x_H} < 0$ and $\frac{\partial \frac{1}{g}}{\partial x_L} > 0$, $\forall x_H, x_L \in (0, 1)$ since $\mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1}$ is a positive definite (pd) matrix.

Note that

$$\begin{aligned} \frac{\partial \frac{1}{g}}{\partial x_H} &= \frac{-2r_{22}\pi_H x_H (\pi_H x_H + \pi_L x_L)^2 - 2r_{22}\pi_H (\pi_H x_H^2 + \pi_L x_L^2) (\pi_H x_H + \pi_L x_L)}{(\pi_H x_H + \pi_L x_L)^4} \\ &= -\frac{x_H (\pi_H x_H + \pi_L x_L) - (\pi_H x_H^2 + \pi_L x_L^2)}{(\pi_H x_H + \pi_L x_L)^3 g^2 / (2r_{22}\pi_H)} \\ &= -\frac{\pi_L x_L (x_H - x_L)}{(\pi_H x_H + \pi_L x_L)^3 g^2 / (2r_{22}\pi_H)} < 0. \end{aligned}$$

Similarly,

$$\frac{\partial \frac{1}{g}}{\partial x_L} = -\frac{\pi_H x_H (x_L - x_H)}{(\pi_H x_H + \pi_L x_L)^3 g^2 / (2r_{22}\pi_L)} > 0. \quad \blacksquare$$

F. Proof of Corollary 1

Note that the two-level degenerate plan

$$\mathcal{P}_d = \begin{pmatrix} \xi_H & \xi_L \\ \pi_H & \pi_L \end{pmatrix}$$

defined in Corollary 1 is **equivalent** to a “single-variable” case with test conditions setting at the lowest and highest stress levels by denoting

$$\mathcal{P}_{d \rightarrow s} = \begin{pmatrix} 1 & \frac{a^*}{\eta_1 + s\eta_2} \\ \pi_H & \pi_L \end{pmatrix}, \text{ for } 0 \leq s \leq 1,$$

and

$$\mathcal{P}_{d \rightarrow s} = \begin{pmatrix} 1 & \frac{a^*}{\eta_1 / s + \eta_2} \\ \pi_H & \pi_L \end{pmatrix}, \text{ for } 1 < s < \infty,$$

where $1/s \stackrel{\text{set}}{=} 0$ when s approaches ∞ .

Then, if denoting both $\frac{a^*}{\eta_1 + s\eta_2}$ and $\frac{a^*}{\eta_1 / s + \eta_2}$ by x_L , the proof of Corollary 1 is **equivalent** to demonstrating that $\text{AVar}(\hat{t}_p)$ is an increasing function of x_L , $0 < x_L < 1$ in general, given any fixed value of π_H (π_L) $\in (0, 1)$. This has already been proved by Appendix E. $\blacksquare$

G. Proof of Theorem 1

Without loss generality, we assume that $s = 1$. Following Appendix F, denote the degenerate (after transforming to the “single-variable” case) and nondegenerate plans by

$$\mathcal{P}_{d \rightarrow s} = \begin{pmatrix} 1 & \frac{a^*}{\eta_1 + \eta_2} \\ \pi_H & \pi_L \end{pmatrix} \text{ and } \mathcal{P}_n = \begin{pmatrix} \xi_H & \xi_{L1} & \xi_{L2} \\ \pi_H & \pi_{L1} & \pi_{L2} \end{pmatrix} = \begin{pmatrix} 1 & x_{1L1} & x_{1L2} \\ 1 & x_{2L1} & x_{2L2} \\ \pi_H & \pi_{L1} & \pi_{L2} \end{pmatrix},$$

where $\pi_H + \pi_L = 1$, $\pi_L = \pi_{L1} + \pi_{L2}$, $\eta_1 x_{1L1} + \eta_2 x_{2L1} = a^*$, and $\eta_1 x_{1L2} + \eta_2 x_{2L2} = a^*$.

① Firstly, we derive $\text{AVar}^n(\hat{t}_p)$ as follows.

Similar to Appendix E, we partition X_i^H , X_i^{L1} , X_i^{L2} , and $\mathbf{c}$ in the form

$$\begin{aligned} X_i^H &= \left(\begin{array}{cc|cc} 1 & \tau_{i1} & \tau_{i1} & \tau_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \tau_{im_i} & \tau_{im_i} & \tau_{im_i} \end{array} \right) \stackrel{\text{def}}{=} (\mathbf{T} \mid \mathbf{t}\mathbf{j}'), \\ X_i^{L1} &= \left(\begin{array}{cc|cc} 1 & \tau_{i1} & x_{1L1}\tau_{i1} & x_{2L1}\tau_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \tau_{im_i} & x_{1L1}\tau_{im_i} & x_{2L1}\tau_{im_i} \end{array} \right) \stackrel{\text{def}}{=} (\mathbf{T} \mid \mathbf{t}\xi'_{L1}), \\ X_i^{L2} &= \left(\begin{array}{cc|cc} 1 & \tau_{i1} & x_{1L2}\tau_{i1} & x_{2L2}\tau_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \tau_{im_i} & x_{1L2}\tau_{im_i} & x_{2L2}\tau_{im_i} \end{array} \right) \stackrel{\text{def}}{=} (\mathbf{T} \mid \mathbf{t}\xi'_{L2}). \end{aligned}$$

Then, $\mathbf{R}_{\mathcal{P}_n}$ is given by

$$\begin{aligned} \mathbf{R}_{\mathcal{P}_n} &= n\pi_H X_i^{H'} \Sigma_i^{-1} X_i^H + n\pi_{L1} X_i^{L1'} \Sigma_i^{-1} X_i^{L1} + n\pi_{L2} X_i^{L2'} \Sigma_i^{-1} X_i^{L2} \\ &= n \left[\pi_H \begin{pmatrix} \mathbf{T}' \Sigma_i^{-1} \mathbf{T} & \mathbf{T}' \Sigma_i^{-1} \mathbf{t}\mathbf{j}' \\ \mathbf{j}\mathbf{t}' \Sigma_i^{-1} \mathbf{T} & \mathbf{j}\mathbf{t}' \Sigma_i^{-1} \mathbf{t}\mathbf{j}' \end{pmatrix} + \pi_{L1} \begin{pmatrix} \mathbf{T}' \Sigma_i^{-1} \mathbf{T} & \mathbf{T}' \Sigma_i^{-1} \mathbf{t}\xi'_{L1} \\ \xi_{L1} \mathbf{t}' \Sigma_i^{-1} \mathbf{T} & \xi_{L1} \mathbf{t}' \Sigma_i^{-1} \mathbf{t}\xi'_{L1} \end{pmatrix} \right. \\ &\quad \left. + \pi_{L2} \begin{pmatrix} \mathbf{T}' \Sigma_i^{-1} \mathbf{T} & \mathbf{T}' \Sigma_i^{-1} \mathbf{t}\xi'_{L2} \\ \xi_{L2} \mathbf{t}' \Sigma_i^{-1} \mathbf{T} & \xi_{L2} \mathbf{t}' \Sigma_i^{-1} \mathbf{t}\xi'_{L2} \end{pmatrix} \right] \\ &= n \begin{pmatrix} \mathbf{R}_{11} & \pi_H \mathbf{r}_{12} \mathbf{j}' + \pi_{L1} \mathbf{r}_{12} \xi'_{L1} + \pi_{L2} \mathbf{r}_{12} \xi'_{L2} \\ \pi_H \mathbf{j} \mathbf{r}'_{12} + \pi_{L1} \xi_{L1} \mathbf{r}'_{12} + \pi_{L2} \xi_{L2} \mathbf{r}'_{12} & \pi_H \mathbf{r}_{22} \mathbf{j}' + \pi_{L1} \mathbf{r}_{22} \xi'_{L1} + \pi_{L2} \mathbf{r}_{22} \xi'_{L2} \end{pmatrix} \\ &= n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12}(\pi_H \mathbf{j}' + \pi_{L1} \xi'_{L1} + \pi_{L2} \xi'_{L2}) \\ (\pi_H \mathbf{j} + \pi_{L1} \xi_{L1} + \pi_{L2} \xi_{L2}) \mathbf{r}'_{12} & (\pi_H \mathbf{J} + \pi_{L1} \xi_{L1} \xi'_{L1} + \pi_{L2} \xi_{L2} \xi'_{L2}) \mathbf{r}_{22} \end{pmatrix} \\ &\stackrel{\text{def}}{=} n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12} \mathbf{u}'_n \\ \mathbf{u}_n \mathbf{r}'_{12} & \mathbf{U}_n \mathbf{r}_{22} \end{pmatrix}, \end{aligned}$$

where $\mathbf{j}$ is a 2×1 vector with ones and $\mathbf{J}$ is a 2×2 matrix with ones.

Thus,

$$\begin{aligned} \text{AVar}^n(\hat{t}_p) &= \mathbf{c}' \begin{pmatrix} \mathbf{R}_{\mathcal{P}_n}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}_n}^{-1} \end{pmatrix} \mathbf{c} \\ &= \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_n}^{-1} \mathbf{c}_1 + \mathbf{c}'_2 \mathbf{M}_{\mathcal{P}_n}^{-1} \mathbf{c}_2 \\ &= \mathbf{c}'_{11} \mathbf{A}_{\mathcal{P}_n}^{-1} \mathbf{c}_{11} + \mathbf{c}'_2 \mathbf{M}_{\mathcal{P}_n}^{-1} \mathbf{c}_2, \end{aligned}$$

where $\mathbf{c} = \left(\frac{\partial t_p}{\partial \beta_0}, \frac{\partial t_p}{\partial \beta_1}, \frac{\partial t_p}{\partial \eta_1}, \frac{\partial t_p}{\partial \eta_2}, \frac{\partial t_p}{\partial \sigma_0}, \dots, \frac{\partial t_p}{\partial \sigma} \right)' \stackrel{\text{def}}{=} (\mathbf{c}'_1, \mathbf{c}'_2)' \stackrel{\text{def}}{=} (\mathbf{c}'_{11}, \mathbf{0}', \mathbf{c}'_2)'$ and

$$\mathbf{A}_{\mathcal{P}_n} = \mathbf{R}_{11} - \frac{1}{r_{22}} \mathbf{r}_{12} \mathbf{u}'_n \mathbf{U}_n^{-1} \mathbf{u}_n \mathbf{r}'_{12}.$$

② Secondly, we derive $\text{AVar}^d(\hat{t}_p)$ as follows.

According to Appendix E, we denote

$$\begin{aligned} \mathbf{R}_{\mathcal{P}_{d \rightarrow s}} &= n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12}(\pi_H + \pi_L \frac{a^*}{\eta_1 + \eta_2}) \\ (\pi_H + \pi_L \frac{a^*}{\eta_1 + \eta_2}) \mathbf{r}'_{12} & (\pi_H + \pi_L (\frac{a^*}{\eta_1 + \eta_2})^2) r_{22} \end{pmatrix} \\ &\stackrel{\text{def}}{=} n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12} u_d \\ u_d \mathbf{r}'_{12} & U_d r_{22} \end{pmatrix} \\ &= n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12} \frac{u'_n \eta}{(\eta_1 + \eta_2)} \\ \frac{u'_n \eta}{(\eta_1 + \eta_2)} \mathbf{r}'_{12} & \frac{\eta' \mathbf{U}_n \eta}{(\eta_1 + \eta_2)^2} r_{22} \end{pmatrix} \end{aligned}$$

where

$$u_d = \pi_H + \pi_L \frac{a^*}{\eta_1 + \eta_2} = \frac{u'_n \eta}{(\eta_1 + \eta_2)}$$

and

$$U_d = \pi_H + \pi_L \left(\frac{a^*}{\eta_1 + \eta_2} \right)^2 = \frac{\eta' \mathbf{U}_n \eta}{(\eta_1 + \eta_2)^2}$$

given that $\boldsymbol{\eta} = (\eta_1, \eta_2)'$.

Thus,

$$\begin{aligned} \text{AVar}^d(\hat{t}_p) &= \mathbf{c}' \begin{pmatrix} \mathbf{R}_{\mathcal{P}_{d \rightarrow s}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\mathcal{P}_{d \rightarrow s}}^{-1} \end{pmatrix} \mathbf{c} \\ &= \mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_{d \rightarrow s}}^{-1} \mathbf{c}_1 + \mathbf{c}'_2 \mathbf{M}_{\mathcal{P}_{d \rightarrow s}}^{-1} \mathbf{c}_2 \\ &= \mathbf{c}'_{11} \mathbf{A}_{\mathcal{P}_{d \rightarrow s}}^{-1} \mathbf{c}_{11} + \mathbf{c}'_2 \mathbf{M}_{\mathcal{P}_{d \rightarrow s}}^{-1} \mathbf{c}_2, \end{aligned}$$

where

$$\begin{aligned} \mathbf{A}_{\mathcal{P}_{d \rightarrow s}} &= \mathbf{R}_{11} - \frac{1}{r_{22}} \frac{u_d^2}{U_d} \mathbf{r}_{12} \mathbf{r}'_{12} \\ &= \mathbf{R}_{11} - \frac{1}{r_{22}} \mathbf{r}_{12} \mathbf{u}'_n \frac{\eta \eta'}{\eta' \mathbf{U}_n \eta} \mathbf{u}_n \mathbf{r}'_{12}. \end{aligned}$$

③ Thirdly, we provide a proof for part a).

First of all,

$$\begin{aligned} \text{AVar}^n(\hat{t}_p) - \text{AVar}^d(\hat{t}_p) \\ = \mathbf{c}'_{11} (\mathbf{A}_{\mathcal{P}_n}^{-1} - \mathbf{A}_{\mathcal{P}_{d \rightarrow s}}^{-1}) \mathbf{c}_{11}, \end{aligned}$$

since $\mathbf{M}_{\mathcal{P}_n} = \mathbf{M}_{\mathcal{P}_{d \rightarrow s}}$.

When $\pi_{L1} \xi_{L1} + \pi_{L2} \xi_{L2} = \pi_L \xi_L$, it is obvious that $u_d = \mathbf{u}'_n \mathbf{j}$ and $U_d^{-1} = \mathbf{j}' \mathbf{U}_n^{-1} \mathbf{j}$, which implies that $\mathbf{A}_{\mathcal{P}_n} = \mathbf{A}_{\mathcal{P}_{d \rightarrow s}}$. Thus, $\text{AVar}^n(\hat{t}_p) = \text{AVar}^d(\hat{t}_p)$.

To show that $\text{AVar}^n(\hat{t}_p) \geq \text{AVar}^d(\hat{t}_p)$, it is **equivalent** to demonstrating (see (Rao et al., 1973), problem 9 on page 70) that $\mathbf{A}_{\mathcal{P}_{d \rightarrow s}} - \mathbf{A}_{\mathcal{P}_n}$ is a nonnegative definite matrix. According to the existing results above, we have

$$\begin{aligned} \mathbf{A}_{\mathcal{P}_{d \rightarrow s}} - \mathbf{A}_{\mathcal{P}_n} &= \frac{1}{r_{22}} \mathbf{r}_{12} \mathbf{u}'_n \left[\mathbf{U}_n^{-1} - \frac{\eta \eta'}{\eta' \mathbf{U}_n \eta} \right] \mathbf{u}_n \mathbf{r}'_{12} \\ &= \frac{1}{r_{22}} \mathbf{r}_{12} \mathbf{u}'_n \mathbf{L}' [\mathbf{I} - \mathbf{v} \mathbf{v}'] \mathbf{L} \mathbf{u}_n \mathbf{r}'_{12}, \end{aligned}$$

where $\mathbf{I}$ is an identity matrix, $\mathbf{L}' \mathbf{L} = \mathbf{U}_n^{-1}$ is the Cholesky decomposition of $\mathbf{U}_n^{-1}$, and $\mathbf{v} = \frac{(\mathbf{L}')^{-1} \eta}{\sqrt{\eta' (\mathbf{L}')^{-1} \eta}}$ is a unit length vector.

Then, it is further **equivalent** to showing that $\mathbf{I} - \mathbf{v} \mathbf{v}'$ is a nonnegative definite matrix. By Cauchy-Schwartz inequality, given any arbitrary vector $\mathbf{a}$, we have

$$\begin{aligned} &\mathbf{a}' (\mathbf{I} - \mathbf{v} \mathbf{v}') \mathbf{a} \\ &= \mathbf{a}' \mathbf{a} - (\mathbf{a}' \mathbf{v})^2 \\ &\geq \mathbf{a}' \mathbf{a} - (\mathbf{a}' \mathbf{a})(\mathbf{v}' \mathbf{v}) = 0. \end{aligned} \quad \blacksquare$$

④ Lastly, we provide a proof for part b).

First of all, according to Appendix D, we have

$$\det(\mathbf{I}(\boldsymbol{\theta}; \mathcal{P}_n)) = \det(\mathbf{R}_{\mathcal{P}_n}) \det(\mathbf{M}_{\mathcal{P}_n}),$$

where $\mathbf{M}_{\mathcal{P}_n}$ doesn't involve stress variables. Thus, the proof is **equivalent** to showing that $\det(\mathbf{R}_{\mathcal{P}_n})$ is maximized when the split is done as much as possible.

Based on the results above, in general, the split is done such that $\pi_{L1} \xi_{L1} + \pi_{L2} \xi_{L2} = \pi_L \xi_L$, i.e.

$$\pi_{L1} \begin{pmatrix} x_{1L1} \\ x_{2L1} \end{pmatrix} + \pi_{L2} \begin{pmatrix} x_{1L2} \\ x_{2L2} \end{pmatrix} = \pi_L \begin{pmatrix} x_{1L} \\ x_{2L} \end{pmatrix},$$

where $\pi_L = \pi_{L1} + \pi_{L2}$, $\eta_1 x_{1L1} + \eta_2 x_{2L1} = a^*$, $\eta_1 x_{1L2} + \eta_2 x_{2L2} = a^*$, and $\eta_1 x_{1L} + \eta_2 x_{2L} = a^*$.

Without loss of generality, we assume that $d_1 = x_{1L1} - x_{1L} > 0$ and $d_2 = x_{1L} - x_{1L2} > 0$. Denote the slope of the minimum-stress level line $-\frac{\eta_1}{\eta_2}$ by m and $\mathbf{m} = (1, m)'$, we have

$$\begin{aligned}\xi_{L1} &= \begin{pmatrix} x_{1L1} \\ x_{2L1} \end{pmatrix} = \begin{pmatrix} d_1 + x_{1L} \\ md_1 + x_{2L} \end{pmatrix} = \xi_L + d_1 \mathbf{m}, \\ \xi_{L2} &= \begin{pmatrix} x_{1L2} \\ x_{2L2} \end{pmatrix} = \begin{pmatrix} x_{1L} - d_2 \\ x_{2L} - md_2 \end{pmatrix} = \xi_L - d_2 \mathbf{m},\end{aligned}$$

and $(\pi_{L1}d_1 - \pi_{L2}d_2)\mathbf{m} = \mathbf{0}$. Then, $\mathbf{R}_{\mathcal{P}_n}$ can be represented using d_1 and d_2 as

$$\begin{aligned}\mathbf{R}_{\mathcal{P}_n} &= n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12}(\pi_H \mathbf{j}' + \pi_{L1} \xi'_{L1} + \pi_{L2} \xi'_{L2}) \\ (\pi_H \mathbf{j} + \pi_{L1} \xi_{L1} + \pi_{L2} \xi_{L2}) \mathbf{r}'_{12} & (\pi_H \mathbf{J} + \pi_{L1} \xi_{L1} \xi'_{L1} + \pi_{L2} \xi_{L2} \xi'_{L2}) \mathbf{r}_{22} \end{pmatrix} \\ &= n \begin{pmatrix} \mathbf{R}_{11} & \mathbf{r}_{12}(\pi_H \mathbf{j}' + \pi_L \xi'_L) \\ (\pi_H \mathbf{j} + \pi_L \xi_L) \mathbf{r}'_{12} & [\pi_H \mathbf{J} + \pi_L \xi_L \xi'_L + (\pi_{L1}d_1^2 + \pi_{L2}d_2^2)\mathbf{m}\mathbf{m}'] \mathbf{r}_{22} \end{pmatrix}.\end{aligned}$$

and

$$\frac{\partial \mathbf{R}_{\mathcal{P}_n}}{\partial d_1} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 2r_{22}\pi_{L1}d_1\mathbf{m}\mathbf{m}' \end{pmatrix}$$

Then, the proof is further **equivalent** to showing that $\det(\mathbf{R}_{\mathcal{P}_n})$ is an increasing function of both d_1 and d_2 . We prove this for d_1 ; the proof for d_2 is similar.

By partitioning $\mathbf{R}_{\mathcal{P}_n}$ into a 2×2 matrix, we denote its corresponding right lower partitioned inverse by $\mathbf{R}^{22}$. Note that $\mathbf{R}_{\mathcal{P}_n}$ and $\mathbf{R}^{22}$ are both pd matrices. By Equation (8.5) on page 309 of (Harville, 1998), the derivative of the determinant of $\mathbf{R}_{\mathcal{P}_n}$ is given by

$$\begin{aligned}\frac{\partial \det(\mathbf{R}_{\mathcal{P}_n})}{\partial d_1} &= \det(\mathbf{R}_{\mathcal{P}_n}) \times \text{tr} \left[\mathbf{R}_{\mathcal{P}_n}^{-1} \frac{\partial \mathbf{R}_{\mathcal{P}_n}}{\partial d_1} \right] \\ &= \det(\mathbf{R}_{\mathcal{P}_n}) \times \text{tr} [2r_{22}\pi_{L1}d_1 \mathbf{R}^{22} \mathbf{m}\mathbf{m}'] \\ &= 2r_{22}\pi_{L1}d_1 \det(\mathbf{R}_{\mathcal{P}_n}) \mathbf{m}' \mathbf{R}^{22} \mathbf{m} > 0.\end{aligned}$$

Thus, $\det(\mathbf{R}_{\mathcal{P}_n})$ is an increasing function of both d_1 and d_2 ; $\det(\mathbf{R}_{\mathcal{P}_n})$ is maximized by choosing d_1 and d_2 as large as possible. It implies that the split should be done such that the two lower levels reach the boundaries of $\mathbb{D}$. ■

H. Proof of Theorem 2

Without loss of generality, we assume that $x_U = 0$ since one can treat $x'_H = x_H - x_U$ and $x'_L = x_L - x_U$ as the transformed higher and lower test condition, respectively, if $x_U \neq 0$.

Following Appendix E, given fixed values of x_H and x_L , minimizing $\text{AVar}(\hat{t}_p)$ is **equivalent** to minimizing

$$\mathbf{c}'_1 \mathbf{R}_{\mathcal{P}_s}^{-1} \mathbf{c}_1 = \frac{1}{n} \mathbf{c}'_{11} \left[\mathbf{R}_{11}^{-1} + \frac{(\pi_H x_H + \pi_L x_L)^2}{f} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \right] \mathbf{c}_{11},$$

where $f = (\pi_H x_H^2 + \pi_L x_L^2)r_{22} - (\pi_H x_H + \pi_L x_L)^2 \mathbf{r}_{12}^{-1} \mathbf{r}_{12}$. Since $\mathbf{c}'_{11} \mathbf{R}_{11}^{-1} \mathbf{c}_{11}$ and $\mathbf{c}'_{11} \mathbf{R}_{11}^{-1} \mathbf{r}_{12} \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \mathbf{c}_{11}$ do not involve the decision variables and are also positive, minimizing the term above is further **equivalent** to finding the solutions of

$$\frac{\partial \frac{(\pi_H x_H + \pi_L x_L)^2}{f}}{\partial \pi_H} = 0 \text{ and } \frac{\partial \frac{(\pi_H x_H + \pi_L x_L)^2}{f}}{\partial \pi_L} = 0.$$

We demonstrate how to obtain the first solution. The second one is found similarly or can be obtained simply by $\pi_L =$

$1 - \pi_H$. By denoting $g = \frac{(\pi_H x_H^2 + \pi_L x_L^2)r_{22}}{(\pi_H x_H + \pi_L x_L)^2} - \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \mathbf{r}_{12}$, the equation of partial derivative becomes

$$\begin{aligned}\frac{\partial \frac{(\pi_H x_H + \pi_L x_L)^2}{f}}{\partial \pi_H} &= \frac{\partial \frac{1}{\frac{(\pi_H x_H^2 + \pi_L x_L^2)r_{22}}{(\pi_H x_H + \pi_L x_L)^2} - \mathbf{r}'_{12} \mathbf{R}_{11}^{-1} \mathbf{r}_{12}}}{\partial \pi_H} \\ &= - \frac{(x_H^2 - x_L^2)(\pi_H x_H + \pi_L x_L) - 2(\pi_H x_H^2 + \pi_L x_L^2)(x_H - x_L)}{g^2(\pi_H x_H + \pi_L x_L)^3 / r_{22}} = 0.\end{aligned}$$

Then, it is **equivalent** to solving the followings:

$$\begin{aligned}(x_H^2 - x_L^2)(\pi_H x_H + \pi_L x_L) - 2(\pi_H x_H^2 + \pi_L x_L^2)(x_H - x_L) &= 0 \\ (x_H + x_L)(\pi_H x_H - \pi_H x_L + x_L) - 2\pi_H(x_H^2 - x_L^2) - 2x_L^2 &= 0 \\ \pi_H(x_H^2 - x_L^2) + x_H x_L + x_L^2 - 2\pi_H(x_H^2 - x_L^2) - 2x_L^2 &= 0 \\ \pi_H(x_H^2 - x_L^2) &= x_L(x_H - x_L) \\ \pi_H &= \frac{x_L}{x_H + x_L}.\end{aligned}$$

One can easily verify that the second derivative is greater than 0, so we omit it here. Thus, we complete the proof. ■

Guanqi Fang recently graduated from the Ph.D. program of Industrial Engineering (IE) at Arizona State University (ASU), where he obtained his concurrent Master degree of Statistics. Previously, he obtained his bachelor and master degree in IE from Hunan University, China and North Carolina State University, respectively. Prior to attending ASU, he worked as a Quality Engineer at Foxconn Technology Group. He will join School of Statistics and Mathematics at Zhejiang Gongshang University, China. His research interests include statistical modeling and data analysis of industrial systems along with data science and machine learning.

Rong Pan is an Associate Professor in the School of Computing, Informatics, and Decision Systems Engineering at Arizona State University. He received his Ph.D. in Industrial Engineering from Penn State University in 2002. His research interests include quality and reliability engineering, design of experiments, time series analysis, and statistical learning theory. He is a senior member of IEEE, ASQ and IISE, and a lifetime member of SRE.

John Stufken holds a Kandidaats and Doctoraal degree in Mathematics from the Radboud University of Nijmegen. He holds a PhD in Mathematics, majoring in Statistics, from the University of Illinois at Chicago. Since 2019 he is a Bank of America Excellence Professor and Director of Informatics and Analytics at UNC Greensboro. Prior to this he was Charles Wexler Professor of Statistics at Arizona State University, Head of the Department of Statistics at the University of Georgia, Program Director for Statistics at the National Science Foundation, and faculty member in the Department of Statistics at Iowa State University. His primary area of research interest is design of experiments.