Optimal Designs for Accelerated Degradation Testing

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von MSc. Helmi Shat

geb. am 15.08.1990 in Khan-Yunis, Palestine

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Rainer Schwabe
             Prof. Dr. Norbert Gaffke
             Prof. Dr. Ewaryst Rafajlowicz

eingereicht am: 14.07.2021

Verteidigung am: 26.01.2022

Abstract

Accelerated degradation tests are intensively used in engineering applications to provide an accurate estimation of lifetime properties of highly reliable products within a relatively short testing time. In this regard, the corresponding data from particular accelerated tests at high levels of stress (e.g., temperature, voltage, or vibration) are extrapolated, through a physically reasonable statistical model, to obtain estimates of lifetime quantiles at normal use conditions of these models. In this thesis we investigate the problem of obtaining optimal designs for accelerated degradation tests under different degradation path models.

After the first two introductory chapters we characterize in the third chapter an analytical approach for obtaining optimal designs for repeated measures accelerated degradation tests with multiple stress variables when the observational times are either fixed in advance or are also to be optimized. Subsequently, we consider the particular case when the degradation path is assumed to follow a linear mixed effects model which is quite common in settings when repeated measures are made.

In the fourth chapter we extend the degradation model of chapter 3 in order to present optimal experimental designs for accelerated degradation tests with repeated measures and competing failure modes that correspond to multiple response components. The corresponding marginal degradation paths are expressed using linear mixed effects models.

The gamma process is a natural model for describing the degradation increments of the degradation path which exhibit a strictly increasing degradation pattern. In chapter 5, we derive first an optimal design for repeated measures accelerated degradation testings with a single failure mode that corresponds to a single response component. The univariate degradation process is expressed using a gamma model in terms of a generalized linear model to facilitate the derivation of an optimal design. Subsequently, we extend the univariate model and characterize optimal designs for accelerated degradation tests with bivariate degradation processes with independent marginal components. The first bivariate model includes two independent gamma processes as marginal degradation models. The second bivariate models is expressed by a gamma process along with a linear mixed effects model.

Finally, in chapter 6 we extend the approach presented in chapter 5 to develop optimal designs for accelerated degradation testing with multiple stress variables and multiple components. The marginal degradation paths are assumed to follow gamma process models, and copula-based dependence between marginal components are considered.

Zusammenfassung

Beschleunigte Degradationstests werden intensiv in technischen Anwendungen genutzt, um eine genaue Abschätzung der Lebensdauereigenschaften von hochzuverlässigen Produkten innerhalb eines relativ kurzen Testzeitraums zu erhalten. Dabei werden die entsprechenden Daten aus beschleunigten Tests bei hohen Belastungen (z. B. Temperatur, Spannung oder Vibration) durch ein physikalisch sinnvolles statistisches Modell extrapoliert, um Schätzungen der Lebensdauerquantile unter normalen Einsatzbedingungen zu erhalten. In dieser Arbeit untersuchen wir das Problem, optimale Designs für beschleunigte Degradationstests unter verschiedenen Degradationspfadmodellen zu erhalten.

Nach den ersten beiden einführenden Kapiteln beschreiben wir im dritten Kapitel einen analytischen Ansatz zur Gewinnung optimaler Designs für beschleunigte Degradationstests mit wiederholten Messungen und mehreren Belastungsvariablen, wenn die Beobachtungszeiten entweder im Voraus festgelegt sind oder ebenfalls optimiert werden sollen. Im Folgenden betrachten wir den besonderen Fall, dass der Degradationspfad einem linearen Mixed-Effects-Modell folgt, was in Einstellungen mit wiederholten Messungen recht häufig angenommen werden kann.

Im vierten Kapitel erweitern wir das Degradationsmodell aus Kapitel 3, um optimale Versuchspläne für beschleunigte Degradationstests mit wiederholten Messungen und konkurrierenden Schadensarten zu präsentieren, die mehreren Wirkungskomponenten entsprechen. Die entsprechenden marginalen Degradationspfade werden mit Hilfe linearer Mixed-Effects-Modelle dargestellt.

Der Gamma-Prozess ist ein natürliches Modell zur Schätzung der Degradationsinkremente über den Degradationspfad, die ein streng monoton zunehmendes Degradationsmuster aufweisen. In Kapitel 5 leiten wir zunächst ein optimales Design für beschleunigte Degradationstests mit wiederholten Messungen mit einem einzigen Fehlermodus her, der einer einzigen Wirkungskomponente entspricht. Der univariate Degradationsprozess wird durch ein Gamma-Modell beschrieben, wobei ein verallgemeinert lineares Modell eingeführt wird, um die Herleitung eines optimalen Designs zu erleichtern. Im Folgenden erweitern wir das univariate Modell und charakterisieren optimale Designs für beschleunigte Degradationstests mit bivariaten Degradationsprozessen mit unabhängigen Randkomponenten. Das erste bivariate Modell beinhaltet zwei unabhängige Gamma-Prozesse als marginale Degradationsmodelle. Das zweite bivariate Modell wird durch einen Gamma-Prozess zusammen durch ein lineares Mixed-Effects Modell dargestellt.

Schließlich erweitern wir in Kapitel 6 den in Kapitel 5 vorgestellten Ansatz, um optimale Designs für beschleunigte Degradationstests mit mehreren Stressvariablen

und mehreren Komponenten zu entwickeln. Es wird angenommen, dass die marginalen Degradationspfade Gamma-Prozessmodellen folgen, und es wird eine Copula-basierte Abhängigkeit zwischen den marginalen Komponenten zu Grunde gelegt.

Contents

Chapter 1

Overview

1.1   Introduction

Reliability testing is a common scientific approach to gain knowledge about the ability of the target system of maintaining the pre-designed performance level for a certain period of time. Nowadays, rapid advances in technology, development of highly sophisticated products, intense world-wide competition and increasing customer expectations force manufacturers to produce highly reliable products. (Kahneman and Tversky, 1986) state that the customers' experience of a product is greatly affected by its ability to function normally over time, and they are willing to pay more for the products which have demonstrated higher reliability. Accordingly, producers need to test the product reliability carefully before introducing it into the market, which is called reliability testing and analysis. Previously, reliability is defined as the probability that a system operates its intended function under certain conditions for a certain period of time. Consequently, reliability analysis depends on the data which are collected from a sample of testing units that belongs to a particular process or population under identical conditions. The normal life test is a straight-forward analysis method that holds the experiment in the designed usage condition, and collects the samples' life time to fit a life time distribution from the testing sample. However, considering this kind of experimental testing to estimate the failure-time distribution or long-term performance of high-reliable systems is particularly difficult and time consuming.

In many industrial fields, the need for highly reliable components and materials are commonly required for long-lasting performance. In particular, the extremely high reliability is essential in aerospace and aviation industries and also strongly required in automobile industry, semiconductor industry, electronic industry and many other fields. However, most of the highly reliable products are designed to operate without failure for years or even tens of years. Thus, few units will fail or degrade in a normal life test of practical length under normal use conditions. In addition, manufacturers cannot afford years of reliability tests before releasing a product into market. Hence, a more time-affordable reliability test plan is highly required. Accelerated testing which

consists of a high stress level (e.g., temperature, voltage, or vibration) that significantly shortens product life or accelerates the degradation process of the product is developed. One can then statistically fit an acceleration model and data from particular tests at high levels of stress are utilized to obtain estimates of lifetime quantiles at normal stress levels.

Accelerated life test (ALT) model is an example of these acceleration models which consists of a theoretical life distribution (such as Weibull, Log-normal and Exponential distribution). The parameters of these distributions are either functions of accelerating stress variables or unknown coefficients to be estimated from the experimental data. However, because of the high-reliable property of products, it is difficult to obtain enough failure time data during the pre-designed time period to satisfy the requirement of ALT. Additionally, many testing units hove to be subjected to a monotone degradation process over the testing period before they fail. Thus, little to no test failure data are available. To solve this problem, it is therefore necessary to increase the sampling capacity and extend the testing interval. When a certain stress level (such as temperature, humidity, voltage, or vibration), of the product is increased, the degradation process will be accelerated (Whitmore and Schenkelberg, 1997). Therefore, researchers have proposed the approach of utilizing product performance degradation data to assess the product's reliability.

This method of using stress that exceeds the normal use levels to accelerate the performance degradation of a product, collect degradation data, and conduct a reliability assessment under a normal stress level is called accelerated degradation test (ADT) see (Meeker, Escobar, and Lu, 1998). ADT is introduced in order to give estimations in relatively short periods of time about the life time and reliability of the system under study. ADT overcomes defects in those products for which ALT only records the failure time and neglects the specific process of the product failure or performance changes. Meanwhile, this method also makes up for a lack of ALT failure test data, thus greatly improving the assessment efficiency (Pang et al., 2020). It has value as an important engineering application for a fast and accurate completion of product reliability assessment, development, and approval (Hong and Ye, 2017). In ADT, the test units are exposed to higher stress levels for degradation than the normal use conditions in order to obtain degradation data in a considerably shorter testing time than ALT. Consequently, these data are extrapolated, through a physically reasonable statistical model, to obtain estimates of lifetime quantiles at normal use conditions.

1.1.1 Review of accelerated degradation testing

ADTs mainly depend on applying higher stress level during the testing process, expecting to hasten the degradation of a product and recognize the product failure time sooner

than under normal use conditions. Since the product life in regular condition is too long to observe, this information can then be extrapolated to estimate the life time properties of the product during the use conditions. This process is called accelerated modeling, which usually includes statistically fitting the degradation data, and extrapolating the fitted model to normal stress level (Zhang, 2013). The method of reliability and statistical analysis and experimental design has a long history, and is still rapidly developing. In general, the majority of these approaches are intended for interpolation, i.e. the estimation is intended to be within the response range. However, ADTs are different in the sense that they are utilized to estimate outside the range of stress variables range, which may result in many different problems and make the estimation method more complicated. Therefore, ADT has become a mainstream technique over the years and is now a research hotspot in the field of reliability , (Nelson, 2009) provided a detailed summary that discuss the purpose, statistical methodologies and validation of the accelerated testing. He presented many statistical models in details and defined the concept of ADT. (Limon, Yadav, and Liao, 2017) introduced a comprehensive review of critical methods for statistical inference and the optimal design of ADT plans. In another study, (Deng et al., 2007) made a simple comparison between ADT and ALT, and reviewed the application of ADT on the basis of the related background, a data processing method, a degradation model, and design and optimization. The authors also stated some fundamental problems that need to be resolved in the application of ADT, as well as its application prospects. (Nelson, 2005) provided a large number of references on statistical plans for ADT, which will support practitioners to choose a testing plan and will motivate researchers to introduce new plans and software.

1.1.2 Fundamentals of ADT and reliability analysis

To overcome the problems of a traditional reliability assessment method during the 1970s, (Gertsbackh, 1969) was the first to use performance degradation data to assess the reliability of a product, which opened up a new way to study the remaining useful life (RUL) prediction of high reliability long-lifetime products. (Shiomi and Yanagisawa, 1979) and (Carey and Tortorella, 1988) were among the first to consider accelerated degradation models based on their practical work with actual industrial problems. (Nelson, 1990) is the first researcher to present applications and statistical models for ADTs with an application of the Arrhenius relationship which is commonly utilized to standardize stress variables and provide a life-stress relationship model based on the Arrhenius equation. The use of the acceleration model depends usually on the type of stress; e.g., the Arrhenius and Eyring models are suitable for a temperature stress variable, the inverse power model is commonly used with nonthermal stresses. Later on, studies for the relationship between failure time and life testing stress level have been

discussed in some literatures; e.g., see (Nelson, 1990) and (Zhao and Elsayed, 2005). Further, (Meeker, Escobar, and Lu, 1998) discuss the reliability analysis and inference based on the estimation of failure time from observed degradation data. However, the main disadvantage of this method is underestimating the variance of the ADT process by ignoring the error in the degradation curve fitting process. Consequently, the efficiency of this method strongly depends on the selection of the degradation regression model. Later, (Boulanger and Escobar, 1994) address the problem of optimizing both the stress levels as well as the sample size for each stress level under a previously determined termination time. ADT is an extension of ALT in which the life estimates of the testing unit is not simply measured by failure time, but utilizes a degradation level instead. This model is generally applicable when the system under study has measurable degradation attributes. In such experimental models, the experimenters consistently observe the degradation level for each testing unit until it reaches a predetermined failure threshold. Another significant advantage of ADT compared with ALT is the ability of dealing with censoring data. In ADT model, even if a test unit is censored, it can still provide much information because we have the trend of its degradation and the information loss is less, see (Zhang, 2013). Following are the primary assumptions when conducting ADTs models:

1. Degradation is not reversible.

2. A model applies to a single or multiple degradation process, mechanism, or failure mode.

3. The nature of the failure at accelerated stress levels is the same as at the design or use stress levels.

Accelerated degradation data mainly have three main elements, namely, the test time, performance degradation and accelerated stress level (Pang et al., 2020). The reliability assessment based on accelerated degradation data mainly considers two aspects involving reliability modeling and a statistical analysis. The aim is to extrapolate the reliability of the product under the normal use conditions through the known accelerated degradation path and the failure time threshold. Firstly, based on an analysis of the product failure mechanism, the statistical degradation model is derived, and the parameters of the degradation model related to the acceleration stress are determined (i.e. the parameter values change with the change in the stress level), and an acceleration model of the relevant parameters is established. Subsequently, using the accelerated degradation data to identify whether the failure mechanism is consistent, the parameters of the model are estimated by subtracting the invalid data. Finally, the reliability assessment results of the product under the working conditions are extrapolated by combining them with the failure threshold of the performance degradation parameter.

Considering the technique of accelerating stress applied, ADT is categorized into

constant-stress, step-stress, progressive-stress and cyclic-stress approaches (Nelson, 2009). Constant-stress ADT (CSADT) has become the most widely used ADT method due to its ease of stress application, the availability of existing theoretical models, and convenient calculations. For example, (Liu et al., 2017) apply the Bayesian model averaging (BMA) method to CSADT to address the issue of model uncertainty. The CSADT method usually requires a longer testing interval at low stress levels to observe a system degradation. Thus, to handle this issue, step-stress ADT (SSADT) results in a quicker degradation than CSADT (Shen et al., 2017). With SSADT, the testing units is subject to a constant stress for a specific period of time, and the stress then increases to a higher level or decreases to a lower level until the test reaches the termination time or the number of failed units reaches a predetermined level. The difficulty of applying a parameter estimation and reliability extrapolation under normal operating conditions, as well as the possibility of introducing new fault modes, are the drawbacks of SSADT, see (Pang et al., 2020). Progressive-stress ADT (PSADT) is another type of ADT, in which the testing units are subjected to a continually increasing stress over the testing interval (Peng and Tseng, 2010). Finally, considering products which are subjected to cyclic stress during normal use conditions, the cyclic-stress ADT, is the most appropriate method, see (Luo et al., 2017). In this type of accelerated testing, the testing units undergo more intense repetitive cyclic stress, such as sinusoidal voltage or fatigue stress, for further details see (Nelson, 2009).

The selection of accelerating stress variables generally depends on the failure mode of the system under study as well as the primary stresses that generate the failure. For instance, mechanical systems usually fail through wear or corrosion, thereby accelerating the occurrence of faults from vibration, see (Pang et al., 2020). Further, random impacts and humidity are other types of stresses applied to various mechanical components, such as springs, shafts and bearings. For electronic systems, electrical current, voltage, humidity, temperature and vibration are common accelerating stress variables. In general, electrical current (Wang and Chu, 2012) and UV radiation (Wen et al., 2018) are among the most utilized accelerating stress variables. According to the actual use conditions and the failure mode of the testing unit, a combination of these stress variables may also be applied. In fact, the number of stress variables applied in ADT is another primary issue to be considered. The single-stress ADT is the most common and favorable method due to its analytical simplicity. However, in order to provide a more realistic simulation of the normal operating conditions and enhance the testing efficiency, single-stress ADT has been developed into multi-stress ADT, see (Sun et al., 2018) and (Tsai et al., 2015). Nevertheless, the multi-stress ADT method also has some obstacles, such as interaction effects between stress variables and the implementation of higher complexity. Meanwhile, with enhancing the data acquisition and analysis techniques of

the degradation process as well as the failure mode analysis capabilities, the test objects have been extended from single component-level systems, e.g. semiconductors (Ye, Hu, and Yu, 2019), to multiple-level systems, e.g. electrical connectors (Ye et al., 2014).

1.2 Description of the thesis

The problem of characterizing optimal designs for accelerated degradation testing has attracted increasing attention in the last few decades. This problem is mainly concerned with deriving some experimental design for obtaining some reliability estimates with minimum variance.

In this work we are aiming to derive analytical as well as algorithm-based optimal designs for accelerated degradation tests under several degradation models. For these models, we are dealing with regression functions that corresponds to Gamma process (GP) models which are intensively used for the analysis of reliability tests as well as general mixed models including linear mixed effects models (LMEM). Under different testing settings and assumptions throughout, we will consider regression models with multiple accelerating stress variables, multiple response components and multiple measurement times.

In particular, after introducing some overview and preliminaries in chapter 2, the third chapter develops locally *c*-optimal designs for accelerated degradation testing with multiple stress variables and multiple observation times. In accordance with (Shat and Schwabe, 2021), we handle in this chapter the particular case when the degradation path follows a linear mixed effects model which is a common approach when longitudinal data are considered.

On the basis of the univariate model introduced in chapter 3, the degradation model of the fourth chapter, as introduced in (Shat, 2021), is an extension to a multivariate degradation model for deriving *c*-optimal designs for accelerated degradation testing with multiple response components. Similar to chapter 3, the corresponding marginal degradation paths are again expressed by linear mixed effects models with repeated measures.

Based on (Shat and Schwabe, 2019), we introduce at the beginning of chapter 5 a *c*-optimal design for Gamma-based accelerated degradation testings with a single response component. Subsequently, we introduce an extension of the univariate model in order to develop optimal designs for bivariate degradation models with independent marginal components. We consider first a bivariate model that includes two independent gamma processes as marginal degradation models. On the other hand the second bivariate models is expressed using a gamma process along with a linear mixed effects model.

Subsequently, on the basis of (Shat and Gaffke, 2021), the work in chapter 6 extends the model presented in chapter 5 in order to characterize c- and D-optimal experimental designs for multivariate degradation models with dependent components. The marginal degradation paths are assumed to follow Gamma process models where Gaussian- as well as Frank-copula functions are introduced to characterize the dependence between marginal components.

Finally, concluding remarks as well as some outlook for future research are introduced in chapter 7.

It should be noted that notation throughout this work, except for some minor notational differences in chapter 6, has been unified in comparison to the corresponding articles mentioned for each chapter.

Chapter 2

Basic concepts

2.1   Statistical modelling of degradation process

In this section, we review the different statistical modeling methods of the system degradation process based on the following three types: failure lifetime, degradation distribution, and stochastic process. Under different accelerating stress levels, the failure lifetime follows the same distribution with different distribution parameters. These parameters are usually functions related to the stress level (Pang et al., 2020). By setting up an acceleration model for these parameters, the distribution parameters of the failure lifetime of the products under normal use conditions can be extrapolated, and thus an design of experiments and reliability analysis of the products can be carried out. In general, the method of fitting the degradation path can be divided into two types which are linear/nonlinear (Lu and Meeker, 1993) and exponential (Wang and Chu, 2012), and the other is the use of a neural network (Gebraeel and Lawley, 2008), time series (Wang et al., 2013), and other intelligent algorithms. Considering the first type of modelling, (Lu and Meeker, 1993) constructed a description method of the degradation path for a nonlinear comprehensive effect model, and utilized the Monte Carlo (MC) simulation method to obtain the failure lifetime data to assess the reliability of the system under study. (Wang and Chu, 2012) introduced an exponential function to fit the degradation path of a light emitting diode (LED) and performed an average lifetime assessment. On the other hand, (Gebraeel and Lawley, 2008) conducted reliability assessment based on a neural network, and constructed a degradation model based on the relationship between the bearing vibration signal and its life as a way to predict the bearing lifetime. Additionally, (Wang et al., 2013) utilized a time series regression analysis method, combined with a neural network in order to conduct reliability assessment for the contact resistance of aerospace electromagnetic relays. (Wu et al., 2009) utilized a least squares support vector machine to model the degradation path of CNC machine tools. The vector machine is trained through historical degradation data and the parameter values are estimated based on the real-time degradation data.

Due to the performance difference between different testing units, the degradation level of different units depends on the stress level applied as well as the observational time points where the parameters of the degradation level distribution are functions of the accelerating stress level and observation times. Hence, obtaining the parameter values of the degradation level distribution leads to constructing the relationship between the distribution parameters and time or stress which can be further used in the design of experiments and reliability analysis. For example, (Sun et al., 2004) introduced a Gauss-Poisson joint distribution for high-energy selfhealing metal film pulse capacitors to express the distribution of their degradation process. (Jiang and Jardine, 2008) investigated the degradation process of different degradation levels following different distribution. The inverse Weibull distribution was utilized to describe the degradation of the initial level, and a Weibull distribution was used to describe the subsequent degradation levels.

Utilizing stochastic processes to describe the degradation path of testing units during ADTs has significant advantages due to the stochastic performance of testing units as well as the uncertainty in their degradation behaviour. Stochastic processes, such as the Wiener process, inverse Gaussian (IG) process and Gamma process (GP) are intensively utilized in systems degradation modeling.

The Wiener process has historically been the most commonly used model (Pang et al., 2020). The Wiener process $\{Z_t, t \geq 0\}$ can be represented as $Z_t = \lambda t + \sigma B(t)$, where λ denotes a certain drift parameter, $\sigma > 0$ refers to the diffusion coefficient. Further, $B(t)$ is the standard Brownian motion and Z_t is commonly utilized to indicate the system.

The Wiener process has independent and normally distributed increments. Hence, for k observational time points $t_1 < ... < t_k$ the degradation increments $Y_j = Z_{t_j} - Z_{t_{j-1}}$ are independent and $Y_j \sim \mathbf{N}(\lambda \Delta_j, \sigma^2 \Delta_j)$, $j = 1, ..., k$, where $\Delta_j = t_j - t_{j-1}$ and $t_0 = 0$. Further, the failure time T can be expressed with the first passage time of $\{Z_t\}$ exceeding a predetermined threshold ϑ as follows

$$T = \inf\{t \geq 0 : Z_t \geq \vartheta\}. \tag{2.1}$$

(Li, Pan, and Chen, 2014) proposed a reliability assessment model for the momentum wheel based on the Wiener process where the expectation maximization (EM) algorithm is utilized to calculate the corresponding reliability. (Jin, Matthews, and Zhou, 2013) expressed the degradation process of spacecraft battery packs using the Wiener process with random parameters. Further, (Zhai et al., 2018) provided a random-effect Wiener process model based on the accelerated failure time principle, and applied an inverse Gaussian distribution to express the unit-specific heterogeneity along the degradation

trajectories.

Despite the fact that the Wiener and Gamma processes have been commonly utilized in degradation modelling, IG is an efficient degradation modelling method with monotone performance paths. For instance, (Ye and Chen, 2014) stated that the IG process has several superb properties compared with the Gamma process when covariates and random effects are considered. The IG process Z_t, $t \geq 0$, has independent degradation increments and $Z_t - Z_s \sim IG(\mu h(s,t), \lambda h^2(s,t))$ for all $0 \leq s < t$, where $h(s,t) = \Lambda(t) - \Lambda(s)$, and Λ is a given increasing function on $[0, \infty)$. Furthermore, $IG(a,b), a, b > 0$ is an IG distribution with the following density function

$$f(y; a, b) = \sqrt{\frac{b}{2\pi y^3}} \exp\left[-\frac{b(y-a)^2}{2ya^2}\right], \qquad y > 0. \tag{2.2}$$

(Peng, 2015) introduce accelerated degradation modeling using the IG process under the assumption that the mean parameter is related to the accelerating stress variable(s), whereas the scale parameter is independent of these variables. (Ye et al., 2014) proposed an ADT scheme with the IG process under the existence of random effects. The authors proved that the considered model parameters have a relatively high robustness. (Duan and Wang, 2018) derived an optimal design of the SSADT given that the degradation process of the system is modelled with the IG process where both the mean and the scale parameters are functions of the accelerating stress variables.

GP is a statistical model describing the natural degradation trajectory in which the degradation is considered to occur gradually over time in a series of tiny positive increments. (Abdel-Hameed, 1975) stated that the Gamma process $\{Z_t, t \geq 0\}$ has the following three characteristics:

1. $Z_0 = 0$, the process has independent increments, and for all $0 \leq s < t$ the increment $Z_t - Z_s$ has a Gamma distribution $Ga(\gamma(t-s), \nu)$ with a given shape rate $\gamma > 0$ and a given scale parameter $\nu > 0$.

2. The increments for any set of disjoint time intervals are independent random variables having the distributions described in (1).

3. The density function of the degradation increments Y_j, where $\Delta_j = t_j - t_{j-1}$, is defined as

$$f(y_j) = \frac{y_j^{\gamma\Delta_j - 1} \exp\left(\frac{-y_j}{\nu}\right)}{\Gamma(\gamma\Delta_j)\nu^{\gamma\Delta_j}}. \tag{2.3}$$

(Tsai, Tseng, and Balakrishnan, 2012) discussed the problem of optimal design for degradation tests based on a Gamma degradation process with random effects. The authors considered several decision variables such as the sample size, inspection frequency, and

measurement numbers in order to find the c-optimal decision variables. (Duan and Wang, 2019) addressed the optimal design problems for constant-stress accelerated degradation test constant stress ADT based on Gamma processes with fixed effect and random effect. (Pan and Balakrishnan*, 2010) considered the modeling method of the Wiener and Gamma degradation processes in multiple-steps ADT and applied the Bayesian Markov chain Monte Carlo (MCMC) method to conduct a reliability analysis.

An Acceleration model indicates the functional relationships between the reliability characteristics of the system under study, such as quantiles of failure time distribution and failure rates, and the corresponding accelerating stress variable(s), say **x**, of that system. The acceleration model is regarded as the basis for the extrapolation, and can be used to derive the reliability features under normal use conditions. This has a direct impact on the accuracy of the extrapolation results, which are the key to ADT technology (Pang et al., 2020). Furthermore, acceleration models can be utilized in certain engineering applications to standardize an applied stress variable resulting a standardized stress variables that is incorporated in the design stage of ADTs. Acceleration models can be divided into three categories, namely an empirical acceleration model, a physical acceleration model and a statistical acceleration model, see (Nelson, 2009).

The empirical acceleration model is mainly utilized in engineering applications to express a life-stress relationship through long-term observations of the system performance. The inverse power law is the most typically utilized empirical acceleration model. (Srivastava and Gupta, 2017) used the inverse power law model to establish the stress-life relationship of solar lighting equipment, and optimized the test plan using *D*-optimal criteria to find the best stress level and the best stress rate change point, which efficiently assess the system reliability. (Azrulhisham, Mohamad, and Hamid, 2013) derived a parametric model of a carbon steel stub axle using an inverse power law model where the fatigue lifetime distribution.

A physical acceleration model was suggested to express the system failure driven by physical changes. The Arrhenius model is a typical example of physical acceleration models. This model describes the relationship between system lifetime characteristics and temperature stress. Based on the Wiener process, (Guan, Tang, and Xu, 2016) and (Lim et al., 2019) discussed the reliability assessment of CSADT and two-phase partial ADT, respectively, where the Arrhenius model is applied in both tests to express the relationship between the drift parameter and the temperature. Further, (Wang, Wang, and Duan, 2016) proposed an optimal optimal of SSADT based on the IG degradation process and assumed that the relationship between the link function of the stress level and the mean IG distribution follows the Arrhenius acceleration model.

Statistical acceleration models are based on statistical analysis methods, which are often used to interpret data that are difficult to be analyzed using physical methods. Accelerated degradation data are processed using correlation methods applying mathematical statistics and stochastic processes. Statistical acceleration models can usually be divided into parametric and nonparametric models. The most typically utilized parametric model is the accelerated failure time model, see (Elsayed, 2012), which assumes that the explanatory variables have multiplier effects on the failure time of the system. The nonparametric model is a model with no distribution assumptions and the number and characteristics of its parameters are flexible and need not to be specified in advance. The proportional hazard model, which is introduced by (Cox, 1972), is the most widely used nonparametric model. This model assumes that the explanatory variables have multiplier effects on the baseline hazard rate function of the products. In other words at any time $t(>0)$ the hazard function of a subject with a certain vector of covariates values is proportional to the hazard function of another subject and therefore their hazard ratio is constant over time.

2.2 Copula-based reliability assessment

The majority of related work on degradation-based reliability assessment and experimental design either assumes response components are mutually independent or they are dependent with a certain multivariate distribution. (Pan and Balakrishnan, 2011), for instance, introduced a bivariate degradation model based on a Birnbaum Saunders distribution with Gamma processes as marginals. (Si et al., 2018) proposed a multivariate general degradation path model considering dynamic measurements. However, (Fang, Pan, and Hong, 2020) stated that assigning a multivariate joint distribution to marginals may not be a suitable solution, as it is not an easy task to obtain a suitable joint distribution in most cases especially when the marginal processes correspond to distinct distributions. Hence, a more flexible multivariate model is required. Recently, the modeling of multiple degradation processes via a copula function has gained a noticeable interest due to the flexibility of copula functions. Copula is a tool to couple correlated marginal distributions to produce a new joint distribution. It is able to handle major multivariate modeling difficulties including the existence of dependence between marginal response components and the lack of closed-form multivariate distribution. For example, (Peng et al., 2016a) introduced a bivariate modeling structure based on an IG process via the Gaussian copula and applied it on a degradation dataset from heavy machine tools. (Wang et al., 2015a) proposed a multivariate-modeling structure based on Gamma process via Frank copula. (Peng et al., 2016b) adopted the Wiener process and the IG process to propose a bivariate degradation model with both monotonic

and nonmonotonic paths. Further, (Pan et al., 2013) utilized the Frank copula with marginal paths that correspond to Wiener processes in terms of the same dataset used in (Pan and Balakrishnan, 2011).

A copula is the function that connects the joint distribution function with individual marginal distribution functions. As described in (Sklar, 1959), a copula function is defined as $C(u_1, ..., u_n) : [0, 1]^n \to [0, 1]$, which is the joint distribution function of n standard uniformly distributed random variables $U_1, ..., U_n$, i.e. $C(u_1, ..., u_n) = \mathrm{P}(U_1 \leq u_1, ..., U_n \leq u_n)$ for all $(u_1, ..., u_n) \in [0, 1]^n$. (Sklar, 1959), expressed the connection between a copula and a general multivariate distribution with the following theorem.

Sklar's Theorem : let $\mathbf{Y} = (Y_1, ..., Y_n)^T$ be a random vector with continuous marginal distribution functions $u_1, ..., u_n$ where $u_i = F_i(y_i)$, $i = 1, ..., n$. Hence, there exists a copula function C such that

$$
\begin{aligned}
F_{\mathbf{Y}}(y_1, ..., y_n) &= \mathrm{P}(Y_1 \leq y_1, ..., Y_n \leq y_n) \\
&= C(F_1(y_1), ..., F_n(y_n)), \text{ for all } (y_1, ..., y_n) \in \mathbb{R}^n.
\end{aligned}
\tag{2.4}
$$

and $C(u_1, ..., u_n) = F_{\mathbf{Y}}(y_1, ..., y_n)$. Consequently, the joint density function is given by

$$
f_{\mathbf{y}}(y_1, ..., y_n) = c(F_1(y_1), ..., F_n(y_n)) \prod_{i=1}^{n} f_i(y_i)
\tag{2.5}
$$

where $f_i(y_i)$ is the ith marginal density function and $c(u_1, ..., u_n) = \frac{\partial^n C(u_1, ..., u_n)}{\partial u_1, ..., \partial u_n}$ is the copula density function. A large number of copula functions exist in the literature which include different dependence relationships between the marginal distribution functions, see, for example, (Nelsen, 2007). Among all available copulas, there is a popular family of copulas called the Archimedean family. This family admits explicit formulas and they allow modeling variable dependence through a dependence parameter. In this section, we consider the following, when $n = 2$, three commonly-used copulas in the Archimedean family.

1. The Clayton Copula (see (Clayton, 1978)) is given by

$$
C_{\varkappa}(u_1, u_2) = (u_1^{-\varkappa} + u_2^{-\varkappa} - 1)^{-1/\varkappa}, \quad \varkappa \in [-1, \infty) \backslash \{0\}.
\tag{2.6}
$$

As the dependence parameter $\varkappa$ approaches 0, the Clayton Copula approaches the product copula $C(u_1, u_2) = u_1 u_2$ and corresponds to independence. On the other hand, as $\varkappa$ approaches ∞, the Copula approaches the Fréchet-Hoeffding upper bound, i.e. $\min\{u_1, u_2\}$, which is an upper bound on all Copula functions. In the bivariate case, the upper bound represents perfect positive dependence

between variates (similarly for the lower bound, i.e. $\max\{u_1 + u_2 - 1, 0\}$, and perfect negative dependence).

2. The Frank Copula (see (Frank, 1979)) is given by

$$C_{\varkappa}(u_1, u_2) = -\frac{1}{\varkappa} \log \left(1 + \frac{\left(e^{-\varkappa u_1} - 1\right)\left(e^{-\varkappa u_2} - 1\right)}{e^{-\varkappa} - 1} \right) \tag{2.7}$$

where $\varkappa \in (-\infty, \infty) \backslash \{0\}$ is the fixed copula dependence parameter. This Copula allows for negative dependence, and the dependence is symmetric in both tails. Further, both Fréchet-Hoeffding bounds are attained through an appropriate choice of $\varkappa$.

3. The Gumbel Copula (see (Gumbel, 1960)) is given by

$$C_{\varkappa}(u_1, u_2) = e^{\left(-((-\log(u_1))^{\varkappa} + (-\log(u_2))^{\varkappa})^{\frac{1}{\varkappa}}\right)}, \quad 1 \leq \varkappa < \infty. \tag{2.8}$$

As the dependence parameter $\varkappa$ approaches 1 and ∞, the Copula approaches independence and the Fréchet-Hoeffding upper bound, respectively. However, there are no values for $\varkappa$ such that the Copula approaches the Fréchet-Hoeffding lower bound. This Copula does not allow negative dependence, and exhibits strong right tail dependence.

In addition to Archimedean copulas, the Gaussian copula is an important parametric copula which is constructed form the multivariate normal distribution and does not admit an explicit formula when compared to the Archimedean copulas. The bivariate Gaussian copula is defined as

$$C(u_1, u_2) = F_{0,\Sigma}\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right) \tag{2.9}$$

where $\Phi(.)$ denotes the standard univariate normal distribution function, and $F_{0,\Sigma}$ is the distribution function of the bivariate normal distribution with expectation 0 and covariance matrix

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad -1 < \rho < 1.$$

2.3 Review of optimal experimental design

The main aim of this section is to provide a general introduction of several primary topics on optimality theory of experimental designs, in particular the classical optimality criteria and an associated General Equivalence Theorem. For a more detailed summary

of these concepts see, for example, (Silvey, 1980) and (Atkinson, Donev, and Tobias, 2007). For the rest of this research we consider experiments where one is interested in estimating the unknown model parameters. Thus, we are dealing with an estimation problem and optimal planning of such experiments is concerned with obtaining the optimal experimental levels and the number of testing units to be assigned to each level so that the parameters are estimated with high precision.

2.3.1   Exact and approximate experimental designs

Before discussing experimental designs we should introduce first $\mathbf{x}$ as experimental setting which has an influence on the observations y and can be adjusted by the experimenter. Experimental designs can be classified into two primary types, exact design and approximate design. Denote $m \in \mathbb{N}$ as the number of mutually distinct support points $\mathbf{x}_1, ..., \mathbf{x}_m$, which is defined over the corresponding design region $\mathcal{X}$, an exact design ξ_E is expressed as

$$\xi_E = \begin{pmatrix} \mathbf{x}_1 & ... & \mathbf{x}_m \\ n_1 & ... & n_m \end{pmatrix} \tag{2.10}$$

where $n_i > 0$, $i = 1, .., m$ is the integer number of testing units to be assigned to the ith experimental level and n is the total number of testing units, i.e. $\sum_{i=1}^{m} n_i = n$. Finding optimal exact designs is, in general, a difficult task of discrete optimization. To circumvent this problem we follow the approach of approximate designs propagated by (Kiefer, 1959) in which the requirement of integer numbers n_i of testing units at a stress level $\mathbf{x}_i$ is relaxed. Then continuous methods of convex optimization can be employed (see e. g. (Silvey, 1980)) and efficient exact designs can be derived by rounding the optimal numbers to nearest integers. This approach is, in particular, of use when the number n of units is sufficiently large. Moreover, the frequencies n_i will be replaced by proportions $w_i = n_i/n$, because the total number n of units does not play a role in the optimization. Thus an approximate design ξ is defined by a finite number of settings $\mathbf{x}_i$, $i = 1, ..., m$, from an experimental region $\mathcal{X}$ with corresponding weights $w_i \geq 0$ satisfying $\sum_{i=1}^{m} w_i = 1$ and is denoted by

$$\xi = \begin{pmatrix} \mathbf{x}_1 & ... & \mathbf{x}_m \\ w_1 & ... & w_m \end{pmatrix}. \tag{2.11}$$

In order to avoid the discrete optimization which is necessary to obtain exact designs, approximate designs will be considered in this research.

2.3.2 Optimality criteria

The optimality criterion, which indicates the purpose of the experiment to be conducted, has the primary role in selecting an optimal experimental design. Considering an approximate design ξ, this is equivalent to maximising the Fisher information matrix. The Fisher information matrix is defined as the covariance matrix (matrix of second moments) of the score function which is defined as the vector of first derivatives of the overall log likelihood function $\ell(\boldsymbol{\theta})$ with respect to the components of the parameter vector $\boldsymbol{\theta}$ where expectation is with respect to the response vector, say $\mathbf{Y} = (Y_1, ..., Y_n)^T$. Moreover, under certain regularity conditions, this is equal to minus the expectations of the second derivatives of the score function. Furthermore, to understand how expectations are taken, it would be preferable to indicate the dependence of the likelihood function on the independent random variables $Y_1, ..., Y_n$, i.e. $\ell(\boldsymbol{\theta}, Y_1, ..., Y_n)$. Under the settings of an exact design with experimental settings $(\mathbf{x}_1, ..., \mathbf{x}_n)$ the corresponding information matrix is defined as

$$\mathbf{M}_{\boldsymbol{\theta}}(\mathbf{x}_1, ..., \mathbf{x}_n) = \mathrm{E}\left(-\frac{\partial^2 \ell(\boldsymbol{\theta}, Y_1, ..., Y_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right). \tag{2.12}$$

Further, it should be noted that the asymptotic covariance matrix of the maximum likelihood estimator is proportional to the inverse of the information matrix. Finding optimal exact designs is, in general, a difficult task of discrete optimization. To handle this issue we follow the approach of approximate designs as introduced earlier. Hence, the per unit, i.e. standardized, information matrix considering the parameter vector $\boldsymbol{\theta}$ and in terms of the approximate design ξ is defined as

$$\mathbf{M}_{\boldsymbol{\theta}}(\xi) = \sum_{i=1}^{m} \omega_i \mathbf{M}_{\boldsymbol{\theta}}(\mathbf{x}_i) \tag{2.13}$$

where $\mathbf{M}_{\boldsymbol{\theta}}(\mathbf{x})$ is the elementary information matrix which puts unit mass at the stress level combination $\mathbf{x}$. Based on the combination of the model parameters we are interested in as well as the statistical properties to be achieved through the optimality criterion, a statistically meaningful function of the information matrix has to be defined. This function, which is called the objective function $\Phi(\mathbf{M}_{\boldsymbol{\theta}}(\xi))$, maps the information matrix to the real line. Consequently, the purpose is then to minimize the objective function according to ξ in order to derive the corresponding optimal design. The optimal designs that correspond to nonlinear models depend on the values of the parameters which complicates the definition of an optimality criterion. A simple approach to handle this issue is to maximize $\Phi(\mathbf{M}_{\boldsymbol{\theta}}(\xi))$ with $\boldsymbol{\theta}$ fixed at a prior guess $\widehat{\boldsymbol{\theta}}$ which leads to locally optimal designs, see (Chernoff, 1953). Optimality criteria are often symbolized by an

alphabetical letter and hence are sometimes called alphabetical optimality criteria, see (Atkinson, Donev, and Tobias, 2007). The most common optimality criteria are listed below under the assumption of a q-dimensional parameter vector $\boldsymbol{\theta}$ and, hence, a $q \times q$ information matrix $\mathbf{M}_{\boldsymbol{\theta}}(\xi)$.

Definition 2.3.1. *(D-optimality criterion.) An approximate design ξ^* is called D-optimal if it maximizes the determinant of $\mathbf{M}_{\boldsymbol{\theta}}(\xi)$ or, equivalently,*

$$- \log(\det(\mathbf{M}_{\boldsymbol{\theta}}(\xi^*))) = \min_{\xi}[- \log(\det(\mathbf{M}_{\boldsymbol{\theta}}(\xi)))].$$

Definition 2.3.2. *A linear criterion Φ is defined as follows*

$$\Phi(\mathbf{M}_{\boldsymbol{\theta}}(\xi)) = \mathrm{tr}[\boldsymbol{L}\,\mathbf{M}_{\boldsymbol{\theta}}(\xi)^{-1}]$$

where $\boldsymbol{L}$ is a $q \times q$ positive semi definite matrix.

The following *A*-optimality criterion is an example of linear criteria.

Definition 2.3.3. *(A-optimality criterion.) An approximate design ξ^* is called A-optimal if it minimizes the trace of the variance covariance matrix, i.e.*

$$\mathrm{tr}(\mathbf{M}_{\boldsymbol{\theta}}(\xi^*)^{-1}) = \min_{\xi}[\mathrm{tr}(\mathbf{M}_{\boldsymbol{\theta}}(\xi)^{-1})].$$

Definition 2.3.4. *(c-optimality criterion.) An approximate design ξ^* is called c-optimal if it minimizes the variance of the estimator of a predetermined linear combination $\boldsymbol{c}^T\boldsymbol{\theta}$ of model parameters, i.e.*

$$\mathbf{c}^T\mathbf{M}_{\boldsymbol{\theta}}(\xi^*)^{-1}\mathbf{c} = \min_{\xi}[\mathbf{c}^T\mathbf{M}_{\boldsymbol{\theta}}(\xi)^{-1}\mathbf{c}].$$

Hence, *c*-optimal criterion is an appropriate criterion when the aim is to estimate some function $m(\boldsymbol{\theta})$ of the model parameters with minimum variance. By the delta-method, the minimum asymptotic variance of $m(\hat{\boldsymbol{\theta}})$ is expressed as follows

$$\min \mathrm{AVar}(m(\hat{\boldsymbol{\theta}})) = \mathbf{c}^T\mathbf{M}_{\boldsymbol{\theta}}(\xi^*)^{-1}\mathbf{c} \tag{2.14}$$

where $m(\boldsymbol{\theta})$ is assumed to be differentiable in a neighborhood of $\boldsymbol{\theta}$ and $\mathbf{c} = \frac{\partial m(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$. The optimality criterion expressed in equation (2.14) is extensively used to obtain optimal experimental designs for ADTs. In this regard, $m(\boldsymbol{\theta})$ is derived as some quantile of the failure time distribution at the normal use conditions of the system under study. Consequently, this optimality criterion, which is a special case of the c-optimal criterion, will be utilized throughout this work to derive optimal experimental designs under various system conditions.

2.3.3 The general equivalence theorem

In terms of approximate designs, the general equivalence theorem is a very useful tool for validating the global optimality of suggested optimal designs. let $\boldsymbol{\Psi}(\xi, \xi_{\mathbf{x}})$ be the directional derivative at the approximate design ξ in the direction of $\xi_{\mathbf{x}}$ given by

$$\boldsymbol{\Psi}(\xi, \xi_{\mathbf{x}}) = \lim_{\alpha \to 0^+} \frac{1}{\alpha}[\Phi\{(1-\alpha)\mathbf{M}_{\boldsymbol{\theta}}(\xi) + \alpha\mathbf{M}_{\boldsymbol{\theta}}(\xi_{\mathbf{x}})\} - \Phi\{(\mathbf{M}_{\boldsymbol{\theta}}(\xi)\}]$$

where $\xi_{\mathbf{x}}$ is the design which puts unit mass at the setting $\mathbf{x}$. (Atkinson, Donev, and Tobias, 2007) state that for a given design ξ^*, the general equivalence theorem ensures, under the assumption that Φ is a convex function (on the set of all positive defi nite matrices), that the following three statements are equivalent.

1. the design ξ^* minimizes $\Phi(\mathbf{M}_{\boldsymbol{\theta}}(\xi))$.

2. The design ξ^* maximizes the minimum over $\mathcal{X}$ of $\boldsymbol{\Psi}(\xi, \xi_{\mathbf{x}})$.

3. The minimum over $\mathcal{X}$ of $\boldsymbol{\Psi}(\xi^*, \xi_{\mathbf{x}})$ is equal to zero.

(Silvey, 1980) states that Caratheodory's theorem can be used to obtain an upper bound for the number of support points of an optimal design on the basis of the additive nature of $\mathbf{M}_{\boldsymbol{\theta}}(\xi)$. There exists an optimal design such that the number of support points is bounded by $q(q+1)/2$ points where q is the dimension of the parameter vector $\boldsymbol{\theta}$. Furthermore, an optimal design which has exactly q support points is said to be minimally supported.

2.3.4 The multiplicative algorithm

Optimal designs do not usually come in closed form (Yu, 2010). Several early research articles as (Wu and Wynn, 1978), and later contributions as (Dette, Pepelyshev, and Zhigljavsky, 2008) proposed novel procedures for numerical computation of optimal experimental designs. Following is the general formula of the multiplicative algorithm as introduced by (Silvey, Titterington, and Torsney, 1978) which is determined through a power parameter $\delta \in (0,1]$ and a finite design region $\mathcal{X} = \{\mathbf{x}_1, ..., \mathbf{x}_m\}$.

1. Set $\delta \in (0,1]$ and $\xi^{(0)} \in \Omega$, with weights $w_i^{(0)} > 0$, where $\Omega = \{\xi : w_i \geq 0, \sum_{i=1}^{m} w_i = 1\}$.

2. For $q = 0, 1, ...$, compute

$$w_i^{(q+1)} = w_i^{(q)} \frac{d_i^{\delta}(\xi^{(q)})}{\sum_{j=1}^{m} w_j^{(q)} d_j^{\delta}(\xi^{(q)})}, \qquad i = 1, ..., m, \tag{2.15}$$

where for a convex, decreasing, and differentiable optimality criterion Φ the quantities $d_i(\xi)$ read as

$$d_i(\xi) = -\mathrm{tr}\Big(\Phi'\big(\mathbf{M}_{\boldsymbol{\theta}}(\xi)\big)\mathbf{M}_{\boldsymbol{\theta}}(\xi_{\mathbf{x}_i})\Big), \qquad \Phi'(\mathbf{M}) \equiv \frac{\partial\Phi(\mathbf{M})}{\partial\mathbf{M}}.$$

3. Iterate until convergence.

For a heuristic explanation, observe that equation (2.15) is equivalent to

$$w_i^{(q+1)} \;\propto\; w_i^{(q)}\left(\frac{\partial\Phi\big(\mathbf{M}_{\boldsymbol{\theta}}(\xi)\big)}{\partial w_i}\Big|_{\xi=\xi^{(q)}}\right)^{\delta}, \qquad i = 1, ..., m. \tag{2.16}$$

The value of the derivative $\frac{\partial\Phi\big(\mathbf{M}_{\boldsymbol{\theta}}(\xi)\big)}{\partial w_i}$ corresponds to the amount of gain in information, as obtained by $\Phi(.)$, by a slight increase in ω_i, the weight on the ith design point. Hence, equation (2.16) is considered as adjusting ξ so that relatively more weight is placed on design points whose increased weight may result in a larger gain in information in $\Phi(.)$. If the optimality criterion is decreasing and convex, then a convenient convergence criterion, based on the general equivalence theorem, see (Kiefer and Wolfowitz, 1960), is

$$\max_{1\leq i\leq m} d_i(\xi^{(q)}) \leq (1 + \epsilon)\sum_{i=1}^{m} w_i d_i(\xi^{(q)}) \tag{2.17}$$

where ϵ is a small positive constant.

c-optimality criterion will be utilized throughout this work to obtain algorithm-based optimal designs, we consider the work of (Yu, 2010) to write the updating rule of the multiplicative algorithm when the power parameter $\delta = 1$ as follows

$$w_i^{(q+1)} = w_i^{(q)}\frac{\mathbf{c}^T\mathbf{M}_{\boldsymbol{\theta}}(\xi^{(q)})^{-1}\mathbf{M}_{\boldsymbol{\theta}}(\boldsymbol{\xi}_{\mathbf{x}_i})\mathbf{M}_{\boldsymbol{\theta}}(\xi^{(q)})^{-1}\mathbf{c}}{\mathbf{c}^T\mathbf{M}_{\boldsymbol{\theta}}(\xi^{(q)})^{-1}\mathbf{c}}. \tag{2.18}$$

As noted by (Yu, 2010), the choice $\delta = 1$ may result in an oscillating behavior in the sense that $\xi^{(q)}$ alternates between two points at which the optimality criterion $\Phi(.)$ takes the same value. Hence, following (Torsney, 1983), a practical recommendation of $\delta = 1/2$ is considered throughout this work. It should be noted that all numerical computations throughout this work were made by using the R programming language(R Core Team, 2020).

Chapter 3

Optimal Designs Based on Linear Mixed Effects Models

3.1   Introduction

In this chapter, we consider repeated measures Accelerated Degradation Testing with multiple stress variables, where the degradation paths are assumed to follow a linear mixed effects model which is quite common in settings when repeated measures are made. We derive optimal experimental designs for minimizing the asymptotic variance for estimating the median failure time under normal use conditions when the time points for measurements are either fixed in advance or are also to be optimized.

A vast amount of literature is devoted to the analysis of Accelerated Degradation Testing, see, for example, (Limon, Yadav, and Liao, 2017) for prominent methods for statistical inference and optimal design of accelerated testing plans. (Li and Kececioglu, 2006) presented an analytical method for the optimum planning of Accelerated Degradation Testing with an application to the reliability of Light-Emitting Diodes. There the author states that the variability of the measured units have a substantial impact on the accuracy of estimation. Therefore these random effects should be encountered in the choice of the experimental settings for the Accelerated Degradation Testing. The general theory of optimal design of experiments is well developed in the mathematical context of approximate designs which allow for analytical solutions (see e. g. (Silvey, 1980) or (Atkinson, Donev, and Tobias, 2007)). In the presence of random effects, (Entholzner et al., 2005) showed that for single samples the optimal designs for fixed effects models retain their optimality for linear optimality criteria. (Debusho and Haines, 2008) show that this also holds for D-optimality in linear models when only the intercept is random. However, in a multi-sample situation (Schmelter, 2007) and (Schwabe and Schmelter, 2008) exhibited that the variability of the intercept has a non-negligible influence on the D-optimal design. In the case of random slope effects this dependence already occurs in single samples as outlines by (Schmelter, Benda, and Schwabe, 2007). (Dette, Pepelyshev, and Holland-Letz, 2010) considered the problem

of constructing *D*-optimal designs for linear and nonlinear random effect models with applications in population pharmacokinetics. These authors present a new approach to determine efficient designs for nonlinear least squares estimation which addresses the problem of additional correlation between observations within units. Based on geometrical arguments, (Graßhoff et al., 2012) derived *D*-optimal designs for random coefficient regression models when only one observation is available per unit, a situation which occurs in destructive testing.

The present approach is based on the discussion paper by (Weaver and Meeker, 2013) in which two case studies are introduced for optimal planning of repeated measures Accelerated Degradation Testing. There the authors consider the influence of a single stress variable and use a criterion based on a large-sample approximation of the precision for estimating a quantile of the failure-time distribution under normal use conditions. We will adopt this approach, generalize the results presented there to more general models, and extend the design optimization also to generate optimal time plans.

The present chapter is organized as follows. Section 3.2 starts with a motivation example based on a case study in (Weaver and Meeker, 2013). In Sections 3.3, 3.4 and 3.5 we state the general model formulation, specify the maximum-likelihood estimation and exhibit the corresponding information matrix. Basic concepts of optimal design theory in the present context are collected in Section 3.6 while Section 3.7 is devoted to the idea of soft failure due to degradation, where we derive the design optimality criterion for estimating a quantile of the failure time distribution under normal use conditions. In Section 3.8 optimal designs are characterized when the time plan for repeated measurements at the testing units is fixed in advance. In Section 3.9 also the measurement times are optimized under the constraint that measurements are taken according to the same time plan for all units, and in Section 3.10 optimal measurement times are determined in the setting of destructive testing. The chapter closes with a short discussion in Section 3.11. Throughout the chapter the theoretical concepts are illustrated by two accompanying running examples. Some technical result is deferred to an Appendix. Finally, it should be mentioned that the content of the current chapter is strongly related to the work of (Shat and Schwabe, 2021)

3.2 Introductory example

Before formulating our general degradation model in Section 3.3, we start in this section for motivation with the description of a simple introductory example based on (Weaver and Meeker, 2013).

Example 1. *The model proposed in (Weaver and Meeker, 2013) is a linear mixed effect model with a single stress variable x. In this model there are n testing units for which*

degradation y_{ij} is observed at k time points t_j, $j = 1, ..., k$. The (standardized) stress variable x can be chosen by the experimenter from the design region $\mathcal{X} = [0, 1]$.

On the unit level the response y_{ij} for the degradation of testing unit i at time t_j is represented by

$$y_{ij} = \beta_{i,1} + \beta_2 x_i + \beta_{i,3} t_j + \beta_4 x_i t_j + \varepsilon_{ij}, \tag{3.1}$$

where the intercept $\beta_{i,1}$ is the mean degradation of unit i at time $t = 0$ under the stress level $x = 0$, β_2 is the common (not unit specific) mean increase in degradation depending on the stress variable x, $\beta_{i,3}$ is the mean increase in degradation of unit i over time t when the stress level is set to $x = 0$, and β_4 is the interaction effect between time and stress. The measurement errors ε_{ij} are assumed to be realizations of normally distributed error variables with mean zero and error variance σ_ε^2.

On the whole experiment level the unit specific parameters $(\beta_{i,1}, \beta_{i,3})^T$ of the units are assumed to be realizations from a bivariate normal distribution with mean $(\beta_1, \beta_3)^T$ and a variance covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. All random effect parameters and measurement errors are assumed to be independent both within as well as between units. Under well controlled measuring testing conditions, the variability of the response is completely described by both the unit to unit variability Σ and the within unit variability of the measurement errors.

To illustrate the situation some virtual degradation paths y_{ij}, $j = 1, ..., k$, are depicted in Figure 3.1 for three different values of the stress variable x. There are three units shown at each value of the stress level ($n = 9$) and $k = 11$ equally spaced measurement times t_j. The roughness of the paths is due to the measurement errors ε_{ij} The corresponding underlying mean degradation paths $\mu_i(x_i, t) = \beta_{i,1} + \beta_2 x_i + \beta_{i,3} t + \beta_4 x_i t$, corrected for the measurement errors, are shown in Figure 3.2. These mean degradation paths are represented by straight lines over time, where both the intercept and the slope may vary across units around an aggregate value determined by the value x_i of the stress variable. For the analysis of degradation under normal use it is assumed that the model equation 3.1 is also valid at the normal use condition x_u, where typically $x_u < 0$ on the standardized scale, i.e. $\mu_u(x_u, t) = \beta_{u,1} + \beta_2 x_u + \beta_{u,3} t + \beta_4 x_u t$ describes the mean degradation of a future unit u at normal use condition x_u and time t, and the unit specific parameters $(\beta_{u,1}, \beta_{u,3})^T$ are bivariate normal with mean $(\beta_1, \beta_3)^T$ and variance covariance matrix Σ.

The degradation is assumed to undergo soft failure, i.e. the unit u will be considered to fail due to degradation when its mean degradation exceeds a predetermined threshold y_0. The corresponding time t_u, for which $\mu_u(x_u, t_u) = y_0$, will be called the failure time of unit u under normal use condition due to degradation. These failure times vary across different unit because of the unit specific parameters $\beta_{u,1}$ and $\beta_{u,3}$.
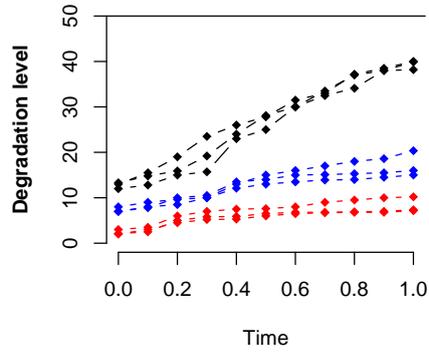
FIGURE 3.1: Observed degradation paths
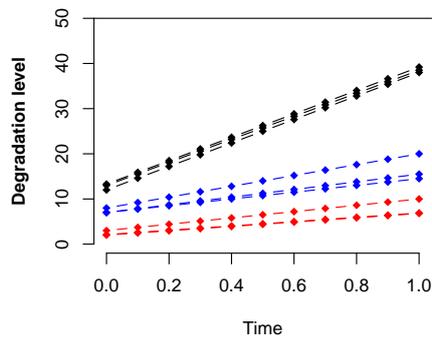
FIGURE 3.2: Mean degradation paths

TABLE 3.1: Nominal values for Example 1

| β_1 | β_2 | β_3 | β_4 | σ_1 | σ_2 | ρ | σ_ε | x_u | y_0 |
1	x	t	xt						
2.397	1.629	1.018	0.0696	0.114	0.105	-0.143	0.048	-0.056	3.912

In Figure 3.1 and Figure 3.2 the predetermined failure threshold $y_0 = 50$ is indicated by a horizontal line. As typical for degradation studies failure does not occur during the time of experiment even for the highest stress level.

(Weaver and Meeker, 2013) derived locally *c*-optimal designs for certain characteristics of the mean degradation curves (for future units) under normal use condition x_u. The optimal design is always supported on the extremal points of the design region $\mathcal{X}$ and the optimal weights depend on the nominal values of the model parameters. In Table 3.1 we reproduce the nominal values of Example 7.2 by (Weaver and Meeker, 2013) on scar width growth after standardization for further use.

In the following section we extend the work of (Weaver and Meeker, 2013) by considering a general mixed effects model with multiple stress variables **x** given that the general model can be applied to linear mixed effect models, linear additive models or exponential models. Further, this generalization allows for degradation models with and without interactions between the multiple design variables **x**.

3.3 Formulation of the model

In this section, we give a general formulation of a mixed effects regression model incorporating a product-type structure with complete interactions between the stress and the time variable. This general formulation can cover the situation described in the introductory Example 1 of the previous section, but may readily be extended to more than one stress variable and to more complex marginal models for both the stress variables and the time variable, as indicated later. To become more specific we assume that there are n testing units $i = 1, ..., n$, for which degradation y_{ij} is to be measured at k subsequent time points t_j, $j = 1, ..., k$, $t_1 < ... < t_k$. Each unit i is observed under a value $\mathbf{x}_i$ of the stress variable(s), which is kept fixed for each unit throughout the degradation process, but may differ from unit to unit. The number k of measurements and the time points are the same for all units. The measurements y_{ij} are regarded as realizations of random variables Y_{ij} which are described by a hierarchical model.

For each unit i the observation Y_{ij} at time point t_j is given by

$$Y_{ij} = \mu_i(\mathbf{x}_i, t_j) + \varepsilon_{ij}, \tag{3.2}$$

where $\mu_i(\mathbf{x}, t)$ is the mean degradation of unit i at time t, when stress $\mathbf{x}$ is applied to unit i, and ε_{ij} is the associated measurement error at time point t_j. The mean degradation $\mu_i(\mathbf{x}, t)$ is assumed to be given by a linear model equation in the stress variable $\mathbf{x}$ and time t,

$$\mu_i(\mathbf{x}, t) = \sum_{r=1}^{p} \beta_{i,r} f_r(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t)^T \boldsymbol{\beta}_i \tag{3.3}$$

where $\mathbf{f}(\mathbf{x}, t) = (f_1(\mathbf{x}, t), ..., f_p(\mathbf{x}, t))^T$ is a p-dimensional vector of known regression functions $f_q(\mathbf{x}, t)$ in both the stress variable(s) $\mathbf{x}$ and the time t, $\boldsymbol{\beta}_i = (\beta_{i,1}, ..., \beta_{i,p})^T$ is a p-dimensional vector of unit specific parameters $\beta_{i,q}$. Hence, the response is given by

$$Y_{ij} = \mathbf{f}(\mathbf{x}_i, t_j)^T \boldsymbol{\beta}_i + \varepsilon_{ij}. \tag{3.4}$$

The measurement error ε_{ij} is assumed to be normally distributed with zero mean and some potentially time dependent error variance $\sigma_{\epsilon,j}^2$ ($\varepsilon_{ij} \sim \mathrm{N}(0, \sigma_{\epsilon,j}^2)$). Moreover, the error terms may be correlated within a unit over time. So, in general the vector $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, ..., \varepsilon_{ik})^T$ of errors associated with the k observations within one unit i is k-dimensional multivariate normally distributed with zero mean and positive definite variance covariance matrix $\boldsymbol{\Sigma}_\varepsilon$ ($\boldsymbol{\varepsilon}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$).

Example (Example 1 cont.)**.** *In the introductory model*

$$Y_{ij} = \beta_{i,1} + \beta_2 x_i + \beta_{i,3} t_j + \beta_4 x_i t_j + \varepsilon_{ij}$$

of Section 3.2 there is only one stress variable $\mathbf{x} = x$*, the vector of regression functions is given by* $\mathbf{f}(x, t) = (1, x, t, xt)^T$*, and the unit specific vector of parameters is* $\boldsymbol{\beta}_i = (\beta_{i,1}, \beta_2, \beta_{i,3}, \beta_4)^T$ *(p = 4). The error terms are assumed to be homoscedastic and independent, i. e.* $\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_k$*, where* $\mathbf{I}_k$ *denotes the* $k \times k$ *identity matrix.*

For the regression functions $\mathbf{f}(\mathbf{x}, t)$ we suppose a product-type structure with complete interactions between the stress variable $\mathbf{x}$ and the time t, i. e. there are marginal regression functions $\mathbf{f}_1(\mathbf{x}) = (f_{11}(\mathbf{x}), ..., f_{1p_1}(\mathbf{x}))^T$ and $\mathbf{f}_2(t) = (f_{21}(t), ..., f_{2p_2}(t))^T$ of dimension p_1 and p_2 which only depend on the stress variable $\mathbf{x}$ and the time t, respectively, and the vector $\mathbf{f}(\mathbf{x}, t) = \mathbf{f}_1(\mathbf{x}) \otimes \mathbf{f}_2(t)$ of regression functions factorizes into its marginal counterparts ($p = p_1 p_2$). Here "$\otimes$" denotes the Kronecker product of matrices or vectors. Then the observation Y_{ij} can be written as

$$Y_{ij} = \sum_{r=1}^{p_1} \sum_{s=1}^{p_2} \beta_{i,rs} f_{1r}(\mathbf{x}_i) f_{2s}(t_j) + \varepsilon_{ij} = (\mathbf{f}_1(\mathbf{x}) \otimes \mathbf{f}_2(t))^T \boldsymbol{\beta}_i + \varepsilon_{ij}, \tag{3.5}$$

where for notational convenience the entries of the vector $\boldsymbol{\beta}_i = (\beta_{i,11}, ..., \beta_{i,1p_2}, ..., \beta_{i,p_1 p_2})$ of parameters are relabeled lexicographically according to their associated marginal regression functions $(q = (r-1)p_2 + s, \ r = 1, ..., p_1, \ s = 1, ..., p_2)$.

Example (Example 1 cont.)**.** *In the introductory model of Section 3.2 the rearranged vector of regression functions* $\mathbf{f}(x,t) = (1, t, x, xt)^T = ((1, x) \otimes (1, t))^T$ *factorizes into the vector* $\mathbf{f}_1(x) = (1, x)^T$ *of a simple linear regression in the stress variable x $(p_1 = 2)$ and the vector* $\mathbf{f}_2(t) = (1, t)^T$ *of a simple linear regression in time t $(p_2 = 2)$. The unit specific vector of parameters* $\boldsymbol{\beta}_i$ *is relabeled by* $\beta_{i,11} = \beta_{i,1}$, $\beta_{i,12} = \beta_{i,3}$, $\beta_{21} = \beta_2$, *and* $\beta_{22} = \beta_4$, *where for* β_{21} *and* β_{22} *the index i of the unit is suppressed as these parameters do not depend on the unit. Thus the model equation can be rewritten as*

$$Y_{ij} = \beta_{i,11} + \beta_{i,12}t_j + \beta_{21}x_i + \beta_{22}x_i t_j + \varepsilon_{ij} = ((1, x_i) \otimes (1, t_j))^T \boldsymbol{\beta}_i + \varepsilon_{ij}.$$

Moreover, as in the introductory example, we will assume throughout that the marginal regression function $\mathbf{f}_1(\mathbf{x}) = (f_{11}(\mathbf{x}), ..., f_{1p_1}(\mathbf{x}))^T$ of the stress variable $\mathbf{x}$ contains a constant term, $f_{11}(\mathbf{x}) \equiv 1$ say, which is a common assumption in the majority of situations, and that only the leading p_2 parameters $\beta_{i,11}, ..., \beta_{i,1p_2}$ of $\boldsymbol{\beta}_i$ associated with this constant term are unit specific. All other parameters in $\boldsymbol{\beta}_i$ are assumed to take the same value β_{rs}, $r = 2, ..., p_1$, $s = 1, ..., p_2$, for all individuals $i = 1, ..., n$. Hence, for unit i the model (3.5) can be rewritten as

$$Y_{ij} = (\mathbf{f}_1(\mathbf{x}_i) \otimes \mathbf{f}_2(t_j))^T \boldsymbol{\beta} + \mathbf{f}_2(t_j)^T \boldsymbol{\gamma}_i + \varepsilon_{ij}, \tag{3.6}$$

where $\boldsymbol{\beta} = (\beta_{11}, ..., \beta_{p_1 p_2})^T$ is the vector of fixed effect (aggregate) parameters (averaged over the units) associated with the constant term in the regression functions of the stress variable $\mathbf{x}$ and $\boldsymbol{\gamma}_i = (\gamma_{i1}, ..., \gamma_{ip_2})^T$ is the p_2-dimensional vector of unit specific deviations $\gamma_{is} = \beta_{i,1s} - \beta_{1s}$, $s = 1, ..., p_2$, from the corresponding aggregate parameters.

On the aggregate level it is assumed that the units are representatives of a larger entity. The deviations of the units from the aggregate value are then modeled as random effects, i.e. they are p_2-dimensional multivariate normal with zero mean and variance-covariance matrix $\boldsymbol{\Sigma}_\gamma$ ($\boldsymbol{\gamma}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\gamma)$). All vectors $\boldsymbol{\gamma}_i$ of random effects and all vectors $\boldsymbol{\varepsilon}_i$ of random errors are assumed to be independent.

Example (Example 1 cont.)**.** *In the introductory model of Section 3.2 the two-dimensional vector of deviations associated with the regression functions* $f_1(x, t) = 1$ *and* $f_2(x, t) = t$ *is* $\boldsymbol{\gamma}_i = (\beta_{i,11} - \beta_{11}, \beta_{i,12} - \beta_{12})^T$ *with general 2×2 variance-covariance matrix* $\boldsymbol{\Sigma}_\gamma = \boldsymbol{\Sigma}$. *This leads to the mixed effects model equation in standard notation*

$$Y_{ij} = \beta_{11} + \beta_{12}t_j + \beta_{21}x_i + \beta_{22}x_i t_j + \gamma_{i1} + \gamma_{i2}t_j + \varepsilon_{ij} = ((1, x_i) \otimes (1, t_j))^T \boldsymbol{\beta} + (1, t_j)^T \boldsymbol{\gamma}_i + \varepsilon_{ij}.$$

In vector notation the k-dimensional vector $\mathbf{Y}_i = (Y_{i1}, ..., Y_{ik})^T$ of observations for unit i can be expressed as

$$\mathbf{Y}_i = (\mathbf{f}_1(\mathbf{x}_i)^T \otimes \mathbf{F}_2)\boldsymbol{\beta} + \mathbf{F}_2\,\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i,$$

where $\mathbf{F}_2 = (\mathbf{f}_2(t_1), ..., \mathbf{f}_2(t_k))^T$ is the $k \times p_2$ marginal design matrix for the time variable. Then $\mathbf{Y}_i$ is k-dimensional multivariate normally distributed with mean $(\mathbf{f}_1(\mathbf{x}_i)^T \otimes \mathbf{F}_2)\boldsymbol{\beta}$ and variance covariance matrix $\mathbf{V} = \mathbf{F}_2\boldsymbol{\Sigma}_\gamma\mathbf{F}_2^T + \boldsymbol{\Sigma}_\varepsilon$. The variance covariance matrix $\mathbf{V}$ is not affected by the choice of the stress level $\mathbf{x}_i$ and, hence, equal for all units i.

Example (Example 1 cont.)**.** *In the introductory model of Section 3.2 the marginal design matrix*

$$\mathbf{F}_2 = \left(\begin{array}{ccc} 1 & ... & 1 \\ t_1 & ... & t_k \end{array} \right)^T$$

is the common design matrix for simple linear regression over time. The variance covariance matrix $\mathbf{V} = \mathbf{F}_2\boldsymbol{\Sigma}\mathbf{F}_2^T + \sigma_\varepsilon^2\mathbf{I}_k$ *has diagonal entries* $v_{jj} = \mathrm{Var}(Y_{ij}) = \sigma_1^2 + 2\rho\sigma_1\sigma_2 t_j + \sigma_2^2 t_j^2 + \sigma_\varepsilon^2$ *and off-diagonal entries* $v_{jj'} = \mathrm{cov}(Y_{ij}, Y_{ij'}) = \sigma_1^2 + \rho\sigma_1\sigma_2(t_j + t_{j'}) + \sigma_2^2 t_j t_{j'}$.

In total, for the observations of all n units the stacked nk-dimensional response vector $\mathbf{Y} = (\mathbf{Y}_1^T, ..., \mathbf{Y}_n^T)^T$ can be represented in matrix notation as

$$\mathbf{Y} = (\mathbf{F}_1 \otimes \mathbf{F}_2)\boldsymbol{\beta} + (\mathbf{I}_n \otimes \mathbf{F}_2)\,\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{3.7}$$

where $\mathbf{F}_1 = (\mathbf{f}_1(\mathbf{x}_1), ..., \mathbf{f}_1(\mathbf{x}_n))^T$ is the $n \times p_1$ marginal design matrix for the stress variables across units, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^T, ..., \boldsymbol{\gamma}_n^T)^T$ is the np_2-dimensional stacked parameter vector of random effects and $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^T, ..., \boldsymbol{\varepsilon}_n^T)^T$ is the nk-dimensional stacked vector of random errors. Such a model equation is sometimes called the "marginal model" for the response $\mathbf{Y}$, but should not be confused with models marginalized for the covariates $\mathbf{x}$ and t, respectively (see the decomposition at the end of Section 3.5). Note that the vectors $\boldsymbol{\gamma} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}_\gamma)$ of all random effects and the vector $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}_\varepsilon)$ are multivariate normal. Hence, the vector $\mathbf{Y}$ of all observations is nk-dimensional multivariate normal, $\mathbf{Y} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{V})$.

Example (Example 1 cont.)**.** *In the introductory model of Section 3.2 the marginal design matrix*

$$\mathbf{F}_1 = \left(\begin{array}{ccc} 1 & ... & 1 \\ x_1 & ... & x_n \end{array} \right)^T$$

is the common design matrix for simple linear regression on the stress variable x.

For the analysis of degradation under normal use we further assume that the general model 3.6 is also valid at the normal use condition $\mathbf{x}_u$, where typically $\mathbf{x}_u \notin \mathcal{X}$, i.e.

$$\mu(\mathbf{x}_u, t) = (\mathbf{f}_1(\mathbf{x}_u) \otimes \mathbf{f}_2(t))^T \boldsymbol{\beta} + \mathbf{f}_2(t)^T \boldsymbol{\gamma}_u \tag{3.8}$$

describes the mean degradation of a future unit u at normal use condition $\mathbf{x}_u$ and time t, and the random effects $\boldsymbol{\gamma}_u$ are p_2-dimensional multivariate normal with mean zero and variance covariance matrix $\boldsymbol{\Sigma}_\gamma$.

In the following we will consider thoroughly a more complex example, where two stress variables are involved.

Example 2. *In this example the degradation is influenced by two standardized accelerating stress variables x_1 and x_2 which act linearly on the response with a potential interaction effect associated with $x_1 x_2$. The two stress variables x_1 and x_2 can be chosen independently from marginal design regions $\mathcal{X}_1 = \mathcal{X}_2 = [0, 1]$, respectively. Also the time is assumed to act linearly on the degradation and all interactions between stress variables and time are present as in Example 1.*

If, for testing unit i, the stress variables are set to x_{i1} and x_{i2} the response y_{ij} at time t_j is given by

$$y_{ij} = \beta_{i,1} + \beta_2 x_{i1} + \beta_3 x_{i2} + \beta_4 x_{i1} x_{i2} + \beta_{i,5} t_j + \beta_6 x_{i1} t_j + \beta_7 x_{i2} t_j + \beta_8 x_{i1} x_{i2} t_j + \varepsilon_{ij}, \tag{3.9}$$

where the intercept $\beta_{i,1}$ is the mean degradation of unit i at time $t = 0$ under the stress levels $x_1 = 0$ and $x_2 = 0$, β_2 is the common (not unit specific) mean increase in degradation depending on the stress variable x_1 when $x_2 = 0$, β_3 is the common mean increase in degradation depending on the stress variable x_2 when $x_1 = 0$, and β_4 is the interaction effect between the two stress variables. Accordingly $\beta_{i,5}$ is the mean increase in degradation of unit i over time t when the stress levels are set to $x_1 = 0$ and $x_2 = 0$, β_6 is the interaction effect between time and the stress variable x_1 when $x_2 = 0$, β_7 is the interaction effect between time and the stress variable x_2 when $x_1 = 0$, and β_8 is the second-order interaction effect between time and the two stress variables. Also here only the parameters $\beta_{i,1}$ and $\beta_{i,5}$ associated with the constant term in the stress variables may vary across units. On the aggregate level these two unit parameters are again assumed to be normally distributed with means $\mathrm{E}(\beta_{i,1}) = \beta_1$ and $\mathrm{E}(\beta_{i,5}) = \beta_5$ and variance covariance matrix $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$.

After rearranging terms and relabeling the parameters the model can be rewritten as

$$Y_{ij} = (\mathbf{f}_{11}(x_{i1}) \otimes \mathbf{f}_{12}(x_{i2}) \otimes \mathbf{f}_2(t_j))^T \boldsymbol{\beta} + \mathbf{f}_2(t_j)^T \boldsymbol{\gamma}_i + \varepsilon_{ij}, \tag{3.10}$$

where $\mathbf{f}_{11}(x_1) = (1, x_1)^T$, $\mathbf{f}_{12}(x_2) = (1, x_2)^T$ *and* $\mathbf{f}_2(t) = (1, t)^T$ *are the marginal regression functions for the stress variables* x_1, x_2 *and the time variable* t, *respectively,* $\boldsymbol{\beta} = (\beta_{111}, \beta_{112}, \beta_{121}, \beta_{122}, \beta_{211}, \beta_{212},$ $\beta_{221}, \beta_{222})^T = (\beta_1, \beta_5, \beta_3, \beta_7, \beta_2, \beta_6, \beta_4, \beta_8)^T$ *is the rearranged vector of aggregate parameters, and* $\boldsymbol{\gamma}_i = (\gamma_{i1}, \gamma_{i2})^T = (\beta_{i,1} - \beta_1, \beta_{i,5} - \beta_5)^T$ *is the vector of parameters for the deviations of unit i from the aggregate values. These deviations constitute again random effects with zero mean and variance covariance matrix* $\boldsymbol{\Sigma}$. *With* $\mathbf{x} = (x_1, x_2)$, $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 = [0, 1]^2$, $\mathbf{f}_1(\mathbf{x}) = (\mathbf{f}_{11}(x_1)^T, \mathbf{f}_{12}(x_2)^T)^T$, $p_1 = 4$, $p_2 = 2$, *and* $p = 8$ *model 3.10 fits into the framework of the general product-type model 3.5.*

Later we will also consider shortly a model in two stress variables without interaction terms between the stress variables.

3.4 Estimation of the model parameters

Under the distributional assumptions of normality for both the random effects and the measurement errors the model parameters may be estimated by means of the maximum likelihood method. Denote by $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \boldsymbol{\varsigma}^T)^T$ the vector of all model parameters, where $\boldsymbol{\varsigma}$ collects all variance covariance parameters from $\boldsymbol{\Sigma}_\gamma$ and $\boldsymbol{\Sigma}_\varepsilon$ For the general model (3.7) the log-likelihood is given by

$$\ell(\boldsymbol{\theta}; \mathbf{y}) = -\tfrac{nk}{2} \log(2\pi) - \tfrac{n}{2} \log(\det(\mathbf{V})) - \tfrac{1}{2}(\mathbf{y} - (\mathbf{F}_1 \otimes \mathbf{F}_2)\boldsymbol{\beta})^T (\mathbf{I}_n \otimes \mathbf{V})^{-1}(\mathbf{y} - (\mathbf{F}_1 \otimes \mathbf{F}_2)\boldsymbol{\beta}),$$

$$(3.11)$$

where the variance covariance matrix $\mathbf{V} = \mathbf{V}(\boldsymbol{\varsigma})$ of measurements per unit depends only on $\boldsymbol{\varsigma}$. The maximum likelihood estimator of $\boldsymbol{\beta}$ can be calculated as

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= ((\mathbf{F}_1 \otimes \mathbf{F}_2)^T (\mathbf{I}_n \otimes \hat{\mathbf{V}})^{-1}(\mathbf{F}_1 \otimes \mathbf{F}_2))^{-1}(\mathbf{F}_1 \otimes \mathbf{F}_2)^T (\mathbf{I}_n \otimes \hat{\mathbf{V}})^{-1}\mathbf{Y} \\ &= ((\mathbf{F}_1^T \mathbf{F}_1)^{-1}\mathbf{F}_1^T) \otimes ((\mathbf{F}_2^T \hat{\mathbf{V}}^{-1}\mathbf{F}_2)^{-1}\mathbf{F}_2^T \hat{\mathbf{V}}^{-1})\mathbf{Y}, \end{aligned} \qquad (3.12)$$

if both $\mathbf{F}_1$ and $\mathbf{F}_2$ are of full column rank p_1 and p_2, respectively, and $\widehat{\mathbf{V}} = \mathbf{V}(\widehat{\boldsymbol{\varsigma}})$, where $\widehat{\boldsymbol{\varsigma}}$ is the maximum likelihood estimator of $\boldsymbol{\varsigma}$. When $\mathbf{V}$ is known, at least up to a multiplicative constant $\mathbf{V} = \sigma^2 \mathbf{V}_0$, then $\widehat{\boldsymbol{\beta}}$ is the best liner unbiased (general least squares) estimator $\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}} = ((\mathbf{F}_1^T \mathbf{F}_1)^{-1}\mathbf{F}_1^T) \otimes ((\mathbf{F}_2^T \mathbf{V}_0^{-1}\mathbf{F}_2)^{-1}\mathbf{F}_2^T \mathbf{V}_0^{-1})\mathbf{Y}$ of $\boldsymbol{\beta}$. In particular, when the measurement errors are uncorrelated and homoscedastic, i.e. $\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_k$, then this estimator reduces to the ordinary least squares estimator $\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}} = ((\mathbf{F}_1^T \mathbf{F}_1)^{-1}\mathbf{F}_1^T) \otimes ((\mathbf{F}_2^T \mathbf{F}_2)^{-1}\mathbf{F}_2^T)\mathbf{Y}$ because $\mathbf{V}\mathbf{F}_2 = \mathbf{F}_2(\boldsymbol{\Sigma}_\gamma \mathbf{F}_2^T \mathbf{F}_2 + \sigma_\varepsilon^2 \mathbf{I}_{p_2})$ by a result of (Zyskind, 1967). Hence, in the case of uncorrelated homoscedastic measurement errors the maximum likelihood estimator of the location parameters $\boldsymbol{\beta}$ does neither depend on the variance covariance parameters nor on their estimates.

In general, the quality of the estimator $\widehat{\boldsymbol{\beta}}$ can be measured in terms of its variance covariance matrix which is given by

$$\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \otimes (\mathbf{F}_2^T \mathbf{V}^{-1} \mathbf{F}_2)^{-1}. \tag{3.13}$$

By using the structure $\mathbf{V} = \mathbf{F}_2 \boldsymbol{\Sigma}_\gamma \mathbf{F}_2^T + \boldsymbol{\Sigma}_\varepsilon$ the last term can be calculated as

$$(\mathbf{F}_2^T \mathbf{V}^{-1} \mathbf{F}_2)^{-1} = (\mathbf{F}_2^T \boldsymbol{\Sigma}_\varepsilon^{-1} \mathbf{F}_2)^{-1} + \boldsymbol{\Sigma}_\gamma \tag{3.14}$$

in terms of the variance covariance matrices $\boldsymbol{\Sigma}_\gamma$ and $\boldsymbol{\Sigma}_\varepsilon$ of the random effects and the measurement errors, respectively.

3.5 Information

In general, the Fisher information matrix is defined as the variance covariance matrix of the score function $\mathbf{U}$ which itself is defined as the vector of first derivatives of the log likelihood with respect to the components of the parameter vector $\boldsymbol{\theta}$. More precisely, let $\mathbf{U} = (\frac{\partial}{\partial \theta_1} \ell(\boldsymbol{\theta}; \mathbf{y}), ..., \frac{\partial}{\partial \theta_q} \ell(\boldsymbol{\theta}; \mathbf{y}))^T$, where q is the dimension of $\boldsymbol{\theta}$. Then for the full parameter vector $\boldsymbol{\theta}$ the Fisher information matrix is defined as $\mathbf{M}_{\boldsymbol{\theta}} = \mathrm{Cov}(\mathbf{U})$, where the expectation is taken with respect to the distribution of $\mathbf{Y}$. The Fisher information can also be computed as the expectations of the second derivatives of the score function U, i.e. $\mathbf{M}_{\boldsymbol{\theta}} = -\mathrm{E}\left(\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \ell(\boldsymbol{\theta}; \mathbf{y})\right)$. Under common regularity conditions the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is consistent and asymptotically normal with asymptotic variance covariance matrix equal to the inverse $\mathbf{M}_{\boldsymbol{\theta}}^{-1}$ of the Fisher information matrix $\mathbf{M}_{\boldsymbol{\theta}}$.

To specify the Fisher information matrix further, denote by $\mathbf{M}_{\boldsymbol{\beta}} = -\mathrm{E}\left(\frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \ell(\boldsymbol{\theta}; \mathbf{y})\right)$, $\mathbf{M}_{\boldsymbol{\varsigma}} = -\mathrm{E}\left(\frac{\partial^2}{\partial \boldsymbol{\varsigma} \partial \boldsymbol{\varsigma}^T} \ell(\boldsymbol{\theta}; \mathbf{y})\right)$, $\mathbf{M}_{\boldsymbol{\beta}\boldsymbol{\varsigma}} = -\mathrm{E}\left(\frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\varsigma}^T} \ell(\boldsymbol{\theta}; \mathbf{y})\right)$ and $\mathbf{M}_{\boldsymbol{\varsigma}\boldsymbol{\beta}} = \mathbf{M}_{\boldsymbol{\beta}\boldsymbol{\varsigma}}^T$ the blocks of the Fisher information matrix corresponding to the second derivatives with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\varsigma}$ and the mixed derivatives, respectively. The mixed blocks can be seen to be zero and the Fisher information matrix is block diagonal,

$$\mathbf{M}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathbf{M}_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\boldsymbol{\varsigma}} \end{pmatrix}. \tag{3.15}$$

Moreover, the block $\mathbf{M}_{\boldsymbol{\beta}}$ associated with the aggregate location parameters $\boldsymbol{\beta}$ can be determined as

$$\mathbf{M}_{\boldsymbol{\beta}} = (\mathbf{F}_1^T \mathbf{F}_1) \otimes (\mathbf{F}_2^T \mathbf{V}^{-1} \mathbf{F}_2) \tag{3.16}$$

which turns out to be the inverse of the variance covariance matrix for the estimator $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, when $\mathbf{V}$ is known. Actually, because the Fisher information matrix for $\boldsymbol{\theta}$ is block

diagonal, the inverse $\mathbf{M}_{\beta}^{-1} = (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \otimes (\mathbf{F}_2^T \mathbf{V}^{-1} \mathbf{F}_2)^{-1}$ of the block associated with $\boldsymbol{\beta}$ is the corresponding block of the inverse of $\mathbf{M}_{\theta}$ and is, hence, the asymptotic variance covariance matrix of $\widehat{\boldsymbol{\beta}}$.

Accordingly the asymptotic variance covariance matrix for estimating the variance parameters $\boldsymbol{\varsigma}$ is the inverse of the block $\mathbf{M}_{\varsigma}$. In the following we will call $\mathbf{M}_{\beta}$ and $\mathbf{M}_{\varsigma}$ the information matrices for $\boldsymbol{\beta}$ and $\boldsymbol{\varsigma}$, respectively, for short. The particular form of $\mathbf{M}_{\varsigma}$ will be not of interest here. However, as the information matrix $\mathbf{M}_{\varsigma}$ for the variance parameters $\boldsymbol{\varsigma}$ is given by

$$\mathbf{M}_{\varsigma} = \frac{n}{2} \left(\frac{\partial^2 \log(\det(\mathbf{V}))}{\partial \boldsymbol{\varsigma} \partial \boldsymbol{\varsigma}^T} + \mathrm{tr}\left(\mathbf{V} \frac{\partial^2 \mathbf{V}^{-1}}{\partial \boldsymbol{\varsigma} \partial \boldsymbol{\varsigma}^T} \right) \right).$$

It is important to note that $\mathbf{M}_{\varsigma}$ does not depend on the settings $\mathbf{x}_1, ..., \mathbf{x}_n$ of the stress variable in contrast to the information matrix $\mathbf{M}_{\beta}$ of the aggregate location parameters $\boldsymbol{\beta}$. For the general product-type model (3.7) the information matrix $\mathbf{M}_{\beta}$ for the aggregate parameters $\boldsymbol{\beta}$ factorizes according to

$$\mathbf{M}_{\beta} = \mathbf{M}_1 \otimes \mathbf{M}_2 \tag{3.17}$$

into the information matrix $\mathbf{M}_1 = \mathbf{F}_1^T \mathbf{F}_1$ in the marginal model

$$Y_i^{(1)} = \mathbf{f}_1(\mathbf{x}_i)^T \boldsymbol{\beta}^{(1)} + \varepsilon_i^{(1)}, \tag{3.18}$$

$i = 1, ..., n$, in the stress variable $\mathbf{x}$ with standardized uncorrelated homoscedastic error terms, $\sigma_{\varepsilon^{(1)}}^2 = 1$, and the information matrix $\mathbf{M}_2 = \mathbf{F}_2^T \mathbf{V}^{-1} \mathbf{F}_2$ in the mixed effects marginal model

$$Y_j^{(2)} = \mathbf{f}_2(t_j)^T \boldsymbol{\beta}^{(2)} + \mathbf{f}_2(t_j)^T \boldsymbol{\gamma}^{(2)} + \varepsilon_j^{(2)}, \tag{3.19}$$

$j = 1, ..., k$, in the time variable t with variance covariance matrices $\boldsymbol{\Sigma}_{\gamma}$ and $\boldsymbol{\Sigma}_{\varepsilon}$ for the random effects $\boldsymbol{\gamma}^{(2)}$ and measurement errors $\boldsymbol{\varepsilon}^{(2)} = (\varepsilon_1^{(2)}, ..., \varepsilon_k^{(2)})^T$, respectively. Then the information matrix $\mathbf{M}_{\theta}$ in the full model depends on the settings $\mathbf{x}_1, ..., \mathbf{x}_n$ of the stress variable only through the information matrix $\mathbf{M}_1$ in the first marginal model.

3.6 Design

The quality of the estimates will be measured in terms of the information matrix and, hence, depends on both the settings of the stress variable and the time points of measurements. When these variables are under the control of the experimenter, then their choice will be called the design of the experiment. Here we assume that the time plan $\mathbf{t} = (t_1, ..., t_k)^T$ for the time points of measurements within units is fixed in advance

and is not under disposition of the experimenter. Then only the settings $\mathbf{x}_1, ..., \mathbf{x}_n$ of the stress variable $\mathbf{x}$ can be adjusted to the units $i = 1, ..., n$. Their choice $(\mathbf{x}_1, ..., \mathbf{x}_n)$ is then called an "exact" design, and their influence on the performance of the experiment is indicated by adding them as an argument to the information matrices, $\mathbf{M}_{\boldsymbol{\theta}}(\mathbf{x}_1, ..., \mathbf{x}_n)$, $\mathbf{M}_{\boldsymbol{\beta}}(\mathbf{x}_1, ..., \mathbf{x}_n)$, and $\mathbf{M}_1(\mathbf{x}_1, ..., \mathbf{x}_n)$, where appropriate. Remind that both $\mathbf{M}_{\boldsymbol{\varsigma}}$ and $\mathbf{M}_2$ do not depend on the design for the stress variable.

As $\mathbf{M}_1(\mathbf{x}_1, ..., \mathbf{x}_n) = \sum_{i=1}^{n} \mathbf{f}_1(\mathbf{x}_i)\mathbf{f}_1(\mathbf{x}_i)^T$ it can easily be seen that the information matrices do not depend on the order of the setting but only on their mutually distinct settings, $\mathbf{x}_1, ..., \mathbf{x}_m$ say, and their corresponding frequencies $n_1, ..., n_m$, such that $\sum_{i=1}^{m} n_i = n$, i.e. $\mathbf{M}_1 = \sum_{i=1}^{m} n_i\mathbf{f}_1(\mathbf{x}_i)\mathbf{f}_1(\mathbf{x}_i)^T$. Finding optimal exact designs is, in general, a difficult task of discrete optimization. To circumvent this problem we follow the approach of approximate designs ξ as introduced in Subsection 2.3.1.The corresponding standardized, per unit information matrices are accordingly defined as

$$\mathbf{M}_1(\xi) = \sum_{i=1}^{m} w_i\mathbf{f}_1(\mathbf{x}_i)\mathbf{f}_1(\mathbf{x}_i)^T \tag{3.20}$$

for the marginal model on itself or by plugging (3.20) in into the standardized, per unit information matrix

$$\mathbf{M}_{\boldsymbol{\beta}}(\xi) = \mathbf{M}_1(\xi) \otimes \mathbf{M}_2 \tag{3.21}$$

for the aggregate parameters $\boldsymbol{\beta}$, where again $\mathbf{M}_2 = \mathbf{F}_2^T\mathbf{V}^{-1}\mathbf{F}_2$, and

$$\mathbf{M}_{\boldsymbol{\theta}}(\xi) = \begin{pmatrix} \mathbf{M}_1(\xi) \otimes \mathbf{M}_2 & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{M}_{\boldsymbol{\varsigma}}} \end{pmatrix} \tag{3.22}$$

or the full parameter vector $\boldsymbol{\theta}$, where now $\widetilde{\mathbf{M}_{\boldsymbol{\varsigma}}} = \frac{1}{n}\mathbf{M}_{\boldsymbol{\varsigma}}$ is the standardized, per unit information for the variance parameters $\boldsymbol{\varsigma}$. If all nw_i are integer, then these standardized versions coincide with the information matrices of the corresponding exact design up to the normalizing factor $1/n$ and are , hence, an adequate generalization.

In order to optimize information matrices, some optimality criterion has to be employed which is a real valued function of the information matrix and reflects the main interest in the experiment.

3.7 Optimality criterion based on the failure time under normal use condition

As in (Weaver and Meeker, 2013) we are interested in some characteristics of the failure time distribution of soft failure due to degradation (see the introductory example in

Section 3.2). Therefore it is assumed that the model equation (3.8) $\mu_u(t) = \mu(\mathbf{x}_u, t) = (\mathbf{f}_1(\mathbf{x}_u) \otimes \mathbf{f}_2(t))^T \boldsymbol{\beta} + \mathbf{f}_2(t)^T \boldsymbol{\gamma}_u$ for the mean degradation paths is also valid under normal use condition $\mathbf{x}_u$, where μ_u denotes the degradation path under normal use condition for short. We further denote by $\mu(t) = \mathrm{E}(\mu_u(t)) = (\mathbf{f}_1(\mathbf{x}_u) \otimes \mathbf{f}_2(t))^T \boldsymbol{\beta}$ the aggregate degradation path under normal use condition and by $\boldsymbol{\delta} = \boldsymbol{\delta}(\boldsymbol{\beta}) = (\delta_1(\boldsymbol{\beta}), ..., \delta_{p_2}(\boldsymbol{\beta}))^T$ the vector of its coefficients $\delta_s = \delta_s(\boldsymbol{\beta}) = \sum_{r=1}^{p_1} f_{1r}(\mathbf{x}_u)\beta_{rs}$, $s = 1, ..., p_2$, in the regression functions f_{2s} in t, i.e. $\mu(t) = \mathbf{f}_2(t)^T \boldsymbol{\delta} = \sum_{s=1}^{p_2} \delta_s f_{2s}(t)$.

For the following it is assumed that the mean degradation paths are strictly increasing over time. Then a soft failure due to degradation is defined as the exceedance of the degradation over a failure threshold y_0. This definition is based on the mean degradation path and not on a "real" path subject to measurement errors. The failure time T under normal use condition is then defined as the first time t the mean degradation path $\mu_u(t)$ reaches or exceeds the threshold y_0, i.e. $T = \min\{t \geq 0; \mu_u(t) \geq y_0\}$. As random effects $\boldsymbol{\gamma}_u$ are involved in the mean degradation path, the failure time T is random. Actually, T may become infinite, if the mean degradation path does not reach the threshold, or may degenerate to $T = 0$, if the degradation already exceeds the threshold at time $t = 0$, because of unfortunate values of the random effects $\boldsymbol{\gamma}_u$, but this will happen only with low probability and will not affect the further argumentation.

In order to describe certain characteristics of the distribution of the failure time, we will determine the distribution function $F_T(t) = \mathrm{P}(T \leq t)$. First note that $T \leq t$ if and only if $\mu_u(t) \geq y_0$. Hence

$$
\begin{aligned}
F_T(t) &= \mathrm{P}(\mu_u(t) \geq y_0) \\
&= \mathrm{P}(\mu(t) + \mathbf{f}_2(t)^T \boldsymbol{\gamma}_u \geq y_0) \\
&= \mathrm{P}(-\mathbf{f}_2(t)^T \boldsymbol{\gamma}_u \leq \mu(t) - y_0) \\
&= \Phi(h(t)),
\end{aligned} \tag{3.23}
$$

where

$$
h(t) = \frac{\mu(t) - y_0}{\sigma_u(t)}, \tag{3.24}
$$

$\sigma_u^2(t) = \mathbf{f}_2(t)^T \boldsymbol{\Sigma}_\gamma \mathbf{f}_2(t)$ is the variance of the mean degradation path $\mu_u(t)$ at time t, and Φ denotes the distribution function of the standard normal distribution. Here it is tacitly assumed that the variance $\sigma_u^2(t)$ of the mean degradation path is greater than zero for every $t \geq 0$. This condition is satisfied, in particular, when the variance covariance matrix $\boldsymbol{\Sigma}_\gamma$ of the random effects is positive definite.

We will be interested in quantiles t_α of the failure time distribution, i.e. $\mathrm{P}(T \leq t_\alpha) = \alpha$. For each α the quantile t_α gives the time up to which under normal use conditions (at least) $\alpha \cdot 100$ percent of the units fail and (at least) $(1 - \alpha) \cdot 100$ percent

of the units persist. The quantiles t_α are increasing in α. Note that this standard definition of quantiles is in contrast to the "upper" quantiles $(t_{1-\alpha})$ used in (Weaver and Meeker, 2013) where percentages of failures and persistence are reversed. Of particular importance is the median $t_{0.5}$ up to which under normal use conditions half of the units fails and half of the units persist ($\alpha = 0.5$). Other characteristics of interest may be the five or ten percent quantiles $t_{0.05}$ and $t_{0.1}$ which give the times up to which 95 or 90 % percent of the units persist, respectively. By (3.24) these quantiles can be determined as the solutions of the equation

$$h(t_\alpha) = z_\alpha, \tag{3.25}$$

where $z_\alpha = \Phi^{-1}(\alpha)$ is the α-quantile of the standard normal distribution. For the median ($\alpha = 1/2$) we have $z_{0.5} = 0$ and, hence the median failure time $t_{0.5}$ is the solution of $\mu(t) = y_0$, i. e. the aggregate degradation path reaches the threshold at time $t_{0.5}$. Note that the function h represents the failure time distribution function F_T on a normal Q-Q-plot scale.

In the particular case of straight lines for the mean degradation paths, i. e. $\mathbf{f}_2(t) = (1, t)^T$ as considered in Examples 1 and 2, the function $h(t)$ specifies to

$$h(t) = \frac{\delta_2 t + \delta_1 - y_0}{\sqrt{\sigma_1^2 + 2\rho\sigma_1\sigma_2 t + \sigma_2^2 t^2}}, \tag{3.26}$$

where $\delta_1 = \sum_{r=1}^{p_1} f_{1r}(\mathbf{x}_u)\beta_{r1}$ and $\delta_2 = \sum_{r=1}^{p_1} f_{1r}(\mathbf{x}_u)\beta_{r2}$ are the intercept and the slope of the aggregate degradation path $\mu(t) = \delta_1 + \delta_2 t$ under normal use condition, respectively. The median failure time is then given by $t_{0.5} = (y_0 - \delta_1)/\delta_2$ which provides a proper solution $t_{0.5} > 0$ under the natural assumptions that the aggregate degradation path is increasing, $\delta_2 > 0$, and that the aggregate degradation at the beginning of the testing at time $t = 0$ is less than the threshold of soft failure, $\delta_1 < y_0$.

Example (Example 1 cont.)**.** *In the introductory example the aggregate degradation path $\mu(t) = \delta_1 + \delta_2 t$ has intercept $\delta_1 = \beta_1 + \beta_2 x_u$ and slope $\delta_2 = \beta_3 + \beta_4 x_u$. Hence, the median failure time is given by $t_{0.5} = (y_0 - \beta_1 - \beta_2 x_u)/(\beta_3 + \beta_4 x_u)$. If we use the standardized nominal values of Table 3.1 for Example 7.2 by (Weaver and Meeker, 2013), the aggregate degradation path becomes $\mu(t) = 2.306 + 1.014t$ under normal use condition, and the median failure time is $t_{0.5} = 1.583$. Note that, as typical for degradation experiments, the median failure time is larger than the maximal experimental time $t_{\max} = 1$.*

Under the additional assumption that the correlation of the random effects is non-negative for the intercept and the slope of the mean degradation path, $\rho \geq 0$, the function $h(t)$ can be seen to be strictly increasing, $h'(t) > 0$, in $t > 0$. This also remains true for small to moderate negative correlations. However, the range of $h(t)$ is bounded

and does not cover the whole real line such that not all quantiles are non-degenerate. For small α the α-quantile to be positive requires $z_\alpha > h(0) = -(y_0 - \delta_1)/\sigma_1$, i.e. the variance σ_1^2 of the intercept of the mean degradation path has to be sufficiently small compared to the distance from its mean δ_1 to the threshold y_0. In particular, in the case of the 5%-quantile $\sigma_1 < 0.608(y_0 - \delta_1)$ is needed for $t_{0.05} > 0$. For large α the α-quantile is finite if $z_\alpha < \lim_{t\to\infty} h(t) = \delta_2/\sigma_2$, i.e. the variance σ_2^2 of the slope of the mean degradation path has to be sufficiently small compared to its mean δ_2. Note that $\Phi(h(0)) = 1 - \Phi((y_0 - \delta_1)/\sigma_1)$ is the probability that under normal use condition the mean degradation path exceeds the threshold y_0 already at the initial time $t = 0$. Note also that formally $1 - \Phi(\lim_{t\to\infty} h(t)) = \Phi(-\delta_2/\sigma_2)$ is the probability that the mean degradation path has a negative slope which may be interpreted as the probability that soft failure due to degradation will not occur at all under normal use condition. When the α-quantile is non-degenerate ($0 < t_\alpha < \infty$), then t_α is a solution of the quadratic equation

$$(\delta_2 t + \delta_1 - y_0)^2 = z_\alpha^2(\sigma_1^2 + 2\rho\sigma_1\sigma_2 t + \sigma_2^2 t^2),$$

as indicated by (Weaver and Meeker, 2013) for the situation of Example 1. In the special case of only a random intercept in the random effects, i.e. $\sigma_2^2 = 0$, all α-quantiles t_α finitely exist for $\alpha \geq \Phi(-(y_0 - \delta_1)/\sigma_1)$ and can be determined as the solution of a linear equation to $t_\alpha = (y_0 - \delta_1 + z_\alpha\sigma_1)/\delta_2$.

Example (Example 1 cont.)**.** *For the introductory example the function $h(t)$ is plotted in Figure 3.4 under the standardized nominal values of (Weaver and Meeker, 2013) given in Table 3.1. The defining function $h(t)$ is seen to be strictly increasing although the correlation is moderately negative ($\rho = -0.143$). Thus the distribution function $F_T(t) = \Phi(h(t))$ is well-defined, and it is represented in Figure 3.3. In both plots the median failure time $t_{0.5} = 1.583$ is indicated by a dashed vertical line. Moreover, as $h(0) = -14.03$ and $\lim_{t\to\infty} h(t) = 9.67$, the range of h covers all reasonable quantiles.*

It has to be noted that in the case of the standardized nominal values of Table 3.1 for high stress levels ($x_h = 1$) the mean degradation path μ_i exceeds the threshold y_0 for soft failure due to degradation with high probability ($\mathrm{P}(\mu_i(1,0) \geq y_0) > 1/2$) already at the initial experimental time $t_{\min} = 0$. Hence, care has to be taken that the model equation for the mean degradation paths is also valid beyond the threshold, i.e. in the case that soft failure has already occurred. To avoid this complication we consider in Example 2 nominal values which guarantee that soft failure occurs during the experiment only with negligible probability.

Example (Example 2 cont.)**.** *For the model with two interacting stress variables x_1 and x_2 we consider the virtual nominal values for the parameters, normal use conditions and threshold given in Table 3.2. The aggregate degradation path $\mu(t)$ is*
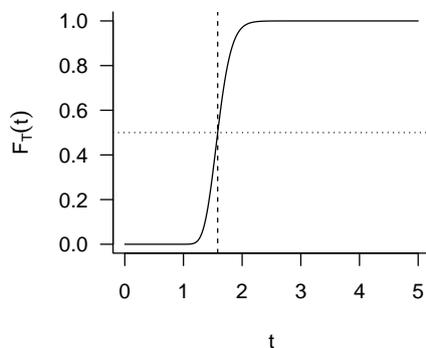
FIGURE 3.3: Failure time distribution F_T for Example 1
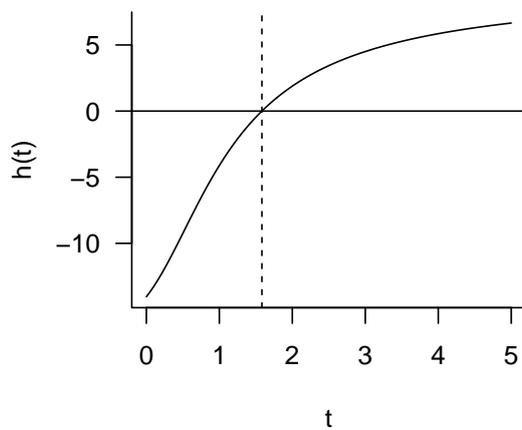
FIGURE 3.4: Defining function h for Example 1

TABLE 3.2: Nominal values for Example 2

$f_j(\mathbf{x}, t)$	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	σ_γ	σ_ε	x_{u1}	x_{u2}	y_0
	1	x_1	x_2	$x_1 x_2$	t	$x_1 t$	$x_2 t$	$x_1 x_2 t$					
	4.0	1.5	0.75	1.8	0.5	0.25	0.25	4.03	0.7	0.85	-0.5	-0.4	14.39
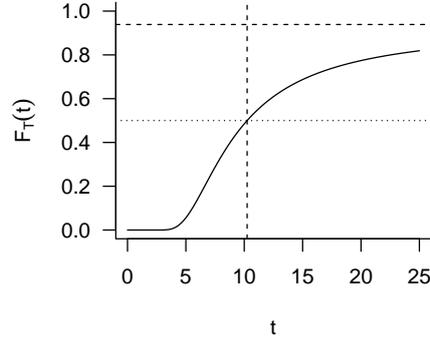
FIGURE 3.5: Failure time distribution F_T for Example 2
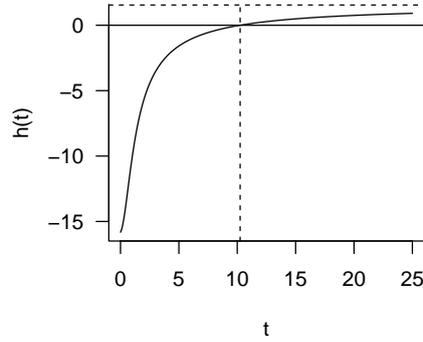
FIGURE 3.6: Defining function h for Example 2

a straight line with intercept $\delta_1 = \beta_1 + \beta_2 x_{u1} + \beta_3 x_{u2} + \beta_4 x_{u1} x_{u2} = 3.31$ and slope $\delta_2 = \beta_5 + \beta_6 x_{u1} + \beta_7 x_{u2} + \beta_8 x_{u1} x_{u2} = 1.08$, i. e. $\mu(t) = 3.31 + 1.08t$ under normal use condition. With a threshold of $y_0 = 14.39$ for soft failure the median failure time results in $t_{0.5} = (y_0 - \delta_1)/\delta_2 = 10.25$ which is substantially larger than the maximal experimental time $t_{\max} = 1$. For the characterization of other quantiles the function $h(t)$ is plotted in Figure 3.6 together with the corresponding distribution function $F_T(t) = \Phi(h(t))$ in Figure 3.5. The median failure time $t_{0.5} = 10.25$ is indicated in both plots by a dashed vertical line. As $\rho = 0$ the function $h(t)$ is strictly increasing and ranges from $h(0) = -15.83$ to $h_{\max} = \lim_{t \to \infty} h(t) = 1.54$. Thus, quantiles t_α are non-degenerate as long as $\alpha \leq \alpha_{\max}$, where $\alpha_{\max} = \Phi(h_{\max}) = 0.939$, and $(1 - \alpha_{\max}) \cdot 100 = 6.1$ percent of the mean degradation paths do not lead to a soft failure. Both $\alpha_{\max}$ and $h_{\max}$ are indicated in the respective plots by a dashed horizontal line. Note that for the nominal values of Table 3.2 the mean degradation $\mu(\mathbf{x}, t)$ under experimental conditions attains its maximum 13.08 for the maximal stress levels ($x_1 = x_2 = 1$) and maximal

experimental time ($t = 1$), and, hence, the mean degradation paths do not exceed the threshold for all experimental settings.

In any case the quantile $t_\alpha = t_\alpha(\boldsymbol{\theta})$ is a function of both the aggregate location parameters $\boldsymbol{\beta}$ and the variance parameters $\boldsymbol{\varsigma}$, in general. Hence, the maximum likelihood estimator of the quantile t_α is given by $\widehat{t}_\alpha = t_\alpha(\widehat{\boldsymbol{\theta}})$ in terms of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$. The task of designing the experiment will now be to provide an as precise estimate of the α-quantile as possible.

By the delta-method $\widehat{t}_\alpha$ is seen to be asymptotically normal with asymptotic variance

$$\mathrm{aVar}(\widehat{t}_\alpha) = \mathbf{c}^T \mathbf{M}_{\boldsymbol{\theta}}^{-1} \mathbf{c}, \tag{3.27}$$

where $\mathbf{c} = \frac{\partial}{\partial \boldsymbol{\theta}} t_\alpha$ is the gradient vector of partial derivatives of t_α with respect to the components of the parameter vector $\boldsymbol{\theta}$. The asymptotic variance depends on the design of the experiment through the information matrix $\mathbf{M}_{\boldsymbol{\theta}}$ and will be chosen as the optimality criterion for the design.

The gradient $\mathbf{c}$ can be seen to be equal to

$$\mathbf{c} = -c_0 \left(\frac{\partial}{\partial \boldsymbol{\theta}} \mu(t)|_{t=t_\alpha} - z_\alpha \frac{\partial}{\partial \boldsymbol{\theta}} \sigma_u(t)|_{t=t_\alpha} \right), \tag{3.28}$$

in view of (3.24) and (3.25) by the implicit function theorem (see e. g. (Krantz and Parks, 2012)), where $c_0 = 1/(\mu'(t_\alpha) - z_\alpha \sigma_u'(t_\alpha))$ is the inverse of the derivative of the defining function h with respect to t.

As the aggregate mean degradation $\mu(t)$ only depends on the aggregate location parameters $\boldsymbol{\beta}$ and the variance $\sigma_u^2(t)$ only depends on the variance parameters $\boldsymbol{\varsigma}$ the gradient simplifies to $\mathbf{c} = -c_0(\mathbf{c}_{\boldsymbol{\beta}}^T, \mathbf{c}_{\boldsymbol{\varsigma}}^T)^T$, where

$$\mathbf{c}_{\boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \mu(t)|_{t=t_\alpha} = \mathbf{f}(\mathbf{x}_u, t_\alpha)$$

is the gradient of $\mu(t)$ with respect to $\boldsymbol{\beta}$ and

$$\mathbf{c}_{\boldsymbol{\varsigma}} = -z_\alpha \frac{\partial}{\partial \boldsymbol{\varsigma}} \sigma_u(t)|_{t=t_\alpha}$$

is $-z_\alpha$ times the gradient of $\sigma_u(t)$ with respect to $\boldsymbol{\varsigma}$. The particular shape of $\mathbf{c}_{\boldsymbol{\varsigma}}$ does not play a role here, in general. But note that $\mathbf{c}_{\boldsymbol{\varsigma}} = \mathbf{0}$ in the case of the median ($\alpha = 0.5$).

By the block diagonal form (3.15) of the information matrix the asymptotic variance (3.27) of $\widehat{t}_\alpha$ becomes

$$\mathrm{aVar}(\widehat{t}_\alpha) = c_0^2 \left(\mathbf{c}_{\boldsymbol{\beta}}^T \mathbf{M}_{\boldsymbol{\beta}}^{-1} \mathbf{c}_{\boldsymbol{\beta}} + \mathbf{c}_{\boldsymbol{\varsigma}}^T \mathbf{M}_{\boldsymbol{\varsigma}}^{-1} \mathbf{c}_{\boldsymbol{\varsigma}} \right) \tag{3.29}$$

which simplifies to

$$\mathrm{aVar}(\widehat{t}_{0.5}) = c_0^2 \mathbf{c}_\beta^T \mathbf{M}_\beta^{-1} \mathbf{c}_\beta \tag{3.30}$$

in the case of the median.

For the product-type model (3.7) the expression related to the aggregate parameters $\boldsymbol{\beta}$ further decomposes,

$$\mathbf{c}_\beta^T \mathbf{M}_\beta^{-1} \mathbf{c}_\beta = \mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1^{-1} \mathbf{f}_1(\mathbf{x}_u) \cdot \mathbf{f}_2(t_\alpha)^T \mathbf{M}_2^{-1} \mathbf{f}_2(t_\alpha), \tag{3.31}$$

as $\mathbf{c}_\beta = \mathbf{f}_1(\mathbf{x}_u) \otimes \mathbf{f}_2(t_\alpha)$ factorizes.

3.8 Optimal designs in the case of predetermined measurement times

From (3.29) and (3.31) it can be seen that for obtaining a minimal asymptotic variance for $\widehat{t}_\alpha$ only $\mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1^{-1} \mathbf{f}_1(\mathbf{x}_u)$ has to be minimized, because all other terms do not depend on the experimental settings $\mathbf{x}_1, ..., \mathbf{x}_n$ of the stress variable, when the measurement times $t_1, ..., t_k$ are predetermined. The optimality criterion of minimization of the asymptotic variance of $\widehat{t}_\alpha$ thus reduces to a c-criterion $\mathbf{c}_1^T \mathbf{M}_1(\xi)^{-1} \mathbf{c}_1$ for extrapolation of the marginal response at normal use condition $\mathbf{x}_u$ in the first marginal model (3.18), $\mathbf{c}_1 = \mathbf{f}_1(\mathbf{x}_u)$, which is a well-known problem from the literature (see (Kiefer and Wolfowitz, 1964)). It is remarkable that this criterion and, hence, the corresponding optimal design is the same whatever the value of α is, as long as there is a proper solution $0 < t_\alpha < \infty$ for the α-quantile of the failure time.

Proposition 3.8.1. *If the design ξ^* is c-optimal for extrapolation of the mean response at the normal use condition in the marginal model (3.18) for the stress variable, then ξ^* minimizes the asymptotic variance for the estimator $\widehat{t}_\alpha$ of the α-quantile of the failure time for every α when $0 < t_\alpha < \infty$ (for predetermined measurement times $t_1, ..., t_k$).*

Although the normal use condition is typically outside the experimental region, the above proposition also would hold for interpolation, i. e. $\mathbf{x}_u \in \mathcal{X}$. The result of Proposition 3.8.1 is next used to derive optimal designs for the situation in Examples 1 and 2.

Example (Example 1 cont.)**.** *In the introductory model of Section 3.2 the marginal model for the stress variable x is given by a simple linear regression, $\mathbf{f}_1(x) = (1, x)^T$. In this marginal model the c-optimal design ξ^* for extrapolation of the mean response $\mu^{(1)}(x_u) = \beta_1^{(1)} + \beta_2^{(1)} x_u$ under normal use condition $x_u < 0$ assigns weight $w^* = |x_u|/(1+2|x_u|)$ to the highest stress level $x_h = 1$ and weight $1 - w^* = (1+|x_u|)/(1+2|x_u|)$*

to the lowest stress level $x_l = 0$ on the standardized scale $\mathcal{X} = [0, 1]$ (see (Kiefer and Wolfowitz, 1964)). Note that larger weight $1 - w^* > w^*$ is assigned to the lowest stress level $x_l = 0$ which is closer to $x_u < 0$ than $x_h = 1$ and that the weight $1 - w^*$ at x_l decreases from 1 to $1/2$, when the distance between the normal use condition and the experimental region gets larger, i. e. x_u decreases. For the standardized value $x_u = -0.056$ of the normal use condition from Table 3.1 the optimal weights for extrapolation at $x_u = -0.056$ are $w^* = 0.05$ at $x_h = 1$ and $1 - w^* = 0.95$ at $x_l = 0$, and the optimal design is

$$\xi^* = \left(\begin{array}{cc} 0 & 1 \\ 0.95 & 0.05 \end{array} \right).$$

Further examples for extrapolation at $x_u = -0.4, -0.5,$ and -1 give optimal weights $w^* = 0.22, 0.25,$ and 0.33 at $x_h = 1$, and $1 - w^* = 0.78, 0.75,$ and 0.67 at $x_l = 0$, respectively.

By Proposition 3.8.1 the design ξ^* is also optimal for minimization of the asymptotic variance for estimating the α-quantile t_α of the failure time for soft failure due to degradation under normal use condition x_u, when $0 < t_\alpha < \infty$ and the measurement times $t_1, ..., t_k$ are predetermined (see (Schwabe, Prus, and Graßhoff, 2014) for estimation of the median, $\alpha = 0.5$). In particular, for the standardized value $x_u = -0.056$ of the normal use condition from Table 3.1 the optimal design for estimating any α-quantile t_α assigns weight 0.95 to $x_l = 0$ and weight 0.05 to $x_h = 1$, as found numerically by (Weaver and Meeker, 2013) in the case of the median $t_{0.5}$.

The so obtained designs for one stress variable can now be used to construct *c*-optimal designs in the presence of two stress variables with interactions.

Example (Example 2 cont.)**.** *In the model with two interacting stress variables x_1 and x_2 the marginal model for the combined stress variable $\mathbf{x} = (x_1, x_2)$ is given itself by a product-type structure, $\mathbf{f}_1(\mathbf{x}) = \mathbf{f}_{11}(x_1) \otimes \mathbf{f}_{12}(x_2)$, where both components x_1 and x_2 are specified by $\mathbf{f}_{1v}(x_v) = (1, x_v)^T$ as simple linear regressions in their corresponding marginal models $Y_i^{(1v)} = \beta_1^{(1v)} + \beta_2^{(1v)} x_v + \varepsilon_i^{(1v)}$, $v = 1, 2$, with standardized homoscedastic and uncorrelated error terms. Moreover, the experimental region $\mathcal{X} = [0, 1]^2$ for the combined stress variable $\mathbf{x}$ is the Cartesian product of the marginal experimental regions $\mathcal{X}_1 = \mathcal{X}_2 = [0, 1]$ for the components x_1 and x_2, respectively. The vector $\mathbf{c}$ for extrapolation of the mean response $\mathbf{c}^T \boldsymbol{\beta}^{(1)} = \mu^{(1)}(\mathbf{x}_u) = \beta_1^{(1)} + \beta_2^{(1)} x_{u1} + \beta_3^{(1)} x_{u2} + \beta_4^{(1)} x_{u1} x_{u2}$ under normal use condition $\mathbf{x}_u = (x_{u1}, x_{u2})$, $x_{u1}, x_{u2} < 0$, is given by $\mathbf{f}_1(\mathbf{x}_u) = \mathbf{f}_{11}(x_{u1}) \otimes \mathbf{f}_{12}(x_{u2})$ and, hence, also factorizes as $\mathbf{c} = \mathbf{c}_1 \otimes \mathbf{c}_2$, where $\mathbf{c}_v = \mathbf{f}_{1v}(x_{uv})$. In this setting the c-optimal design ξ^* for extrapolation at $\mathbf{x}_u$ can be obtained as the product $\xi^* = \xi_1^* \otimes \xi_2^*$ of the c-optimal designs ξ_v^* for extrapolation at x_{uv} in the marginal models (see Theorem 4.4 in (Schwabe, 1996a)).*

The marginal c-optimal designs ξ_v^* can be derived as in Example 1. They assign weight $w_v^* = |x_{uv}|/(1 + 2|x_{uv}|)$ to $x_v = 1$ and weight $1 - w_v^* = (1 + |x_{uv}|)/(1 + 2|x_{uv}|)$ to $x_v = 0$. Hence, the c-optimal design $\xi^* = \xi_1^* \otimes \xi_2^*$ for extrapolation at $\mathbf{x}_u$ is given by

$$\xi^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ (1-w_1^*)(1-w_2^*) & (1-w_1^*)w_2^* & w_1^*(1-w_2^*) & w_1^*w_2^* \end{pmatrix}.$$

Then, by Proposition 3.8.1, the design ξ^* is also optimal for minimization of the asymptotic variance for estimating the α-quantile t_α of the failure time for soft failure due to degradation, when $0 < t_\alpha < \infty$ and the measurement times $t_1, ..., t_k$ are predetermined. For example, when the normal use conditions are $x_{u1} = -0.5$ for the first component and $x_{u2} = -0.4$ for the second component as specified in Table 3.2, then by the results in Example 1 the optimal marginal weights are $w_1^* = 0.25$ and $w_2^* = 0.22$, and the optimal design $\xi^* = \xi_1^* \otimes \xi_2^*$ is given by

$$\xi^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.58 & 0.17 & 0.19 & 0.06 \end{pmatrix}.$$

The *c*-optimal extrapolation designs can be obtained in both Examples 1 and 2 by Elfving's theorem ((Elfving, 1952)) which provides a geometrical construction of a *c*-optimal design (see (Schwabe, 1996a), Theorem 2.13). To give a rough idea of this construction one has to consider the Elfving set which is the convex hull

$$\mathcal{E} = \text{conv}(\{\mathbf{f}(\mathbf{x}); \mathbf{x} \in \mathcal{X}\} \cup \{-\mathbf{f}(\mathbf{x}); \mathbf{x} \in \mathcal{X}\})$$

of the union of the so-called induced design region $\{\mathbf{f}(\mathbf{x}); \mathbf{x} \in \mathcal{X}\}$ and its image $\{-\mathbf{f}(\mathbf{x}); \mathbf{x} \in \mathcal{X}\}$ under reflection at the origin $\mathbf{0}$ in $\mathbb{R}^p$. Here $\mathbf{x}$ and $\mathbf{f}$ denote variables and regression functions associated with a generic model $Y_i = \mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta} + \varepsilon_i$. In Example 1 we have $\mathbf{x} = x$, $\mathbf{f}(\mathbf{x}) = (1, x)^T$ and $\mathcal{X} = [0, 1]$, and the Elfving set $\mathcal{E}$ is given as a parallelogram in $\mathbb{R}^2$ with one edge from $(1, 0)^T$ to $(1, 1)^T$ representing the induced design region and the opposite edge from $(-1, 0)^T$ to $(-1, -1)^T$ representing its image under reflection. For another illustration of an Elfving set see Figure 3.7 in Section 3.10, where an optimal design is sought for the time variable t.

The *c*-optimal design for estimating $\mathbf{c}^T \boldsymbol{\beta}$ can then be constructed as follows: Determine the intersection point of the ray $\lambda \mathbf{c}$, $\lambda > 0$, with the boundary of the Elfving set, $\lambda_c \mathbf{c}$ say. This point can be represented as a convex combination of (extremal) points $\pm \mathbf{f}(\mathbf{x}_i)$ of the induced design region and its reflection,

$$\lambda_c \mathbf{c} = \sum_{i=1}^m w_i z_i \mathbf{f}(\mathbf{x}_i),$$

where $z_i = 1$, when the (extremal) point $\mathbf{f}(\mathbf{x}_i)$ is from the induced design region, and $z_i = -1$, when the point $-\mathbf{f}(\mathbf{x}_i)$ is from the reflection, and the weights w_i of the convex combination satisfy $w_i > 0$ and $\sum_{i=1}^m w_i = 1$. Then Elfving's theorem states that the design ξ^* which assigns weights w_i to the settings $\mathbf{x}_i$ is c-optimal (for $\mathbf{c}$). Moreover, this construction also provides the value of the c-criterion, $\mathbf{c}^T \mathbf{M}(\xi^*)^{-1}\mathbf{c} = 1/\lambda_c^2$.

In Example 1 the ray $\lambda(1, x_u)^T$ intersects the boundary of the Elfving set $\mathcal{E}$ at the connecting line from $(1,0)^T = \mathbf{f}(0)$ to $(-1,-1)^T = -\mathbf{f}(1)$ at $\lambda_c(1, x_u)^T = w_1(-1,-1)^T + w_2(1,0)^T$ with $\lambda_c = 1/(1 - 2x_u)$, $w_1 = -x_u/(1 - 2x_u) > 0$ and $w_2 = 1 - w_1 = (1 - x_u)/(1 - 2x_u) > 0$. Hence, the optimality of the given design follows.

We will use Elfving's theorem next to characterize optimal designs for the situation with two non-interacting stress variables.

Example 3. *In the case of two non-interacting stress variables x_1 and x_2 we consider the model equation*

$$y_{ij} = \beta_{i,1} + \beta_2 x_{i1} + \beta_3 x_{i2} + \beta_{i,4} t_j + \beta_5 x_{i1} t_j + \beta_6 x_{i2} t_j + \varepsilon_{ij} \tag{3.32}$$

for the combined stress variable $\mathbf{x} = (x_1, x_2)$ and the time variable t. This model contains all terms of the full interaction model 4.23) with the exception of the terms $x_1 x_2$ and $x_1 x_2 t$ related to potential interactions between the stress variables. The interpretation of all other terms in (3.32) is the same as in Example 2. Model (3.32) is constructed from the marginal model

$$Y_i^{(1)} = \beta_1^{(1)} + \beta_2^{(1)} x_{i1} + \beta_3^{(1)} x_{i2} + \varepsilon_i^{(1)}$$

which is additive in the effects of the stress variables x_1 and x_2, i.e. $\mathbf{f}_1(\mathbf{x}) = (1, x_1, x_2)^T$.

For this marginal model of multiple regression, the Elfving set is an oblique prism with quadratic base with vertices $(1,0,0)^T$, $(1,0,1)^T$, $(1,1,0)^T$, $(1,1,1)^T$ and quadratic top with vertices $(-1,0,0)^T$, $(-1,0,-1)^T$, $(-1,-1,0)^T$, $(-1,-1,-1)^T$. To find the c-optimal extrapolation design at the normal use condition $\mathbf{x}_u = (x_{u1}, x_{u2})$ by Elfving's theorem we have to determine the intersection point of the ray $\lambda(1, x_{u1}, x_{u2})^T$ with the surface of the Elfving set. For $x_{u1} < x_{u2} < 0$ the ray intersects the surface at the quadrangular face of the prism spanned by $(1,0,0)^T$, $(1,0,1)^T$, $(-1,-1,0)^T$, and $(-1,-1,-1)^T$ when $\lambda_c = 1/(1 + 2|x_{u1}|)$. The representation of the intersection point $\lambda_c(1, x_{u1}, x_{u2})^T$ by the vertices of the quadrangle is not unique. There are two c-optimal designs

$$\xi_0^* = \begin{pmatrix} (0,0) & (0,1) & (1,1) \\ (1 + |x_{u2}|)\lambda_c & (|x_{u1}| - |x_{u2}|)\lambda_c & |x_{u1}|\lambda_c \end{pmatrix}$$

and

$$\xi_1^* = \begin{pmatrix} (0,0) & (1,0) & (1,1) \\ (1 + |x_{u1}|)\lambda_c & (|x_{u1}| - |x_{u2}|)\lambda_c & |x_{u2}|\lambda_c \end{pmatrix}$$

*which are supported on three vertices. As a consequence, also for all coefficients a,
$0 < a < 1$, the convex combination*

$$\xi_a^* = (1-a)\xi_0^* + a\xi_1^* =$$

$$\begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ (1 + a|x_{u1}| + (1-a)|x_{u2}|)\lambda_c & (1-a)(|x_{u1}| - |x_{u2}|)\lambda_c & a(|x_{u1}| - |x_{u2}|)\lambda_c & ((1-a)|x_{u1}| + a|x_{u2}|)\lambda_c \end{pmatrix}$$

*supported on all four vertices is c-optimal for extrapolation at $\mathbf{x}_u$, $0 < t_\alpha < \infty$. Then,
by Proposition 3.8.1, the designs ξ_a^* are also optimal for minimization of the asymptotic
variance for estimating the α-quantile t_α of the failure time for soft failure due to
degradation, when $0 < t_\alpha < \infty$ and the measurement times $t_1, ..., t_k$ are predetermined,
$0 \le \alpha \le 1$. For example, when the normal use conditions are $x_{u1} = -0.5$ for the first
component and $x_{u2} = -0.4$ for the second component as in Example 2, then the optimal
design ξ_a^* is given by*

$$\xi_a^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.70 + 0.05a & 0.05 - 0.05a & 0.05a & 0.25 - 0.05a \end{pmatrix}$$

with the special cases

$$\xi_0^* = \begin{pmatrix} (0,0) & (0,1) & (1,1) \\ 0.70 & 0.05 & 0.25 \end{pmatrix} \quad \text{and} \quad \xi_1^* = \begin{pmatrix} (0,0) & (1,0) & (1,1) \\ 0.75 & 0.05 & 0.20 \end{pmatrix}$$

supported on three vertices.

*Note that there are also other designs which are c-optimal for extrapolation at $\mathbf{x}_u$,
but which are not solely supported on the vertices. For example, for $x_{u1} < x_{u2} < 0$
the two-point design which assigns weight $w = |x_{u1}|/(1 + 2|x_{u1}|)$ to $(0,0)$ and weight
$1 - w = (1 + |x_{u1}|)/(1 + 2|x_{u1}|)$ to $(1, x_{u2}/x_{u1})$ is c-optimal by Elfving's theorem. However,
these designs can be used for estimating t_α by means of maximum-likelihood only when
the resulting information matrix is non-singular, i. e. when the design has, at least, three
distinct support points.*

*For $x_{u2} < x_{u1} < 0$ optimal designs can be obtained from the above case by inter-
changing the roles of the two components x_1 and x_2.*

*In the case $x_{u1} = x_{u2} < 0$ there is only one c-optimal design for extrapolation. This
design is supported on two vertices and assigns weight $w = |x_{u1}|/(1 + 2|x_{u1}|)$ to $(0,0)$
and weight $1 - w = (1 + |x_{u1}|)/(1 + 2|x_{u1}|)$ to $(1,1)$. As the resulting information matrix
is singular, this design cannot be used for estimating the α-quantile t_α of the failure
time for soft failure due to degradation. Hence, no suitable optimal design exists in this*

case, but the c-optimal design may serve as a benchmark for judging the quality of a competing design in terms of efficiency.

To quantify the quality of a standard design ξ_0 for estimating the quantile t_α of the mean failure time under normal use condition we make use of the efficiency

$$\text{eff}_{\text{aVar}}(\xi_0) = \frac{\text{aVar}(\widehat{t}_\alpha; \xi^*)}{\text{aVar}(\widehat{t}_\alpha; \xi_0)}, \tag{3.33}$$

where $\text{aVar}(\widehat{t}_\alpha; \xi) = \mathbf{c}^T \mathbf{M}_{\boldsymbol{\theta}}(\xi)^{-1}\mathbf{c}$ denotes the standardized asymptotic variance for estimating t_α by equation (3.22) when design ξ is used, and ξ^* is the corresponding optimal design. The efficiency gives the proportion of units to be used under the optimal design ξ^* which provides (asymptotically) the same accuracy (in terms of the asymptotic variance) compared to the standard design ξ_0. For example, if the efficiency is 0.5 twice the number of units have to be used under ξ_0 than under the optimal design ξ^* to get the same accuracy. Note that both the asymptotic variance and the efficiency may also depend on the parameter vector $\boldsymbol{\theta}$, at least, through t_α and are, hence, local quantities (at $\boldsymbol{\theta}$) without explicitly stated in the notation.

In the case of estimating the median $t_{0.5}$ the standardized asymptotic variance factorizes as

$$\text{aVar}(\widehat{t}_{0.5}; \xi) = \frac{1}{n}c_0^2 \mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1(\xi)^{-1}\mathbf{f}_1(\mathbf{x}_u) \cdot \mathbf{f}_2(t_\alpha)^T \mathbf{M}_2^{-1}\mathbf{f}_2(t_\alpha) \tag{3.34}$$

by equations (3.30) and (3.31) for the general product-type model (3.7). Thus the efficiency defined in (3.33) reduces to the c-efficiency

$$\text{eff}_c(\xi_0) = \frac{\mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1(\xi^*)^{-1}\mathbf{f}_1(\mathbf{x}_u)}{\mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1(\xi_0)^{-1}\mathbf{f}_1(\mathbf{x}_u)}$$

for extrapolation at the normal use condition $\mathbf{x}_u$ in the first marginal model with uncorrelated homoscedastic errors and does not depend on $\boldsymbol{\theta}$.

Example (Example 1 cont.)**.** *In the introductory model of Section 3.2 the c-criterion for extrapolation at x_u is defined by $\Phi_c(\xi) = \mathbf{f}_1(x_u)^T \mathbf{M}_1(\xi)^{-1}\mathbf{f}_1(x_u)$. For $x_u < 0$ it attains its minimal value value $\Phi_c(\xi^*) = (1 + 2|x_u|)^2$ for the optimal design ξ^* which assigns weight $w^* = |x_u|/(1+2|x_u|)$ to $x_h = 1$ and weight $1 - w^* = (1 + |x_u|)/(1+2|x_u|)$ to $x_l = 0$. Common alternatives would be uniform designs $\bar{\xi}_m$ which assign equal weights $w = 1/m$ to m experimental settings on an equidistant grid $\{x_1, ..., x_m\} = \{0, 1/(m-1), ..., 1\}$ of the experimental region $\mathcal{X} = [0, 1]$. For these designs the c-criterion for extrapolation at $x_u < 0$ can be calculated as $\Phi_c(\bar{\xi}_m) = 1 + a_m(1 + 2|x_u|)^2$, where $a_m = 3(m-1)/(m+1)$. Their c-efficiency for extrapolation at x_u and, hence, their efficiency for estimating the median failure time under normal use condition x_u is*

TABLE 3.3: Efficiency of uniform designs $\bar{\xi}_m$ for various normal use conditions x_u in Example 1

m	x_u					
	0	-0.056	-0.400	-0.500	-1.000	$-\infty$
2	0.50	0.55	0.76	0.80	0.90	1.00
3	0.40	0.43	0.55	0.57	0.62	0.67
4	0.36	0.38	0.47	0.49	0.52	0.56
5	0.33	0.36	0.43	0.44	0.47	0.50
∞	0.25	0.26	0.30	0.31	0.32	0.33

equal to $\mathrm{eff}_c(\bar{\xi}_m) = \Phi_c(\xi^*)/\Phi_c(\bar{\xi}_m) = (1 - 1/(1 + a_m(1 + 2|x_u|)^2))/a_m$ *which increases from* $1/(1 + a_m) = (m + 1)/(4m - 2)$ *for* x_u *close to the lowest stress level* $x_l = 0$ *to* $1/a_m = (m + 1)/(3m - 3)$ *when* x_u *tends to minus infinity. Moreover, for fixed* x_u*, the efficiency decreases when* m *increases, i.e. when the grid becomes more dense. For selected values of the normal use condition* x_u *and numbers* m *of grid points numerical values of the efficiency are reported in Table 3.3 Note that in Table 3.3 the row* $m = \infty$ *corresponds to a continuous uniform design as an approximation to large numbers* m *of grid points, while the columns* $x_u = 0$ *and* $x_u = -\infty$ *give approximations for normal use conditions* x_u *close to the lowest experimental stress level or far away, respectively.*

For the particular case $m = 2$*, where the design* $\bar{\xi}_2$ *assigns equal weights* $w = 1 - w = 1/2$ *to both the highest and the lowest stress level* $x_h = 1$ *and* $x_l = 0$*, we have* $a_2 = 1$ *and, hence,* $\Phi_c(\xi_0) = 1 + (1 + 2|x_u|)^2$ *for the c-criterion. The c-efficiency of* $\bar{\xi}_2$ *for extrapolation at* x_u *and, thus, its efficiency for estimating the median failure time under normal use condition* x_u *is equal to* $\mathrm{eff}_c(\bar{\xi}_2) = 1 - 1/(1 + (1 + 2|x_u|)^2)$ *which ranges from* $1/2$ *for* x_u *close to the lowest stress level* $x_l = 0$ *to 1 when* x_u *tends to minus infinity.*

For the nominal value $x_u = -0.056$ *of the normal use condition in Table 3.1 the efficiency of the equidistant grid designs* $\bar{\xi}_m$ *is reported in the third column of Table 3.3. In particular, for the uniform design* $\bar{\xi}_2$ *on the endpoints of the experimental region this efficiency is 0.55 which means that* $\mathrm{eff}_c(\bar{\xi}_2)^{-1} - 1 = 1/(1 + 2|x_u|)^2 = 81\%$ *more units would have to be used for design* $\bar{\xi}_2$ *to obtain the same quality for estimating the median failure time than for the optimal design* ξ^**.*

The efficiencies in the case of one stress variable can be used to compute the efficiency in the presence of two stress variables with interactions.

Example (Example 2 cont.)**.** *In the model with two interacting stress variables* x_1 *and* x_2 *the c-criterion* $\Phi_c(\xi) = \mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1(\xi)^{-1} \mathbf{f}_1(\mathbf{x}_u)$ *for extrapolation at* $\mathbf{x}_u = (x_{u1}, x_{u2})$ *factorizes into its counterparts in the marginal models,* $\Phi_c(\xi) = \mathbf{f}_{11}(x_{u1})^T \mathbf{M}_{11}(\xi_1)^{-1} \mathbf{f}_{11}(x_{u1}) \cdot \mathbf{f}_{12}(x_{u2})^T \mathbf{M}_{12}(\xi_2)^{-1} \mathbf{f}_{12}(x_{u2})$*. Because also the c-optimal design* $\xi^* = \xi_1^* \otimes \xi_2^*$ *has product-type structure, the c-efficiency for extrapolation at* $\mathbf{x}_u$ *and, hence, the efficiency for*

estimating the median failure time factorizes, $\text{eff}_c(\xi_1 \otimes \xi_2) = \text{eff}_{c1}(\xi_1) \cdot \text{eff}_{c2}(\xi_2)$, where $\text{eff}_{cv}(\xi_v)$ is the corresponding efficiency in the vth marginal model, $v = 1, 2$.

The design $\bar{\xi}$ which assigns equal weights $1/4$ to the four vertices $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ of the experimental region serves as a natural standard design. This design can be seen to be the product $\bar{\xi} = \bar{\xi}_2 \otimes \bar{\xi}_2$ of marginal designs $\bar{\xi}_2$ which assign equal weights $1/2$ to the lowest and highest stress level x_{vl} and x_{vh} in the marginal models, $v = 1, 2$. Hence, from Example 1 we get the efficiency of $\bar{\xi}$ as $\text{eff}_c(\bar{\xi}) = ((1 + 2|x_{u1}|)(1 + 2|x_{u2}|))^2 / ((1 + (1 + 2|x_{u1}|)^2)(1 + (1 + 2|x_{u2}|)^2))$ which ranges from $1/4$ for $\mathbf{x}_u$ close to the combination (x_{1l}, x_{2l}) of lowest stress levels $x_{1l} = 0$ and $x_{2l} = 0$ to 1 when both normal use conditions x_{u1} and x_{u2} tend to minus infinity.

For example, when the normal use conditions are $x_{u1} = -0.5$ for the first component and $x_{u2} = -0.4$ for the second component as specified in Table 3.2, then according to Table 3.3 the efficiency $\text{eff}_{cv}(\bar{\xi}_2)$ of $\bar{\xi}_2$ is 0.80 and 0.76 in the respective marginal models, $v = 1, 2$. By the above considerations the efficiency of $\bar{\xi}$ is $\text{eff}_c(\bar{\xi}) = 0.80 \cdot 0.76 = 0.61$. This means that $\text{eff}_c(\bar{\xi})^{-1} - 1 = 0.39/0.61 = 64\%$ more units have to be used for design $\bar{\xi}$ to obtain the same quality for estimating the median failure time than for the optimal design ξ^*. Hence, the optimal design ξ^* performs much better than the standard design $\bar{\xi}$ in this situation.

Even more prominent results can be obtained for the marginal model without interactions between the stress variables.

Example (Example 3 cont.)**.** *In the model with two non-interacting stress variables x_1 and x_2 the value of the c-criterion for the locally c-optimal design ξ^* for extrapolation at $\mathbf{x}_u = (x_{u1}, x_{u2})$, $x_{u1} < x_{u2} < 0$, is given by $\Phi_c(\xi^*) = 1/\lambda_c^2 = (1 + 2|x_{u1}|)^2$ as seen before. The uniform design $\bar{\xi}$ which assigns equal weights $1/4$ to the four vertices $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ of the experimental region has a value of $\Phi_c(\bar{\xi}) = 1 + (1 + 2|x_{u1}|)^2 + (1 + 2|x_{u2}|)^2$. Hence, the uniform design $\bar{\xi}$ has efficiency $\text{eff}_c(\bar{\xi}) = ((1 + 2|x_{u1}|)^2 / (1 + (1 + 2|x_{u1}|)^2 + (1 + 2|x_{u2}|)^2)$ which ranges from $1/3$ for $\mathbf{x}_u$ close to the combination (x_{1l}, x_{2l}) of lowest stress levels $x_{1l} = 0$ and $x_{2l} = 0$ to 1 when the lower normal use condition x_{u1} tends to minus infinity while x_{u2} remains fixed. Moreover, the efficiency approaches $1/2$ when $x_{u2} \approx x_{u1}$ and both normal use conditions tend to minus infinity simultaneously.*

For example, when the normal use conditions are $x_{u1} = -0.5$ for the first component and $x_{u2} = -0.4$ for the second component as specified in Table 3.2, then the values of the c-criterion are $\Phi_c(\xi^*) = 4.00$ for the optimal design ξ^* and $\Phi_c(\bar{\xi}) = 8.24$ for the uniform design $\bar{\xi}$, respectively. Hence, the efficiency of the uniform design $\bar{\xi}$ is $\text{eff}_c(\bar{\xi}) = 4.00/8.24 = 0.49$. This means that more than twice as many units have to be used for design $\bar{\xi}$ to obtain the same quality for estimating the median failure time than for the optimal design ξ^*. This highlights that the optimal design ξ^* performs

substantially better than the standard design $\bar{\xi}$ in the current model of two non-interacting stress variables.

The above efficiency calculations are all related to estimating the median failure time for soft failure due to degradation under normal use conditions $\mathbf{x}_u$. For estimating any other quantile t_α of the failure time distribution, the efficiency of a design ξ can be written as

$$\text{eff}_{\text{aVar}}(\xi) = \text{eff}_c(\xi) + (1 - \text{eff}_c(\xi))\mathbf{c}_\varsigma^T \widetilde{\mathbf{M}}_\varsigma^{-1} \mathbf{c}_\varsigma / \text{aVar}(\widehat{t}_\alpha; \xi) \tag{3.35}$$

by equations (3.29) and (3.22). This efficiency depends on the variance parameters, but it is bounded from below by the *c*-efficiency $\text{eff}_c(\xi)$ of ξ for extrapolation at $\mathbf{x}_u$. Hence, designs with a high efficiency for estimating the median failure time are also suitable for estimating any other reasonable quantile t_α, $0 < t_\alpha < \infty$.

3.9 Optimization of the time plan

In Section 3.8, we considered the situation when there is a fixed time plan $\mathbf{t} = (t_1, ..., t_k)^T$, and only the stress variables $\mathbf{x} \in \mathcal{X}$ are to be optimized across units. In contrast to that we consider now the situation when also the settings $t_1, ..., t_k$ for the standardized time variable $t \in [0, 1]$ may be optimized within units. As in Section 3.3 we still assume that the same time plan $\mathbf{t} = (t_1, ..., t_k)^T$ is used for all units.

Similar to the case of fixed time points $t_1, ..., t_k$, the standardized asymptotic variance is given by

$$\begin{aligned}
\text{aVar}(\widehat{t}_\alpha; \xi, t_1, ..., t_k) = &c_0^2 (\mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1(\xi)^{-1} \mathbf{f}_1(\mathbf{x}_u) \cdot \mathbf{f}_2(t_\alpha)^T \mathbf{M}_2(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_\alpha) \\
&+ \mathbf{c}_\varsigma^T \widetilde{\mathbf{M}}_\varsigma(t_1, ..., t_k)^{-1} \mathbf{c}_\varsigma)
\end{aligned} \tag{3.36}$$

by equations (3.29) and (3.31). But here the dependence of the asymptotic variance on the settings $t_1, ..., t_k$ for the time variable is explicitly stated which comes through the information matrix $\mathbf{M}_2(t_1, ..., t_k)$ in the second marginal model as well as the standardized information matrix $\widetilde{\mathbf{M}}_\varsigma(t_1, ..., t_k)$ for the variance parameters.

As has been seen in Section 3.8 the optimization with respect to the stress variables $\mathbf{x}_u$ can be done independently of the second marginal model and the settings $t_1, ..., t_k$ of the time variable. The reverse does not hold true for the time points $t_1, ..., t_k$, in general, by the extra dependence of the asymptotic variance on the information matrix $\widetilde{\mathbf{M}}_\varsigma(t_1, ..., t_k)$ for the variance parameters.

To circumvent this problem we restrict to the case of estimating the median failure time $t_{0.5}$. There the second term on the right hand side of equation (3.36) vanishes and

the standardized asymptotic variance simplifies to

$$\text{aVar}(\widehat{t}_{0.5}; \xi, t_1, ..., t_k) = c_0^2 \mathbf{f}_1(\mathbf{x}_u)^T \mathbf{M}_1(\xi)^{-1} \mathbf{f}_1(\mathbf{x}_u) \cdot \mathbf{f}_2(t_{0.5})^T \mathbf{M}_2(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_{0.5}).$$
(3.37)

Hence, for the median failure time also the optimization of the measurement times $t_1, .., t_k$ can be performed independently of the marginal model of the stress variables and their settings. Then, only the marginal c-criterion $\mathbf{f}_2(t_{0.5})^T \mathbf{M}_2(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_{0.5})$ has to be minimized.

Remember that the marginal information matrix $\mathbf{M}_2(t_1, ..., t_k)$ is given by $\mathbf{M}_2(t_1, ..., t_k) = \mathbf{F}_2(t_1, ..., t_k)^T \mathbf{V}(t_1, ..., t_k)^{-1} \mathbf{F}_2(t_1, ..., t_k)$ in the second marginal model for the time variable t, where the variance covariance matrix $\mathbf{V}(t_1, ..., t_k)$ of the vector $\mathbf{Y}_i$ of measurements for each unit i is given by $\mathbf{V}(t_1, ..., t_k) = \mathbf{F}_2(t_1, ..., t_k) \boldsymbol{\Sigma}_\gamma \mathbf{F}_2(t_1, ..., t_k)^T + \boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\gamma$ and $\boldsymbol{\Sigma}_\varepsilon$ are the variance covariance matrices for the random effects $\boldsymbol{\gamma}$ and for the measurement errors $\boldsymbol{\varepsilon}$, respectively.

(Schmelter, 2007) proposed a representation of the inverse of the information matrix in random effects models in the case of uncorrelated homoscedastic errors ($\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_k$). This can be readily extended to a general, non-singular variance covariance structure $\boldsymbol{\Sigma}_\varepsilon$ for the measurement errors $\boldsymbol{\varepsilon}$ (see Appendix A.1). From this we obtain for the marginal information matrix $\mathbf{M}_2(t_1, ..., t_k)$ that its inverse can be decomposed to $\mathbf{M}_2(t_1, ..., t_k)^{-1} = \mathbf{F}_2(t_1, ..., t_k)^T \boldsymbol{\Sigma}_\varepsilon^{-1} \mathbf{F}_2(t_1, ..., t_k) + \boldsymbol{\Sigma}_\gamma$.

As a consequence, the marginal c-criterion $\mathbf{f}_2(t_{0.5})^T \mathbf{M}_2(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_{0.5})$ can be split up into

$$\mathbf{f}_2(t_{0.5})^T \mathbf{M}_2(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_{0.5}) = \mathbf{f}_2(t_{0.5})^T \mathbf{M}_2^{(0)}(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_{0.5}) + \mathbf{f}_2(t_{0.5})^T \boldsymbol{\Sigma}_\gamma \mathbf{f}_2(t_{0.5}),$$

where $\mathbf{M}_2^{(0)}(t_1, ..., t_k) = \mathbf{F}_2(t_1, ..., t_k)^T \boldsymbol{\Sigma}_\varepsilon^{-1} \mathbf{F}_2(t_1, ..., t_k)$ is the information matrix in the marginal fixed effect model

$$Y_j^{(2,0)} = \mathbf{f}_2(t_j)^T \boldsymbol{\beta}^{(2)} + \varepsilon_j^{(2)},$$
(3.38)

$j = 1, ..., k$, in the time variable t with variance covariance matrix $\boldsymbol{\Sigma}_\varepsilon$ for the vector $\boldsymbol{\varepsilon}^{(2)} = (\varepsilon_1^{(2)}, ..., \varepsilon_k^{(2)})^T$ of error terms. Hence, as for any linear criterion, the optimization of the asymptotic variance (3.37) of the median failure time $t_{0.5}$ with respect to the time plan $t_1, ..., t_k$ does not depend on the variance covariance matrix $\boldsymbol{\Sigma}_\gamma$ of the random effects. For optimization only the term $\mathbf{f}_2(t_{0.5})^T \mathbf{M}_2^{(0)}(t_1, ..., t_k)^{-1} \mathbf{f}_2(t_{0.5})$ has to be minimized which is the c-criterion for extrapolation of the mean response at $t_{0.5}$ in the marginal fixed effect model in t with $\text{Cov}(\boldsymbol{\varepsilon}^{(2)}) = \boldsymbol{\Sigma}_\varepsilon$. This leads to the following result which is similar to Proposition 3.8.1 for optimization with respect to the stress variable $\mathbf{x}$.

Proposition 3.9.1. *If the time plan $t_1^*, ..., t_k^*$ is c-optimal for extrapolation of the mean response at $t_{0.5}$ in the marginal fixed effect model (3.38) with covariance $\mathbf{\Sigma}_\varepsilon$ for the time variable t, then $t_1^*, ..., t_k^*$ minimize the asymptotic variance for the estimator $\hat{t}_{0.5}$ of the median failure time $t_{0.5}$ under normal use condition.*

Note that in degradation experiments the median failure time $t_{0.5}$ under normal use condition is typically much larger than the time horizon of the experiment, i. e. $t_{0.5} > 1$ on the standardized time scale. However, the above proposition also holds for interpolation, $t_{0.5} \in [0, 1]$.

The optimal time plan depends on the location parameters $\boldsymbol{\beta}$ through $t_{0.5}$, but also on the variance covariance structure $\mathbf{\Sigma}_\varepsilon$ of the measurement errors within units, and are, hence, local. In Examples 1 and 2 we assumed uncorrelated homoscedastic measurement errors, $\mathbf{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_k$. In that case, the c-criterion for extrapolation can be reduced to $\mathbf{f}_2(t_{0.5})^T (\mathbf{F}_2(t_1, ..., t_k)^T \mathbf{F}_2(t_1, ..., t_k))^{-1} \mathbf{f}_2(t_{0.5})$, and the optimization does not depend on the error variance σ_ε^2.

If the number k of measurement times $t_1, ..., t_k$ is large, one might be tempted to use the concept of approximate designs also here. In that case, the approximate design τ will be defined by mutually distinct time points $t_1, ..., t_\ell$ from the standardized experimental time interval $\mathcal{T} = [0, 1]$ with corresponding proportions $\pi_1, ..., \pi_\ell > 0$ satisfying $\sum_{j=1}^\ell \pi_j = 1$ with the interpretation that (approximately) $\pi_j k$ measurements are performed at time point t_j, $j = 1, ..., \ell$, for each unit. In the situation of a linear time trend, $\mathbf{f}_2(t) = (1, t)^T$, as in Examples 1 and 2, this leads to essentially the same extrapolation problem as for the stress variable in Example 1. The c-optimal design for extrapolation at $t_{0.5} > 1$ is concentrated on the $\ell = 2$ endpoints $t_1 = 0$ and $t_2 = 1$ of the standardized experimental time interval $\mathcal{T} = [0, 1]$ with corresponding proportions $\pi_1 = \frac{t_{0.5} - 1}{2t_{0.5} - 1}$ and $\pi_2 = \frac{t_{0.5}}{2t_{0.5} - 1}$ (see (Schwabe, 1996a), Example 2.1). Similarly, as for extrapolation of the stress variable in Example 1, the proportion π_2 at the endpoint $t_2 = 1$ in the direction of the extrapolation time $t_{0.5}$ decreases from 1, when $t_{0.5}$ is close to 1, to 0.5, when the distance gets large ($t_{0.5} \to \infty$).

However, from a practical point of view, it does not seem meaningful to have replications, i. e. to have more than one measurement at a time at the same unit. Moreover, even if this would be possible, these measurements would be expected to be strongly correlated with a correlation beyond that caused by the random effects. To avoid these technical and modeling problem additional constraints have to be imposed on the time plan to guarantee independence of the measurement errors. For instance, the different measurement times $t_1, ..., t_k$ have to be at least some sufficiently large $\Delta t > 0$ apart, or more specifically, they are restricted to a grid on the standardized experimental time interval with grid size Δt. Nevertheless approximate design theory can be used to determine optimal time plans. To do so the standardized experimental time interval

is discretized to a sufficiently coarse grid, $\mathcal{T} = \{j\Delta t; \; j = 0, 1, ..., J\}$, where $\Delta t = 1/J$, i. e. $\mathcal{T} = \{0, \Delta t, 2\Delta t, ..., 1\}$. Additionally constraints are imposed on the proportions π_j that none of these proportions is larger than $1/k$, i. e. the number of measurements at a time is bounded by one and, hence, there are at least k different time points. Optimal approximate time plans can then be obtained under these constraints by using standard algorithms for design optimization. Actually, the so obtained proportions may be smaller than one for some of the time points. But a theoretical result based on an equivalence theorem under constraints (convex optimization with Kuhn-Tucker conditions) guarantees that this only occurs for a small number of time points: (Sahm and Schwabe, 2001) proved for the D-optimality criterion that the experimental region splits up in subregions where the weight of the optimal design attains the upper bound $1/k$ and where it is 0. Only at the boundary of these subregions weights different from 0 and $1/k$ may occur. In particular, for straight line regression on $[0, 1]$, the subregions with maximal weight $1/k$ are adjacent to the endpoints 0 and 1 of the interval while in the interior the optimal design has zero weight. This result carries over directly to other criteria like the c-criterion under appropriate conditions which are met in the present case of extrapolation. From this approach efficient exact designs can be obtained by adjusting the weights to the admissible values 0 and $1/k$ under the constraint of total weight 1.

Example (Example 2 cont.). *In the setup of Example 2 we deal with straight line regression*

$$Y_j^{(2,0)} = \beta_1^{(2)} + \beta_2^{(2)} t_j + \varepsilon_j^{(2)}$$

for the fixed effects marginal model in the time variable t. The measurement errors $\varepsilon_j^{(2)}$ are supposed to be uncorrelated and homoscedastic ($\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_k$). We are searching for a locally c-optimal design τ^* for extrapolation at the median failure time $t_{0.5}$ which is equal to 10.25 under the nominal values of Table 3.2. As constraints for the design we assume that $k = 6$ observation can be taken on a grid with increment $\Delta t = 0.05$ of the standardized experimental time interval $[0, 1]$, i. e. $J = 20$ and $\mathcal{T} = \{0, 0.05, 0.10, ..., 1\}$.

By the multiplicative algorithm (see e.g. (Silvey, Titterington, and Torsney, 1978)) adapted to the present constraint situation the following numerical solution is obtained for the locally c-optimal design

$$\tau^* = \begin{pmatrix} 0.00 & 0.05 & 0.10 & 0.85 & 0.90 & 0.95 & 1.00 \\ 0.166 & 0.166 & 0.130 & 0.040 & 0.166 & 0.166 & 0.166 \end{pmatrix}.$$

The optimal approximate design τ^* is supported on seven time points $t_1^*, ..., t_7^*$ which are concentrated to the ends of the standardized experimental time interval $[0, 1]$ with maximal admitted weight $1/k$ for all but the two boundary points $t_3^* = 0.10$ and $t_4^* = 0.85$

separating grid points $t_1^* = 0.00$, $t_2^* = 0.05$, $t_5^* = 0.90$, $t_6^* = 0.95$, and $t_7^* = 1.00$ with full weight $1/k$ from those with zero weight. This shape of the optimal design is in accordance with the findings of (Sahm and Schwabe, 2001) in the case of D-optimality. In view of Proposition 3.9.1 the design τ^* is also optimal for the estimation of the median failure time.

For practical use, the optimal approximate design τ^* may be adjusted to

$$\tau_0 = \left(\begin{array}{cccccc} 0.00 & 0.05 & 0.10 & 0.90 & 0.95 & 1.00 \\ 0.166 & 0.166 & 0.166 & 0.166 & 0.166 & 0.166 \end{array} \right)$$

which is supported on exactly $k = 6$ time points by transferring the weight from the boundary point $t_4^* = 0.85$ with the lower weight to the boundary point $t_3^* = 0.10$ with the higher weight (see (Dorfleitner and Klein, 1999)). As a consequence all weights in the adjusted design τ_0 are equal to $1/6$. Then the design τ_0 can be realized as an exact design by taking one measurement at each of the six time points 0.00, 0.05, 0.10, 0.90, 0.95, and 1.00. To quantify what might have got lost, the quality of the adjusted design τ_0 may be measured in terms of the local c-efficiency

$$\text{eff}_c(\tau_0) = \frac{\mathbf{f}_2(t_{0.5})^T \mathbf{M}_2(\tau^*)^{-1} \mathbf{f}_2(t_{0.5})}{\mathbf{f}_2(t_{0.5})^T \mathbf{M}_2(\tau_0)^{-1} \mathbf{f}_2(t_{0.5})} = \frac{464.42}{469.91} = 98.70\%$$

for extrapolation at $t_{0.5}$ in the marginal mixed effects model and, hence, for estimation of the median failure time. This indicates that the adjusted design τ_0 is highly efficient and can be recommended as the time plan for conducting an accelerated degradation testing experiment.

Note that by Proposition 3.9.1 the optimal design τ^* does not depend on the variance covariance structure $\boldsymbol{\Sigma}_\gamma$ of the random effects, but the efficiency of the adjusted design τ_0 may be affected by the random effects. Nevertheless, similar to equation (3.35), also here the c-efficiency for extrapolation at $t_{0.5}$ provides a lower bound for the efficiency in estimating the median failure time. The assumption $\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_k$ of uncorrelated measurement errors could be replaced by equal correlations ρ_ε between all measurements, i. e. $\boldsymbol{\Sigma}_\varepsilon$ is compound symmetric with all diagonal entries equal to σ_ε^2 and all off-diagonal entries equal to $\rho_\varepsilon \sigma_\varepsilon^2$. For estimating the variance covariance parameters an additional identifiability condition would be required in this case to distinguish the correlation ρ_ε of the measurement errors from the variance σ_1^2 of the random intercept. Then estimation of the location parameters $\boldsymbol{\beta}$ is not affected, and the design τ^* retains its optimality.

For further approaches allowing for correlations depending on the distance of the measurement times, e. g. when measurements follow a random process, we refer to (Näther, 1985) and (Müller, 2007) for optimization of the measurement times.

3.10 Experimental design with a cross-sectional time plan

In contrast to the previous sections we will allow here for time plans $t_{i1}, ..., t_{ik}$ differing across units,

$$Y_{ij} = (\mathbf{f}_1(\mathbf{x}_i) \otimes \mathbf{f}_2(t_{ij}))^T \boldsymbol{\beta} + \mathbf{f}_2(t_{ij})^T \boldsymbol{\gamma}_i + \varepsilon_{ij}\,, \qquad (3.39)$$

$j = 1, ..., k$, $i = 1, ..., n$, where all other expressions have the same meaning as in the general model equation 3.6. Differing time plans are, for example, required when the number k of time points is less than the number p_2 of parameters in the marginal model for the degradation paths. In particular, in the case of destructive testing only $k = 1$ measurement will be available per unit,

$$Y_i = (\mathbf{f}_1(\mathbf{x}_i) \otimes \mathbf{f}_2(t_i))^T \boldsymbol{\beta} + \mathbf{f}_2(t_i)^T \boldsymbol{\gamma}_i + \varepsilon_i\,, \qquad (3.40)$$

$i = 1, ..., n$. We will restrict to the case of $k = 1$ measurements per unit in the remainder of this section. In that case all measurements Y_i are assumed to be independent. Denote by

$$\sigma^2(t) = \mathbf{f}_2(t)^T \boldsymbol{\Sigma}_\gamma \mathbf{f}_2(t) + \sigma_\varepsilon^2 > 0$$

the variance function for measurements at time t, i.e. $\mathrm{Var}(Y_i) = \sigma^2(t_i)$. As discussed at the end of the previous section also here an identifiability condition has to be imposed on the variance parameters $\boldsymbol{\varsigma}$ to distinguish between the variance σ_ε^2 of the measurement error and the variance σ_1^2 of the random intercept, see (Graßhoff et al., 2012) who derived D-optimal designs.

Under the assumption of normally distributed random effects and measurement errors the information matrix $\mathbf{M}_\beta$ for the location parameter $\boldsymbol{\beta}$ can be written as

$$\mathbf{M}_\beta = \sum_{i=1}^n (\mathbf{f}_1(\mathbf{x}_i)\mathbf{f}_1(\mathbf{x}_i)^T) \otimes \left(\tfrac{1}{\sigma^2(t_i)} \mathbf{f}_2(t_i)\mathbf{f}_2(t_i)^T \right)$$

according to the product-type structure of the model (3.40). For estimating the median failure time $t_{0.5}$ the asymptotic variance is given by $\mathrm{aVar}(\widehat{t}_{0.5}) = c_0^2 \mathbf{c}^T \mathbf{M}_\beta^{-1} \mathbf{c}$ as in Section 3.7. Hence, we are searching for a c-optimal design, where the vector $\mathbf{c}$ factorizes according to $\mathbf{c} = \mathbf{c}_1 \otimes \mathbf{c}_2$ into components $\mathbf{c}_1 = \mathbf{f}_1(\mathbf{x}_u)$ and $\mathbf{c}_2 = \mathbf{f}_2(t_{0.5})$ associated with the marginal models for the stress and the time variables, see (3.31). Moreover, we suppose, as before, that the experimental settings $\mathbf{x}_i$ for the stress variable and t_i for the time of measurement may be chosen independently from their standardized experimental regions $\mathcal{X}$ and $\mathcal{T}$, respectively, i.e. $(\mathbf{x}_i, t_i) \in \mathcal{X} \times \mathcal{T}$.

Also here we make use of the concept of approximate designs denoted by ζ which

are specified by m mutually distinct combinations $(\mathbf{x}_i, t_i) \in \mathcal{X} \times \mathcal{T}$ with corresponding weights $\eta_i > 0$, $i = 1, ..., \nu$, $\sum_{i=1}^{\nu} \eta_i = 1$. Accordingly the corresponding normalized information matrix is defined as

$$\mathbf{M}(\zeta) = \sum_{i=1}^{\nu} \eta_i (\mathbf{f}_1(\mathbf{x}_i)\mathbf{f}_1(\mathbf{x}_i)^T) \otimes (\tilde{\mathbf{f}}_2(t_i)\tilde{\mathbf{f}}_2(t_i)^T),$$

where $\tilde{\mathbf{f}}_2(t) = \mathbf{f}_2(t)/\sigma(t)$ is the weighted version of the marginal regression function for the time variable (standardized by the standard deviation for measurement at time t). Then a c-optimal design for $\mathbf{c} = \mathbf{c}_1 \otimes \mathbf{c}_2$ can be obtained as the product-type design generated from c-optimal designs for $\mathbf{c}_v$ in the associated marginal models.

To be more specific, let

$$\xi = \begin{pmatrix} \mathbf{x}_1 & ... & \mathbf{x}_m \\ w_1 & ... & w_m \end{pmatrix} \qquad \text{and} \qquad \tau = \begin{pmatrix} t_1 & ... & t_\ell \\ \pi_1 & ... & \pi_\ell \end{pmatrix}$$

be approximate designs in the marginal models for the stress variables and the time, respectively. Then the product-type design $\zeta = \xi \otimes \tau$ is supported on the $\nu = m\ell$ mutually distinct combinations $(\mathbf{x}_i, t_j)$ with corresponding weights $\eta_{ij} = w_i \pi_\ell$, $i = 1, ..., m$, $j = 1, ..., \ell$, and its standardized information matrix

$$\mathbf{M}(\xi \otimes \tau) = \mathbf{M}_1(\xi) \otimes \tilde{\mathbf{M}}_2(\tau)$$

factorizes into the standardized information matrix $\mathbf{M}_1(\xi) = \sum_{i=1}^{m} w_i \mathbf{f}_1(\mathbf{x}_i)\mathbf{f}_1(\mathbf{x}_i)^T$ of ξ in the marginal model for the stress variables and the standardized information matrix $\tilde{\mathbf{M}}_2(\tau) = \sum_{j=1}^{k} \pi_j \tilde{\mathbf{f}}_2(t_j)\tilde{\mathbf{f}}_2(t_j)^T$ of τ in the weighted marginal model

$$Y_j^{(2)} = \tilde{\mathbf{f}}_2(t_j)^T \boldsymbol{\beta}^{(2)} + \varepsilon^{(2)} \tag{3.41}$$

for the time variable. Then also the c-criterion

$$\mathbf{c}^T \mathbf{M}(\xi \otimes \tau)^{-1} \mathbf{c} = \mathbf{c}_1^T \mathbf{M}_1(\xi)^{-1} \mathbf{c}_1 \cdot \mathbf{c}_2^T \tilde{\mathbf{M}}_2(\tau)^{-1} \mathbf{c}_2$$

for $\mathbf{c} = \mathbf{c}_1 \otimes \mathbf{c}_2$ factorizes into the c-criteria for $\mathbf{c}_1$ and $\mathbf{c}_2$ in the corresponding marginal models. This expression is minimized if ζ^* is the product $\zeta^* = \xi^* \otimes \tau^*$ of the c-optimal marginal designs ξ^* and τ^* for $\mathbf{c}_1$ and $\mathbf{c}_2$ in their marginal models, respectively. The design $\zeta^* = \xi^* \otimes \tau^*$ is not only c-optimal for $\mathbf{c} = \mathbf{c}_1 \otimes \mathbf{c}_2$ in the class of product-type designs $\xi \otimes \tau$, but also compared to all designs ζ on $\mathcal{Z}$ (see (Schwabe, 1996b), Theorem 4.4) which establishes the following result.

Theorem 3.10.1. *If ξ^* is c-optimal for extrapolation at normal use condition $\mathbf{x}_u$ in the marginal model for the stress variables $\mathbf{x}$ and τ^* is c-optimal for extrapolation at*
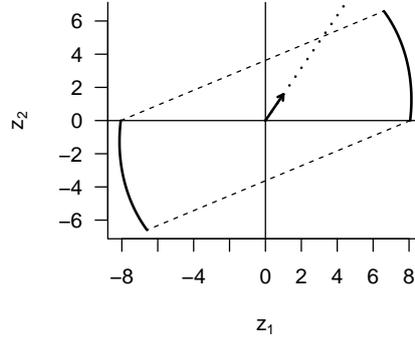
FIGURE 3.7: Elfving set for t in Example 1: Induced design region (right solid line), negative image (left solid line), boundary (dashed lines), $\mathbf{c}_2 = (1, t_{0.5})^T$ (arrow) and corresponding ray (dotted line)

the median failure time $t_{0.5}$ in the weighted marginal model (3.41) for the time variable t, then the design $\zeta^* = \xi^* \otimes \tau^*$ is optimal for estimating the median failure time $t_{0.5}$ in the combined model (3.40).

Note that for the second marginal model we utilize that $\tilde{\mathbf{f}}_2(t_{0.5})$ and $\mathbf{c}_2 = \mathbf{f}_2(t_{0.5})$ only differ by the factor $1/\sigma(t_{0.5}) > 0$ which does not affect c-optimality.

The c-optimal designs ξ^* for extrapolation at $\mathbf{x}_u$ are the same as obtained in Section 3.8, and the c-optimal designs for extrapolation at $t_{0.5}$ can be obtained similarly by Elfving's theorem applied to the weighted regression functions $\tilde{\mathbf{f}}_2(t) = \mathbf{f}_2(t)/\sigma(t)$.

Example (Example 1 cont.). *In the introductory model of Section 3.2 we have straight line regression in both the stress variable x and the time variable t with an interaction term xt. For the stress variable, the c-optimal marginal design ξ^* for extrapolation at $x_u < 0$ assigns weight $w^* = |x_u|/(1 + 2|x_u|)$ to $x_h = 1$ and weight $1 - w^* = (1 + |x_u|)/(1 + 2|x_u|)$ to $x_l = 0$, see Section 3.8. For the time variable t, the c-optimal marginal design τ^* for extrapolation at $t_{0.5} > 1$ can be similarly obtained by Elfving's theorem applied to the vector of weighted regression functions $\tilde{\mathbf{f}}_2(t)$. The shape of the Elfving set is exhibited in Figure 3.7. Note that the induced design region $\{\tilde{\mathbf{f}}_2(t); t \in [0, 1]\}$ and its negative image constitute non-overlapping arc segments of the ellipse defined by $\mathbf{z}^T \boldsymbol{\Sigma} \mathbf{z} = 1$, $\mathbf{z} \in \mathbb{R}^2$, centered at $(0, 0)^T$, where $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_\gamma + \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}$ is the variance covariance matrix of the random effects augmented by the variance of the measurement error. The ray $\lambda \mathbf{f}_2(t_{0.5})$ intersects the boundary of the Elfving set at the line segment connecting $\tilde{\mathbf{f}}_2(1) = (1, 1)^T/\sigma(1)$ and $-\tilde{\mathbf{f}}_2(0) = (1, 0)^T/\sigma(0)$ when $t_{0.5} > 1$. Hence, the c-optimal design τ^* is supported by the endpoints of the standardized experimental time interval $\mathcal{T} = [0, 1]$, and the optimal weights can be*

calculated by Elfving's theorem to $\pi^* = t_{0.5}\sigma(1)/(t_{0.5}\sigma(1) + (t_{0.5} - 1)\sigma(0))$ at $t_{\max} = 1$ and $1 - \pi^* = (t_{0.5} - 1)\sigma(0)/(t_{0.5}\sigma(1) + (t_{0.5} - 1)\sigma(0))$ at $t_{\min} = 0$.

By Theorem 3.10.1, the design

$$\zeta^* = \xi^* \otimes \tau^* = \begin{pmatrix} (x_l, t_{\min}) & (x_l, t_{\max}) & (x_h, t_{\min}) & (x_h, t_{\max}) \\ (1-w^*)(1-\pi^*) & (1-w^*)\pi^* & w^*(1-\pi^*) & w^*\pi^* \end{pmatrix}$$

is optimal for estimating the median failure time $t_{0.5}$. Under the nominal values of Table 3.1 the optimal weights are $w^* = 0.05$ for extrapolation at $x_u = -0.056$ in the marginal model for the stress variable, see Section 3.8, and $\pi^* = 0.77$ for extrapolation at $t_{0.5} = 1.583$ in the weighted marginal model for the time variable. Hence, the optimal design for estimating the median failure time is given by

$$\zeta^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.22 & 0.73 & 0.01 & 0.04 \end{pmatrix}.$$

According to this design 73% of the units should be exposed to the lowest stress level $x_l = 0$ and measured at the maximum allowed time $t_{\max} = 1$ while only 1% of the units should be exposed to the highest stress level $x_h = 1$ and measured at the initial time $t_{\min} = 0$. Accordingly 22% should be exposed to low stress and measured initially, and 4% should be measured at the end of the experimental time interval under high exposure.

A similar results can be obtained for the situation of Example 2.

Example (Example 2 cont.)**.** *In the model with two interacting stress variables (Example 2) the c-optimal marginal design ξ^* for extrapolation at $\mathbf{x}_u = (x_{u1}, x_{u2})^T$ is the same as in Section 3.8, $\xi^* = \xi_1^* \otimes \xi_2^*$. For the time variable t, the c-optimal marginal design τ^* for extrapolation at $t_{0.5} > 1$ is obtained as in Example 1: τ^* assigns weights $\pi^* = t_{0.5}\sigma(1)/(t_{0.5}\sigma(1) + (t_{0.5} - 1)\sigma(0))$ to $t_{\max} = 1$ and $1 - \pi^*$ to $t_{\min} = 0$. By Theorem 3.10.1, the optimal design for estimation of the median failure time is then $\zeta^* = \xi_1^* \otimes \xi_2^* \otimes \tau^*$.*

Under the nominal values of Table 3.2 the optimal weights are $w_1^* = 0.25$ and $w_2^* = 0.22$ for extrapolation at $x_{u1} = -0.5$ and $x_{u2} = -0.4$ in the marginal models of the stress variables, respectively, see Section 3.8. For the time variable the optimal weight results in $\pi^* = 0.57$ for extrapolation at $t_{0.5} = 10.25$ in the weighted marginal model. Hence, the optimal design for estimating the median failure time is given by

$$\xi_1^* \otimes \xi_2^* \otimes \tau^* = \begin{pmatrix} (0,0,0) & (0,0,1) & (0,1,0) & (0,1,1) & (1,0,0) & (1,0,1) & (1,1,0) & (1,1,1) \\ 0.25 & 0.33 & 0.07 & 0.10 & 0.09 & 0.11 & 0.02 & 0.03 \end{pmatrix}.$$

The locally optimal designs for estimating the median failure time are influenced by both the location and the variance parameters β and ς. Therefore we make a
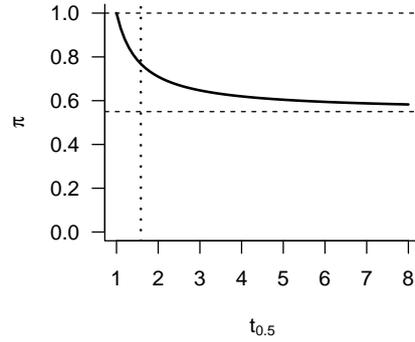
FIGURE 3.8: Optimal weights π^* in dependence on $t_{0.5}$ for Example 1

sensitivity analysis, how the optimal designs change with the parameters and how well they perform under parameter misspecification. For this we note first that the optimal marginal design ξ^* for extrapolation at $\mathbf{x}_u$ is globally optimal and does not depend on the parameters. Moreover, in the case of straight line degradation paths as in Examples 1 and 2 only the weight $\pi^* = t_{0.5}\sigma(1)/(t_{0.5}\sigma(1) + (t_{0.5} - 1)\sigma(0))$ of the optimal marginal design τ^* depends on the location parameters $\boldsymbol{\beta}$ through $t_{0.5}$ and on the variance parameters $\boldsymbol{\varsigma}$ through the ratio $\sigma(1)/\sigma(0)$ of the standard deviations at the endpoints of the experimental time region. Note that the c-efficiency for extrapolation at $t_{0.5}$ in the weighted marginal model for t coincides with the efficiency for estimating the median failure time when the optimal design ξ^* is used for the stress variables.

Example (Example 1 cont.)**.** *For the introductory model with one stress variable (Example 1), the optimal weight $\pi^* = t_{0.5}\sigma(1)/(t_{0.5}\sigma(1) + (t_{0.5} - 1)\sigma(0))$ is plotted in Figure 3.8 as a function of $t_{0.5}$ while the ratio $\sigma(1)/\sigma(0)$ is held fixed to 1.22 induced by the nominal values in Table 3.1, and in Figure 3.9 as a function of the ratio $\sigma(1)/\sigma(0)$ while $t_{0.5}$ is held fixed to the nominal value 1.583.*

When $t_{0.5}$ increases, the optimal weight π^* decreases from 1 for $t_{0.5}$ close to the maximal experimental time $t_{\max} = 1$ to $\sigma(1)/(\sigma(1) + \sigma(0)) = 0.55$ for $t_{0.5} \to \infty$. On the other hand the optimal weight π^* increases in the ratio $\sigma(1)/\sigma(0)$ of standard deviations from 0 for the ratio close to 0 to 1 when the ratio tends to infinity. The limits for π^* are indicated in Figure 3.8 and Figure 3.9 by horizontal dashed lines while the nominal values for $t_{0.5}$ and $\sigma(1)/\sigma(0)$ are indicated by vertical dotted lines, respectively.

To assess the influence of the variation of the optimal weights we consider the efficiency of a locally optimal design when the underlying parameters are misspecified. Define by

$$\text{eff}_{\text{aVar}}(\zeta; \boldsymbol{\theta}) = \frac{\text{aVar}_{\boldsymbol{\theta}}(\widehat{t}_{0.5}; \zeta)}{\text{aVar}_{\boldsymbol{\theta}}(\widehat{t}_{0.5}; \zeta_{\boldsymbol{\theta}}^*)}$$
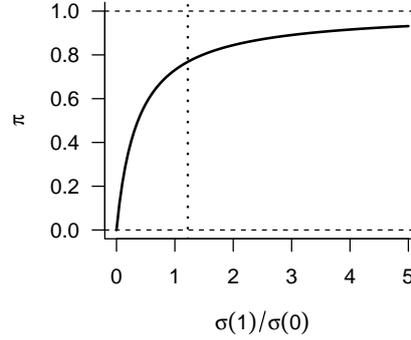
FIGURE 3.9: Optimal weights π^* in dependence on $\sigma(1)/\sigma(0)$ for Example 1

the asymptotic efficiency of the design ζ for estimating $t_{0.5}$, where $\mathrm{aVar}_{\boldsymbol{\theta}}(\widehat{t}_{0.5}; \zeta)$ denotes the asymptotic variance of $\widehat{t}_{0.5}$, when ζ is used and when $\boldsymbol{\theta}$ is the true vector of underlying parameters, and by $\zeta_{\boldsymbol{\theta}}^*$ the locally optimal design at $\boldsymbol{\theta}$. Similar to the situation of marginal models in Example 2 (see Section 3.8), the asymptotic efficiency

$$\mathrm{eff}_{\mathrm{aVar}}(\xi \otimes \tau; \boldsymbol{\theta}) = \mathrm{eff}_{c1}(\xi) \cdot \mathrm{eff}_{c2}(\tau; \boldsymbol{\theta}),$$

factorizes for product-type designs $\zeta = \xi \otimes \tau$ into the marginal efficiency $\mathrm{eff}_{c1}(\xi)$ of ξ for extrapolation at $\mathbf{x}_u$ in the marginal model for the stress variable and the marginal efficiency $\mathrm{eff}_{c2}(\tau; \boldsymbol{\theta})$ of τ for extrapolation at $t_{0.5}$ in the weighted marginal model for the time variable. Remember that, for the stress variable, the marginal information matrix $\mathbf{M}_1(\xi)$ and, hence, the marginal efficiency $\mathrm{eff}_{c1}(\xi)$ does not depend on $\boldsymbol{\theta}$. Thus, also the c-optimal marginal design ξ^* does not rely on the parameters and can be used throughout in the comparison. The asymptotic efficiency of a product-type design $\zeta = \xi^* \otimes \tau$ with optimal marginal ξ^* reduces to the c-efficiency of the second marginal τ,

$$\mathrm{eff}_{\mathrm{aVar}}(\xi^* \otimes \tau; \boldsymbol{\theta}) = \mathrm{eff}_{c2}(\tau; \boldsymbol{\theta}) = \frac{\mathbf{c}_2(\boldsymbol{\theta})^T \mathbf{M}_2(\tau; \boldsymbol{\theta}) \mathbf{c}_2(\boldsymbol{\theta})}{\mathbf{c}_2(\boldsymbol{\theta})^T \mathbf{M}_2(\tau_{\boldsymbol{\theta}}^*; \boldsymbol{\theta}) \mathbf{c}_2(\boldsymbol{\theta})},$$

where $\mathbf{c}_2(\boldsymbol{\theta}) = \mathbf{f}_2(t_{0.5})$ and $\tau_{\boldsymbol{\theta}}^*$ is the locally c-optimal design at $\boldsymbol{\theta}$ for extrapolation at $t_{0.5}$ in the weighted marginal model for the time variable. The c-vector $\mathbf{c}_2(\boldsymbol{\theta}) = \mathbf{f}_2(t_{0.5})$ depends on the location parameters $\boldsymbol{\beta}$ merely through $t_{0.5}$, but not on $\boldsymbol{\varsigma}$. In contrast to that, the information matrices $\mathbf{M}_2(\tau; \boldsymbol{\theta})$ are only affected by the variance parameters $\boldsymbol{\varsigma}$, but not by $\boldsymbol{\beta}$. For straight line regression in the time variable (cf. Examples 1 and 2), only the ratio $\sigma(1)/\sigma(0)$ of the standard deviations for measurements at the endpoints of the experimental time interval has an effect on the information matrix $\mathbf{M}_2(\tau; \boldsymbol{\theta})$, when the design τ is supported by these endpoints, as is the case for the locally c-optimal
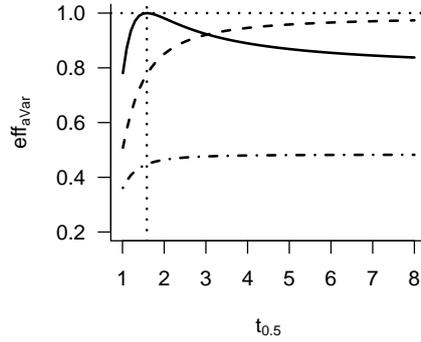
FIGURE 3.10: Efficiency of $xi^* \otimes \tau^*$ (solid line), $\xi^* \otimes \bar{\tau}_2$ (dashed line) and $\xi^* \otimes \bar{\tau}_6$ (dashed and dotted line) in dependence on $t_{0.5}$ for Example 1
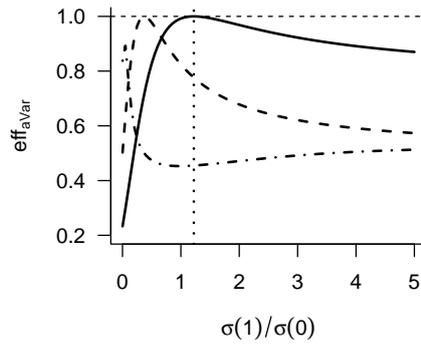
FIGURE 3.11: Efficiency of $xi^* \otimes \tau^*$ (solid line), $\xi^* \otimes \bar{\tau}_2$ (dashed line) and $\xi^* \otimes \bar{\tau}_6$ (dashed and dotted line) in dependence on $\sigma(1)/\sigma(0)$ for Example 1

design $\tau_{\boldsymbol{\theta}}^*$.

Example (Example 1 cont.)**.** *For the setting of the introductory model with one stress variable (Example 1), we examine the efficiency of the design $\zeta* = \xi^* \otimes \tau^*$ which is locally optimal for estimation of the median failure time under the nominal values of Table 3.1 when the nominal parameters are misspecified. In Figure 3.10 and Figure 3.11 we plot the efficiency of the locally optimal design $xi^* \otimes \tau^*$ at the nominal values (solid line) together with the efficiency of the designs $xi^* \otimes \bar{\tau}_2$ (dashed line) and $xi^* \otimes \bar{\tau}_6$ (dashed and dotted line) for which the marginal design $\bar{\tau}_k$ is uniform on k equally spaced time points, i. e. $\bar{\tau}_2$ assigns weight $1/2$ to the endpoints 0 and 1 of the time interval, and τ_6 assigns weight $1/6$ to the time points 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 (cf. the definition of $\bar{\xi}_m$ in Section 3.8).*

In Figure 3.10 the efficiency is displayed in dependence on the true value of the median failure time $t_{0.5}$ while the ratio $\sigma(1)/\sigma(0)$ of standard deviations is held fixed to the nominal value 1.22. In Figure 3.11 the efficiency is shown in dependence on the true value of the ratio $\sigma(1)/\sigma(0)$ of standard deviations while the median failure time $t_{0.5}$ is held fixed to 1.583 derived from the nominal values. Also here the nominal values for $t_{0.5}$ and $\sigma(1)/\sigma(0)$ are indicated by vertical dotted lines in the corresponding panel, respectively.

The efficiency seems to be more sensitive with respect to deviations in the ratio than in the median failure time $t_{0.5}$. However, commonly neither small values of $t_{0.5}$ nor small values of $\sigma(1)/\sigma(0)$ seem to be reasonable. In particular, we may expect that the variance $\sigma(1)^2$ at the end of the experimental time interval is larger than the variance $\sigma(0)^2$ at the initial time, i. e. $\sigma(1)/\sigma(0) \geq 1$. This is satisfied when $\rho \geq -\sigma_1/(2\sigma_2)$. In total, the design $\xi^* \otimes \tau^*$ which is locally optimal for the nominal values seems to perform quite well over a wide range of parameters while the design $\xi^* \otimes \bar{\tau}_2$ with the same number of units at the endpoints of the intervals is only preferable for larger values of the median failure time $t_{0.5}$. The design $\xi^* \otimes \bar{\tau}_6$ with six equally spaced time points performs substantially worse throughout for reasonable parameter values. It has to be noted in this context that the efficiency of $\bar{\tau}_6$ depends on the variance parameters ς not only through the ratio $\sigma(1)/\sigma(0)$ because measurements are also to be taken in the interior of the interval. The current plot has been generated by fixing $\sigma_2^2 = \sigma_1^2 + \sigma_\varepsilon^2$ and letting ρ vary. However, other choices of the variance parameters do not change much in the performance of the design.

Similar results can be obtained for the settings of Example 2.

3.11   Concluding remarks

During the design stage of highly reliable systems it is extremely important to assess the reliability related properties of the product. One method to handle this issue is to conduct accelerated degradation testing. Accelerated degradation tests have the advantage to provide an estimation of lifetime and reliability of the system under study in a relatively short period of time. To account for variability between units in accelerated degradation tests, it is assumed that the degradation function can be described by a mixed-effects linear model. This also leads to a non-degenerate distribution of the failure time, due to soft failure by exceedance of the expected (conditionally per unit) degradation path over a threshold, under normal use conditions. Therefore it is desirable to estimate certain quantiles of this failure time distribution as a characteristic of the reliability of the product. In this context we discussed the existence of non-degenerate solutions for the quantiles. The purpose of optimal experimental design is then to find

the best settings for the stress variable and/or the time variable to obtain most accurate estimates for these quantities.

In the present model for accelerated degradation testing, it is further assumed that stress remains constant within each testing unit during the whole period of experimental measurements but may vary between units. Hence, in the corresponding experiment a cross-sectional design between units has to be chosen for the stress variable while for repeated measurements the time variable varies according to a longitudinal design within units.

In the present chapter we assumed a model with complete interactions between the time and the stress variables and random effects only associated with time but not with stress. Then the cross-sectional design for the stress variables and the longitudinal design for the time variable can be optimized independently, and the resulting common optimal design can be generated as the cross-product of the optimal marginal designs for stress and time, respectively. In particular, the same time plan for measurements can be used for all units in the test. Moreover, the marginal optimal design for the stress variables can be chosen independently of any model parameters. Optimal time plans may depend on the aggregate location parameters via the median failure time, but do not depend on which quantile of the failure distribution is to be estimated. These results were extended to a model of destructive testing in which also the time variable has to be chosen cross-sectionally. There the optimal choice of measurement times may also be affected by the variance covariance parameters of the random effects. In both cases (longitudinal and cross-sectional time settings) the efficiency of the designs considered factorizes which facilitates to assess their performance when the nominal values for these parameters are misspecified at the design stage.

Finding optimal designs may become more complicated when the above assumptions are not met. In particular, the designs for stress and time variables may no longer be optimized independently if there are only additive effects in the model (lacking interaction terms xt, cf. Example 3 for a similar situation in the marginal stress model) or when also the stress variables are accompanied by random effects. The impact of these deviations from the model assumptions on optimal designs are object of further research as well as the construction of designs which are robust against misspecification of the nominal parameters, such as maximin efficient or weighted ("Bayesian") optimal designs. Of further interest would be to consider optimality criteria accounting for simultaneous estimation of various characteristics of the failure time distribution.

Chapter 4

Extensions to Multivariate Linear Mixed Effects Models

4.1 Introduction

In this chapter we propose optimal experimental designs for repeated measures Accelerated Degradation Testing with competing failure modes that correspond to multiple response components. The observation time points are assumed to be fixed and known in advance. The marginal degradation paths are expressed using linear mixed effects models. The optimal design is obtained by minimizing the asymptotic variance of the estimator of some quantile of the failure time distribution at the normal use conditions. Numerical examples are introduced to ensure the robustness of the proposed optimal designs and compare their efficiency with standard experimental designs.

The degradation process in complicated systems may occur due to multiple operating components, where theses components may be independent or have a certain level of interaction. Hence, Accelerated Degradation Testing in the presence of competing failure modes is an important reliability area to be addressed. Hence, the study of the statistical inference of Accelerated Degradation Testing with competing failures is of great significance and have been considered by many authors. For instance, (Haghighi, 2014) presented a step-stress test in the presence of competing risks and using degradation measurements where the underlying degradation process is represented with a concave degradation model under the assumption that the intensity functions corresponding to competing risks depend only on the level of degradation. In order to obtain the maximum likelihood estimates of intensity functions at normal use conditions, the author extrapolates the information from step-stress test at high level of stress through a tempered failure rate model. For linear models with nuisance parameters, (Filipiak, Markiewicz, and Szczepańska, 2009) gave relationships between Kiefer optimality of designs in univariate models and in their multivariate extensions with known and partially known dispersion matrices. With an application in plastic substrate active matrix light-emitting diodes, (Haghighi and Bae, 2015) proposed a modeling

approach to simultaneously analyze linear degradation data and traumatic failures with competing risks in a step stress Accelerated Degradation Testing experiment. In their research, the authors investigate the convergence criteria with a power law failure rate under step-stress Accelerated Degradation Testings. (Wang et al., 2015b) utilized Monte Carlo simulation to derive a cost-constrained optimal design for a constant stress accelerated degradation test with multiple stresses and multiple degradation measures. The authors assume that the degradation measures follow multivariate normal distribution with an application in a pilot-operated safety valve. Furthermore, the theory of optimal designs of experiments for multivariate models is well developed in the mathematical context of approximate designs. For instance, (Mukhopadhyay and Khuri, 2008) discussed response surface designs for multivariate generalized linear models (GLMs) considering a special case of the bivariate binary distribution. (Markiewicz and Szczepańska, 2007) discussed the optimality of an experimental design under multivariate linear models with a known or unknown dispersion matrix. The authors utilize Kiefer optimality to derive optimal designs for these linear models. (Dror and Steinberg, 2006) proposed a simple heuristic for constructing robust experimental designs for multivariate generalized linear models. The authors incorporate a method of clustering a set of local experimental designs to derive local D-optimal designs. (Schwabe, 1996a) treated in his monograph the theory of optimal designs for multi-factor models and provides an excellent review on the optimal design theory, i.e. the optimality criteria and the general equivalence theorems, up to that time. Considering random intercept models, (Schmelter and Schwabe, 2008) derived D-optimal designs for single and multiple treatments situations. The authors show that in a multi-sample situation the variability of the intercept has substantial influence on the choice of the optimal design.

The rest of this chapter is organized as follows. In Section 4.2, we formulate a multivariate degradation path on the basis of marginal linear mixed effects models (LMEMs). Section 4.3 is devoted to characterize the possibly estimated parameter vector, the resulting information matrix, and the proposed approximate design for the optimization. The considered optimality criterion for deriving c-optimal design based on the failure time distribution is introduced in Section 4.4. Section 4.5 addresses two numerical examples under two different testing conditions where the robustness of the proposed optimal designs along with their efficiency were investigated in comparison to some standard experimental design. Finally, we summarize with some concluding remarks in Section 4.6.

4.2   Model description

In this section, we introduce a formulation of a mixed effects degradation model with r response components where each of these components is observed under a value $\mathbf{x}$ of the experimental stress variable(s). The stress variable(s) is defined over the design region $\mathcal{X}$ and kept fixed for each unit throughout the degradation process, but may differ from unit to unit. The observed degradation data for this model are approximately fitted by a multivariate LMEM. The number k of measurements and the time points are the same for all units $i = 1, ..., n$. The measurements y_{ijl}, which are realizations of random variables Y_{ijl} at response component l, are described by a hierarchical model. For each unit i the observation Y_{ijl} for response component l at time point t_j is given by

$$Y_{ijl} = \mu_{il}(\mathbf{x}_i, t_j) + \varepsilon_{ijl}, \tag{4.1}$$

where $\mu_{il}(\mathbf{x}, t)$ is the mean degradation of unit i at response component l and time t, when stress $\mathbf{x}$ is applied to unit i, and ε_{ijl} is the associated measurement error at time point t_j. The measurement error ε_{ijl} is assumed to be independent from $\mathbf{x}$ and t, and normally distributed with zero mean and error variance $\sigma_\varepsilon^2 > 0$ ($\varepsilon_{ijl} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$). The mean degradation $\mu_{il}(\mathbf{x}, t)$ is assumed to be given by a linear model equation in the stress variable $\mathbf{x}$ and time t,

$$\mu_{il}(\mathbf{x}_i, t_j) = \sum_{s=1}^{p_l} \beta_{ils} f_{ls}(\mathbf{x}_i, t_j) = \mathbf{f}_l(\mathbf{x}_i, t_j)^T \boldsymbol{\beta}_{il} \tag{4.2}$$

where $\mathbf{f}_l(\mathbf{x}, t) = (f_{l1}(\mathbf{x}, t), ..., f_{lp_l}(\mathbf{x}, t))^T$ is the p_l-dimensional vector of regression functions $f_{ls}(\mathbf{x}, t)$ in both the stress variable(s) $\mathbf{x}$ and the time t considering the lth response component and denote by $\boldsymbol{\beta}_{il} = (\beta_{il1}, ..., \beta_{ilp_l})^T$ the p_l-dimensional vector of unit specific parameters $\beta_{ils}, s = 1, ..., p_l$, at response component l. Moreover, we assume throughout this work that the regression functions $\mathbf{f}_l(\mathbf{x}, t)$, $l = 1, ..., r$, include a constant term, $f_{l1}(\mathbf{x}, t) \equiv 1$ say. Denote by $\mathbf{g}_l(t)$ the q_l-dimensional random effect regression function which only depends on the time t. Hence, for unit i the model (4.1) can be rewritten as

$$Y_{ijl} = \mathbf{f}_l(\mathbf{x}_i, t_j)^T \boldsymbol{\beta}_l + \mathbf{g}_l(t_j)^T \boldsymbol{\gamma}_{il} + \varepsilon_{ijl}, \tag{4.3}$$

where $\boldsymbol{\gamma}_{il} = (\gamma_{il1}, ..., \gamma_{ilq_l})^T$ is the q_l-dimensional vector of unit specific deviations $\gamma_{ils} = \beta_{ils} - \beta_{ls}$, $s = 1, ..., q_l$, from the corresponding aggregate parameters such that $f_{ls}(\mathbf{x}, t) = g_{ls}(t)$ for all $s = 1, ..., q_l$. Hence, $\boldsymbol{\gamma}_{il}$ has q_l-dimensional multivariate normal distribution with zero mean and variance-covariance matrix $\boldsymbol{\Sigma}_{\gamma_l}$ ($\boldsymbol{\gamma}_{il} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\gamma_l})$) where $\boldsymbol{\Sigma}_{\gamma_l}$ is the corresponding $q_l \times q_l$ positive definite variance covariance matrix. Denote by $\boldsymbol{t} = (t_1, ..., t_k)^T$ the k-dimensional time points of measurements within units which

is fixed in advance and is not under disposition of the experimenter. Further, denote by $\mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t}) = \big(\mathbf{f}_l(\mathbf{x}_i, t_1), ..., \mathbf{f}_l(\mathbf{x}_i, t_k)\big)^T$ the $k \times p_l$ fixed effect design matrix for the marginal response component l of unit i. In vector notation the k-dimensional vector $\mathbf{Y}_{il} = (Y_{i1l}, ..., Y_{ikl})^T$ can be represented as

$$\mathbf{Y}_{il} = \mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})\boldsymbol{\beta}_l + \mathbf{G}_l(\boldsymbol{t})\boldsymbol{\gamma}_{il} + \boldsymbol{\varepsilon}_{il} \tag{4.4}$$

where $\mathbf{G}_l(\boldsymbol{t}) = (\mathbf{g}_l(t_1), ..., \mathbf{g}_l(t_k))^T$ is the $k \times q_l$ random effects design matrix. The k-dimensional vector $\boldsymbol{\varepsilon}_{il}$ is normally distributed as $\boldsymbol{\varepsilon}_{il} \sim \mathrm{N}(\mathbf{0}_k, \sigma_\varepsilon^2 \mathbf{I}_k)$, and $\mathbf{I}_k$ refers to the k-dimensional identity matrix. Hence, the k-dimensional vector of observations $\mathbf{Y}_{il}$ has a multivariate normal distributions as $\mathbf{Y}_{il} \sim \mathrm{N}(\mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})\boldsymbol{\beta}_l, \mathbf{V}_l)$ where $\mathbf{V}_l = \mathbf{G}_l(\boldsymbol{t})\boldsymbol{\Sigma}_{\gamma_l}\mathbf{G}_l(\boldsymbol{t})^T + \sigma_\varepsilon^2 \mathbf{I}_k$. Let $\mathbf{F}(\mathbf{x}_i, \boldsymbol{t}) = \mathrm{diag}\big(\mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})\big)_{l=1,...,r}$ be the $kr \times p$ fixed effect design matrix, and let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, ..., \boldsymbol{\beta}_r^T)^T$ be the p-dimensional overall vector of fixed effects parameters for r response components where $p = \sum_{l=1}^r p_l$. Further, the random effects $\boldsymbol{\gamma}_{il}$ of the components $\mathbf{Y}_{il}$ are assumed to be independent within units which implies independence of the components $\mathbf{Y}_{il}$ themselves within units. Hence, the per unit random effects parameter vector $\boldsymbol{\gamma}_i = (\boldsymbol{\gamma}_{i1}^T, ..., \boldsymbol{\gamma}_{ir}^T)^T$ is normally distributed with a zero mean and a covariance matrix $\boldsymbol{\Sigma}_\gamma = \mathrm{diag}\big(\boldsymbol{\Sigma}_{\gamma_l}\big)_{l=1,...,r}$ where $q = \sum_{l=1}^r q_l$. Denote $\boldsymbol{\varepsilon}_i = (\boldsymbol{\varepsilon}_{i1}^T, ..., \boldsymbol{\varepsilon}_{ir}^T)^T$ as the cumulative vector of random errors which is considered to be normally distributed with mean zero and variance covariance matrix $\sigma_\varepsilon^2 \mathbf{I}_{kr}$. Hence, the stacked kr-dimensional response vector $\mathbf{Y}_i = (\mathbf{Y}_{i1}^T, ..., \mathbf{Y}_{ir}^T)^T$ is given by

$$\mathbf{Y}_i = \mathbf{F}(\mathbf{x}_i, \boldsymbol{t})\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i \tag{4.5}$$

where $\mathbf{G} = \mathrm{diag}(\mathbf{G}_l(\boldsymbol{t}))_{l=1,...,r}$ is the $kr \times q$ block diagonal random effects design matrix. Then $\mathbf{Y}_i$ is kr-dimensional multivariate normally distributed with mean $\mathbf{F}(\mathbf{x}_i, \boldsymbol{t})\boldsymbol{\beta}$ and variance covariance matrix $\mathbf{V} = \mathbf{G}\boldsymbol{\Sigma}_\gamma\mathbf{G}^T + \sigma_\varepsilon^2 \mathbf{I}_{kr}$. Further, $\mathbf{G}\boldsymbol{\Sigma}_\gamma\mathbf{G}^T = \mathrm{diag}(\mathbf{G}_l\boldsymbol{\Sigma}_{\gamma_l}\mathbf{G}_l^T)$ and, hence, $\mathbf{V} = \mathrm{diag}(\mathbf{V}_l)$, which illustrates the independence of $\mathbf{Y}_{il}$ within units. It can be noted that the variance covariance matrix $\mathbf{V}$ is not affected by the choice of the stress level $\mathbf{x}_i$ and, hence, equal for all units i. For the observations of all n units the stacked nkr-dimensional response vector $\mathbf{Y} = (\mathbf{Y}_1^T, ..., \mathbf{Y}_n^T)^T$ can be expressed as

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + (\mathbf{I}_n \otimes \mathbf{G})\,\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{4.6}$$

where $\mathbf{F} = (\mathbf{F}(\mathbf{x}_1, \boldsymbol{t})^T, ..., \mathbf{F}(\mathbf{x}_n, \boldsymbol{t})^T)^T$ is the $nkr \times p$ design matrix for the stress variables across units, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^T, ..., \boldsymbol{\gamma}_n^T)^T$ is the nq-dimensional stacked parameter vector of random effects. The vector $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^T, ..., \boldsymbol{\varepsilon}_n^T)^T$ is the nkr-dimensional stacked vector of random errors which is normally distributed with mean zero and variance covariance matrix $\sigma_\varepsilon^2 \mathbf{I}_{nkr}$ ($\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_{nkr})$ and the vector $\boldsymbol{\gamma} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}_\gamma)$ of all random effects is

multivariate normal. In total, the vector $\mathbf{Y}$ of all observations is nkr-dimensional multivariate normal, $\mathbf{Y} \sim \mathrm{N}(\mathbf{F}\boldsymbol{\beta}, \mathbf{I}_n \otimes \mathbf{V})$.

4.3   Estimation, information and design

Under the distributional assumptions of normality for both the random effects and the measurement errors the model parameters may be estimated by means of the maximum likelihood method. Denote by $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \boldsymbol{\varsigma}^T)^T$ the vector of all model parameters where $\boldsymbol{\varsigma}$ indicates the variance covariance parameter vector. The log-likelihood for the current model is given by

$$\ell(\boldsymbol{\theta}; \mathbf{y}) = -\tfrac{nkr}{2}\log(2\pi) - \tfrac{n}{2}\log(\det(\mathbf{V})) - \tfrac{1}{2}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^T(\mathbf{I}_n \otimes \mathbf{V})^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}), \quad (4.7)$$

where the variance covariance matrix $\mathbf{V} = \mathbf{V}(\boldsymbol{\varsigma})$ of measurements per unit depends only on $\boldsymbol{\varsigma}$. The maximum likelihood estimator of $\boldsymbol{\beta}$ can be calculated as

$$\widehat{\boldsymbol{\beta}} \;=\; (\mathbf{F}^T(\mathbf{I}_n \otimes \widehat{\mathbf{V}})^{-1}\mathbf{F})^{-1}\mathbf{F}^T(\mathbf{I}_n \otimes \widehat{\mathbf{V}})^{-1}\mathbf{Y} \tag{4.8}$$

if $\mathbf{F}$ is of full column rank p, and $\widehat{\mathbf{V}} = \mathbf{V}(\widehat{\boldsymbol{\varsigma}})$, where $\widehat{\boldsymbol{\varsigma}}$ is the maximum likelihood estimator of $\boldsymbol{\varsigma}$. We note further that $\widehat{\boldsymbol{\beta}}$ can be represented by

$$\widehat{\boldsymbol{\beta}} \;=\; \left(\sum_{i=1}^{n}\mathbf{F}(\mathbf{x}_i, \mathbf{t})^T\widehat{\mathbf{V}}^{-1}\mathbf{F}(\mathbf{x}_i, t)\right)^{-1}\sum_{i=1}^{n}\mathbf{F}(\mathbf{x}_i, \mathbf{t})^T\widehat{\mathbf{V}}^{-1}\mathbf{Y}_i \tag{4.9}$$

Based on the definition of the Fisher information matrix in Subsection 3.5 denote by $\mathbf{M}_{\boldsymbol{\beta}} = -\mathrm{E}\left(\frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T}\ell(\boldsymbol{\theta}; \mathbf{Y})\right)$, $\mathbf{M}_{\boldsymbol{\varsigma}} = -\mathrm{E}\left(\frac{\partial^2}{\partial\boldsymbol{\varsigma}\partial\boldsymbol{\varsigma}^T}\ell(\boldsymbol{\theta}; \mathbf{Y})\right)$, $\mathbf{M}_{\boldsymbol{\beta}\boldsymbol{\varsigma}} = -\mathrm{E}\left(\frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\varsigma}^T}\ell(\boldsymbol{\theta}; \mathbf{Y})\right)$ and $\mathbf{M}_{\boldsymbol{\varsigma}\boldsymbol{\beta}} = \mathbf{M}_{\boldsymbol{\beta}\boldsymbol{\varsigma}}^T$ the blocks of the Fisher information matrix corresponding to the second derivatives with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\varsigma}$ and the mixed derivatives, respectively. The mixed blocks can be seen to be zero due to the independence property that arises for the normal distribution, and the Fisher information matrix is block diagonal,

$$\mathbf{M}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathbf{M}_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\boldsymbol{\varsigma}} \end{pmatrix}. \tag{4.10}$$

Moreover, the block $\mathbf{M}_{\boldsymbol{\beta}}$ associated with the aggregate location parameters $\boldsymbol{\beta}$ turns out to be the inverse of the variance covariance matrix for the estimator $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, when $\mathbf{V}$ is known. Actually, because the Fisher information matrix for $\boldsymbol{\theta}$ is block diagonal, the inverse $\mathbf{M}_{\boldsymbol{\beta}}^{-1}$ of the block associated with $\boldsymbol{\beta}$ is the corresponding block of the inverse of $\mathbf{M}_{\boldsymbol{\theta}}$ and is, hence, the asymptotic variance covariance matrix of $\widehat{\boldsymbol{\beta}}$. Accordingly the asymptotic variance covariance matrix for estimating the variance parameters $\boldsymbol{\varsigma}$

is the inverse of the block $\mathbf{M}_\varsigma$. In the following we will refer to $\mathbf{M}_\beta$ and $\mathbf{M}_\varsigma$ as the information matrices for β and ς, respectively, for short. The particular form of $\mathbf{M}_\varsigma$ will be not of interest here. It is important to note that $\mathbf{M}_\varsigma$ does not depend on the settings $\mathbf{x}_1, ..., \mathbf{x}_n$ of the stress variable in contrast to the information matrix $\mathbf{M}_\beta$ of the aggregate location parameters β.

As mentioned earlier the time points of measurements within units is fixed in advance and is not considered for optimization process. Then only the settings $\mathbf{x}_1, ..., \mathbf{x}_n$ of the stress variable $\mathbf{x}$ can be adjusted to the units $i = 1, ..., n$.

$$\mathbf{M}_\beta(\mathbf{x}_1, ..., \mathbf{x}_n) = \sum_{i=1}^{n} \mathbf{F}(\mathbf{x}_i, \boldsymbol{t})^T \mathbf{V}^{-1} \mathbf{F}(\mathbf{x}_i, \boldsymbol{t}). \tag{4.11}$$

By the independence of the components the information matrix $\mathbf{M}_\beta$ decomposes into its marginal counterparts $\mathbf{M}_\beta = \text{diag}(\mathbf{M}_{\beta_l})$ where $\mathbf{M}_{\beta_l} = \sum_{i=1}^{n} \mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})^T \mathbf{V}_l^{-1} \mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})$. In general, β can be estimated by MLE or, more usually, by restricted maximum likelihood (Debusho and Haines, 2008). In addition, the variance covariance matrix of the estimator $\widehat{\beta}$ of the location parameters β can be asymptotically approximated by the inverse of the information matrix $\mathbf{M}_\beta(\mathbf{x}_1, ..., \mathbf{x}_n)$. In accordance with the experimental design introduced in Subsection 3.6 we follow again the approach of approximate designs ξ as explained in Subsection 2.3.1 in which the requirement of integer numbers n_i of testing units at a stress level $\mathbf{x}_i$ is relaxed. The corresponding standardized, per unit information matrices are accordingly defined as

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^{m} w_i \mathbf{F}(\mathbf{x}_i, \boldsymbol{t})^T \mathbf{V}^{-1} \mathbf{F}(\mathbf{x}_i, \boldsymbol{t}) \tag{4.12}$$

for the aggregate parameters β. By the independence of the components $\mathbf{M}_\beta(\xi)$ decomposes accordingly $\mathbf{M}_\beta(\xi) = \text{diag}(\mathbf{M}_{\beta_l}(\xi))$ where $\mathbf{M}_{\beta_l}(\xi) = \sum_{i=1}^{m} w_i \mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})^T \mathbf{V}_l^{-1} \mathbf{F}_l(\mathbf{x}_i, \boldsymbol{t})$. For the full parameter vector $\boldsymbol{\theta}$ the information matrix $\mathbf{M}_{\boldsymbol{\theta}}(\xi)$ is expressed as

$$\mathbf{M}_{\boldsymbol{\theta}}(\xi) = \begin{pmatrix} \mathbf{M}_\beta(\xi) & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{M}}_\varsigma \end{pmatrix} \tag{4.13}$$

where now $\widetilde{\mathbf{M}}_\varsigma = \frac{1}{n}\mathbf{M}_\varsigma$ is the standardized, per unit information for the variance parameters ς. In order to optimize information matrices, we define in Subsection 4.4 the intended optimality criterion as a real valued function of the information matrix.

4.4   Optimal design based on failure times

In accordance with the work of (Weaver and Meeker, 2013) we consider some characteristics of the failure time distribution of soft failure due to degradation. For the analysis of degradation under normal use we further assume that the general model 4.3 is also valid at the normal use condition $\mathbf{x}_u$, where typically $\mathbf{x}_u \notin \mathcal{X}$, i.e.

$$\mu_{ul}(t) = \mu_l(\mathbf{x}_u, t) = \mathbf{f}_l(\mathbf{x}_u, t)^T \boldsymbol{\beta}_l + \mathbf{g}_l(t)^T \boldsymbol{\gamma}_{ul} \tag{4.14}$$

describes the mean degradation of a future unit u at normal use condition $\mathbf{x}_u$, time t and response component l where μ_{ul} denotes the degradation path under normal use condition for short. The random effects parameter vector $\boldsymbol{\gamma}_{ul}$ is normally distributed with mean zero and variance $\boldsymbol{\Sigma}_{\gamma_l}$. Further, denote $\mu_l(t) = \mathrm{E}(\mu_{ul}(t)) = \mathbf{f}_l(\mathbf{x}_u, t)^T \boldsymbol{\beta}_l$ as the aggregate degradation path under normal use condition for response component l. For the following it is assumed that the r mean degradation paths are strictly increasing over time. Then a soft failure at component l due to degradation is defined as the exceedance of the degradation over a failure threshold y_{l0}. This definition is based on the mean degradation path and not on a "real" path subject to measurement errors. The marginal failure time T_l under normal use condition is then defined as the first time t the mean degradation path $\mu_{ul}(t)$ reaches or exceeds the threshold y_{l0}, i.e. $T_l = \min\{t \geq 0; \mu_{ul}(t) \geq y_{l0}\}$. As the random effect γ_{ul} is involved in the mean degradation path, the marginal failure time T_l is random. Actually, T_l may become infinite, if the l-th mean degradation path does not reach the threshold, or may degenerate to $T_l = 0$, if the degradation already exceeds the threshold at time $t = 0$, because of unfortunate values of the random effects vector $\boldsymbol{\gamma}_{ul}$, but this will happen only with low probability and will not affect the further argumentation. In order to describe certain characteristics of the distribution of the failure time, we will determine first the marginal distribution function $F_{T_l}(t) = \mathrm{P}(T_l \leq t)$. First note that $T_l \leq t$ if and only if $\mu_{ul}(t) \geq y_{l0}$. Hence

$$
\begin{aligned}
F_{T_l}(t) &= \mathrm{P}(\mu_{ul}(t) \geq y_{l0}) \\
&= \mathrm{P}(\mu_l(t) + \boldsymbol{\gamma}_{ul} \geq y_{l0}) \\
&= \mathrm{P}(-\boldsymbol{\gamma}_{ul} \leq \mu_l(t) - y_{l0}) \\
&= \Phi(h_l(t)),
\end{aligned}
\tag{4.15}
$$

where

$$h_l(t) = \frac{\mu_l(t) - y_{l0}}{\sigma_{ul}(t)}, \tag{4.16}$$

$\sigma_{ul}^2(t) = \mathbf{g}_l(t)^T \mathbf{\Sigma}_{\gamma_l} \mathbf{g}_l(t)$ is the variance of the mean degradation path $\mu_{ul}(t)$ at time t. In the particular case of straight lines for the mean degradation paths, i.e. $\mathbf{g}_l(t) = (1, t)^T$, $l = 1, .., r$, the variance covariance matrix is given by $\mathbf{\Sigma}_{\gamma_l} = \begin{pmatrix} \sigma_{l1}^2 & \rho_l \sigma_{l1} \sigma_{l2} \\ \rho_l \sigma_{l1} \sigma_{l2} & \sigma_{l2}^2 \end{pmatrix}$, and, hence, the function $h_l(t)$ specifies to

$$h_l(t) = \frac{\mu_l(t) - y_{l0}}{\sqrt{\sigma_{l1}^2 + 2\rho_l \sigma_{l1} \sigma_{l2} t + \sigma_{l2}^2 t^2}}, \tag{4.17}$$

The joint failure time T is defined consequently as a function, say ψ, of the marginal failure times, $T = \psi(T_1, ..., T_r)$. For instance, a failure for an s-out-of-r system occurs if, at least, s of its r components exceed their corresponding failure thresholds. Hence, for the special case 1-out-of-r system, the joint failure time T might be defined as $T = \min(T_1, ..., T_r)$ so that a failure of the system occurs if, at least, one of its components fail . Quantiles t_α of the joint failure time distribution, i. e. $P(T \le t_\alpha) = \alpha$, are considered for further calculations. For each α the quantile t_α indicates the time up to which under normal use conditions (at least) $\alpha \cdot 100$ percent of the units fail and (at least) $(1 - \alpha) \cdot 100$ percent of the units survive. The quantiles t_α are increasing in α. Further, the current standard definition of quantiles is in contrast to the "upper" quantiles $(t_{1-\alpha})$ used in (Weaver and Meeker, 2013) where percentages of failures and persistence are reversed. Of particular interest is the median $t_{0.5}$ up to which under normal use conditions half of the units fails and half of the units persist ($\alpha = 0.5$). The quantile $t_\alpha = t_\alpha(\boldsymbol{\theta})$ is a function of both the aggregate location parameters $\boldsymbol{\beta}$ and the variance parameters $\boldsymbol{\varsigma}$. Hence, the maximum likelihood estimator of the quantile t_α is given by $\widehat{t}_\alpha = t_\alpha(\widehat{\boldsymbol{\theta}})$ in terms of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$. The task of designing the experiment will now be to provide an as precise estimate of the α-quantile as possible. By the delta-method $\widehat{t}_\alpha$ is seen to be asymptotically normal with asymptotic variance

$$\text{aVar}(\widehat{t}_\alpha) = \mathbf{c}^T \mathbf{M}_{\boldsymbol{\theta}}^{-1} \mathbf{c}, \tag{4.18}$$

where $\mathbf{c} = \frac{\partial}{\partial \boldsymbol{\theta}} t_\alpha$ is the gradient vector of partial derivatives of t_α with respect to the components of the parameter vector $\boldsymbol{\theta}$. The asymptotic variance depends on the design of the experiment through the information matrix $\mathbf{M}_{\boldsymbol{\theta}}$ and will be chosen as the optimality criterion for the design.

As the marginal aggregate mean degradation $\mu_l(t)$ only depend on the aggregate location parameters $\boldsymbol{\beta}$, and σ corresponds to the variance parameter in $\boldsymbol{\varsigma}$ the gradient simplifies to $\mathbf{c} = (\mathbf{c}_{\boldsymbol{\beta}}^T, \mathbf{c}_{\boldsymbol{\varsigma}}^T)^T$, where

$$\mathbf{c}_{\boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} t_\alpha(\boldsymbol{\theta})$$

is the gradient of t_α with respect to $\boldsymbol{\beta}$ and

$$\mathbf{c}_\varsigma = \tfrac{\partial}{\partial \varsigma} t_\alpha(\boldsymbol{\theta})$$

is the gradient of t_α with respect to ς where the particular shape of $\mathbf{c}_\varsigma$ does not play a role here, in general. Due to the block diagonal form of the information matrix in equation (4.10) the asymptotic variance (4.18) of $\hat{t}_\alpha$ specifies to

$$\mathrm{aVar}(\hat{t}_\alpha) = \mathbf{c}_\beta^T \mathbf{M}_\beta^{-1} \mathbf{c}_\beta + \mathbf{c}_\varsigma^T \mathbf{M}_\varsigma^{-1} \mathbf{c}_\varsigma \tag{4.19}$$

where the second term in the right hand side is an additive constant and does not depend on ξ.

Due to the complexity of deriving an explicit formula of t_α, the following equality is ensured by the implicit function theorem, see (Krantz and Parks, 2012)

$$\frac{\partial t_\alpha}{\partial \boldsymbol{\theta}} = \left(\frac{\partial F_T(t_\alpha)}{\partial t_\alpha} \right)^{-1} \frac{\partial F_T(t_\alpha)}{\partial \boldsymbol{\theta}}. \tag{4.20}$$

such that $\frac{\partial F_T(t_\alpha)}{\partial t_\alpha}$ is a scaling constant that is irrelevant to the design. The gradient vector $\mathbf{c}_\beta$ can be replaced by $\mathbf{c}_\beta = \frac{\partial}{\partial \boldsymbol{\beta}} F_T(t_\alpha)$ such that $\mathbf{c}_\beta$ decomposes into marginal components $\mathbf{c}_\beta = (\mathbf{c}_1^T, ..., \mathbf{c}_r^T)^T$ where $\mathbf{c}_l = \partial F_T(t_\alpha)/\partial \boldsymbol{\beta}_l$. Because the components are assumed to be independent within units, the information matrix $\mathbf{M}_\beta(\xi)$ is block diagonal with diagonal entries $\mathbf{M}_{\beta_1}(\xi), ..., \mathbf{M}_{\beta_r}(\xi)$ as noted in the previous section. Accordingly, based on the optimality criterion defined in equation (4.19), the gradient vector depends only on the parameter vector $\boldsymbol{\beta}$ and the locally *c*-optimal design ξ^* can be defined by

$$\xi^* = \arg\min_\xi \left(\sum_{l=1}^r \mathbf{c}_l^T \mathbf{M}_{\beta_l}^{-1}(\xi) \mathbf{c}_l \right). \tag{4.21}$$

In order to assess the influence of the variation of the optimal weights we consider the efficiency of the resulting optimal design optimal design ξ^* when the underlying nominal values are varied. where the asymptotic efficiency of the design ξ for estimating t_α is defined by

$$\mathrm{eff}_{\mathrm{aVar}}(\xi) = \frac{\mathbf{c}_\varsigma^T \widetilde{\mathbf{M}}_\varsigma^{-1} \mathbf{c}_\varsigma + \sum_{l=1}^r \mathbf{c}_l^T \mathbf{M}_{\beta_l}^{-1}(\xi^*) \mathbf{c}_l}{\mathbf{c}_\varsigma^T \widetilde{\mathbf{M}}_\varsigma^{-1} \mathbf{c}_\varsigma + \sum_{l=1}^r \mathbf{c}_l^T \mathbf{M}_{\beta_l}^{-1}(\xi) \mathbf{c}_l}. \tag{4.22}$$

If $\boldsymbol{\Sigma}_{\gamma_1} = ... = \boldsymbol{\Sigma}_{\gamma_r}$, $\mathbf{g}_1 = ... = \mathbf{g}_r = \mathbf{g}$, and $\mathbf{f}_1 = ... = \mathbf{f}_r = \mathbf{f}$ such that $\mathbf{f}$ has product type structure $\mathbf{f}(\mathbf{x}, t) = \mathbf{f}^{(1)}(\mathbf{x}) \otimes \mathbf{g}(t)$, then the terms in the criterion (4.21) factorize, $\sum_{l=1}^r \tilde{c}_l \mathbf{f}^{(1)}(\mathbf{x}_u)^T \mathbf{M}_{(1)}(\xi)^{-1} \mathbf{f}^{(1)}(\mathbf{x}_u)$ where $\mathbf{M}_{(1)}(\xi) = \sum_{i=1}^m w_i \mathbf{f}^{(1)}(\mathbf{x}_i) \mathbf{f}^{(1)}(\mathbf{x}_i)^T$ is the (fixed effect) information matrix in the first marginal model related to the stress variable and $\tilde{c}_l = c_l^2 \mathbf{g}(t_\alpha)^T \mathbf{M}_{(2)}(\mathbf{t}) \mathbf{g}(t_\alpha)$. Hence, under this assumption, if ξ is optimal

for extrapolation at $\mathbf{x}_u$ in the first marginal model, it is also optimal for estimating t_α in the system. Actually, this holds not only for series systems but also for *s*-out-of-*r* systems in which case the constants c_l are more complicated (see below in Subsection 4.5).

4.5 Numerical Examples

In this section we present two examples of optimal designs of accelerated degradation testing. We consider first an example for a series system in accordance with the work of (Shat and Schwabe, 2021) with full interaction between stress and time variables. We propose further another example for an *s*-out-of-*r* system with partial interaction of explanatory variables with the time variable.

Example 4. *We derive in this example a locally c-optimal design for the degradation model in Section 4.2 under the standardized time plan $\boldsymbol{t} = (0, 0.5, 1)$ which is identical for all testing units. The degradation is influenced by two standardized accelerating stress variables x_1 and x_2 which are defined over the design region $\mathcal{X} = [0,1]^2$ and act linearly on the response with a potential interaction effect associated with $x_1 x_2$. As in the univariate situation described in Chapter 3, for each testing unit i, the stress variables are set to $\mathbf{x}_i = (x_{i1}, x_{i2})$, and for each component l the response y_{ilj} at time t_j is given by*

$$
\begin{aligned}
y_{ijl} =& \beta_{il1} + \beta_{l2}x_{i1} + \beta_{l3}x_{i2} + \beta_{l4}x_{i1}x_{i2} + \beta_{il5}t_j + \beta_{l6}x_{i1}t_j + \beta_{l7}x_{i2}t_j + \beta_{l8}x_{i1}x_{i2}t_j + \varepsilon_{ijl}, \\
=& \mathbf{f}_l(\mathbf{x}_i, t_j)^T \boldsymbol{\beta}_{il} + \varepsilon_{ijl}
\end{aligned}
$$

(4.23)

where the vector of regression functions $\mathbf{f}_l(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t) = (1, x_1, x_2, x_1x_2, t, x_1t, x_2t, x_1x_2t)^T$ is the same for all components l and $\boldsymbol{\beta}_{il} = (\beta_{il1}, \beta_{l2}, \beta_{l3}, \beta_{l4}, \beta_{il5}, \beta_{l6}, \beta_{l7}, \beta_{l8})^T$. Consequently, it should be further noted that here $\mathbf{g}_1 = ... = \mathbf{g}_r = \mathbf{g}$ where $\mathbf{g}(t) = (1, t)^T$. As noted in Section 4.4 the model equation (4.2) for the mean degradation paths is also assumed to be valid under normal use condition $\mathbf{x}_u = (x_{u1}, x_{u2})$. Hence, the aggregate degradation path under normal use conditions is given by

$$
\mu_l(t) = \mathbf{f}_l(\mathbf{x}_u, t)^T \boldsymbol{\beta}_l = \delta_{l1} + \delta_{l2}t
\tag{4.24}
$$

where $\delta_{l1} = \beta_{l1} + \beta_{l2}x_{u1} + \beta_{l3}x_{u2} + \beta_{l4}x_{u1}x_{u2}$ and $\delta_{l2} = \beta_{l5} + \beta_{l6}x_{u1} + \beta_{l7}x_{u2} + \beta_{l8}x_{u1}x_{u2}$ are the intercept and the slope of the aggregate degradation path $\mu_l(t)$ under normal use conditions, respectively. For the particular case of a series system with two response components, i.e. $r = 2$, the joint failure time distribution function can be expressed as
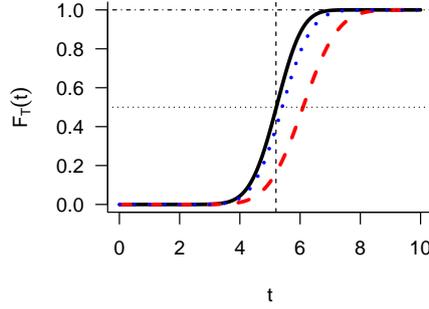
FIGURE 4.1: Distribution function $F_T(t)$ (solid line) at the bivariate linear model with random intercept in Example 4, dashed line: $F_{T_1}(t)$, dotted line: $F_{T_2}(t)$

$$F_T(t) = \mathrm{P}\bigg(\min(T_1, T_2) \leqslant t\bigg) = 1 - \bigg(1 - \Phi\bigg(\frac{\delta_{11} + \delta_{12}t - y_{10}}{\sigma_1(t)}\bigg)\bigg) \times \bigg(1 - \Phi\bigg(\frac{\delta_{21} + \delta_{22}t - y_{20}}{\sigma_2(t)}\bigg)\bigg).$$
$$(4.25)$$

where $\sigma_l^2(t) = Var(\mu_{ul}(t)) = \mathbf{g}_l(t)^T \boldsymbol{\Sigma}_{\gamma l} \mathbf{g}_l(t)$ is the variance function of the mean degradation path of component l. For illustration, the distribution function $F_T(t)$ is plotted in Figure 4.1 under the nominal values given in Table 4.1, the normal use conditions $x_{u1} = -0.40$ and $x_{u2} = -0.20$, and the failure thresholds $y_{10} = 5.4$ and $y_{20} = 5.8$. The median failure time $t_{0.5} = 5.2$ is indicated in Figure 4.1 by a dashed vertical line. Consequently, in view of (4.20) and (4.25) the gradient vector of the parameter vector $\boldsymbol{\beta}$ can be expressed as $\mathbf{c}_\beta = (c_1 \mathbf{f}(\mathbf{x}_u, t_\alpha)^T, c_2 \mathbf{f}(\mathbf{x}_u, t_\alpha)^T)^T$ where the constants c_1 and c_2 are given by

$$c_1 = \phi\bigg(\frac{\delta_{11} + \delta_{12}t_\alpha - y_{10}}{\sqrt{\sigma_{11}^2 + 2\rho_1\sigma_{11}\sigma_{12}t + \sigma_{12}^2 t^2}}\bigg)\bigg(1 - \Phi\bigg(\frac{\delta_{21} + \delta_{22}t_\alpha - y_{20}}{\sqrt{\sigma_{21}^2 + 2\rho_2\sigma_{21}\sigma_{22}t + \sigma_{22}^2 t^2}}\bigg)\bigg),$$

$$c_2 = \phi\bigg(\frac{\delta_{21} + \delta_{22}t_\alpha - y_{20}}{\sqrt{\sigma_{21}^2 + 2\rho_2\sigma_{21}\sigma_{22}t + \sigma_{22}^2 t^2}}\bigg)\bigg(1 - \Phi\bigg(\frac{\delta_{11} + \delta_{12}t_\alpha - y_{10}}{\sqrt{\sigma_{11}^2 + 2\rho_1\sigma_{11}\sigma_{12}t + \sigma_{12}^2 t^2}}\bigg)\bigg).$$

It can be concluded that the optimal design for each of the two univariate components is also optimal for the joint bivariate model under the condition of i.i.d response components. Further, it should be noted that the resulting design remains optimal with regard to the aggregate (fixed) effects parameter vector. Consequently, the problem has been reduced now to finding an optimal design for any of the univariate models with two explanatory variables. As depicted in Chapter 3 the degradation model in equation

TABLE 4.1: Nominal values of the bivariate linear model in Example 4.

β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	β_{16}	β_{17}	β_{18}
2.30	1.60	1.30	0.02	0.70	0.07	0.08	0.03

β_{21}	β_{22}	β_{23}	β_{24}	β_{25}	β_{26}	β_{27}	β_{28}
2.17	1.10	0.84	0.01	0.80	0.03	0.02	0.02

$\sigma_{11}^2 = \sigma_{21}^2$	$\sigma_{12}^2 = \sigma_{22}^2$	σ_ε^2	$\rho_1 = \rho_2$	$t_{0.5}$
0.36	0.10	0.10	0.00	5.2

(4.23) *after rearranging terms can be rewritten as a Kronecker product model*

$$Y_{ijl} = (\mathbf{f}_1(x_{i1}) \otimes \mathbf{f}_2(x_{i2}) \otimes \mathbf{g}(t_j))^T \boldsymbol{\beta}_l + \mathbf{g}(t_j)^T \boldsymbol{\gamma}_{il} + \varepsilon_{ijl}, \qquad (4.26)$$

where $\mathbf{f}_1(x_1) = (1, x_1)^T$, $\mathbf{f}_2(x_2) = (1, x_2)^T$ and $\mathbf{g}(t) = (1, t)^T$ are the marginal regression functions for the stress variables x_1, x_2 and the time variable t, respectively. In the model with two interacting stress variables x_1 and x_2 the marginal model for the combined stress variable $\mathbf{x} = (x_1, x_2)$ is given itself by a product-type structure given both components x_1 and x_2 are specified as simple linear regressions in their corresponding marginal models. Moreover, the experimental region $\mathcal{X} = [0,1]^2$ for the combined stress variable $\mathbf{x}$ is the Cartesian product of the marginal experimental regions $\mathcal{X}_1 = \mathcal{X}_2 = [0,1]$ for the components x_1 and x_2, respectively. In this setting the c-optimal design ξ^* for extrapolation at $\mathbf{x}_u$ can be obtained as the product $\xi^* = \xi_1^* \otimes \xi_2^*$ of the c-optimal designs ξ_v^* for extrapolation at $x_{uv}, v = 1, 2$, in the marginal models (see Theorem 4.4 in (Schwabe, 1996a)).

As specified in (Shat and Schwabe, 2021) the marginal c-optimal designs ξ_v^* assign weight $w_v^* = |x_{uv}|/(1 + 2|x_{uv}|)$ to $x_v = 1$ and weight $1 - w_v^* = (1 + |x_{uv}|)/(1 + 2|x_{uv}|)$ to $x_v = 0$. Hence, the c-optimal design $\xi^* = \xi_1^* \otimes \xi_2^*$ for extrapolation at $\mathbf{x}_u$ is given by

$$\xi^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ (1-w_1^*)(1-w_2^*) & (1-w_1^*)w_2^* & w_1^*(1-w_2^*) & w_1^*w_2^* \end{pmatrix}.$$

Accordingly, the design ξ^* is also optimal for minimization of the asymptotic variance for estimating the α-quantile t_α of the failure time for soft failure due to degradation, when $0 < t_\alpha < \infty$. For instance, under the normal use conditions $x_{u1} = -0.40$ and $x_{u2} = -0.20$ the optimal marginal weights are $w_1^* = 0.222$ and $w_2^* = 0.143$, and the optimal design $\xi^* = \xi_1^* \otimes \xi_2^*$ is given by

$$\xi^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.67 & 0.11 & 0.19 & 0.03 \end{pmatrix} \qquad (4.27)$$
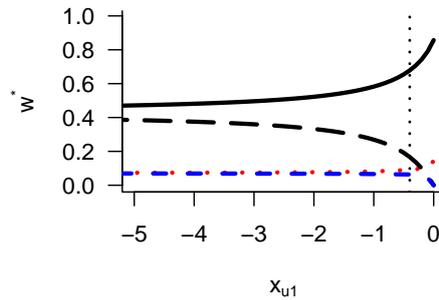
FIGURE 4.2: Optimal weights in dependence on x_{u1} in Example 4, solid line: w_1^*, dotted line: w_2^*, long-dashed line: w_3^*, dashed line: w_4^*
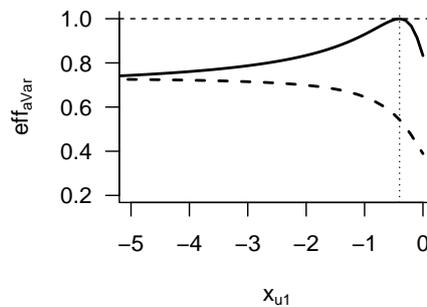
FIGURE 4.3: Efficiency of ξ^* (solid line) and $\bar{\xi}$ (dashed line) in dependence on x_{u1} in Example 4

where the indices of the optimal support points in ξ^* correspond to the design variables x_1 and x_2, respectively. Sensitivity analysis proved that ξ^* is robust against misspecification of the parameter vector $\boldsymbol{\beta}$. To exhibit the dependence on the normal use condition, the optimal weights $w_1^*, ..., w_4^*$ are plotted in Figure 4.2 as a function of x_{u1} where all parameters are held fixed to their nominal values in Table 4.1. It should be noted that similar results are obtained with regard to x_{u2}, and omitted for brevity. As depicted in Figure 4.2 the optimal weights w_3^* and w_4^* which correspond to the maximum testing setting of the first design variable, i.e. $x_{1\max} = 1$, degenerate to zero when the normal use condition approaches the lower bound of $\mathcal{X}_1$, i.e. $x_{u1} \to 0$.   For the setting of the present model with two stress variables, we examine the efficiency of the design ξ^* which is locally optimal for estimation of the median failure time under the nominal values of Table 4.1 when the nominal value of x_{u1} is changed. In Figure 4.3 we plot the efficiency of the locally optimal design ξ^* at the nominal values (solid line) together with

the efficiency of the design $\bar{\bar{\xi}}$ (dashed line) which assigns equal weights $1/4$ to the four vertices $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ where the design $\bar{\bar{\xi}}$ is a standard experimental designs for comparison. In Figure 4.3 the efficiency is illustrated in dependence on the nominal value of x_{u1} while all remaining parameters and constants are held fixed to their nominal value in Table 4.1. The nominal value for x_{u1} is indicated by vertical dotted lines in the corresponding figures. In total, the design ξ^* which is locally optimal for the nominal values seems to perform quite well and is preferred over $\bar{\bar{\xi}}$ throughout.

Example 5. *In this example we consider the model in Subsection 4.2 for an s-out-of-r system to attain a locally c-optimal design. The resulting optimal design is attained in regards to the time plan $\boldsymbol{t} = (0, 0.5, 1)^T$ which is unified for all testing units. We assume here, again, that each of the marginal degradation paths is influenced by two standardized accelerating stress variables x_1 and x_2 which are defined over the design region $\mathcal{X} = [0, 1]^2$. For some testing unit i, the stress variables are set to x_{i1} and x_{i2} the response y_{ijl} of the response component l at time t_j is given by*

$$
\begin{aligned}
y_{ijl} &= \beta_{il1} + \beta_{l2}x_{i1} + \beta_{l3}x_{i2} + \beta_{il4}t_j + \beta_{l5}x_{i2}t_j + \varepsilon_{ijl}, \\
&= \mathbf{f}_l(\mathbf{x}_i, t_j)\boldsymbol{\beta}_{il} + \varepsilon_{ijl}
\end{aligned}
\tag{4.28}
$$

where $\mathbf{f}_l(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t) = (1, x_1, x_2, t, x_2t)^T$, $\boldsymbol{\beta}_{il} = (\beta_{il1}, \beta_{l2}, \beta_{l3}, \beta_{il4}, \beta_{l5})^T$, and $\mathbf{g}_l = \mathbf{g} = (1, t)^T$.

On the basis of the marginal distribution functions $F_{T_l}(t)$ defined in (4.14), the model is extended in terms of the general model (4.15) under normal use conditions. In accordance with Example 4, we consider further an s-out-of-r system with statistically independent response components with equal regression functions and identical covariances. In accordance with the work of (Koucký, 2003), we denote by $F_D(t) = \mathrm{P}(T_l \leq t \; \forall l \in D) = \prod_{l \in D} F_{T_l}(t)$ the probability of joint failure of the components $D \subseteq \{1, ..., r\}$. Consequently, the joint failure time distribution function for a s-out-of-r system is expressed as,

$$
F_T(t) = \sum_{l=0}^{s-1} (-1)^l \binom{l+r-s}{l} \sum_{D; |D|=l+1+r-s} \prod_{d \in D} F_{T_d}(t). \tag{4.29}
$$

The aggregate degradation path under normal use condition for response component l is given by

$$
\mu_l(t) = \mathbf{f}_l(\mathbf{x}_u, t)^T \boldsymbol{\beta}_l = \delta_{l1} + \delta_{l2}t \tag{4.30}
$$

where $\delta_{l1} = \beta_{l1} + \beta_{l2}x_{u1} + \beta_{l3}x_{u2}$ and $\delta_{l2} = \beta_{l4} + \beta_{l5}x_{u2}$ are the intercept and the slope of the aggregate degradation path $\mu_l(t)$ under normal use condition, respectively. Based on equation (4.29) the joint failure time distribution for the particular case of a 2-out-of-3
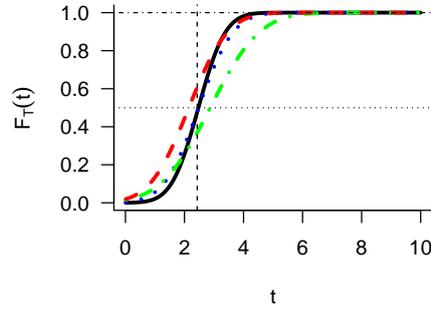
FIGURE 4.4: Distribution function $F_T(t)$ (solid line) for the model in
Example 5, dotted line:$F_{T_1}(t)$, dotted-dashed line:$F_{T_2}(t)$, dashed
line:$F_{T_3}(t)$
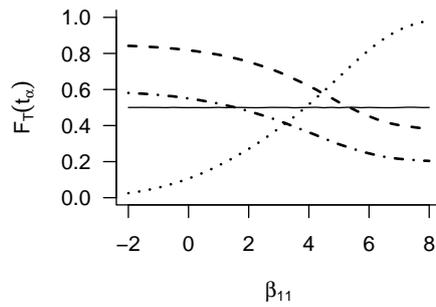
FIGURE 4.5: Dependence of $F_{T_1}(t_\alpha)$ (dotted line), $F_{T_2}(t_\alpha)$
(dotted-dashed line) and $F_{T_3}(t_\alpha)$ (dashed line) on β_{11} for the model in
Example 5

can be expressed as

$$F_T(t) = F_{T_1}(t) \, F_{T_2}(t) + F_{T_1}(t) \, F_{T_3}(t) + F_{T_2}(t) \, F_{T_3}(t) - 2F_{T_1}(t) \, F_{T_2}(t) \, F_{T_3}(t). \quad (4.31)$$

*For further calculation we assume that $\rho_l = 0, \sigma_{l1} = \sigma_1, \sigma_{l2} = \sigma_2, l = 1, 2, 3,$ and, hence,
the variance covariance matrix $\boldsymbol{\Sigma}_\gamma$ turns to a block diagonal matrix with diagonal blocks
$\boldsymbol{\Sigma}_{\gamma 0} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$*

*The distribution function $F_T(t)$ is plotted in Figure 4.4 under the nominal values
given in Table 4.2 and the median failure time $t_{0.5} = 2.43$ is indicated by a dashed
vertical line. The normal use conditions correspond to $x_{u1} = -0.50$ and $x_{u2} = -0.40$,
and the failure thresholds $y_{10} = 7.5$, $y_{20} = 5.2$ and $y_{30} = 4.25$. In view of (4.20)
and (4.31) the gradient vector of the parameter vector $\boldsymbol{\beta}$ can be expressed as $\mathbf{c}_\beta =$*

TABLE 4.2: Nominal values of the multivariate linear model with random effect.

β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	σ_ε^2
3.80	0.52	0.72	2.00	0.67	0.15

β_{21}	β_{22}	β_{23}	β_{24}	β_{25}	σ_1^2
2.20	0.44	0.64	1.50	0.63	0.40

β_{31}	β_{32}	β_{33}	β_{34}	β_{35}	σ_2^2
1.33	0.30	0.92	1.91	0.80	0.32

$(c_1\mathbf{f}(\mathbf{x}_u, t_\alpha)^T, c_2\mathbf{f}(\mathbf{x}_u, t_\alpha)^T, c_3\mathbf{f}(\mathbf{x}_u, t_\alpha)^T)^T$ where the constants c_l, $l = 1, ..., r$ for a general s-out-of-r system are given by

$$\sigma_l(t_\alpha)^{-1}\phi((\delta_{l1} + \delta_{l2}t_\alpha - y_{l0})/\sigma_l(t)) \sum_{m=0}^{s-1}(-1)^m \binom{m+r-s}{m} \sum_{D;|D|=m+r-s;l\notin D} \prod_{d\in D} F_{T_d}(t_\alpha)$$

$$(4.32)$$

and, hence, constants c_1, c_2 and c_3 for the current 2-out-of-3 system are expressed as

$$c_1 = \sigma_1(t_\alpha)^{-1}\phi\left(\frac{\delta_{11} + \delta_{12}t_\alpha - y_{10}}{\sqrt{\sigma_1^2 + \sigma_2^2 t_\alpha^2}}\right)(F_{T_2}(t_\alpha) + F_{T_3}(t_\alpha) - 2F_{T_2}(t_\alpha)F_{T_3}(t_\alpha)),$$

$$c_2 = \sigma_2(t_\alpha)^{-1}\phi\left(\frac{\delta_{21} + \delta_{22}t_\alpha - y_{20}}{\sqrt{\sigma_1^2 + \sigma_2^2 t_\alpha^2}}\right)(F_{T_1}(t_\alpha) + F_{T_3}(t_\alpha) - 2F_{T_1}(t_\alpha)F_{T_3}(t_\alpha)),$$

$$c_3 = \sigma_3(t_\alpha)^{-1}\phi\left(\frac{\delta_{31} + \delta_{32}t_\alpha - y_{30}}{\sqrt{\sigma_1^2 + \sigma_2^2 t_\alpha^2}}\right)(F_{T_1}(t_\alpha) + F_{T_2}(t_\alpha) - 2F_{T_1}(t_\alpha)F_{T_2}(t_\alpha)).$$

As mentioned earlier in this example the marginal response components are independent with equal regression functions and identical covariances. Hence, in accordance with Example 4, the optimization is reduced to finding an optimal design of the first response component under the nominal values in Table 4.2. It should be further noted that the resulting locally c-optimal design will be optimal for any s-out-of-r system under the assumption of independent marginal response components with equal regression functions and identical covariances. In other words, under the assumptions $\mathbf{f}_l = \mathbf{f}$, $\mathbf{g}_l = \mathbf{g}$, $\Sigma_{\gamma l} = \Sigma_{\gamma 0}$ the c-optimal design for extrapolation at $(\mathbf{x}_u, t_\alpha)$ in the LMEM with fixed time plan $\mathbf{t}$ is optimal for estimating t_α. In contrast to Example 4, it should be mentioned that the optimal design for the current experimental settings depends on the given time plan $\boldsymbol{t}$ as well as the nominal values of $\boldsymbol{\beta}$, through the value of t_α, due to the particular form of the gradient c_β as well as the degradation path in

equation (4.28). In particular, the optimal design may also vary with α in contrast to the situation in Example 4. In order to derive a locally c-optimal design ξ^* that minimizes the asymptotic variance of $\widehat{t}_{0.5}$, the multiplicative algorithm with a grid of marginal 0.05 increments over the standardized design region $\mathcal{X} = [0,1]^2$ is used. The resulting optimal design is given by

$$\xi^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.60 & 0.03 & 0.13 & 0.24 \end{pmatrix} \tag{4.33}$$

where the general equivalence theorem is applied to prove the optimality of the numerically obtained design. The locally optimal designs for estimating the median failure time are influenced by the parameter vector $\boldsymbol{\beta}$ as well as the normal use conditions $\mathbf{x}_u$. For brevity, we consider β_{11} and x_{u1} for further analysis procedures. Sensitivity analysis procedures are conducted to demonstrate how the optimal designs change with the parameters and how well they perform under variations of the nominal values. The optimal weights $w_1^*, ..., w_4^*$ are depicted in Figure 4.6 as a function of β_{11} where the variations of $t_{0.5}$ have been generated by letting β_{11} vary over the range -2 to 5 and fixing all remaining parameters to their nominal values in Table 4.2. The analysis indicated that the optimal weights in (4.33) slightly change under variations of β_{11}. On the other hand the optimal weights $w_1^*, ..., w_4^*$ are plotted in Figure 4.7 as a function of x_{u1}, while all remaining parameters are held fixed to their nominal values in Table 4.2. The results exhibit that the optimal weights are more sensitive to variations of x_{u1} when compared to the misspecifications of β_{11}. Further, Figure 4.5 illustrates the dominance of the marginal failure components where the marginal distribution functions $F_{T_1}(t_\alpha)$, $F_{T_2}(t_\alpha)$, and $F_{T_3}(t_\alpha)$ are shown in dependence on β_{11}. Figure 4.5 depicts that the first component dominates for large values of its intercept β_{11} while the second and third components dominate for small values of β_{11}.

For the present setting of the mixed effects model with two stress variables we examine further, based on equation (4.22), the efficiency of the design ξ^* which is locally optimal for estimation of the median failure time under the nominal values of Table 4.2 when the nominal values are misspecified. In Figure 4.8 and Figure 4.9 the efficiencies of ξ^* along with the efficiency of $\bar{\xi}$ are displayed in dependence on the true value of β_{11} and the normal use condition x_{u1}, respectively. The results indicate that ξ^* performs generally well under misspecifications of β_{11} and x_{u1} with more robustness with regard to variations of β_{11}. In total, the optimal design ξ^* is quite preferable over the standard design $\bar{\xi}$ throughout.
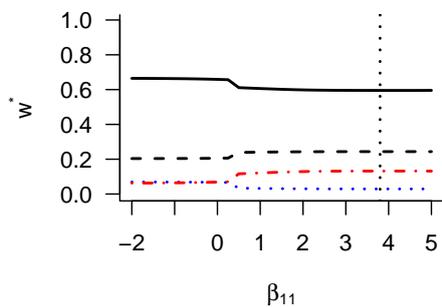
FIGURE 4.6: Optimal weights in dependence on β_{11} for the model in Example 5, solid line: w_1^*, dotted line: w_2^*, dotted-dashed line: w_3^*, dashed line: w_4^*
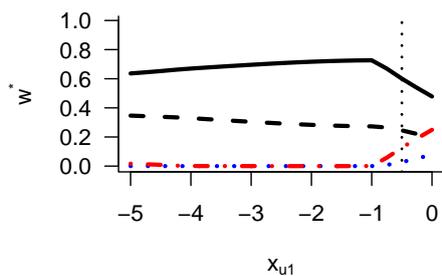
FIGURE 4.7: Optimal weights in dependence on x_{u1} for the model in Example 5, solid line: w_1^*, dotted line: w_2^*, dotted-dashed line: w_3^*, dashed line: w_4^*
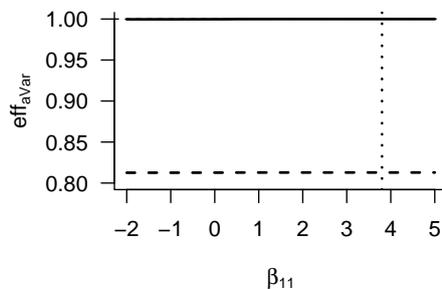
FIGURE 4.8: Efficiency of ξ^* (solid line) and $\bar{\xi}$ (dashed line) in dependence on $t_{0.5}$ for the model in Example 5
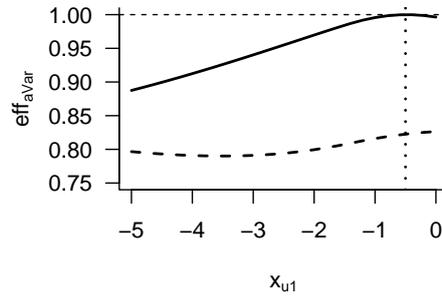
FIGURE 4.9: Efficiency of ξ^* (solid line) and $\bar{\xi}$ (dashed line) in dependence on x_{u1} for the model in Example 5

4.6 Conclusion

Designing highly reliable systems needs a sufficient assessment of the reliability related characteristics. A common approach to handle this issue is to conduct accelerated degradation testing which provides an estimation of lifetime and reliability of the system under study in a relatively short testing time. To account for variability between units in accelerated degradation tests, we assume int this work that the marginal degradation functions can be described by a mixed-effects linear model. This also leads to a non-degenerate distribution of the failure time, due to soft failure by exceedance of the expected (conditionally per unit) degradation path over a threshold, under normal use conditions. Therefore we are aiming to estimate certain quantiles of the joint failure time distribution as a property of the reliability of the product. In this regard we considered the availability of non-degenerate solutions for the quantiles. The purpose of optimal experimental design is then to find the best settings for the stress variables to obtain most accurate estimates for these quantities.

For the existing degradation models in this chapter it is further assumed that stress remains constant within each testing unit during the whole period of experimental measurements but may vary between units. Hence, in the corresponding experiment a cross-sectional design between units has to be specified for the stress variable while for repeated measures.

In the present chapter we presented optimal designs for accelerated degradation testing under bivariate LMEMs with full as well as partial interactions between the time and stress variables.

For all models the efficiency of the corresponding optimal design is considered to assess its performance when nominal values are varied at the design stage.

Chapter 5

Optimal Designs in Accelerated Degradation Testing for Various Degradation Models

5.1 Introduction

We characterize in this chapter optimal designs for accelerated degradation tests with various bivariate degradation models. The first bivariate model includes two gamma processes as marginal degradation models. The second bivariate models is expressed by a gamma process along with a LMEM. We derive optimal designs for minimizing the asymptotic variance for estimating some quantile of the failure time distribution at the normal use conditions. Sensitivity analysis is conducted to study the behavior of the resulting optimal designs under misspecifications of adopted nominal values. Considering Gamma process models, (Pan and Sun, 2014) introduced reliability model of the degradation products with two performance characteristics based on a Gamma process. In order to predict the lifetime of the population from Accelerated Degradation Testing, (Wang, Xu, and Mi, 2015) considered Gamma process with a time transformation and random effects. They present a deducing method for determining the relationships between the shape and scale parameters of Gamma process and accelerated stresses. (Lim, 2015) developed statistical methods for optimal designing Accelerated Degradation Testing plans under the total experimental cost constraint and assuming that the degradation characteristic follows a Gamma process model. Accelerated Degradation Testing with the presence of Competing failure modes is an important reliability area to be addressed. Therefore, the study of the statistical inference of Accelerated Degradation Testing with competing failures is of great significance. (Haghighi and Bae, 2015) introduced a modeling approach to simultaneously analyze linear degradation data and traumatic failures with competing risks in an SSADT experiment. Moreover, methodology for ALT planning when there are two or more independent failure modes was discussed by (Pascual, 2007). The author assumed that the failure modes have

respective latent failure times, and the minimum of these times corresponds to the product lifetime. Considering accelerated destructive degradation tests (ADDT), (Shi and Meeker, 2014) proposed methods to find unconstrained and constrained optimum test plans for competing risk applications under a *V*-optimality criterion that aim to minimize the large-sample approximate variance of a failure-time distribution quantile at use conditions. The authors consider linearly degraded response models with an application for an adhesive bond. In regards to nonparameteric methods of evaluation, (Balakrishnan and Qin, 2019) introduced some approximation techniques of the first passage time distribution of the degradation processes incorporating random effects if the process type is unknown. The authors approximate the density function of some stochastic degradation processes by inverting the empirical Laplace transform using the empirical saddlepoint method. (Palayangoda, Ng, and Butler, 2020) extended the work of (Balakrishnan and Qin, 2019) and proposed some improved approximation techniques where numerical examples and Monte Carlo simulation studies are used to illustrate the advantages of the proposed techniques. The rest of this chapter is organized as follows. Section 5.2 is devoted to develop optimal experimental designs for a univariate gamma model. In Section 5.3 we introduce an optimal design considering a bivariate Gamma process where the corresponding failure modes do not interact. In Section 5.4, we characterize a *c*-optimal design for an ADT with a bivariate degradation model with repeated measures given that one of the marginal response components follows a Gamma process model where the other follows a LMEM. The paper closes with a short discussion in Section 5.5. It should be noted that the content of the current chapter is directly related to the paper of (Shat and Schwabe, 2019).

5.2 Accelerated degradation testing based on a gamma process

The gamma process is a natural stochastic model for degradation processes in which degradation is assumed to occur gradually over time in a sequence of independent increments. In this section, we assume that the testing unit has a single dependent failure mode where the degradation path is characterized by a Gamma process model in terms of a standardized time variable t. In addition, it is assumed that there is a single stress variable and its (standardized) stress level x can be chosen by the experimenter from the experimental region $\mathcal{X} = [0, 1]$. The subsequent subsections clarify the approximation of the Gamma model with a generalized linear model approach. Further, we explain the derivation of the corresponding information matrix in order to

obtain an algorithm based optimal experimental design with respect to the asymptotic variance of a quantile of the failure time distribution.

5.2.1   Model formulation

A gamma process Z_t is a stochastic process with independent gamma distributed increments. The process can be parameterized by the rate γ and a scale parameter ν. If the process is observed at k subsequent time points t_j, $0 < t_1 < ... < t_k$, then the jth degradation increment $Y_j = Z_{t_j} - Z_{t_{j-1}}$ is Gamma distributed with shape $\gamma\Delta_j$ and scale ν, where $\Delta_j = t_j - t_{j-1}$ is the length of the jth time interval and $Z_{t_0} = 0$ at $t_0 = 0$.

We assume that the stress variable x only affects the rate $\gamma = \gamma(x)$ of the Gamma process and, hence, the shape parameters $\gamma(x)\Delta_j$ of the increments while the scale parameter ν is fixed and known. We further assume that the rate $\gamma(x)$ is given by a linear trend in the stress variable with a logarithmic link,

$$\gamma(x) = e^{\beta_0 + \beta_1 x}, \tag{5.1}$$

where the intercept β_0 and the slope β_1 are to be estimated. When a unit is tested under stress level x during a time interval of length Δ, the degradation increment Y has density

$$f_Y(y) = \frac{y^{\gamma(x)\Delta - 1} e^{-y/\nu}}{\Gamma(\gamma(x)\Delta)\nu^{\gamma(x)\Delta}}, \tag{5.2}$$

where $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} \mathrm{d}z$ is the Gamma function. The mean of the increment is given by

$$\mu(x) = \mathrm{E}(Y) = \gamma(x)\Delta\nu = e^{\beta_0 + \beta_1 x}\Delta\nu. \tag{5.3}$$

Thus the mean $\mu(x)$ is linked to the linear predictor $\beta_0 + \beta_1 x$ by a scaled log link. Hence, the model assumptions fit into the concept of generalized linear models.

To be more specific, in accelerated degradation testing n distinct testing units are tested at potentially different stress settings x_i which are held fixed over time for each unit $i = 1, ..., n$. Measurements are made at predetermined time points $t_1, ..., t_k$ which are identical for all units. The degradation increments Y_{ij} when testing unit i during the jth time interval of length $\Delta_j = t_j - t_{j-1}$ are independent Gamma distributed with shape $\gamma(x_i)\Delta_j$ and scale ν.

5.2.2 Estimation and information

Denote by $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ the vector of unknown parameters. By (5.2) the log-likelihood of a single degradation increment Y is given by

$$\ell(\boldsymbol{\beta}; y) = (e^{\beta_0 + \beta_1 x}\Delta - 1)\log(y) - y/\nu - \log\left(\Gamma(e^{\beta_0 + \beta_1 x}\Delta)\right) - e^{\beta_0 + \beta_1 x}\Delta\log(\nu) \quad (5.4)$$

when the stress level x is applied and the increment is measured over a time interval of length Δ. The elemental Fisher information matrix $\mathbf{M}_{\boldsymbol{\beta}}(x, \Delta)$ related to a single increment can be calculated as minus the matrix of expected second order derivatives of the log-likelihood,

$$\mathbf{M}_{\boldsymbol{\beta}}(x, \Delta) = q(\beta_0 + \beta_1 x + \log(\Delta)) \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}, \quad (5.5)$$

where q is defined by $q(z) = e^{2z}\psi_1(e^z)$ and $\psi_1(\alpha) = \mathrm{d}^2 \ln\left(\Gamma(\alpha)\right)/\mathrm{d}\alpha^2$ is the trigamma function.

Because the increments $Y_{i1}, ..., Y_{ik}$ measured at times $t_1, ..., t_k$ are statistically independent within a unit i, the log-likelihood $\ell(\boldsymbol{\beta}; y_{i1}, ..., y_{ik}) = \sum_{j=1}^{k} \ell(\boldsymbol{\beta}; y_{ij})$ of a unit i is the sum of the log-likelihoods for the single observations Y_{ij}. Thus also the information matrix $\mathbf{M}_{\boldsymbol{\beta}}(x_i)$ of a unit is the sum of the information of the single increments,

$$\mathbf{M}_{\boldsymbol{\beta}}(x_i) = \sum_{j=1}^{k} \mathbf{M}_{\boldsymbol{\beta}}(x_i, \Delta_j) = \lambda(\beta_0 + \beta_1 x_i) \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}, \quad (5.6)$$

where the "intensity" $\lambda(z) = \sum_{j=1}^{k} q\left(z + \log(\Delta_j)\right)$ accounts for the contribution of the non-linearity at $z = \beta_0 + \beta_1 x_i$ to the information.

Furthermore, because measurements are statistically independent between units, both the log-likelihood $\ell(\boldsymbol{\beta}; y_{11}, ..., y_{nk}) = \sum_{i=1}^{n} \ell(\boldsymbol{\beta}; y_{i1}, ..., y_{ik})$ and the information

$$\mathbf{M}_{\boldsymbol{\beta}}(x_1, ..., x_n) = \sum_{i=1}^{n} \mathbf{M}_{\boldsymbol{\beta}}(x_i) \quad (5.7)$$

for the whole experiment summarize the log-likelihood and the information of the units. This information matrix $\mathbf{M}_{\boldsymbol{\beta}}(x_1, ..., x_n)$ provides a measure for the performance of the experiment as its inverse is proportional to the asymptotic variance covariance matrix for the maximum likelihood estimator of $\boldsymbol{\beta}$.

In an accelerated degradation experiment the stress variable x is under control of the experimenter. For each unit i, the setting x_i of the stress variable adjusted to i may be chosen from an experimental region $\mathcal{X}$. The collection $x_1, ..., x_n$ of these settings is called the design of the experiment. An optimal design then aims at minimizing an

optimality criterion which is a function of the information matrix. As explained earlier, finding optimal designs $x_1, ..., x_n$ is, in general, a difficult task of discrete optimization. Hence, we consider for further calculations the approach of approximate designs ξ as introduced in Subsection 2.3.1. Accordingly, the corresponding standardized, per unit information matrix is defined as

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^{m} w_i \mathbf{M}_\beta(x_i) \tag{5.8}$$

so that "exact" designs $x_1, ..., x_n$ are properly embedded by $\mathbf{M}_\beta(\xi) = (1/n)\mathbf{M}_\beta(x_1, ..., x_n)$. As the information matrix depends on the parameter vector $\boldsymbol{\beta}$ only through the linear predictor $\beta_0 + \beta_1 x$, a canonical transformation can be employed which simultaneously maps experimental settings x to $z = \beta_0 + \beta_1 x$ and the parameters β_0 and β_1 to the standardized value $\beta_0 = 0$ and $\beta_1 = 1$ for analytical solutions, see (Ford, Torsney, and Wu, 1992). When all time intervals have the same length $\Delta_j = \Delta$, $j = 1, ..., k$, the influence of the repeated measurements reduces to $\lambda(z) = k\, q\big(z + \exp(\Delta)\big)$ for the intensity and, hence, to a multiplicative factor k in the information matrix. Thus, for common design criteria, the number k of measurements is immaterial for design optimization.

5.2.3   Optimality criterion based on the failure time distribution

In degradation testing we are interested in characteristics of the failure time distribution of soft failure due to degradation under normal use condition x_u. It is supposed that the gamma process $Z_{u,t}$ describing the degradation under normal use condition has the rate $\gamma(x_u) = \exp(\beta_0 + \beta_1 x_u)$ as in equation (5.1) and scale ν. Typically the normal use condition x_u is not contained in the experimental region $\mathcal{X}$, $x_u < 0$. Further it is natural to assume that the degradation paths are strictly increasing over time. Then a soft failure due to degradation is defined as exceedance of the degradation path over a failure threshold z_0. The failure time T under normal use condition is defined as the first time t the degradation path $Z_{u,t}$ reaches or exceeds the threshold z_0, i.e., $T = \inf\{t \geq 0; Z_{u,t} \geq z_0\}$. In order to derive certain characteristics of the distribution of the failure time, we determine its distribution function $F_T(t) = \mathrm{P}(T \leq t)$. For this note that $T \leq t$ if and only if $Z_{u,t} \geq z_0$. The degradation $Z_{u,t}$ at time t is gamma distributed with shape $\gamma(x_u)t$ and scale ν). Hence, the distribution function of the

failure time T can be expressed as

$$
\begin{aligned}
F_T(t) &= \mathrm{P}(Z_{u,t} \geq z_0) \\
&= \frac{1}{\Gamma(\gamma(x_u)t)} \int_{z_0}^{\infty} (z/\nu)^{\gamma(x_u)t-1} e^{-z/\nu} \nu^{-1} \mathrm{d}z \\
&= Q(\gamma(x_u)t, z_0/\nu)
\end{aligned}
\tag{5.9}
$$

where $Q(s,z) = \Gamma(s,z)/\Gamma(s)$ is the regularized gamma function and $\Gamma(s,z) = \int_z^\infty x^{s-1} e^{-x} \mathrm{d}x$ the incomplete gamma function.

We will be interested in some quantile t_α of the failure time distribution. In the case of a continuous distribution function $F_T(t)$, the α-quantile t_α satisfies $F_T(t_\alpha) = \alpha$, i. e., it represents the time up to which under normal use conditions $\alpha \cdot 100$ percent of the units fail and $(1-\alpha) \cdot 100$ percent of the units persist. The distribution function and, hence, the quantile $t_\alpha = t_\alpha(\boldsymbol{\beta})$ depends on the parameter vector $\boldsymbol{\beta}$ in which the quantile t_α is a decreasing functions of the linear predictor $\beta_0 + \beta_1 x_u$.

With this functional relationship the maximum likelihood estimator for the quantile t_α is given by $\widehat{t}_\alpha = t_\alpha(\widehat{\boldsymbol{\beta}})$, where $\widehat{\boldsymbol{\beta}}$ is the maximum likelihood estimator of $\boldsymbol{\beta}$. The performance of these estimators is measured by their asymptotic variance $\mathrm{aVar}(\widehat{t}_\alpha)$, and design optimization will be conducted with respect to the minimum asymptotic variance criterion, i. e. an optimal design minimizes $\mathrm{aVar}(\widehat{t}_\alpha)$. This criterion is commonly used in planning degradation tests when experimenters are interested in accurately estimating reliability properties of a system over its life cycle.

If the distribution function $F_T(t)$ is strictly increasing with continuous density $f_T(t) = F_T'(t)$, the asymptotic variance can be derived by the delta method from the information matrix in Section 5.2.2 as

$$
\mathrm{aVar}(\widehat{t}_\alpha) = \mathbf{c}^T \mathbf{M}_\beta(\xi)^{-1} \mathbf{c},
\tag{5.10}
$$

where $\mathbf{c} = \partial t_\alpha(\boldsymbol{\beta})/\partial \boldsymbol{\beta}$ is the vector of partial derivatives of $t_\alpha = F_T^{-1}(\alpha)$ with respect to the components of the parameter vector $\boldsymbol{\beta}$ evaluated at the true values of $\boldsymbol{\beta}$. Let $g(s) = Q(s, z_0/\nu)$ be the regularized Gamma function with the second argument fixed to z_0/ν, then $t_\alpha = g^{-1}(\alpha)/\gamma(x_u)$ by (5.9) and the vector $\mathbf{c}$ of partial derivatives can be written as $\mathbf{c} = -t_\alpha(1, x_u)^T$, where the minus sign and the scaling factor t_α do not affect the optimization problem. Hence, the minimum asymptotic variance criterion is equivalent to a c-criterion with $\mathbf{c} = (1, x_u)^T$, i. e., extrapolation of the linear component $\beta_0 + \beta_1 x_u$ at the normal use condition x_u, and standard optimization methods for c-criteria can be employed. In particular, the design optimization does not depend on which quantile t_α is to be estimated, and the obtained design is simultaneously optimal for all α.

Because the information matrix $\mathbf{M}_\beta(\xi)$ depends on the parameter vector $\boldsymbol{\beta}$, this affects the design optimization. Hence, nominal values have to be assumed for these parameters, and locally optimal designs can be obtained for those nominal values. Numerical calculations indicate that the locally optimal designs ξ^* are supported on the endpoints of the design region $\mathcal{X}$, i.e., they are of the form $\xi^* = \xi_{w^*}$, where ξ_w denotes a design with weight $w_1 = w$ on $x_1 = 0$ and weight $w_2 = 1 - w$ on $x_2 = 1$. Under this premise the optimal weight w^* can be determined analytically by Elfving's theorem (Elfving, 1952),

$$w^* = \frac{(1 + |x_u|)\sqrt{\lambda(\beta_0 + \beta_1)}}{(1 + |x_u|)\sqrt{\lambda(\beta_0 + \beta_1)} + |x_u|\sqrt{\lambda(\beta_0)}} \tag{5.11}$$

for (standardized) normal use condition $x_u < 0$. This optimal weight w^* is a decreasing function in the distance $|x_u|$ between the normal use condition and the lowest stress level $x_1 = 0$, and it decrease from $w^* = 1$ when formally letting $x_u = 0$ to $\sqrt{\lambda(\beta_0 + \beta_1)}/(\sqrt{\lambda(\beta_0 + \beta_1)} + \sqrt{\lambda(\beta_0)})$ for $x_u \to -\infty$, where this lower bound is larger than 0.5 since $\beta_1 > 0$ and the intensity $\lambda(z)$ is an increasing function in z.

Concerning the parameters β_0 and β_1 the optimal weight w^* is increasing in the slope parameter β_1 while it does not seem to be sensitive with respect to the intercept parameter β_0 as will be illustrated in Figure 5.3 and Figure 5.4 below for some nominal values. Therefore it is of interest to check how a misspecification of the nominal values for $\boldsymbol{\beta}$ may affect the performance of a locally optimal design $\xi^* = \xi_{w^*}$. To measure the performance we make use of the concept of efficiency

$$\text{eff}_{\text{aVar}}(\xi; \boldsymbol{\beta}) = \frac{\text{aVar}_\beta(\widehat{t}_\alpha; \xi^*_\beta)}{\text{aVar}_\beta(\widehat{t}_\alpha; \xi)} \tag{5.12}$$

of a design ξ with respect to the asymptotic variance for estimating t_α when $\boldsymbol{\beta}$ is the true value of the parameter, where $\text{aVar}_\beta(\widehat{t}_\alpha; \xi)$ denotes the asymptotic variance of $\widehat{t}_\alpha$ at $\boldsymbol{\beta}$, when the design ξ is used, and ξ^*_β is the locally optimal design at $\boldsymbol{\beta}$. This efficiency attains a value between 0 and 1. It can be interpreted as the proportion of units needed, when the locally optimal design ξ^*_β is used, to obtain the same precision in the asymptotic variance as for the design ξ under consideration. Thus high values of the efficiency are advantageous for a design to be used.

5.2.4 Numerical example

In this example we consider an accelerated degradation experiment as described in Subsection 5.2.1 with standardized normal use condition $x_u = -0.4$, underlying gamma process with scale parameter $\nu = 1$ and degradation threshold $z_0 = 5.16$. We will be interested in estimating the median $t_{0.5}$ of the failure time T due to degradation. The
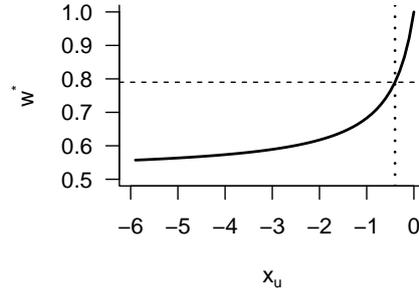
FIGURE 5.1: Optimal weight w^* in dependence on the normal use condition x_u for the univariate Gamma process in the example of Subsection 5.2.4

standardized observation times are $t_j = 0.25$, 0.5, 0.75 and 1, i.e., there are $k = 4$ degradation increments measured on time intervals of constant length $\Delta = 0.25$. With respect to the location parameters we assume the nominal values $\beta_0 = 0.23$ for the intercept and $\beta_1 = 0.53$ for the slope. For these parameter values, the distribution function $F_T(t)$ of the failure time T is exhibited in Figure 5.7 below as F_{T_1}. The corresponding median failure time $t_{0.5} = 5.39$ for which $F_T(t_{0.5}) = 1/2$ is indicated in Figure 5.7 by a dashed vertical line.

To find the optimal design ξ^* for estimating the median failure time t_α, we apply the multiplicative algorithm following (Torsney and Martín-Martín, 2009) for the standardized stress parameter x on a grid with increments of size 0.01 on the design region $\mathcal{X} = [0, 1]$. The optimal design ξ^* is found to be of the form ξ_w with optimal weight $w^* = 0.79$ at the lowest stress level $(x = 0)$ and weight $1 - w^* = 0.21$ at the highest stress level $(x = 1)$ in the experiment.

For illustrative purposes, the optimal weight w^* is plotted in Figure 5.1 as a function of the normal use condition x_u when the nominal values of the other parameters are held fixed.

For the normal use condition x_u close to the lowest value $x = 0$ of the design region, the optimal weight w^* approaches 1, and w^* decreases to 0.516 when x_u is far away from the design region $(x_u \to -\infty)$. The nominal value $x_u = -0.4$ and the corresponding optimal weight $w^* = 0.79$ are indicated in Figure 5.1 by a vertical and a horizontal dashed line, respectively. To imagine the gain in applying the optimal design ξ^*, we exhibit the efficiency (5.12) of commonly used standard designs $\bar{\xi}_2$ and $\bar{\xi}_3$ in Figure 5.2, when the nominal values for the other parameters are held fixed. In this comparison the designs $\bar{\xi}_m$ are uniform on m equidistant stress values $x_1, ..., x_m$ covering the whole range of the design region $0 \le x \le 1$. In particular, $\bar{\xi}_2$ is of the form ξ_w with $w = 1/2$,
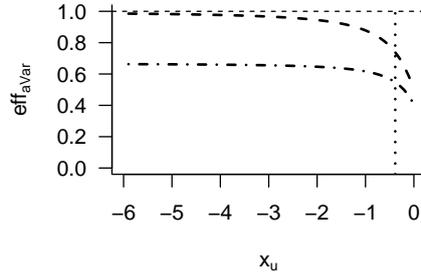
FIGURE 5.2: Efficiency of $\bar{\xi}_2$ (dashed line) and $\bar{\xi}_3$ (dashed-dotted line) in dependence on the normal use condition x_u in the example of Subsection 5.2.4

and $\bar{\xi}_3$ assigns weight $1/3$ to each of the endpoints $(x = 0)$ and $(x = 1)$ and to the midpoint $(x = 0.5)$ of the design region.

The nominal value $x_u = -0.4$ of the normal use condition is indicated in Figure 5.2 by a vertical dotted line. The uniform two-point design $\bar{\xi}_2$ shows a high efficiency at values x_u which are sufficiently far from the standardized design region. This is in accordance with the similarity of the weights in $\bar{\xi}_2$ and in the optimal design ξ^* for such values. For x_u close to the lowest experimental stress level $x = 0$ the efficiency of $\bar{\xi}_2$ drops to $50\,\%$. The uniform three-point design $\bar{\xi}_3$ shows a much lower efficiency throughout. At the nominal value $x_u = -0.4$ the efficiency of the uniform two- and three-point designs $\bar{\xi}_2$ and $\bar{\xi}_3$ is $\mathrm{eff}_{\mathrm{aVar}}(\bar{\xi}_2; \boldsymbol{\beta}) = 75\,\%$ and $\mathrm{eff}_{\mathrm{aVar}}(\bar{\xi}_3; \boldsymbol{\beta}) = 55\,\%$, respectively. That means that for the optimal design ξ^* only 75 percent of units are required compared to the design $\bar{\xi}_2$ and 55 percent of units compared to the design $\bar{\xi}_3$ to achieve the same accuracy for estimating the median failure time.

To assess the sensitivity of the locally optimal design $\xi^* = \xi_{w^*}$ we plot the optimal weights w^* in dependence on the intercept and slope parameters β_0 and β_1 in Figure 5.3 and Figure 5.4, respectively, while the other nominal values are held fixed. As mentioned earlier, the

Finally, to judge the performance of the locally optimal design $\xi^* = \xi_{\boldsymbol{\beta}}^*$ at the proposed nominal values $\beta_0 = 0.23$ and $\beta_1 = 0.53$ of the location parameters under misspecifications, we show the efficiency in dependence on the intercept and slope parameters β_0 and β_1 in Figure 5.5 and Figure 5.6, respectively, while the other parameters are held fixed to their nominal values. Figure 5.5 displays that the optimal design ξ^* maintain its efficiency under misspecifications of β_0. In contrast, Figure 5.6 depicts that misspecifications of the slope β_1 substantially affects the efficiency of the design ξ^* and more attention should be paid to a correct specification of the nominal value for the slope parameter opposite to the intercept.
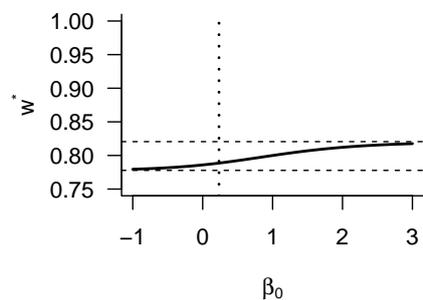
FIGURE 5.3: Dependence of the optimal weight w^* on β_0 in the example of Subsection 5.2.4
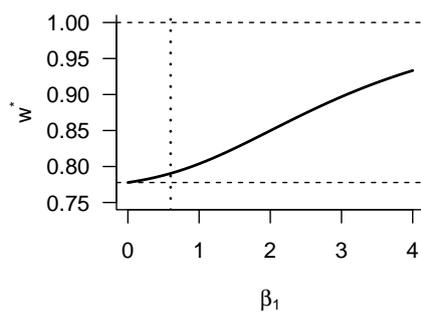
FIGURE 5.4: Dependence of the optimal weight w^* on β_1 in the example of Subsection 5.2.4
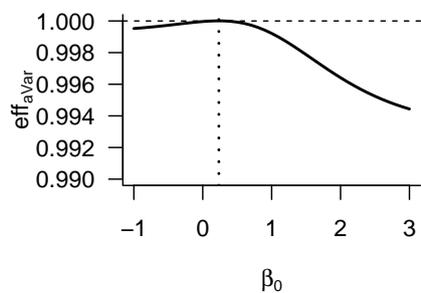
FIGURE 5.5: Efficiency of ξ^* in dependence on β_0 in the example of Subsection 5.2.4
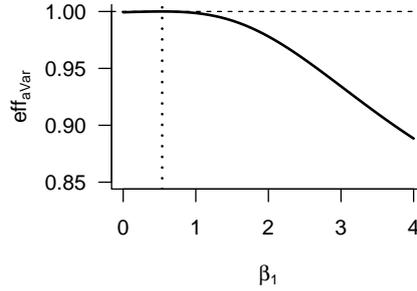
FIGURE 5.6: Efficiency of ξ^* in dependence on β_1 in the example of Subsection 5.2.4

5.3 Bivariate accelerated degradation testing with two gamma processes

We consider now the optimal design problem for a bivariate degradation process incorporating serially two independent failure modes which means that a failure of the system occurs when one of the two components fail.

5.3.1 Model formulation

We assume that degradation in the two components takes place according to independent Gamma processes $Z_t^{(1)}$ and $Z_t^{(2)}$, respectively, as described in Section 5.2, where for both processes $Z_t^{(l)}$ the rate $\gamma_l(x) = e^{\beta_{l0} + \beta_{l1}x}$ depends on the same standardized accelerating stress variable $x \in \mathcal{X} = [0,1]$ via a linear trend $\beta_{l0} + \beta_{l1}x$ under the log link as in (5.1). By assumption the degradation increments $Y_{ilj} = Z_{t_j}^{(l)} - Z_{t_{j-1}}^{(l)}$ of both components during the jth time interval of length Δ_j are all gamma distributed with shape $\gamma_l(x_i)\Delta_j$ and scale ν_l, $l = 1, 2$, and independent.

The failure times T_1 and T_2 of the components for soft failure due to degradation are defined as in Subsection 5.2.3. The failure of the system occurs when either of the two components fail, and the failure time T of the system is defined by $T = \min\{T_1, T_2\}$. Because of the independence of the underlying processes, the failure times T_1 and T_2 of the components are independent.

5.3.2 Information

Denote by $\boldsymbol{\beta}_l = (\beta_{l1}, \beta_{l2})^T$ the marginal parameter vector associated with the lth failure mode. Because of the independence of the components the joint log-likelihood of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ is the sum $\ell(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \ell(\boldsymbol{\beta}_1; y_{111}, ..., y_{n1k}) + \ell(\boldsymbol{\beta}_2; y_{121}, ..., y_{n2k})$ of the log-likelihoods

$\ell(\boldsymbol{\beta}_l; y_{1l1}, ..., y_{nlk})$ of the components given by 5.4. Hence, the maximum likelihood estimators $\widehat{\boldsymbol{\beta}}_l$ of $\boldsymbol{\beta}_l$ in the whole system coincides with those in the marginal models and the joint information matrix $\mathbf{M}_{\beta_1, \beta_2}(x_1, ..., x_n)$ for all parameters is block diagonal,

$\mathbf{M}_{\beta_1, \beta_2}(x_1, ..., x_n) = \begin{pmatrix} \mathbf{M}_{\beta_1}(x_1, ..., x_n) & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\beta_2}(x_1, ..., x_n) \end{pmatrix}$, where the diagonal blocks $\mathbf{M}_{\beta_l}$ are the marginal information matrices for the single failure modes as specified in Subsection 5.2.2. Accordingly, for approximate designs ξ the standardized information matrix

$$\mathbf{M}_{\beta_1, \beta_2}(\xi) = \begin{pmatrix} \mathbf{M}_{\beta_1}(\xi) & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\beta_2}(\xi) \end{pmatrix}. \tag{5.13}$$

is also block diagonal with the marginal information matrices $\mathbf{M}_{\beta_l}(\xi)$ on the diagonal.

5.3.3 Optimality criterion based on the failure time distribution

As in Section 5.2.3 we are interested in characteristics of the failure time distribution of soft failure due to degradation under normal use condition x_u. The marginal failure times T_l under normal use condition are defined as the first time t the degradation path $Z_{u,t}^{(l)}$ reaches or exceeds the corresponding threshold z_{l0}, i. e., $T_l = \inf\{t \geq 0; Z_{u,t}^{(l)} \geq z_{l0}\}$. A failure of the system occurs if one of the components fail. Hence, the failure time T of the system is defined by $T = \min\{T_1, T_2\}$. Because of the independence of the components the survival function $1 - F_T(t) = \mathrm{P}(T_1 > t, T_2 > t)$ factorizes into the marginal survival functions $1 - F_{T_l}(t)$. Hence, the failure time distribution of the system can be expressed as

$$F_T(t) = 1 - (1 - F_{T_1}(t))(1 - F_{T_2}(t)), \tag{5.14}$$

where $F_{T_l}(t) = Q(\gamma_l(x_u)t, z_{l0}/\nu_l)$ by (5.9).

As in Subsection 5.2.3, we will consider quantiles t_α of the failure time distribution. Also here the distribution function F_T and, hence, the quantile $t_\alpha = t_\alpha(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ is a function of the parameters and the maximum likelihood estimate $\widehat{t}_\alpha = t_\alpha(\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2)$ of the quantile t_α is based on the maximum likelihood estimates $\widehat{\boldsymbol{\beta}}_l$ of $\boldsymbol{\beta}$ for the components.

The task of designing the experiment is to provide an as precise estimate of the α-quantile as possible, i. e., to minimize the asymptotic variance $\mathrm{aVar}(\widehat{t}_\alpha)$ of $\widehat{t}_\alpha$ at the normal use condition. As in Subsection 5.2.3 the asymptotic variance can be obtained as $\mathrm{aVar}(\widehat{t}_\alpha) = \mathbf{c}^T \mathbf{M}_{\beta_1, \beta_2}(\xi)^{-1} \mathbf{c}$, where $\mathbf{c} = (\mathbf{c}_1^T, \mathbf{c}_2^T)^T$ and $\mathbf{c}_l = \partial t_\alpha(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)/\partial \boldsymbol{\beta}_l$ is the vector of partial derivatives of t_α with respect to the parameter vector $\boldsymbol{\beta}_l$ evaluated at the true values of $\boldsymbol{\beta}_l$. Differently from the univariate case in Subsection 5.2.3 there is no explicit formula for t_α. Therefore, the gradient vectors $\mathbf{c}_l$ will be derived by the

implicit function theorem as

$$\frac{\partial t_\alpha}{\partial \boldsymbol{\beta}_l} = -\frac{1}{f_T(t_\alpha)} \frac{\partial F_T(t_\alpha)}{\partial \boldsymbol{\beta}_l} \tag{5.15}$$

in terms of the failure time distribution $F_T(t)$, where $f_T(t) = \partial F_T(t)/\partial t$ is the density of T. The common scaling factor $c_0 = -1/f_T(t_\alpha)$ is irrelevant for the optimization problem. Hence, the components of the c-criterion vector $\mathbf{c} = (\mathbf{c}_1^T, \mathbf{c}_2^T)^T$ can be reduced to $\mathbf{c}_l = \partial F_T(t_\alpha)/\partial \boldsymbol{\beta}_l$. Based on equation 5.14, the gradient vectors $\mathbf{c}_l$ can be expressed as $\mathbf{c}_l = c_l(1, x_u)^T$ similar to the univariate case, where the constant c_l can be expressed as, see (Tsai, Tseng, and Balakrishnan, 2012),

$$\begin{aligned}
c_l =& \kappa_l \left(1 - F_{T_{l'}}(t_\alpha)\right) \left(\frac{\Gamma(\kappa_l)}{\Gamma(\kappa_l + 1)^2} (z_{l0}/\nu_l)^{\kappa_l} \, {}_2F_2(\kappa_l, \kappa_l; \kappa_l + 1, \kappa_l + 1; -z_{l0}/\nu_l) \right. \\
& + \left. (Q(\kappa_l, z_{l0}/\nu_l) - 1) \left(\log(z_{l0}/\nu_l) - \psi(\kappa_l)\right) \right)
\end{aligned} \tag{5.16}$$

is a positive constant depending on $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, $\kappa_l = \gamma_l(x_u)t_\alpha$ is the shape parameter for an increment of the lth marginal process during time t_α, ${}_2F_2$ denotes the generalized hypergeometric function

$$ {}_2F_2(\kappa, \kappa; \kappa + 1, \kappa + 1; -z) = 1 + \sum_{k=1}^{\infty} \left(\frac{\kappa}{\kappa + k} \right)^2 \frac{(-z)^k}{k!}, $$

and l' is the index of the respective other component, i.e., $l' = 2$ if $l = 1$ and vice versa.

Since the information matrix in (5.13) is block-diagonal, the optimality criterion

$$\mathrm{aVar}(\widehat{t}_\alpha) = c_0^2 (c_1^2 (1, x_u) \mathbf{M}_{\boldsymbol{\beta}_1}(\xi)^{-1}(1, x_u)^T + c_2^2 (1, x_u) \mathbf{M}_{\boldsymbol{\beta}_2}(\xi)^{-1}(1, x_u)^T) \tag{5.17}$$

is a weighted sum of the optimality criteria for the single components stated in Subsection 5.2.3 and constitutes, hence, a compound criterion.

In the special case that the nominal values are identical for both components, i.e., $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, the optimal design ξ^* for a single component will also be optimal for the bivariate failure process, independent of α. In general, however, the optimal design for the bivariate failure process has to be a compromise of the marginal optimal designs for the components.

5.3.4   Numerical example

In this example we consider an accelerated degradation experiment with two failure components following two independent Gamma processes. The first process is specified as in Subsection 5.2.4 with scale parameter $\nu_1 = 1$, degradation threshold $z_{10} = 5.16$ and nominal values $\beta_{10} = 0.23$ for the intercept and $\beta_{11} = 0.53$ for the slope. For
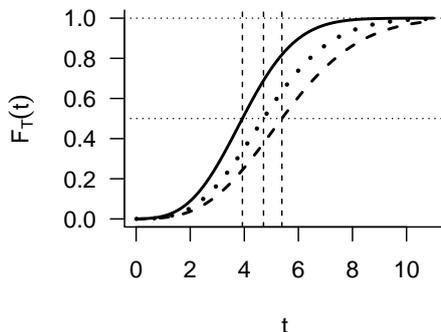
FIGURE 5.7: Failure time distributions $F_T(t)$ (solid line), $F_{T_1}(t)$ (dashed line), and $F_{T_2}(t)$ (dotted line) for the bivariate gamma process in the example of Subsection 5.3.4

the second process we assume a scale parameter $\nu_2 = 0.88$, a degradation threshold $z_{20} = 4.60$ and nominal values $\beta_{20} = 0.31$ for the intercept and $\beta_{21} = 0.35$ for the slope. As in Subsection 5.2.4 the standardized normal use condition is $x_u = -0.40$ and the processes are measured at $k = 4$ standardized time points $t_j = 0.25, 0.5, 0.75$ and 1 with time intervals of constant length $\Delta = 0.25$. Also here we will be interested in estimating the median failure time $t_{0.5}$.

The distribution function $F_T(t)$ of the combined failure time T given by (5.14) is plotted in Figure 5.7 together with the distribution functions $F_{T_1}(t)$ and $F_{T_2}(t)$ of the failure times T_1 and T_2 in the components.

The median failure time $t_{0.5} = 3.93$ satisfying $F_T(t_{0.5}) = 1/2$ is indicated there together with the median failure times for the single components by dashed vertical lines.

For estimating the median failure time, also here the optimal design is sought numerically by means of the multiplicative algorithm on an equidistant grid of step size 0.01 on the design region. The locally optimal design obtained is of the form $\xi^* = \xi_{w^*}$ assigning optimal weights $w^* = 0.78$ to the lowest stress level $x = 0$ and $1 - w^* = 0.22$ to the highest stress level $x = 1$ in the design region. This optimal weight is close to the solution for the first component (see Subsection 5.2.4) and shows a similar behavior, when the value ot the normal use condition is altered. Similar considerations hold for the sensitivity with respect to misspecifications of the nominal values of the parameters.

5.4 Bivariate accelerated degradation testing with a gamma process and a linear mixed model

In Section 5.3 we considered a degradation process with two response components where each is modeled by a gamma model. In this section we consider a bivariate process with two different degradation models. The first degradation mode is modeled by a gamma process as in Section 5.2. As in Section 5.3 the degradation increments of this component are denoted by Y_{i1j} for unit i during a time intervals of length $\delta_j = t_j - t_{j-1}$, $j = 1, ..., k$. The second degradation mode is given by a linear model with random intercept which will be described in the subsequent subsection and is a special case of the model treated in (Shat and Schwabe, 2021). Both failure modes are influenced by the same standardized accelerating stress variable $x \in \mathcal{X} = [0, 1]$. Apart from that the degradation modes are assumed to be independent and, hence, do not have an interactive effect. As before, also here a failure of the system occurs when at least one of the marginal degradation paths exceeds its corresponding failure threshold.

5.4.1 Model formulation of the second degradation component: Linear mixed model

Here we consider a linear regression model, similar to the model presented by (Weaver and Meeker, 2013), for a single stress variable x. Measurements Y_{i2j} of the second component at unit i are taken at the same time points $t_1, ..., t_k$ as for the first component and additionally at the beginning of the degradation experiment, $t_0 = 0$, $j = 0, ..., k$. These measurements are described by a hierarchical model. For each unit i the observation Y_{i2j} at time point t_j is given by

$$Y_{i2j} = \mu_i(x_i, t_j) + \varepsilon_{ij}, \tag{5.18}$$

where $\mu_i(x, t)$ is the mean degradation path of the second marginal response of unit i at time t, when stress x is applied to unit i, and ε_{ij} is the associated measurement error at time point t_j. The mean degradation $\mu_i(x, t)$ is given by a linear model equation in the stress variable x and in time t with stress-time interaction,

$$\mu_i(x, t) = \beta_{i20} + \beta_{21}x + \beta_{22}t + \beta_{23}xt \tag{5.19}$$

where only the intercept is unit specific and the time and stress effects are the same for all units. Hence, the response is given by

$$Y_{i2j} = \beta_{i20} + \beta_{21}x_i + \beta_{22}t_j + \beta_{23}x_i t_j + \varepsilon_{ij}. \tag{5.20}$$

The measurement error ε_{ij} is assumed to be normally distributed with zero mean and a time independent error variance $\sigma_\varepsilon^2 > 0$. Moreover, the error terms are assumed to be independent within a unit over time.

On the aggregate level it is assumed that the units are representatives of a larger entity. The unit specific intercept β_{i20} is modeled as a random effect, i.e., β_{i20} is normally distributed with mean β_{20} and variance $\sigma_0^2 > 0$. All measurement errors ε_{ij} and random effects β_{i20} are assumed to be independent. For transferring the results, it is assumed that the model defined in equation (5.20) also holds for units under normal use condition x_u.

5.4.2 Information for the second degradation component: Linear mixed model

To derive the information matrix in the mixed effects model we first write the model in vector notation. Denote by $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21}, \beta_{22}, \beta_{23})^T$ the vector of fixed effect (aggregate) location parameters and by $\boldsymbol{\varsigma} = (\sigma_0^2, \sigma_\varepsilon^2)^T$ the vector of variance parameters. The $(k+1)$-dimensional vector of observations $\mathbf{Y}_{i2} = (Y_{i20}, ..., Y_{i2k})^T$ at unit i is multivariate normal with expectation $\mathrm{E}(\mathbf{Y}_{i2}) = (\mathbf{D} \otimes (1, x_i)) \boldsymbol{\beta}_2$, where $\mathbf{D} = \left((1, t_0)^T, ..., (1, t_k)^T\right)^T$ is the "design" matrix for the time variable and "$\otimes$" denotes the Kronecker product, and compound symmetric covariance matrix $\mathrm{Cov}(\mathbf{Y}_{i2}) = \mathbf{V}$ with diagonal entries $\sigma_0^2 + \sigma_\varepsilon^2$ and off-diagonals σ_0^2. The elemental information matrix (per unit) $\mathbf{M}_{\boldsymbol{\beta}_2, \boldsymbol{\varsigma}}(x_i) = \begin{pmatrix} \mathbf{M}_{\boldsymbol{\beta}_2}(x_i) & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\boldsymbol{\varsigma}} \end{pmatrix}$ of the linear mixed model component is block diagonal with the elemental information matrices $\mathbf{M}_{\boldsymbol{\beta}_2}(x) = \mathbf{D}^T \mathbf{V}^{-1} \mathbf{D} \otimes \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}$ for the location parameters and $\mathbf{M}_{\boldsymbol{\varsigma}}$ for the variance parameters on the diagonal, where $\mathbf{M}_{\boldsymbol{\varsigma}}$ does not depend on the setting x of the stress variable.

Accordingly, also for an approximate design ξ, the standardized information matrix

$$\mathbf{M}_{\boldsymbol{\beta}_2, \boldsymbol{\varsigma}}(\xi) = \begin{pmatrix} \mathbf{D}^T \mathbf{V}^{-1} \mathbf{D} \otimes \mathbf{M}(\xi) & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\boldsymbol{\varsigma}} \end{pmatrix} \tag{5.21}$$

of the second component is block diagonal, where $\mathbf{M}(\xi) = \sum_{i=1}^m w_i \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}$ is the standardized information matrix of linear fixed effect regression model which does not depend on the parameters. For further details of the linear mixed model see (Shat and Schwabe, 2021).

5.4.3 Failure time distribution for the second degradation component: Linear mixed model

As mentioned in Subsection 5.2.3 we are interested in characteristics of the failure time distribution of soft failure due to degradation. Therefore it is assumed that the model equation $\mu_u(t) = \beta_{u20} + \beta_{21}x_u + \beta_{22}t + \beta_{23}x_u t$ is also valid under the normal use condition, where μ_u denotes the mean degradation path for a unit "u" under the normal use condition x_u and β_{u20} is the random intercept of u. We further denote by $\mu(t) = \mathrm{E}(\mu_u(t)) = \beta_{20} + \beta_{21}x_u + \beta_{22}t + \beta_{23}x_u t$ the aggregate degradation path under normal use condition.

A soft failure due to degradation for the second response component is defined as the exceedance of the degradation over a failure threshold y_{20}. This definition is based on the mean degradation path $\mu_u(t)$ and not on a "real" path subject to measurement errors. The failure time T_2 under normal use condition is then defined as the first time t the mean degradation path $\mu_u(t)$ reaches or exceeds the threshold y_{20}, i. e. $T_2 = \min\{t \geq 0; \mu_u(t) \geq y_{20}\}$. Because the random intercept β_{u20} is involved in the mean degradation path, the failure time T_2 is random.

As in the previous sections, we will describe the characteristics of the failure time T_2 by its distribution function $F_{T_2}(t)$. We note that $T_2 \leq t$ if and only if $\mu_u(t) \geq y_{20}$ and, hence, we can derive

$$F_{T_2}(t) = \mathrm{P}(\mu_u(t) \geq y_{20}) = \Phi((\mu(t) - y_{20})/\sigma_0), \tag{5.22}$$

where Φ denotes the distribution function of the standard normal distribution. For later use we also state the gradient

$$\partial F_{T_2}(t)/\partial\boldsymbol{\beta}_2 = \sigma_0^{-1}\varphi\left((\mu(t_\alpha) - y_{20})/\sigma_0\right)(1, t)^T \otimes (1, x_u)^T \tag{5.23}$$

of $F_{T_2}(t)$ with respect to the vector $\boldsymbol{\beta}_2$ of location parameters (cf. (Shat and Schwabe, 2021)), where φ denotes the density of the standard normal distribution.

5.4.4 Estimation and information in the combined model

The combined model parameters $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and $\boldsymbol{\varsigma}$ can be estimated by means of the maximum likelihood method. As stated in Subsection 5.3.3, in the combined model the maximum likelihood estimates coincide with those for the single components because of the independence between the failure modes. Accordingly, the combined information matrix for all parameters is block diagonal with the information matrices for the components on the diagonal. In view of (5.21) the information matrix of an

approximate design ξ is given by

$$
\mathbf{M}_{\beta_1,\beta_2,\varsigma}(\xi) = \begin{pmatrix} \mathbf{M}_{\beta_1}(\xi) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\beta_2}(\xi) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_{\varsigma}(\xi) \end{pmatrix}, \tag{5.24}
$$

where $\mathbf{M}_{\beta_2}(\xi) = \mathbf{D}^T\mathbf{V}^{-1}\mathbf{D} \otimes \mathbf{M}(\xi)$ and $\mathbf{M}_{\beta_1}(\xi)$ as in Subsection 5.2.2.

5.4.5 Optimality criterion based on the joint failure time

The combined failure time T is defined as the minimum of the marginal failure times T_1 and T_2 for the single components derived in Subsections 5.2.3 and 5.4.3. As in Subsection 5.4.5, the survival function of the joint failure time T factorizes and, hence, the distribution function $F_T(t)$ can be expressed as $F_T(t) = 1 - (1 - F_{T_1}(t))(1 - F_{T_2}(t))$. The quantiles $t_\alpha = t_\alpha(\beta_1, \beta_2, \varsigma)$ are functions of both the location parameters β_1 and β_2 as well as on the variance parameters ς, in general. Consequently, the maximum likelihood estimate of a quantile t_α is given by $\widehat{t}_\alpha = t_\alpha(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\varsigma})$ in terms of the maximum likelihood estimates $\widehat{\beta}_1$, $\widehat{\beta}_2$ and $\widehat{\varsigma}$ of the parameters β_1, β_2 and ς in the components. The asymptotic variance of $\widehat{t}_\alpha$ can again be obtained by the delta method and the implicit function theorem. By the block diagonal structure of the information matrix and the decomposition of the distribution function of the failure time we get

$$
\mathrm{aVar}(\widehat{t}_\alpha) = f_T(t_\alpha)^{-2} \left(c_1^2 (1, x_u)\mathbf{M}_{\beta_1}(\xi)^{-1}(1, x_u)^T + c_2^2 (1, x_u)\mathbf{M}(\xi)^{-1}(1, x_u)^T + c_\varsigma^2 \right), \tag{5.25}
$$

where $f_T(t)$ is the density of T, c_1 is defined as in (5.16) with the distribution function $F_{T_2}(t_\alpha) = \Phi((\mu(t_\alpha) - y_{20})/\sigma_0)$ of the linear mixed effect model inserted,

$$
c_2 = (1 - Q\left(\gamma(x_u)t_\alpha, z_{10}/\nu_1\right)) \sigma_0^{-1}\varphi\left((\mu(t_\alpha) - y_{20})/\sigma_0\right) \left((1, t_\alpha) \left(\mathbf{D}^T\mathbf{V}^{-1}\mathbf{D}\right)^{-1} (1, t_\alpha)^T \right)^{1/2} \tag{5.26}
$$

by (5.21) and (5.23), and $c_\varsigma^2 = \mathbf{c}_\varsigma^T\mathbf{M}_\varsigma^{-1}\mathbf{c}_\varsigma$ is a constant independent of ξ in which $\mathbf{c}_\varsigma = \partial F_{T_2}(t_\alpha)/\partial\varsigma$ is the gradient of $F_{T_2}(t_\alpha)$ with respect to the vector ς of variance parameters.

The criterion (5.25) is a weighted sum of the optimality criteria for the single components and constitutes, hence, a compound criterion, where the weights depend on both vectors β_1 and β_2 of location parameters as well as on the variance parameters ς of the linear mixed model component, in general. Due to convexity the optimal weight w^* for the system lies in the range of the optimal weights w_1^* and w_2^* for the components, $\min\{w_1^*, w_2^*\} \leq w^* \leq \max\{w_1^*, w_2^*\}$.
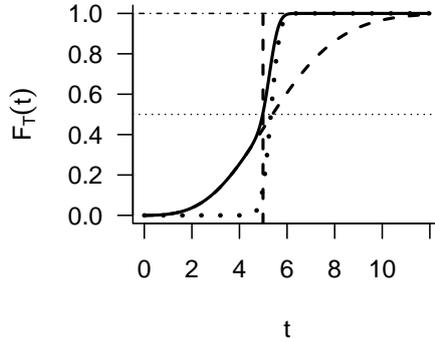
FIGURE 5.8: Failure time distributions $F_T(t)$ (solid line), $F_{T_1}(t)$ (dashed line), and $F_{T_2}(t)$ (dotted line) for the bivariate model with a gamma process (T_1) and a linear mixed effect (T_2) component for the example of Subsection 5.4.6

5.4.6   Numerical example

In this example we consider an accelerated degradation experiment with two independent failure components in which the first component follows a gamma process and the second is described by a linear mixed model with random intercept as described in Subsection 5.4.1. The gamma process is specified as in Subsections 5.2.4 and 5.3.4 with scale parameter $\nu_1 = 1$, degradation threshold $z_{10} = 5.16$ and nominal values $\beta_{10} = 0.23$ for the intercept and $\beta_{11} = 0.53$ for the slope. For the linear model we assume a degradation threshold $y_{20} = 3.73$ and nominal values $\beta_{20} = 2.35$ for the aggregate intercept, $\beta_{21} = 0.06$ for the slope in the stress variable x, $\beta_{22} = 0.28$ for the slope in time t, $\beta_{23} = 0.04$ for the stress-time interaction xt, $\sigma_0 = 0.08$ for the standard deviation of the random intercept, and $\sigma_\varepsilon = 0.09$ for the standard deviation of measurement errors. As before the standardized normal use condition is $x_u = -0.40$ and both degradation processes are measured at the $k = 4$ standardized time points $t_j = 0.25$, 0.5, 0.75 and 1 with time intervals of constant length $\Delta = 0.25$. Additionally, the degradation of the second component is measured initially at $t_0 = 0$, i.e., at the beginning of the experiment. Also in the present setting we will be interested in estimating the median failure time $t_{0.5}$. The distribution function $F_T(t)$ of the combined failure time T is plotted in Figure 5.8 together with the distribution functions $F_{T_1}(t)$ and $F_{T_2}(t)$ of the failure times T_1 and T_2 in the components. The median failure time $t_{0.5} = 4.99$ satisfying $F_T(t_{0.5}) = 1/2$ is indicated there by a dashed vertical line.

For estimating the median failure time, also here the optimal design is sought numerically by means of the multiplicative algorithm on an equidistant grid of step size 0.01 on the design region. As in the univariate case the algorithm indicates that the
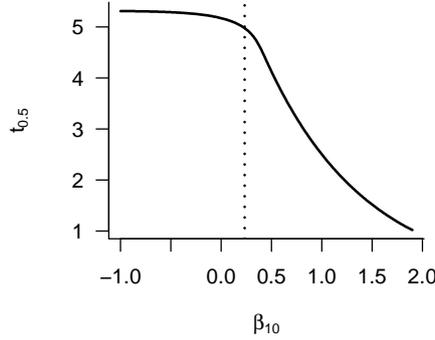
FIGURE 5.9: Dependence of $t_{0.5}$ on β_{10} for the example in Subsection 5.4.6

optimal design ξ^* is of the form ξ_w. Under this premise the optimal value of w^* can be determined by a simple line search on a sufficiently dense grid. The resulting optimal designs which assigns optimal weights $w^* = 0.78$ to the lowest stress level $x = 0$ and $1 - w^* = 0.22$ to the highest stress level $x = 1$ in the design region.

To assess the robustness of the locally optimal design we examine how the optimal weight w^* varies when the underlying parameter values are modified. Computations indicate that the optimal weight does not change much in the nominal values of the parameters $\boldsymbol{\beta}_2$ and $\boldsymbol{\varsigma}$ for the linear mixed effects degradation model in the second component. This property is in accordance with the fact that the design criterion depends on the values of $\boldsymbol{\beta}_2$ and $\boldsymbol{\varsigma}$ only through the weighting factors c_1^2 and c_2^2 while the marginal information matrix $\mathbf{M}(\xi)$ does not. However, similar to the univariate case, there may be moderate changes with respect to the parameters $\boldsymbol{\beta}_1$ of the gamma degradation model in the first component. Additionally, the optimal weight w^* may switch between the marginal optimal weights w_1^* and w_2^* for the marginal failure models depending on which of the marginal failure modes is dominant in the bivaraite system.

We will demonstrate this behavior in the case when the intercept β_{10} of the gamma model component varies while all other parameters are fixed to their nominal values. In Figure 5.9 we plot the median failure time $t_{0.5}$ and the weighting coefficients c_1 and c_2, respectively, in dependence on β_{10}. For negative values of β_{10}, the failure time T_2 of the first component decreases and the failure of the bivariate system is dominated by the second component. Then the median failure time $t_{0.5}$ approaches its marginal counterpart 5.32 in the second component. For increasing values of β_{10}, the failure of the first component becomes dominant and the median failure time $t_{0.5}$ behaves as in the marginal model for the first component. In particular, $t_{0.5}$ is decreasing in β_{10} and becomes smaller than 1 for $\beta_1 > 1.92$. Hence, only values $\beta_{10} \leq 1.92$ are reasonable to
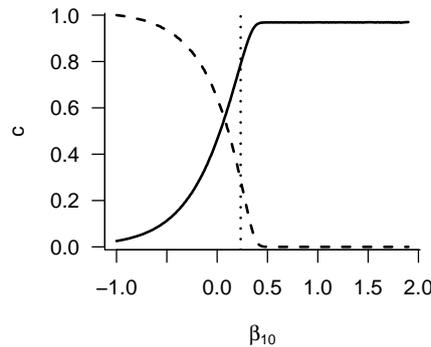
FIGURE 5.10: Dependence of the coefficients c_1 (solid line) and c_2 (dashed line, standardized) on β_{10} for the example in Subsection 5.4.6
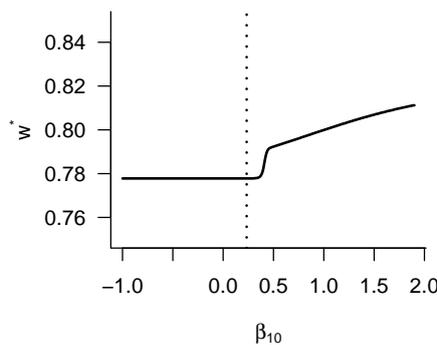
FIGURE 5.11: Dependence of w^* on β_{10} for the example in Subsection 5.4.6

be considered because otherwise no acceleration would be required to obtain failure due to degradation under normal use conditions. The dominance of the failure components is also reflected in Figure 5.10 where the weighting coefficients c_1 and c_2 are shown in dependence on β_{10}. There the second coefficient c_2 is standardized by its maximum for purposes of comparison. For negative values of β_{10}, the coefficient c_2 of the second component dominates the asymptotic variance (5.25) while the dominance is is reversed for $\beta_{10} > 0.5$.

This change in dominance has also an impact on the optimal weights w^* as exhibited in Figure 5.11.

For negative values of β_{10}, the optimal weight w^* coincides with its marginal counterpart in the second component while, for $\beta_{10} > 0.5$, the optimal weight w^* is as in the univariate model for the first component (see Subsection 5.2.4). Caused by the change in dominance, there is a small, but pronounced change in the optimal weight
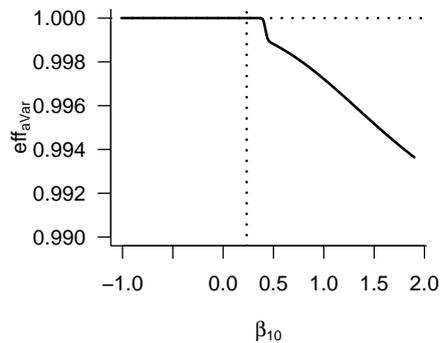
FIGURE 5.12: Efficiency of ξ^* in dependence on β_{10} for the example in Subsection 5.4.6

when β_{10} varies from 0.35 to 0.50. This shift is also visible in the efficiency of the locally optimal design ξ^* at the given nominal values when the intercept parameter β_{10} is misspecified, as shown in Figure 5.12. For values of β_{10} less than the nominal value $\beta_{10} = 0.23$, the locally optimal design ξ^* has an efficiency of nearly 1, up to $\beta_{10} = 0.35$. Then there is a small, instantaneous decrease in efficiency to 0.999 for β_{10} between 0.35 and 0.50. For larger values of β_{10} the efficiency smoothly decreases as in the univariate model for the first component (see Subsection 5.2.4). For the maximal value $\beta_{10} = 1.92$ the efficiency of ξ^* is still remarkably high with a value of about 0.9936. In all of Figures 5.9 to 5.12 the nominal value $\beta_{10} = 0.23$ is indicated by a dotted vertical line.

With respect to the slope parameter $\beta_{20} \leq 0$ of the Gamma component, the failure is dominated by the second component. Hence, neither the optimal weight is affected by β_{20}, nor the efficiency of the locally optimal design ξ^* differs reasonably from 1. In total, the locally optimal design ξ^* at the given nominal values appears to be robust against misspecifications of the parameters within a meaningful range.

5.5 Concluding remarks

The design stage of highly reliable systems requires a sophisticated assessment of the reliability related properties of the product. One approach to handle this issue is to conduct accelerated degradation testing. Accelerated degradation tests have the advantage to provide an estimation of lifetime and reliability of the system under study in a relatively short testing time. The majority of existing literature deals with this issue by considering a single failure mode, which may not be sufficiently representative in many cases.

In this chapter, we propose optimal experimental designs for ADTs with a single response components and extend it to the case of multiple response components with repeated measures. Two bivariate degradation models are considered. The marginal degradation functions are described by two Gamma process models in the first bivariate model, and a Gamma process with a linear model with a random intercept in the second one. In this context it is desirable to estimate certain quantiles of the joint failure time distribution as a characteristic of the reliability of the product. The purpose of optimal experimental design is to find the best settings for the stress variable to obtain most accurate estimates for these quantities.

In the present model for accelerated degradation testing, it is assumed that stress remains constant within each testing unit during the whole period of experimental measurements but may vary between units. Hence, in the corresponding experiment a cross-sectional design between units has to be chosen for the stress variable(s) while for repeated measurements the time variable varies according to a longitudinal time plan within units. In particular, the same time plan for measurements is used for all units in the test. It is further assumed that the marginal response components are uncorrelated.

The multiplicative algorithm is utilized to obtain optimal experimental designs for the single response case as well as the two bivariate degradation models. The sensitivity analysis shows that the optimal designs of the univariate model as well as the bivariate model with two marginal Gamma processes are robust against misspecifications of the corresponding parameter vectors and depend mainly on the normal use condition of the stress variable. For the bivariate model with two different marginal models the sensitivity analysis establishes that the resulting optimal design is slightly dependent on the nominal parameter values.

Although only Gamma processes and LMEM are considered as marginal degradation models here, the underlying methods can be extended to other marginal failure modes. Another object of interest would be to consider optimality criteria accounting for simultaneous estimation of various characteristics of the failure time distribution.

Chapter 6

Optimal Designs in Copula-based Gamma Models

6.1 Introduction

Modern products usually have complex structure with multiple failure mechanisms as well as multiple degradation measures. Thus, it is realistic to assume some kind of dependence among different failure components. In the past decade, copula-based modelling has become an efficient tool in many areas of applied statistics, see (AghaKouchak, Bárdossy, and Habib, 2010). For instance, (Perrone and Müller, 2016) has provided an equivalence theorem for binary bivariate copula models that allows applications of efficient design algorithms and quick checks of whether a design is optimal or at least efficient. The Archimedean, Clayton, Frank and Gumbel copulas are intensively used to describe the dependence among different failure components when the marginal degradation paths correspond to Levy stochastic processes, see (Mireh, Khodadadi, and Haghighi, 2019). For example, (Zhou, Pan, and Sun, 2010) and (Guo and Li, 2017) followed a similar approach through considering a system with multiple failure components where the marginal degradation paths are governed by Gamma processes. They utilized the Frank copula to describe the dependence of failure components. (Adegbola and Yuan, 2019) proposed a multivariate Gamma process to model dependent deterioration phenomena that collectively define the service life of infrastructure assets. (Liu et al., 2014) developed a reliability model for systems with s-dependent degradation processes using several Archimedean copulas. The marginal degradation processes were assumed to be inverse Gaussian with a time scale transformation. Furthermore, the authors incorporated a random drift to account for a possible heterogeneity in population. (Pan et al., 2013) and (Pan, Balakrishnan*, and Sun, 2011) presented a bivariate stochastic process where the dependence of the performance characteristics were described by a Frank copula. In order to provide a more flexible dependence structure between competing failure modes, (Wang and Pham, 2011) introduced time-varying copulas to

develop an *s*-dependent competing risk model for systems subject to multiple degradation processes and random shocks. With an application to toxicity trials, (Denman et al., 2011) derived locally *D*-optimal designs for dependent bivariate binary data, where several Archimedean copulas were utilized to describe the dependence among the marginal regression models. Further, (Mireh, Khodadadi, and Haghighi, 2019) proposed a simulation-based reliability analysis for systems with dependent Gamma degradation processes and Weibull distributed hard failure times. The authors used the Frank copula to represent the dependence between failure modes.

In this chapter, optimal experimental designs are derived for Accelerated Degradation Testing with two response components. We consider the situations of independent as well as dependent marginal responses where the observational times are assumed to be fixed and known. The marginal degradation paths are assumed to follow a Gamma process where copula functions are utilized to express the dependence between marginal components. For the case of independent response components the optimal design minimizes the asymptotic variance of an estimated quantile of the failure time distribution at the normal use conditions. For the case of dependent response components the *D*-criterion is adopted to derive *D*-optimal designs. Further, *D*- and *c*-optimal designs are developed when the copula-based models are reduced to bivariate binary outcomes.

The rest of the present chapter is organized as follows. In Section 6.2 we obtain an optimal experimental design for a bivariate Gamma model with independent marginal components. In Section 6.3 we develop *D*-optimal designs for bivariate Gamma models with dependent responses based on the Frank copula function or the Gaussian copula function, respectively. Section 6.4 introduces *D*- and *c*-optimal designs for Accelerated Degradation Testing with dependent failure modes when the copula-based model is reduced to bivariate binary outcomes.

6.2 Bivariate Gamma process with independent components

6.2.1 Model construction

The Gamma process is a natural stochastic model for degradation processes in which degradation occurs gradually over time in a sequence of independent increments. In this section, we assume that the testing unit has two failure modes where the marginal degradation paths are given by Gamma processes in terms of a standardized continuous time variable $t \geq 0$, and the two marginal Gamma processes are independent. It is further assumed that each of the marginal (standardized) stress levels x_l, $l = 1, 2,$

is a scalar in the standardized interval $[0, 1]$. The joint stress variable $\mathbf{x} = (x_1, x_2)$ can be chosen by the experimenter from the experimental region $\mathcal{X} = [0, 1]^2$. For (locally) optimal design, the information matrix as a function of $\mathbf{x}$ (at given values of the model parameters) is of basic interest and will be considered in Subsection 6.2.2. Locally c-optimal designs will be presented in Subsection 6.2.3, where the particular c-criterion expresses the asymptotic variance of an estimated quantile of the failure time distribution.

A Gamma process $Z_t^{(l)}$, $t \geq 0$, considering the response component $l = 1, 2$ is a stochastic process with independent and Gamma distributed increments. The marginal univariate processes $Z_t^{(l)}$, $l = 1, 2$, coincide with the univariate Gamma processes explained in Subsection 5.2.1.

When an accelerated degradation test is run under a stress setting $\mathbf{x} = (x_1, x_2)$, measurements of the bivariate degradation process at the prescribed time points t_j, $j = 1, \dots, k$, are made. So the increments $\boldsymbol{y}_j = (y_{j1}, y_{j2})$, $j = 1, \dots, k$, $l = 1, 2$, of the bivariate degradation path are obtained, which follow the model of independent bivariate random variables $\mathbf{Y}_j = (Y_{j1}, Y_{j2})$, $j = 1, \dots, k$, with Gamma distributed components Y_{jl} according to (5.2). Thus, under the stress setting $\mathbf{x}$ and given the incremental data $\boldsymbol{y} = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_k)$, the log-likelihood of the parameter vector $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T\right)^T$, where $\boldsymbol{\beta}_l = (\beta_{1l}, \beta_{2l})^T$, $l = 1, 2$, is given by

$$
\begin{aligned}
\ell(\boldsymbol{\beta}; \mathbf{x}, \boldsymbol{y}) &= \sum_{j=1}^{k} \sum_{l=1}^{2} \log\left(f_{jl}(y_{jl})\right) = \sum_{j=1}^{k} \sum_{l=1}^{2} \left(\gamma_l(x_l)\Delta_j - 1\right)\log(y_{jl}) - \frac{y_{jl}}{\nu_l} \\
&\quad - \log\left(\Gamma(\gamma_l(x_l)\Delta_j)\right) - \gamma_l(x_l)\Delta_j \log(\nu_l),
\end{aligned}
\tag{6.1}
$$

where $\gamma_l(x_l) = \exp(\beta_{1l} + \beta_{2l}x_l)$, $l = 1, 2$. Usually, an accelerated degradation test is conducted at n distinct testing units $i = 1, \dots, n$ at stress settings $\mathbf{x}_1, \dots, \mathbf{x_n}$, respectively. Note that the stress settings $\mathbf{x}_i$, $i = 1, \dots, n$ may not all be distinct. Under the assumption of independence of the testing units, the joint log-likelihood equals the sum of the log-likelihoods over the units,

$$
\ell(\boldsymbol{\beta}; \mathbf{x}_1, \dots, \mathbf{x}_n, \boldsymbol{y}_1, \dots, \boldsymbol{y_n}) = \sum_{i=1}^{n} \ell(\boldsymbol{\beta}; \mathbf{x}_i, \boldsymbol{y}_i).
$$

The collection $\mathbf{x}_1, \dots, \mathbf{x}_n$ constitutes the experimental design of the test. Since the ordering of the design points $\mathbf{x}_i$ (along with the response vector $\boldsymbol{y}_i$) is of no importance, a design is usually described by the set of of *distinct* points $\mathbf{x}_1', \dots, \mathbf{x}_m'$ among the collection $\mathbf{x}_1, \dots, \mathbf{x}_n$ and the corresponding frequencies $n_1, \dots, n_m$ of their occurrence among $\mathbf{x}_1, \dots, \mathbf{x}_n$. Hence, in what follows, we will employ the approximate design explained in Subsection 2.3.1 for deriving optimal designs.

6.2.2 Information matrix

By the log-likelihood $\ell(\boldsymbol{\beta}; \mathbf{x}, \boldsymbol{y})$ from (6.1) the elemental Fisher information matrix of $\mathbf{x}$ at $\boldsymbol{\beta}$ is given by either of following two representations,

$$\mathbf{M}_\beta(\mathbf{x}) = \mathrm{E}\left(\left[\frac{\partial\ell(\boldsymbol{\beta}; \mathbf{x}, \boldsymbol{Y})}{\partial\boldsymbol{\beta}}\right]\left[\frac{\partial\ell(\boldsymbol{\beta}; \mathbf{x}, \boldsymbol{Y})}{\partial\boldsymbol{\beta}}\right]^T\right) = -\mathrm{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta}; \mathbf{x}, \boldsymbol{Y})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T}\right). \qquad (6.2)$$

Using the latter representation, direct calculations yield, in view with equation (5.6) and the presentation of the information matrix in Subsection 5.3.2, a block structure because of the independence of the components as

$$\mathbf{M}_\beta(\mathbf{x}) = \begin{pmatrix} \mathbf{M}_{\beta_1}(x_1) & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\beta_2}(x_2) \end{pmatrix}, \qquad (6.3)$$

where $\mathbf{x} = (x_1, x_2)$, $\boldsymbol{\beta}_l = (\beta_{1l}, \beta_{2l})^T$, $l = 1, 2$, and

$$\mathbf{M}_{\beta_l}(x_l) = \lambda_l(x_l, \boldsymbol{\beta}_l) \begin{pmatrix} 1 & x_l \\ x_l & x_l^2 \end{pmatrix}, \quad l = 1, 2, \qquad (6.4)$$

with $\lambda_l(x_l, \boldsymbol{\beta}_l) = \gamma_l^2(x_l) \sum_{j=1}^k \Delta_j^2 \psi_1\big(\gamma_l(x_l)\Delta_j\big)$,

As usual in the approximate design theory, for any approximate design ξ the information matrix of ξ at a parameter point $\boldsymbol{\beta}$ is given by

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^m w_i \mathbf{M}_\beta(\mathbf{x}_i).$$

By the block-diagonal structure of the elemental information matrices (6.3), the information matrix of ξ is again block-diagonal where the blocks are given by the information matrices of the marginal designs w.r.t. the marginal models,

$$\mathbf{M}_\beta(\xi) = \begin{pmatrix} \mathbf{M}_{\beta_1}(\xi_1) & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\beta_2}(\xi_2) \end{pmatrix}. \qquad (6.5)$$

where $\mathbf{M}_{\beta_l}(\xi_l) = \sum_{i=1}^m w_i \mathbf{M}_{\beta_l}(x_{il}), \quad l = 1, 2.$ $\qquad (6.6)$

Recall that $\mathcal{X} = [0, 1]^2$, hence $\mathbf{x}_i = (x_{i1}, x_{i2})$ with $x_{1i}, x_{2i} \in [0, 1]$, $i = 1, \ldots, m$. The designs ξ_1 and ξ_2 on $[0, 1]$ are the marginal designs of ξ, which are defined as the projections on the corresponding components (in a measure theoretic sense).

6.2.3 Optimality criterion based on failure time distribution

In Accelerated Degradation Testing one considers some characteristics of the failure time distribution due to degradation under normal use condition $\mathbf{x}_u = (x_{u1}, x_{u2})$. Note that typically the normal use conditions x_{u1} and x_{u2} are outside the normalized interval $[0,1]$ of the possible stress values x_1 and x_2 in Accelerated Degradation Testing. Usually, one has $x_{ul} < 0$, $l = 1, 2$. It is assumed that the marginal Gamma process $Z_{u,t}^{(l)}$ describing the degradation under normal use condition x_{ul} has the rate $\gamma_l(x_{ul}) = \exp(\beta_{1l} + \beta_{2l}x_{ul})$ and scale ν_l. A soft failure due to degradation is defined by exceedance of the marginal degradation paths over some failure thresholds. The marginal failure time T_l under normal use condition x_{ul} are defined as in Subsection 5.2.3.

As opposed to the approach in Subsection 5.3.3 we assume here a parallel system, that is, the system fails as soon as both marginal components have failed. Denote by T the joint failure time, $T = \max\{T_1, T_2\}$. By independence of the components its distribution function is given by

$$
\begin{aligned}
F_T(t) &= \mathrm{P}(Z_{u,t}^{(1)} \geq z_{10}, Z_{u,t}^{(2)} \geq z_{20}) \\
&= Q(\gamma_1(x_{u1})t, z_{10}/\nu_1)\, Q(\gamma_2(x_{u2})t, z_{20}/\nu_2).
\end{aligned}
\tag{6.7}
$$

In accordance with 5.3.3, the performance of the maximum likelood estimator $\widehat{t}_\alpha$ of the failure time quantile t_α is measured by its asymptotic variance $\mathrm{aVar}(\widehat{t}_\alpha)$, and design optimization will be conducted with respect to minimizing $\mathrm{aVar}(\widehat{t}_\alpha)$. This c-criterion is commonly used in planning degradation tests when experimenters are interested in accurately estimating reliability properties of a system over its life cycle. However, it should be noted that the asymptotic variance will depend on $\boldsymbol{\beta}$, and thus, as a common feature of non-linear models, one is concerned with *local* design optimality at some given parameter point $\boldsymbol{\beta}$. Under a design ξ the asymptotic variance of $\widehat{t}_\alpha$ is given by

$$
\mathrm{aVar}(\widehat{t}_\alpha) = \boldsymbol{c}(\boldsymbol{\beta})^T \mathbf{M}_\beta(\xi)^{-1} \boldsymbol{c}(\boldsymbol{\beta}),
\tag{6.8}
$$

$$
\text{where} \quad \boldsymbol{c}(\boldsymbol{\beta}) = \frac{\partial t_\alpha}{\partial \boldsymbol{\beta}}.
\tag{6.9}
$$

A criterion given by the r.h.s. of (6.8) is called a (local) c-criterion. Efficient algorithms have been developed to compute a c-optimal design, see the numerical example in Subsection 6.2.4 below. However, a more explicit formula of the coefficient vector $\boldsymbol{c}(\boldsymbol{\beta})$ of the criterion has to be provided. Due to the implicit definition of t_α as the unique solution of $F_T(t_\alpha) = \alpha$, the following identity is ensured by the implicit function theorem, see (Krantz and Parks, 2012)

$$
\frac{\partial t_\alpha}{\partial \boldsymbol{\beta}} = \frac{\partial F_T(t_\alpha)}{\partial \boldsymbol{\beta}} \Big/ f_T(t_\alpha), \quad \text{where} \quad f_T(t_\alpha) = \frac{\partial F_T(t)}{\partial t}\Big|_{t=t_\alpha} > 0.
\tag{6.10}
$$

From (6.7) one obtains, denoting $Q_1(s, z) = \frac{\partial Q(s,z)}{\partial s}$ as derived in (5.16),

$$
\begin{aligned}
\frac{\partial F_T(t_\alpha)}{\partial \beta_{11}} &= Q_1\big(\gamma_1(x_{u1})t_\alpha,\, z_{10}/\nu_1\big)\, Q\big(\gamma_2(x_{u2})t_\alpha,\, z_{20}/\nu_2\big)\, \gamma_1(x_{u1})t_\alpha, \\
\frac{\partial F_T(t_\alpha)}{\partial \beta_{21}} &= x_{u1}\frac{\partial F_T(t_\alpha)}{\partial \beta_{11}}, \\
\frac{\partial F_T(t_\alpha)}{\partial \beta_{12}} &= Q\big(\gamma_1(x_{u1})t_\alpha,\, z_{10}/\nu_1\big)\, Q_1\big(\gamma_2(x_{u2})t_\alpha,\, z_{20}/\nu_2\big)\, \gamma_2(x_{u2})t_\alpha, \\
\frac{\partial F_T(t_\alpha)}{\partial \beta_{22}} &= x_{u2}\frac{\partial F_T(t_\alpha)}{\partial \beta_{21}}.
\end{aligned}
$$

Hence, the coefficient vector from (6.9) reads as

$$
\boldsymbol{c}(\boldsymbol{\beta}) = \big(f_T(t_\alpha)\big)^{-1}\big(c_1(\boldsymbol{\beta})\,(1\,,\,x_{u1})\,,\, c_2(\boldsymbol{\beta})\,(1\,,\,x_{u2})\,\big)^T, \quad \text{where} \qquad (6.11)
$$
$$
c_l(\boldsymbol{\beta}) = \partial F_T(t_\alpha)/\partial \beta_{1l} \,>\, 0,\, l = 1, 2.
$$

Together with the block-diagonal structure (6.5) of the information matrices, the c-criterion from (6.8) becomes

$$
\boldsymbol{c}(\boldsymbol{\beta})^T \mathbf{M}_\beta(\xi)^{-1}\boldsymbol{c}(\boldsymbol{\beta}) = \big(f_T(t_\alpha)\big)^{-2}\sum_{l=1}^{2} c_l^2\,(1\,,\,x_{ul})\,\mathbf{M}_{\beta_l}(\xi_l)^{-1}\,(1\,,\,x_{ul})^T. \qquad (6.12)
$$

It follows that a design ξ^* is c-optimal w.r.t. the coefficient vector $\boldsymbol{c}(\boldsymbol{\beta})$, that is, ξ^* minimizes (6.12) over all designs ξ on $\mathcal{X} = [0\,,\,1]^2$, if and only if its marginal designs ξ_l^*, $l = 1, 2$, are c-optimal w.r.t. the coefficient vectors $\boldsymbol{c}_l = (1\,,\,x_{ul})^T$, $l = 1, 2$, respectively, that is ξ_l^* minimizes $(1\,,\,x_{ul})\,\mathbf{M}_{\beta_l}(\xi_l)^{-1}\,(1\,,\,x_{ul})^T$ over all designs ξ_l on $[0, 1]$, $l = 1, 2$. In particular, c-optimality w.r.t. the coefficient vector $\boldsymbol{c}(\boldsymbol{\beta})$ does not depend on α. It should be noted that, under the assumption of independent components, the result can be readily extended to $r > 2$ components and to any s-out-of-r system, see (Shat and Gaffke, 2021) for further details in this regard. Under the premise that the locally optimal designs ξ_l^* are supported on the endpoints of the design region $[0, 1]$, i.e., they are of the form $\xi_l^* = \xi_{w_l^*}$, where ξ_{w_l} denotes a design with weight $w_{1l} = w_l$ on $x_{1l} = 0$ and weight $w_{2l} = 1 - w_{1l}$ on $x_{2l} = 1$, (Shat and Schwabe, 2019) stated that the marginal optimal weight w_l^* can be determined analytically by Elfving's theorem (Elfving, 1952), as depicted in equation (5.11).

6.2.4 Numerical example

The distribution function $F_T(t)$ from (6.7) is plotted for illustration in Figure 6.1 under the nominal values given in Table 6.1, the normal use conditions $x_{u1} = -0.60$ and $x_{u2} = -0.50$, and the failure thresholds $z_{10} = 4.6$ and $z_{20} = 6.25$. The median $t_{0.5} = 2.11$

TABLE 6.1: Nominal values of the Gamma model with independent marginal components

β_{11}	β_{12}	ν_1	β_{21}	β_{22}	ν_2
1.80	1.60	1.24	2.80	3.13	1.17

is indicated by a dashed vertical line. Also, the distribution functions $F_{T_l}(t)$ from (5.9) are shown in the figure. We assume that units are observed according to a time plan with $k = 4$ time points, and $t_1 = 0.02$, $t_2 = 0.04$, $t_3 = 0.06$, $t_4 = 0.1$. For computing optimal marginal designs ξ_l^* minimizing $(1, x_{ul}) \, \mathbf{M}_{\beta_l}(\xi_l)^{-1}(1, x_{ul})^T$, $l = 1, 2$, with nominal values of parameters and constants from Table 6.1, the multiplicative algorithm (Torsney and Martín-Martín, 2009) was applied. The marginal design interval $[\,0\,,\,1\,]$ was replaced by an equidistant grid with increment equal to 0.05. The obtained optimal marginal designs ξ_1^* and ξ_2^* are as follows,

$$\xi_1^* = \begin{pmatrix} 0 & 1 \\ 0.79 & 0.21 \end{pmatrix} \quad \text{and} \quad \xi_2^* = \begin{pmatrix} 0 & 1 \\ 0.91 & 0.09 \end{pmatrix}. \tag{6.13}$$

So the locally c-optimal designs at $\boldsymbol{\beta}$ are given by those designs ξ^* on $\mathcal{X} = [\,0\,,\,1\,]^2$ (actually on the product grid of the employed marginal grids) whose marginal designs are equal to ξ_1^* and ξ_2^* from (6.13). One of them is the product design

$$\xi^* = \xi_1^* \otimes \xi_2^* = \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.72 & 0.07 & 0.19 & 0.02 \end{pmatrix}. \tag{6.14}$$

Note that the locally c-optimal design is not unique: the set of all designs with marginal designs given by (6.13) consists of all designs ξ^* supported by the points $(0,0)$, $(0,1)$, (1.0), and $(1,1)$ with weights

$$\xi^*(0,0) = \omega, \ \ \xi^*(0,1) = 0.79 - \omega, \ \ \xi^*(1,0) = 0.91 - \omega, \ \ \xi^*(1,1) = \omega - 0.70, \text{where} \ \ 0.70 \le \omega \le 0.79.$$

For $0.70 < \omega < 0.79$ the four weights of ξ^* are positive and ξ^* is actually a four-point design. The particular value $\omega = 0.72$ yields the above product design. The boundary values $\omega = 0.70$ and $\omega = 0.79$ yield three-point designs supported by $(0,0)$, $(0,1)$, $(1,0)$ and by $(0,0)$, $(1,0)$, $(1,1)$, respectively.

When the value of normal use conditions x_{ul}, $l = 1, 2$ are altered within some in intervals of the negative half-line, while keeping all other parameters fixed to their nominal values in Table 6.1, the optimal marginal designs ξ_l^*, $l = 1, 2$, computed by the algorithm are again supported by the boundary values 0 and 1. The optimal weight $\omega_1 = \xi_1^*(0)$ as a function of x_{u1} is plotted in Figure 6.2, and the optimal weight
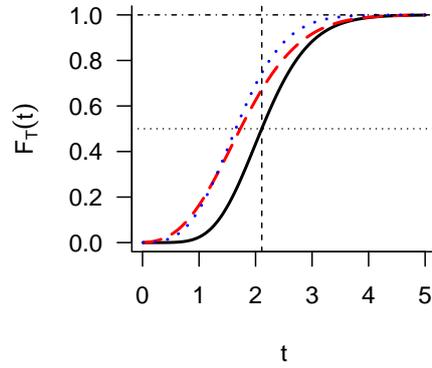
FIGURE 6.1: Failure time distribution $F_T(t)$ at the bivariate Gamma model for Example 6.2.4, dashed line: $F_{T_1}(t)$, dotted line: $F_{T_2}(t)$
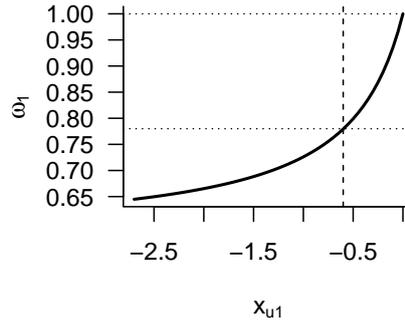
FIGURE 6.2: Optimal weight in dependence on x_{u1} for Example 6.2.4

$\omega_2 = \xi_2^*(0)$ as a function of x_{u2} is plotted in Figure 6.3.

Finally, we examine the influence of varying normal use conditions on the efficiencies of some particular marginal designs ξ_l, $l = 1, 2$. The efficiency of a marginal design ξ_l at a normal use condition x_{ul}, where all other parameters are kept fixed according to Table 6.1, is defined by

$$\text{eff}(\xi_l; \boldsymbol{\beta}_l, x_{ul}) = \frac{(1, x_{ul}) \mathbf{M}_{\beta_l}(\xi_l^{(x_{ul})})^{-1} (1, x_{ul})^T}{(1, x_{ul}) \mathbf{M}_{\beta_l}(\xi_l)^{-1} (1, x_{ul})^T},$$

where $\xi_l^{(x_{ul})}$ denotes a locally optimal design at $\boldsymbol{\beta}_l$, that is, $\xi_l^{(x_{ul})}$ minimizes

$$(1, x_{ul}) \mathbf{M}_{\beta_l}(\widetilde{\xi}_l)^{-1} (1, x_{ul})^T$$
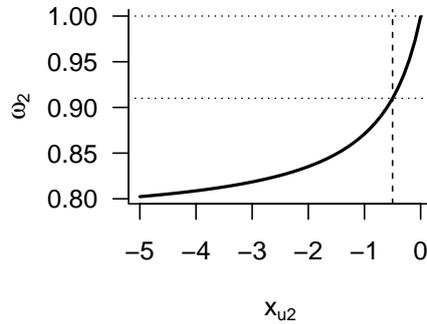
FIGURE 6.3: Optimal weight in dependence on x_{u2} for Example 6.2.4
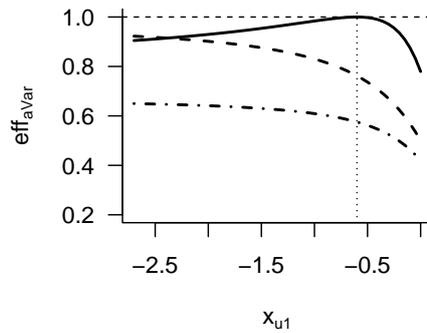
FIGURE 6.4: Efficiency of ξ_1^* (solid line), $\bar{\xi}_2$ (dashed line) and $\bar{\xi}_3$ (dashed and dotted line) in dependence on x_{u1} for Example 6.2.4

over all marginal designs $\widetilde{\xi}_l$ on $[0, 1]$, and the present marginal efficiencies may serve as lower bounds for the combined efficiency $\mathrm{eff}(\xi; \boldsymbol{\beta}, \mathbf{x}_u)$ of the combined design ξ. In Figure 6.4 and Figure 6.5 we plot, respectively, the efficiencies of the locally optimal designs ξ_1^* and ξ_2^* from (6.13) (solid line), the efficiencies of the design $\bar{\xi}_2$ (dashed line) which assigns equal weights $1/2$ to the points 0 and 1, and the design $\bar{\xi}_3$ (dashed line) which assigns equal weights $1/3$ to the marginal stress levels $0, 0, 5$ and 1. Note that the latter designs $\bar{\xi}_2$ and $\bar{\xi}_3$ may serve as standard designs. The nominal values for x_{u1} and x_{u2} from Table 6.1 are indicated in the figures by vertical dotted lines. The efficiencies of the optimal designs ξ_1^* and ξ_2^* from (6.13) seem to perform quite well over the ranges of x_{u1} and x_{u2}, respectively. The design $\bar{\xi}_2$ is preferable for small values of x_{u1} while the design $\bar{\xi}_3$ performs worse throughout for reasonable values of both x_{u1} and x_{u2}.

111
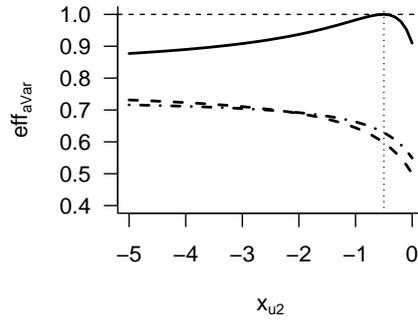
FIGURE 6.5: Efficiency of ξ_2^* (solid line), $\bar{\xi}_2$ (dashed line) and $\bar{\xi}_3$ (dashed and dotted line) in dependence on x_{u2} for Example 6.2.4

6.3 Bivariate Gamma model with dependent components

Again, let the system under study have two failure modes corresponding to two degradation components, but independence of the components will no longer be assumed. How to model the case of *dependent* degradation components. One would like to have as a model like the following. Each marginal degradation component should follow a Gamma process $Z_t^{(l)}$, $l = 1, 2$, as explained in Subsection 6.2.1. The joint degradation path of both failure modes $\boldsymbol{Z}_t = \left(Z_t^{(1)}, Z_t^{(2)} \right)$ should be a process with independent increments, and the distribution function $F^{(h)}(y_1, y_2)$ of an increment $\boldsymbol{Z}_{t+h} - \boldsymbol{Z}_t$, $t \geq 0$, $h > 0$, should be given by a fixed copula $C(r, s)$, $0 \leq r, s \leq 1$, describing the dependence structure between the marginal processes,

$$F^{(h)}(y_1, y_2) = C\left(F_1^{(h)}(y_1), F_2^{(h)}(y_2) \right), \quad y_1, y_2 > 0, \tag{6.15}$$

where $F_l^{(h)}(y_l)$, $l = 1, 2$, denotes the distribution function of the increment $Z_{t+h}^{(l)} - Z_t^{(l)}$ of the marginal Gamma process. Note that (6.15) implies that the bivariate process has stationary increments. The reason for using a copula is its ability to provide a flexible and convenient method for combining marginal distributions in a multivariate distribution, see (Pan, Balakrishnan*, and Sun, 2011), see also (Sklar, 1959) for Sklar's Theorem. Two particular copulas are the Frank copula and the Gaussian copula, employed in recent work on degradation modelling, see the corresponding definitions in Subsection 6.3.1 below. However, a copula $C(r, s)$ such that a bivariate process as described exists, is unknown, unless the independence copula $C(r, s) = rs$ which retrieves the case of independent components. Note that, by the assumption of independent increments of the bivariate process, the family of bivariate distributions Q_h, $h > 0$, given by (6.15)

must form a convolution semi-group which, however, is unknown and even not known to exist (unless, of course, in case of the independence copula). As a way out, we do no longer consider *processes* (marginal or bivariate processes), but restrict to a simple model considering degradations and their increments only at k fixed time points.

6.3.1   A simple bivariate copula model

Let $k \geq 1$ time points be given, $0 < t_1 < \cdots < t_k$. Denote $\Delta_j = t_j - t_{j-1}$, $j = 1, \ldots, k$, where $t_0 = 0$. Consider the degradation $Z_j^{(l)}$ at time t_j of the lth component and the increments $Y_{jl} = Z_j^{(l)} - Z_{j-1}^{(l)}$, $j = 1, \ldots, k$, $l = 1, 2$, where $Z_0^{(l)} = 0$. For each $l = 1, 2$, the increments $Y_{1l}, \ldots, Y_{kl}$ are independent and Gamma distributed with parameters as in Section 6.2. In particular, the shape parameter of the Gamma distribution of Y_{jl} is given by $\gamma_l(x_l) \, \Delta_j$, where $\mathbf{x} = (x_1, x_2)$ is a normalized bivariate stress variable chosen from the experimental region $[\, 0 \,, 1\,]^2$, and

$$\gamma_l(x_l) = \exp\Big(\beta_{1l} + \beta_{2l} x_l\Big).$$

The bivariate increments $\mathbf{Y}_j = (Y_{j1}, Y_{j2})$ of the bivariate degradations $\boldsymbol{Z} = \Big(Z_j^{(1)}, Z_j^{(2)}\Big)$, $j = 1, \ldots, k$, are assumed to be independent and follow a distribution according to (6.15), that is, the distribution function of $\mathbf{Y}_j$ is given by

$$F_j(y_1, y_2) = C\Big(F_{j1}(y_1), F_{j2}(y_2)\Big), \quad y_1, y_2 > 0, \tag{6.16}$$

where C is a given copula and F_{jl} denotes the distribution function of the Gamma distribution with shape parameter $\gamma(x_l) \, \Delta_j$ and scale ν_l. The copula is assumed to be smooth (sufficiently often continuously differentiable), and thus it has a density

$$c(r, s) = \frac{\partial^2 C(r, s)}{\partial r \partial s}, \quad 0 < r, s < 1. \tag{6.17}$$

Hence it follows that the bivariate increment $\mathbf{Y}_j$ has a density

$$f_j(\boldsymbol{y}) = c\Big(F_{j1}(y_1), F_{j2}(y_2)\Big) f_{j1}(y_1) \, f_{j2}(y_2), \quad \boldsymbol{y} = (y_1, y_2) \in (\, 0 \,, \infty)^2, \tag{6.18}$$

where f_{jl} denotes the Gamma density with shape $\gamma(x_l) \, \Delta_j$ and scale ν_l.

By (6.18) and by independence of the increments, the log-likelihood for the parameter vector $\boldsymbol{\beta} = (\beta_{11}, \beta_{21}, \beta_{12}, \beta_{22})^T$ given the values $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ of the increments $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$ and under the stress condition $\mathbf{x} = (x_1, x_2) \in [\, 0 \,, 1\,]^2$, reads as

$$\ell\Big(\boldsymbol{\beta}; \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \mathbf{x}\Big) = \sum_{j=1}^{k} \left[\log\Big(c\big(F_{j1}(y_{j1}), F_{j2}(y_{j2})\big)\Big) + \sum_{l=1}^{2} \log\Big(f_{jl}(y_{jl})\Big) \right]. \tag{6.19}$$

The following definitions present two particular copulas (in two dimensions) to be considered in further applications: the Frank copula and the Gaussian copula.

Definition 6.3.1. *The Frank copula, which is a very common Archimedean copula for bivariate data, is utilized to describe the dependence relation between marginal failure modes. The bivariate Frank copula is defined as*

$$C(r, s) = -\frac{1}{\varkappa} \log \left(1 + \frac{\left(e^{-\varkappa r} - 1\right)\left(e^{-\varkappa s} - 1\right)}{e^{-\varkappa} - 1} \right) \tag{6.20}$$

where $\varkappa \in (-\infty, \infty) \backslash \{0\}$ is a fixed copula dependence parameter. The density from (6.17) becomes

$$c(r, s) = \frac{\varkappa \left(1 - e^{-\varkappa}\right) e^{-\varkappa(r+s)}}{\left(1 - e^{-\varkappa} - \left(1 - e^{-\varkappa r}\right)\left(1 - e^{-\varkappa s}\right)\right)^2} \tag{6.21}$$

Definition 6.3.2. *The Gaussian copula employs a correlation parameter ρ defining a positive definite correlation matrix*

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad -1 < \rho < 1.$$

Denote by Φ the standard normal distribution function, and denote by $F_{\mathbf{0},\mathbf{\Sigma}}$ the distribution function of the bivariate normal distribution with expectation $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$, that is,

$$F_{\mathbf{0},\mathbf{\Sigma}}(a, b) = (2\pi)^{-1}(\det(\Sigma))^{-1/2} \int_{-\infty}^{a} \int_{-\infty}^{b} \exp\left(-\tfrac{1}{2}\mathbf{z}^T \Sigma^{-1} \mathbf{z}\right) d\mathbf{z}. \tag{6.22}$$

Then, the Gaussian copula reads as

$$C(r, s) = F_{\mathbf{0},\mathbf{\Sigma}}\left(\Phi^{-1}(r), \Phi^{-1}(s)\right), \quad r, s \in (0, 1). \tag{6.23}$$

Its density according to (6.17) is given by

$$c(r, s) = \frac{(2\pi)^{-1}(\det(\Sigma))^{-1/2} \exp\left(-\tfrac{1}{2}\left(\Phi^{-1}(r), \Phi^{-1}(s)\right)\Sigma^{-1}\left(\Phi^{-1}(r), \Phi^{-1}(s)\right)^T\right)}{\phi\left(\Phi^{-1}(r)\right) \phi\left(\Phi^{-1}(s)\right)},$$
$$\tag{6.24}$$

where ϕ denotes the standard normal density. The normal copula space provides a flexible and convenient method for combining marginal distributions in a multivariate distribution, see (Pan, Balakrishnan*, and Sun, 2011). Using the Gaussian copula in our bivariate Gamma model, the resulting density (6.18) of a bivariate increment was employed in (Adegbola and Yuan, 2019).

6.3.2 Information matrix

From the log-likelihood (6.19) we calculate the elemental Fisher information matrix of $\mathbf{x}$ at $\boldsymbol{\beta}$,

$$\mathbf{M}_\beta(\mathbf{x}) = -\mathrm{E}\Big(\frac{\partial^2 \ell(\boldsymbol{\beta}; \mathbf{Y}_1, \ldots, \mathbf{Y}_k, \mathbf{x})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T}\Big), \tag{6.25}$$

The symbol $\mathbf{M}_\beta^{(\mathrm{ind})}(\mathbf{x})$ will be used for the elemental information matrix from the model with independent components studied in Section 6.2. In fact, on the r.h.s. of (6.19), the second term (double sum over $j = 1, \ldots, k$ and $l = 1, 2$) yields, after (twice) partial differentiation, taking the expectation and putting a minus sign in front, the information matrix from (6.3) since the expectation of that term depend only on the marginal distributions of $\mathbf{Y}_j$, $j = 1, \ldots, k$, which are the same Gamma distributions as introduced in Section 6.2. It remains to calculate the matrix

$$\mathrm{E}\Big(\frac{\partial^2 \log c\big(F_{j1}(Y_{j1}), F_{j2}(Y_{j2})\big)}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^T}\Big). \tag{6.26}$$

Here F_{jl} denotes the distribution function of the Gamma distribution with shape $\gamma_l(x_l)\,\Delta_j$ and scale ν_l, f_{jl} denotes its density, and $\gamma_l(x_l) = \exp(\beta_{1l} + \beta_{2l}x_l)$. Formulas for (6.26) are derived in A.2, which involve two-dimensional integrals. From this, the information matrix (6.25) reads as

$$\mathbf{M}_\beta(\mathbf{x}) = \mathbf{H}_\beta(\mathbf{x}) + \mathbf{M}_\beta^{(\mathrm{ind})}(\mathbf{x}), \tag{6.27}$$

$$\text{where } \mathbf{H}_\beta(\mathbf{x}) = \left[\begin{array}{cc} \mathbf{H}_1(\mathbf{x}, \boldsymbol{\beta}) & \mathbf{H}_{12}(\mathbf{x}, \boldsymbol{\beta}) \\ \mathbf{H}_{12}^T(\mathbf{x}, \boldsymbol{\beta}) & \mathbf{H}_2(\mathbf{x}, \boldsymbol{\beta}) \end{array} \right],$$

$$\mathbf{H}_l(\mathbf{x}, \boldsymbol{\beta}) = \varphi_l(\mathbf{x}, \boldsymbol{\beta}) \left(1, x_l\right)^T \left(1, x_l\right), \ l = 1, 2, \tag{6.28}$$

$$\mathbf{H}_{12}(\mathbf{x}, \boldsymbol{\beta}) = \varphi_{12}(\mathbf{x}, \boldsymbol{\beta}) \left(1, x_1\right)^T \left(1, x_2\right),$$

$$\varphi_l(\mathbf{x}, \boldsymbol{\beta}) = \gamma_l^2(x_l) \sum_{j=1}^k \int_0^\infty \int_0^\infty \frac{c_l^2\big(F_{j1}(y_1), F_{j2}(y_2)\big)}{c\big(F_{j1}(y_1), F_{j2}(y_2)\big)} \left(\frac{\partial F_{jl}(y_l)}{\partial \gamma_l}\right)^2 \tag{6.29}$$

$$f_{j1}(y_1)\, f_{j2}(y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2, \quad l = 1, 2,$$

and

$$\varphi_{12}(\mathbf{x}, \boldsymbol{\beta}) =$$

$$\gamma_1(x_1)\, \gamma_2(x_2) \sum_{j=1}^k \int_0^\infty \int_0^\infty \Big[\frac{c_1\big(F_{j1}(y_1), F_{j2}(y_2)\big)\, c_2\big(F_{j1}(y_1), F_{j2}(y_2)\big)}{c\big(F_{j1}(y_1), F_{j2}(y_2)\big)} \frac{\partial F_{j1}(y_1)}{\partial \gamma_1}$$

$$\frac{\partial F_{j2}(y_2)}{\partial \gamma_2} f_{j1}(y_1)\, f_{j2}(y_2) - c\big((F_{j1}(y_1), F_{j2}(y_2)\big) \frac{\partial f_{j1}(y_1)}{\partial \gamma_1} \frac{\partial f_{j2}(y_2)}{\partial \gamma_2} \Big] \mathrm{d}y_1\, \mathrm{d}y_2.$$

such that $c_1(r,s)$ and $c_2(r,s)$ denote the first order partial derivatives of the copula density $c(r,s)$, that is,

$$c_1(r,s) = \frac{\partial c(r,s)}{\partial r} \quad \text{and} \quad c_2(r,s) = \frac{\partial c(r,s)}{\partial s}, \quad 0 < r, s < 1.$$

Formulas for the partial derivatives $\partial F_{jl}(y_1)/\partial \gamma_l$ and $\partial f_{jl}(y_1)/\partial \gamma_l$ are given in A.2. Note that in case of equidistant time points $t_1, \ldots, t_k$, that is, $\Delta_j = \Delta$ for $j = 1, \ldots, k$, the distribution functions and densities F_{jl} and f_{jl}, respectively, are independent of j, and the above formulas simplify in that case.

As usual, if ξ is an (approximate) design on $[0, 1]^2$ with support points $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and corresponding weights w_i, $i = 1, \ldots, m$, the information matrix of ξ at a parameter point $\boldsymbol{\beta}$ is given by

$$\mathbf{M}_{\boldsymbol{\beta}}(\xi) = \sum_{i=1}^{m} w_i \, \mathbf{M}_{\boldsymbol{\beta}}(\mathbf{x}_i). \tag{6.30}$$

In contrast to the settings of independent response components in Section 6.2, the D-optimality criterion will be applied, instead of the c-criterion, for the current settings of Copula-based bivariate degradation models. The main reason behind that is the difficulty to define the continuous failure time variable T, and, hence, the quantile t_α, under the assumptions of dependent marginal failure modes based on Copula functions with k fixed time points. Accordingly, we are adopting the D-criterion for the numerical calculations in Example 6 and Example 7 where the general equivalence theorem is utilized to validate the optimality of the numerically obtained designs.

6.3.3 Local D-optimality

For a given parameter point $\boldsymbol{\beta}$, a design ξ^* is called locally D-optimal at $\boldsymbol{\beta}$ if ξ^* maximizes $\det\big(\mathbf{M}_{\boldsymbol{\beta}}(\xi)\big)$ over all designs ξ. For numerical computation of a locally D-optimal design we used the multiplicative algorithm, where the design region $[0, 1]^2$ is discretized by a grid with 0.05 increments in both dimensions. The elemental information matrices from (6.27) were computed by numerical integration in two dimensions. We employed the Frank copula and the Gaussian copula from based on Definitions 6.3.1 and 6.3.2, respectively.

Example 6. *Let $C(r,s)$ be the Frank copula from (6.20). Its density $c(r,s)$ is given by (6.21). By straightforward calculations, one obtains the first order partial derivatives*

TABLE 6.2: Nominal values of the bivariate Gamma model with Copula function

β_{11}	β_{12}	ν_1	β_{21}	β_{22}	ν_2	$\varkappa$	ρ
0.30	0.90	1.17	0.80	0.10	1.15	-0.40	-0.10

$c_1(r,s) = \partial c(r,s) \big/ \partial r$ *and* $c_2(r,s) = \partial c(r,s) \big/ \partial s$,

$$c_1(r,s) = \frac{\varkappa^2\left(1 - e^{-\varkappa}\right)e^{-\varkappa(r+s)}\left[(1 + e^{-\varkappa r})(1 - e^{-\varkappa s}) - (1 - e^{-\varkappa})\right]}{\left[1 - e^{-\varkappa} - \left(1 - e^{-\varkappa r}\right)\left(1 - e^{-\varkappa s}\right)\right]^3}, \quad c_2(r,s) = c_1(s,r).$$

(6.31)

Choosing $k = 4$ equidistant time points $t_1 = 0.05$, $t_2 = 0.10$, $t_3 = 0.15$, $t_4 = 0.20$, and the nominal values of the parameter vector $\boldsymbol{\beta}$ in Table 6.2, numerical computations with the multiplicative algorithm were done for a locally D-optimal design. The obtained locally D-optimal design is a uniformly weighted 6-point design,

$$\xi_D^* = \begin{pmatrix} (0,0) & (0,1) & (0.5,0) & (0.5,1) & (1,0) & (1,1) \\ 0.166 & 0.166 & 0.166 & 0.166 & 0.166 & 0.166 \end{pmatrix}$$

(6.32)

Example 7. *Let $C(r,s)$ be the Gaussian copula from (6.23) with parameter value $\rho = -0.1$. Its density is given by (6.24), and the first order partial derivatives of the latter are given by*

$$c_1(r,s) = \frac{\rho}{1 - \rho^2}\, c(r,s)\, \frac{\Phi^{-1}(s) - \rho\Phi^{-1}(r)}{\phi\left(\Phi^{-1}(r)\right)}, \quad c_2(r,s) = c_1(s,r).$$

As is the preceeding example, we choose $k = 4$ equidistant time points $t_1 = 0.05$, $t_2 = 0.10$, $t_3 = 0.15$, $t_4 = 0.20$, and the nominal values of the parameter vector $\boldsymbol{\beta}$ from Table 6.2. The locally D-optimal design obtained with the multiplicative algorithm has the same six support points as that for Example 6, with non-uniform weights, as

$$\xi_D^* = \begin{pmatrix} (0,0) & (0,1) & (0.5,0) & (0.5,1) & (1,0) & (1,1) \\ 0.20 & 0.20 & 0.16 & 0.16 & 0.18 & 0.09 \end{pmatrix}$$

(6.33)

Due to the difficulty of accurately deriving the information matrix 6.27 for the Copula-based models 6.3.1 and 6.3.2 with multiple observations, we consider in Section 6.4 a simplified approach with binary outcomes which facilitates the derivations of the corresponding information matrix and, hence, considerably reduce the calculations time.

6.4 Copula-based gamma model with binary outcomes

6.4.1 Model formulation

In this section, we consider the model from Section 6.3, but now the measurements of bivariate degradations $\boldsymbol{Z}_j = (Z_j^{(1)}, Z_j^{(2)})$, $j = 1, ..., k$, are reduced to the information on whether or not the marginal degradation paths have reached or exceeded given thresholds $z_{10} > 0$ and $z_{20} > 0$, respectively, at each time t_j, $j = 1, ..., k$. This information is equivalently reflected by two discrete variables U and V with values in $\{1, ..., k, k + 1\}$, where U (resp. V) gives the first time label j such that the marginal degradation $Z_j^{(1)}$ (resp. $Z_j^{(2)}$) has reached or exceeded the threshold z_{01} (resp. z_{02}), and the value $k + 1$ expresses that failure did not occur until time t_k. That is, we define

$$
\begin{aligned}
U &= \min\left\{j \in \{1, ..., k\} : Z_j^{(1)} \geq z_{10}\right\}, \\
V &= \min\left\{j \in \{1, ..., k\} : Z_j^{(2)} \geq z_{20}\right\},
\end{aligned}
$$

where the minimum of the empty set is defined to be $k+1$. The joint distribution of U, V is given by the probabilities $P_{u,v} = \Pr(U = u, V = v)$, $u, v \in \{1, ..., k, k + 1\}$. Below we will see that their calculation involves multi-dimensional integrals over polyhedral regions which are difficult to handle theoretically as well as numerically. A slight simplification of the integration regions is gained by considering the probabilities

$$
Q_{u,v} = \Pr(U \leq u, V \leq v) \quad \text{for} \quad 1 \leq u, v \leq k + 1.
$$

Note that $Q_{k+1,v} = \Pr(V \leq v)$ and $Q_{u,k+1} = \Pr(U \leq u)$, and especially $Q_{k+1,k+1} = 1$. The probabilities $P_{u,v}$ are obtained from the $Q_{u,v}$ by

$$
P_{u,v} = Q_{u,v} - Q_{u,v-1} - Q_{u-1,v} + Q_{u-1,v-1} \quad \text{for } 1 \leq u, v \leq k + 1 \qquad (6.34)
$$

where $Q_{0,0} = Q_{0,v} = Q_{u,0} = 0$ for $1 \leq u, v \leq k + 1$. By the two equivalences, for any $u, v \in \{1, ..., k\}$,

$$
U \leq u \iff Z_u^{(1)} \geq z_{10}, \quad V \leq v \iff Z_v^{(2)} \geq z_{20},
$$

and writing the degradations as sums of increments, $Z_u^{(1)} = \sum_{j=1}^{u} Y_{j1}$ and $Z_v^{(2)} = \sum_{j=1}^{v} Y_{j2}$, we get for all $u, v \in \{1, ..., k\}$,

$$Q_{u,v} = \int_{A_{u,v}} \prod_{j=1}^{k} f_j(\boldsymbol{y}_j) \, d\boldsymbol{y}_1 \cdots \boldsymbol{y}_k, \tag{6.35}$$

where

$$A_{u,v} = \left\{ (\boldsymbol{y}_1, ..., \boldsymbol{y}_k) \in (0, \infty)^{2k} : \sum_{j=1}^{u} y_{j1} \geq z_{10}, \sum_{j=1}^{v} y_{j2} \geq z_{20} \right\},$$

and f_j denotes the density of the bivariate increment $\mathbf{Y}_j = (Y_{j1}, Y_{j2})$ from (6.18). For $u = k + 1$ or $v = k + 1$, a calculation of $Q_{k+1,v}$ or $Q_{u,k+1}$ involves only the marginal degradations, which are Gamma distributed,

$$Q_{k+1,v} = \Pr\left(Z_v^{(2)} \geq z_{20}\right) = \frac{\Gamma\left(\gamma_2(x_2) \, t_v, y_2/\nu_2\right)}{\Gamma\left(\gamma_2(x_2) \, t_v\right)}, \quad 1 \leq v \leq k,$$

$$Q_{u,k+1} = \frac{\Gamma\left(\gamma_1(x_1) \, t_u, y_1/\nu_1\right)}{\Gamma\left(\gamma_1(x_1) \, t_u\right)}, \quad 1 \leq u \leq k.$$

6.4.2 Information matrix

The log likelihood of the bivariate discrete variable (U, V) is given by

$$\ell(\boldsymbol{\beta}; u, v, \mathbf{x}) = \log P_{u,v}(\mathbf{x}, \boldsymbol{\beta}) \tag{6.36}$$

where now we observe the dependence of the probabilities $P_{u,v}, 1 \leq u, v \leq k + 1$, on the design variable $\mathbf{x} = (x_1, x_2)$ and the parameter vector $\boldsymbol{\beta} = (\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})^T$. The elemental information matrix of $\mathbf{x}$ at a parameter point $\boldsymbol{\beta}$ is given by

$$\mathbf{M}_{\beta}(\mathbf{x}) = \mathrm{E}\left[\left(\frac{\partial \ell(\boldsymbol{\beta}; U, V, \boldsymbol{x})}{\partial \boldsymbol{\beta}} \right) \left(\frac{\partial \ell(\boldsymbol{\beta}; U, V, \boldsymbol{x})}{\partial \boldsymbol{\beta}} \right)^T \right]. \tag{6.37}$$

We can decompose ℓ, as a function of $\boldsymbol{\beta}$, according to

$$\boldsymbol{\beta} \longrightarrow \boldsymbol{\gamma} = (\gamma_1, \gamma_2)^T \longrightarrow \boldsymbol{P} = (P_{1,1}, ..., P_{u,v}, ..., P_{k+1,k+1})^T \longrightarrow \ell,$$

where the $P_{u,v}$, $1 \leq u, v \leq k + 1$, have been arranged in lexicographic order, say, to form the vector $\boldsymbol{P}$. By the chain rule a factorization of the gradient $\partial \ell(\boldsymbol{\beta}; u, v, \boldsymbol{x})/\partial \boldsymbol{\beta}$ results,

$$\frac{\partial \ell(\boldsymbol{\beta}; u, v, \boldsymbol{x})}{\partial \boldsymbol{\beta}} = \boldsymbol{ABC},$$

with matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and a column vector $\boldsymbol{C}$,

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{A}(\mathbf{x}, \boldsymbol{\beta}) = \frac{\partial \boldsymbol{\gamma}}{\partial \boldsymbol{\beta}} =
\begin{bmatrix}
\gamma_1 & 0 \\
x_1\gamma_1 & 0 \\
0 & \gamma_2 \\
0 & x_2\gamma_2
\end{bmatrix}, \\[2mm]
\boldsymbol{B} &= \boldsymbol{B}(\mathbf{x}, \boldsymbol{\beta}) = \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\gamma}} =
\begin{bmatrix}
\frac{\partial P_{u,v}}{\partial \gamma_1}(1 \le u, v \le k+1) \\
\frac{\partial P_{u,v}}{\partial \gamma_2}(1 \le u, v \le k+1)
\end{bmatrix}, \\[2mm]
\boldsymbol{C} &= \boldsymbol{C}(u, v) = \frac{\partial \ell}{\partial \boldsymbol{P}_{(1 \le u,v \le k+1)}} = \left(\frac{1}{P_{uv}}\right) \mathbf{1}_{(u,v)},
\end{aligned}
$$

where $\mathbf{1}_{(u,v)}$ is the unit vector with entry "1" at position (u,v). Note that the two rows of $\boldsymbol{B}$ and the column vector $\boldsymbol{C}$ have components indexed by the pairs (u,v) arranged in lexicographic order. It follows that

$$
\mathbf{M}_{\beta}(\mathbf{x}) = \boldsymbol{A}\boldsymbol{B}\,\mathrm{E}(\boldsymbol{C}\boldsymbol{C}^T)\,\boldsymbol{B}^T\boldsymbol{A}^T, \tag{6.38}
$$

and

$$
\mathrm{E}(\boldsymbol{C}\boldsymbol{C}^T) = \mathrm{diag}\left(\frac{1}{P_{u,v}}(1 \le u, v \le k+1)\right).
$$

Again, for a design ξ with support points $\mathbf{x}_i$ and weights w_i, $i = 1, \ldots, m$, the information matrix of ξ at $\boldsymbol{\beta}$ is given by

$$
\mathbf{M}_{\beta}(\xi) = \sum_{i=1}^{m} w_i\,\mathbf{M}_{\beta}(\mathbf{x}_i).
$$

In order to obtain explicit formulas for the entries of $\boldsymbol{B}(\mathbf{x}, \boldsymbol{\beta})$, that is, the partial derivatives $\partial P_{uv}/\partial \gamma_l$, we consider the corresponding partial derivatives of the probabilities $Q_{u,v}$ from (6.35). One gets

$$
\begin{aligned}
\frac{\partial Q_{u,v}}{\partial \gamma_l} &= \int_{A_{u,v}} \frac{\partial}{\partial \gamma_l} \prod_{j=1}^{k} f_j(\boldsymbol{y}_j)\,\mathrm{d}\boldsymbol{y}_1...\mathrm{d}\boldsymbol{y}_k \\[2mm]
&= \int_{A_{u,v}} \sum_{i=1}^{k} \left[\prod_{j \ne i} f_j(\boldsymbol{y}_j)\right] \frac{\partial f_i(\boldsymbol{y}_i)}{\partial \gamma_l}\,\mathrm{d}\boldsymbol{y}_1...\boldsymbol{y}_k, \tag{6.39}
\end{aligned}
$$

and by (6.18),

$$
\begin{aligned}
\frac{\partial f_j(\boldsymbol{y}_j)}{\partial \gamma_1} &= \left[c_1\Big(F_{j1}(y_{j1}), F_{j2}(y_{j2})\Big)\frac{\partial F_{j1}(y_{j1})}{\partial \gamma_1} f_{j1}(y_{j1}) + c\Big(F_{j1}(y_{j1}), F_{j2}(y_{j2})\Big)\frac{\partial f_{j1}(y_{j1})}{\partial \gamma_1}\right] f_{j2}(y_{y2}), \\[2mm]
\frac{\partial f_j(\boldsymbol{y}_j)}{\partial \gamma_2} &= \left[c_2\Big(F_{j1}(y_{j1}), F_{j2}(y_{j2})\Big)\frac{\partial F_{j2}(y_{j2})}{\partial \gamma_2} f_{j2}(y_{j2}) + c\Big(F_{j1}(y_{j1}), F_{j2}(y_{j2})\Big)\frac{\partial f_{j2}(y_{j2})}{\partial \gamma_2}\right] f_{j1}(y_{j1}).
\end{aligned}
$$
$$
\tag{6.40}
$$

However, due to the $2k$-dimensional integration in (6.39) the calculation of information matrices is not tractable when $k > 1$. Therefore, we consider now the simple case $k = 1$ of a single measurement. Then, we have one bivariate increment $\boldsymbol{Y} = (Y_1, Y_2)$, and the distribution function of $\boldsymbol{Y}$ is given by

$$C(F_1(y_1); F_2(y_2)), \quad y_1, y_2 \in (0, \infty).$$

The probabilities $P_{u,v}$, $u, v \in \{1, 2\}$, can be expressed by the latter joint distribution function and the marginal distribution functions F_1 and F_2,

$$
\begin{aligned}
P_{2,2} &= \mathrm{P}\Big(Y_1 < z_{10}, Y_2 < z_{20}\Big) = C\Big(F_1(z_{10}), F_2(z_{20})\Big), \\
P_{1,2} &= \mathrm{P}\Big(Y_1 \geq z_{10}, Y_2 < z_{20}\Big) = F_2(z_{20}) - C\Big(F_1(z_{10}), F_2(z_{20})\Big), \\
P_{2,1} &= \mathrm{P}\Big(Y_1 < z_{10}, Y_2 \geq z_{20}\Big) = F_1(z_{10}) - C\Big(F_1(z_{10}), F_2(z_{20})\Big), \\
P_{1,1} &= \mathrm{P}\Big(Y_1 \geq z_{10}, Y_2 \geq z_{20}\Big) = 1 - F_1(z_{10}) - F_2(z_{20}) + C\Big(F_1(z_{10}), F_2(z_{20})\Big).
\end{aligned}
\tag{6.41}
$$

The partial derivatives $\partial P_{u,v}/\partial \gamma_l$ are easily obtained from the partial derivatives $\partial F_l(z_{l0})/\partial \gamma_l$ and the partial derivatives of the copula, $C_1(r, s) = \partial C(r, s)/\partial r$ and $C_2(r, s) = \partial C(r, s)/\partial s$, since by the chain rule

$$\frac{\partial}{\partial \gamma_l} C\Big(F_1(z_{10}), F_2(z_{20})\Big) = C_l\Big(F_1(z_{10}), F_2(z_{20})\Big) \frac{\partial F_l(z_{l0})}{\partial \gamma_l}, \quad l = 1, 2. \tag{6.42}$$

In particular, when C is the Frank copula with parameter $\varkappa$, then by straightforward calculation,

$$C_1(r, s) = \frac{e^{-\varkappa r}\left(e^{-\varkappa s} - 1\right)}{e^{-\varkappa} - 1 + \left(e^{-\varkappa r} - 1\right)\left(e^{-\varkappa s} - 1\right)} \quad \text{and} \quad C_2(r, s) = C_1(s, r). \tag{6.43}$$

When C is the Gaussian copula with correlation parameter ρ, then one obtains (see A.2)

$$C_1(r, s) = \Phi\left(\frac{\Phi^{-1}(s) - \rho\Phi^{-1}(r)}{\sqrt{1 - \rho^2}}\right) \quad \text{and} \quad C_2(r, s) = C_1(s, r). \tag{6.44}$$

6.4.3 Locally D- or c-optimal designs when $k = 1$

For our simple binary model ($k = 1$) employing the Frank copula or the Gaussian copula, locally D- or c-optimal designs are presented in the example below. A locally D-optimal design ξ_D^* at a given parameter point $\boldsymbol{\beta}$ maximizes $\det\big(\mathbf{M}_\beta(\xi)\big)$ over all designs ξ. A locally c-optimal design ξ_c^* at $\boldsymbol{\beta}$ minimizes $\boldsymbol{c}^T \mathbf{M}_\beta(\xi)^{-1} \boldsymbol{c}$ over all designs ξ, where $\boldsymbol{c}$ is a given nonzero column vector of dimension four. Here the coefficient

vector $\boldsymbol{c}$ is chosen such that the c-criterion represents the asymptotic variance of the maximum likelihood estimator $\widehat{P}_{1,1}$ of the joint failure probability $P_{1,1} = P_{1,1}(\mathbf{x}_u, \boldsymbol{\beta})$ at normal use conditions $\mathbf{x}_u = (x_{u1}, x_{u2})$. That is,

$$\boldsymbol{c} = \frac{\partial P_{1,1}(\mathbf{x}_u, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \Big(c_1.(1, x_{u1})\,,\, c_2.(1, x_{u2})\Big)^T,$$

where

$$c_l = \gamma_l(x_{ul})\,\Delta\frac{\partial P_{1,1}}{\partial \gamma_l}, \quad l = 1, 2.$$

The partial derivatives $\partial P_{1,1}/\partial \gamma_l$ can be evaluated using formulas (6.41), (6.42), and (A.13).

Example 8. *For obtaining numerically optimal designs, the multiplicative algorithm with an equidistant grid of 0.05 marginal increments over the standardized design region $\mathcal{X} = [0, 1]^2$ is employed. The single point time plan is chosen as $t_1 = \Delta = 0.3$. The resulting optimal designs are derived in regard to the nominal values of parameters given in Table 6.2, the normal use conditions $x_{u1} = -0.40$ and $x_{u2} = -0.60$, and the failure thresholds $z_{10} = 2.56$ and $z_{20} = 2.37$.*

The D-optimal designs computed by the algorithm are the following four-point designs, which nearly coincide for the two copulas,

$$\text{Frank copula:}\quad \xi_D^* \;=\; \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.24 & 0.24 & 0.26 & 0.26 \end{pmatrix};$$

$$\text{Gaussian copula:}\quad \xi_D^* \;=\; \begin{pmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0.22 & 0.23 & 0.27 & 0.28 \end{pmatrix}.$$

The c-optimal designs from the algorithm are again four-point designs. Under the condition of an equidistant grid of 0.5 marginal increments over the standardized design region $\mathcal{X} = [0, 1]^2$ the multiplicative algorithm is employed and the resulting designs coincide on the location of support points and nearly coincide on the optimal weights with minor differences in the optimal weights of the two middle points,

$$\text{Frank copula:}\quad \xi_c^* \;=\; \begin{pmatrix} (0,0) & (0,1) & (0.5,1) & (1,1) \\ 0.07 & 0.19 & 0.47 & 0.27 \end{pmatrix};$$

$$\text{Gaussian copula:}\quad \xi_c^* \;=\; \begin{pmatrix} (0,0) & (0,1) & (0.5,1) & (1,1) \\ 0.11 & 0.22 & 0.39 & 0.28 \end{pmatrix}.$$

Further, the general equivalence theorem was used to prove the optimality of the resulting designs over $\mathcal{X}$. To evaluate the behaviour of the resulting optimal designs we consider the variations of the optimal weights when the underlying nominal values are misspecified.
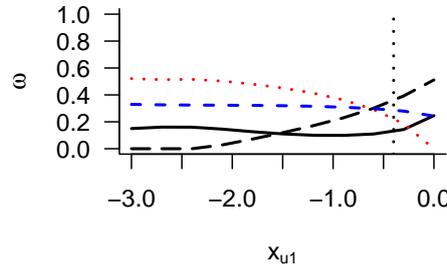
FIGURE 6.6: Dependence of the optimal weights on x_{u1} for Example 8, w_1^*: solid line, w_2^*: dotted line, w_3^*: long dashed line, w_4^*: dashed line

For brevity we consider the c-optimal design ξ_c^* on the basis of the Gaussian copula function under deviations of the normal use condition x_{u1}, and the correlation parameter ρ. The four optimal weights w_1^*, w_2^*, w_3^* and w_4^*, which belong respectively to the stress settings (0,0), (0,1), (0.5,1) and (1,1), are plotted in Figure 6.6 in dependence on x_{u1} where all parameters are held fixed to their nominal values and in Figure 6.8 in dependence on ρ where all parameters are held fixed to their nominal values. Figure 6.6 shows that the optimal weights of the middle points, i.e. w_2^* and w_3^*, considerably vary under changes of x_{u1} where the optimal weights of the extremal points, i.e. w_1^* and w_4^*, are nearly constant throughout. Figure 6.8 indicates that the resulting optimal design is more robust against misspecification of the correlation parameter ρ. The nominal values for x_{u1} and ρ are indicated by vertical dotted lines in the corresponding figure. Define by

$$\mathrm{eff}(\xi; \boldsymbol{\beta}) = \frac{\boldsymbol{c}^T \mathbf{M}_\beta(\xi_{c,\beta}^*)^{-1}\boldsymbol{c}}{\boldsymbol{c}^T \mathbf{M}_\beta(\xi)^{-1}\boldsymbol{c}},$$

the efficiency of a design ξ in regard to the asymptotic variance for estimating $P_{1,1}$ when $\boldsymbol{\beta}$ is the true value of the parameter where $\boldsymbol{c}^T\mathbf{M}_\beta(\xi)^{-1}\boldsymbol{c}$ indicates the asymptotic variance for estimating $P_{1,1}$ when the design ξ is used and $\xi_{c,\beta}^*$ is the locally optimal design at $\boldsymbol{\beta}$. Figure 6.7 and Figure 6.9 show, respectively, the efficiencies in dependence on x_{u1} and ρ together with the efficiency of the design $\bar{\xi}_2$ which assigns equal weights $1/4$ to the same support points of ξ_c^*, and the design $\bar{\xi}_3$ which assigns equal weights $1/4$ to the vertices (0,0), (0,1), (1,0) and (1,1). Again, the nominal values for x_{u1} and ρ are indicated by vertical dotted lines in the corresponding figure. In total, Figure 6.7 and Figure 6.9 indicate that the optimal design ξ_c^* performs quite well over the range of x_{u1} and ρ when compared to $\bar{\xi}_2$ and $\bar{\xi}_3$, which indicates that the optimal design is robust against changes of the normal use conditions as well as the nominal values. The existing results of the sensitivity analysis of $\xi_{c,\beta}^*$ on the basis of the Frank copula nearly
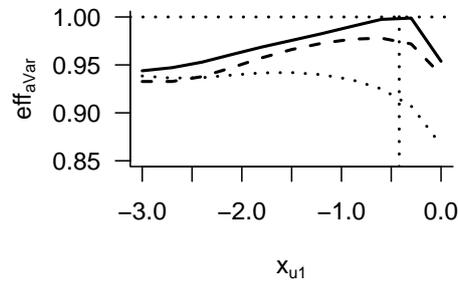
FIGURE 6.7: Efficiency of ξ_c^* (solid line), $\bar{\xi}_2$ (dashed line) and $\bar{\xi}_3$ (dotted line) in dependence on x_{u1} for Example 8

FIGURE 6.8: Dependence of the optimal weights on ρ for Example 8, w_1^*: solid line, w_2^*: dotted line, w_3^*: long dashed line, w_4^*: dashed line
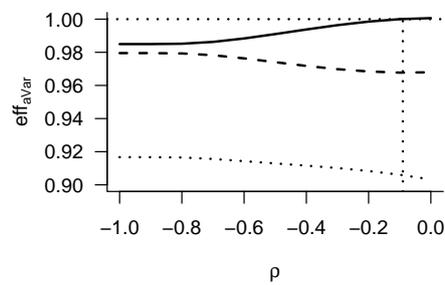
FIGURE 6.9: Efficiency of $\xi_{c,\beta}^*$ (solid line), $\bar{\xi}_2$ (dashed line) and $\bar{\xi}_3$ (dotted line) in dependence on ρ for Example 8

coincides with the obtained results in regard to the Gaussian copula, and, hence, the latter results have been removed to avoid redundancy.

6.5   Concluding remarks

Reliability engineers are demanded to provide a sophisticated assessment of the reliability related properties during the design stage of highly reliable systems. Accelerated Degradation Testing is a common approach to handle this issue. This approach has the advantage to give an estimation of lifetime and reliability characteristics of the system under study in a relatively short testing time. In this chapter, we introduced optimal experimental designs for accelerated degradation tests with two response components and repeated measures with or without dependence between marginal components. The marginal degradation paths are expressed using Gamma process models. In the current models for Accelerated Degradation Testing, we assume that stress remains constant within each unit during the whole test but may vary between units. Further, the same time plan for measurements is used for all units in the test. In the case of independent components, it is desirable to estimate certain quantiles of the joint failure time distribution as a characteristic of the reliability of the product. Hence, the purpose of optimal experimental design is to find the best settings for the stress variable to obtain most accurate estimates of the quantiles. On the other hand, the Frank copula as well as the Gaussian copula are separately adopted to represent the dependence relation in bivariate Gamma models when dependence is assumed between response components. The *D*-criterion is considered for locally optimal designs in both cases. The resulting optimal designs coincide in terms of the optimal support points but differ in their weights allocated to the points. We developed further *D*- and *c*-optimal designs when the two Copula-based models are reduced to binary responses. A sensitivity analysis showed that the resulting locally optimal designs are quite efficient against deviations from the assumed nominal values.

Chapter 7

Conclusion and Outlook

Accelerated Degradation Testing is an effective approach to assess the reliability characteristics of products with high reliability and long life. The optimal design of Accelerated Degradation Testing can considerably shorten the test duration and improve the estimation accuracy of lifetime properties. The main purpose of this thesis is to obtain optimal designs for Accelerated Degradation Testing in the presence of various testing conditions in regards to response components, explanatory variables and observation times. LMEMs as well as Gamma process models, which are frequently introduced to model the degradation process in Accelerated Degradation Testing, are considered in this thesis to express the corresponding degradation paths. The minimum asymptotic variance optimality criterion, which is close to the c-optimality criterion, is mainly considered throughout this research to derive optimal designs.

For the case of modelling Accelerated Degradation Testing with LMEMs, analytical as well as algorithm based c-optimal designs were presented for univariate and multivariate tests, respectively, under different testing conditions. Both the explanatory (independent) variables as well as the time variable are separately considered for the optimization process in the univariate case to analytically derive c-optimal designs. Regarding the multivariate case, the marginal regression functions, which correspond to LMEMs, are assumed to be dependent where the dependence between marginal models is expressed using appropriate correlation matrices. Consequently, the multiplicative algorithm is utilized to obtain locally c-optimal designs under various testing conditions. A sensitivity analysis is conducted to study the quality of the resulting experimental designs under various parameters specifications. The resulting optimal designs, which are always supported on the vertices or portion of the vertices of the design region, are relatively robust in terms of parameter variation.

Regarding the case of Accelerated Degradation Testing with a univariate Gamma model, an approach was introduced based on a generalized linear model to derive a c-optimal design in terms of a single explanatory variable. The model is extended to the situation of bivariate Gamma process with a single independent variable. We proposed further an optimal experimental design for an Accelerated Degradation Testing with

multiple failure modes given that the marginal failure modes are independent and uncorrelated. For this model, a Gamma process model and a LMEM are considered as the marginal degradation models such that the multiplicative algorithm is utilized to obtain an optimal experimental design. The sensitivity analysis showed that the optimal design for the latter case strongly depends on the corresponding nominal values.

Considering Accelerated Degradation Testing with bivariate Gamma process models and multiple stress variables, algorithm-based optimal designs were presented for the cases of independent as well as dependent marginal response components. For the case of dependent response components, a Frank copula function and a Gaussian copula function are separately considered to express the dependence relation between marginal response components from two different distributional perspectives. Subsequently, the *c*- and *D*-optimality criteria are used to obtain optimal experimental designs. The results show that for the case of independent response components the optimal design is totally supported on the vertices of the marginal design regions where the optimal designs regarding the two approaches of dependent response components are partially supported on a portion of the vertices of the combined design region along with an additional inner support point. Finally, the conducted sensitivity analysis indicates that the optimal designs in regard to Gamma process-based Accelerated Degradation Testing are more sensitive to inaccurate parameter specifications when compared to their counterparts of Accelerated Degradation Testing with linear mixed effects degradation models.

In future research, we may investigate the extension of the proposed approach of constant-stress Accelerated Degradation Testing to step-stress Accelerated Degradation Testing and/or ramp-stress Accelerated Degradation Testing along with various stochastic processes which correspond to monotonic degradation behaviour over testing time, i.e. the IG process or Wiener process. A further extension of the existing research may consider multivariate non linear mixed models with dependent components for modelling the degradation path instead of multivariate LMEMs. This approach may allow to the coverage of a wider class of complicated engineering systems with considerable efficiency. In addition to the Frank- and Gaussian-copula functions, further Copula functions, i.e. the Clayton- and Gumbel-copula functions, might be applied in order to investigate the resulting experimental designs under various dependence models.

Furthermore, several other optimality criteria can be employed for deriving optimal experimental designs for Accelerated Degradation Testing. For instance, *A*-criterion, the integrated mean squared error (IMSE)-criterion, and a minimax criterion are highly recommended for further research prospects under GP models as well as LMEMs.

Additionally, it may be of interest to incorporate cost constraints as well as design bounds into the design process and, hence, let the optimal configuration of Accelerated Degradation Testing be more accurate and realistic.

Appendix A

Appendix: Some Technical Results

A.1 Extension of $\mathbf{M}^{-1}$ in Section 3.9

Lemma A.1.1. *Let $\mathbf{F}$ be a $k \times p$ matrix of rank p, $\boldsymbol{\Sigma}_\gamma$ a non-negative definite $p \times p$ matrix, $\boldsymbol{\Sigma}_\varepsilon$ a positive definite $k \times k$ matrix, $\mathbf{V} = \mathbf{F}\boldsymbol{\Sigma}_\gamma\mathbf{F}^T + \boldsymbol{\Sigma}_\varepsilon$, and $\mathbf{M} = \mathbf{F}^T\mathbf{V}^{-1}\mathbf{F}$. Then*

$$\mathbf{M}^{-1} = (\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})^{-1} + \boldsymbol{\Sigma}_\gamma. \tag{A.1}$$

Proof. We prove the statement of the Lemma by showing that multiplication of $\mathbf{M}$ with the right hand side $\mathbf{C} = (\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})^{-1} + \boldsymbol{\Sigma}_\gamma$ results in the $k \times k$ identity matrix $\mathbf{I}_k$. For this note first that after premultiplication with $\mathbf{F}$ the right hand side can be expanded to

$$\mathbf{FC} = \boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F}(\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})^{-1} + \mathbf{F}\boldsymbol{\Sigma}_\gamma(\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})(\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})^{-1} = \mathbf{V}\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F}(\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})^{-1}.$$

Hence, by straightforward multiplication of $\mathbf{M}$ from the right by $\mathbf{C}$

$$\mathbf{MC} = \mathbf{F}^T\mathbf{V}^{-1}\mathbf{FC} = \mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F}(\mathbf{F}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{F})^{-1} = \mathbf{I}_k$$

which proofs the lemma. $\square$

A.2 Entries of the information matrix in Subsection 6.3.2

We derive formulas (6.25) by developing the double integral in (6.26). Since the index j will be fixed in our derivations, we simply write F_l, f_l, Y_l instead of F_{jl}, f_{jl}, Y_{jl}, respectively, $l = 1, 2$. Recall the partitioning of $\boldsymbol{\beta}$ as $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T\right)^T$, where $\boldsymbol{\beta}_l = \left(\beta_{1l}, \beta_{2l}\right)^T$, $l = 1, 2$. By $c_1(r, s)$, $c_2(r, s)$, $c_{11}(r, s)$, $c_{22}(r, s)$, and $c_{12}(r, s)$ we denote the

partial derivatives of $c(r, s)$,

$$c_1(r, s) = \frac{\partial c(r, s)}{\partial r}, \quad c_2(r, s) = \frac{\partial c(r, s)}{\partial s},$$
$$c_{11}(r, s) = \frac{\partial^2 c(r, s)}{\partial r^2}, \quad c_{22}(r, s) = \frac{\partial^2 c(r, s)}{\partial s^2}, \quad c_{12}(r, s) = \frac{\partial^2 c(r, s)}{\partial r\, \partial s}.$$

By straightforward calculation,

$$\frac{\partial \log c\big(F_1(y_1), F_2(y_2)\big)}{\partial \boldsymbol{\beta}_l} = \frac{c_l\big(F(y_1), F_2(y_2)\big)}{c\big(F_1(y_1), F_2(y_2)\big)} \frac{\partial F_l(y_l)}{\partial \boldsymbol{\beta}_l}, \quad l = 1, 2;$$

$$\frac{\partial^2 \log c\big(F_1(y_1), F_2(y_2)\big)}{\partial \boldsymbol{\beta}_l \boldsymbol{\beta}_l^T} = \tag{A.2}$$
$$\left[\frac{c_{ll}\big(F_1(y_1), F_2(y_2)\big)}{c\big(F_1(y_1), F_2(y_2)\big)} - \frac{c_l^2\big(F_1(y_1), F_2(y_2)\big)}{c^2\big(F_1(y_1), F_2(y_2)\big)} \right] \frac{\partial F_l(y_l)}{\partial \boldsymbol{\beta}_l} \left(\frac{\partial F_l(y_l)}{\partial \boldsymbol{\beta}_l} \right)^T$$
$$+ \frac{c_l\big(F_1(y_1), F_2(y_2)\big)}{c\big(F_1(y_1), F_2(y_2)\big)} \frac{\partial^2 F_l(y_l)}{\partial \boldsymbol{\beta}_l \partial \boldsymbol{\beta}_l^T}, \quad \text{\it l} = 1, 2; \tag{A.3}$$

$$\frac{\partial^2 \log c\big(F_1(y_1), F_2(y_2)\big)}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}_2^T} = \left[\frac{c_{12}\big(F_1(y_1), F_2(y_2)\big)}{c\big(F_1(y_1), F_2(y_2)\big)} \right.$$
$$\left. - \frac{c_1\big(F_1(y_1), F_2(y_2)\big) c_2\big(F_1(y_1), F_2(y_2)\big)}{c^2\big(F_1(y_1), F_2(y_2)\big)} \right] \frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} \left(\frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2} \right)^T \tag{A.4}$$

We show that

$$\mathrm{E}\left(\frac{c_{ll}\big(F_1(Y_1), F_2(Y_2)\big)}{c\big(F_1(Y_1), F_2(Y_2)\big)} \frac{\partial F_l(Y_l)}{\partial \boldsymbol{\beta}_l} \left(\frac{\partial F_l(Y_l)}{\partial \boldsymbol{\beta}_l} \right)^T \right) = \mathbf{0}, \quad l = 1, 2. \tag{A.5}$$

Using the joint density of $\boldsymbol{Y} = (Y_1, Y_2)$ from (6.18), the expectation on the l.h.s. of (A.5) rewrites, when $l = 1$, as

$$\int_0^\infty \int_0^\infty c_{11}\big(F_1(y_1), F_2(y_2)\big) \frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} \left(\frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} \right)^T f_1(y_1)\, f_2(y_2)\, \mathrm{d}y_2 \mathrm{d}y_2$$
$$= \int_0^\infty \frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} \left(\frac{\partial F_l(y_l)}{\partial \boldsymbol{\beta}_l} \right)^T f_1(y_1) \left\{ \int_0^\infty c_{11}\big(F_1(y_1), F_2(y_2)\big) f_2(y_2)\, \mathrm{d}y_2 \right\} \mathrm{d}y_1.$$

For any fixed y_1, the inner integral becomes, by substituting $s = F_2(y_2)$ and interchanging integral and derivatives,

$$\int_0^1 c_{11}\Big(F_1(y_1), s\Big)\,\mathrm{d}s = \frac{\partial^2}{\partial r^2} \int_0^1 c(r, s)\,\mathrm{d}s\bigg|_{r=F_1(y_1)} = 0,$$

where the last equation follows from $\int_0^1 c(r, s)\,\mathrm{d}s = 1$ for all $0 < r < 1$. Hence (A.5) follows for $l = 1$, and the case $l = 2$ can be proved analoguously. Next we show that

$$\mathrm{E}\left(\frac{c_l\Big(F_1(Y_1), F_2(Y_2)\Big)}{c\Big(F_1(Y_1), F_2(Y_2)\Big)} \frac{\partial^2 F_l(Y_l)}{\partial\boldsymbol{\beta}_l\partial\boldsymbol{\beta}_l^T}\right) = \mathbf{0}, \quad l = 1, 2. \tag{A.6}$$

Again using the density from (6.18) and restricting to $l = 1$ (the case $l = 2$ is analogous), the expectation on the l.h.s. of (A.6) rewrites as

$$\int_0^\infty \int_0^\infty c_1\Big(F_1(y_1), F_2(y_2)\Big) \frac{\partial^2 F_1(y_1)}{\partial\boldsymbol{\beta}_1\partial\boldsymbol{\beta}_1^T} f_1(y_1)\, f_2(y_2)\,\mathrm{d}y_2\mathrm{d}y_1$$
$$= \int_0^\infty \frac{\partial^2 F_1(y_1)}{\partial\boldsymbol{\beta}_1\partial\boldsymbol{\beta}_1^T} f_1(y_1) \left\{\int_0^\infty c_1\Big(F_1(y_1), F_2(y_2)\Big) f_2(y_2)\,\mathrm{d}y_2\right\}\mathrm{d}y_1,$$

and for any fixed y_1 the inner integral is equal to

$$\int_0^1 c_1\Big(F_1(y_1), s\Big)\,\mathrm{d}s = \frac{\partial}{\partial r} \int_0^1 c(r, s)\,\mathrm{d}s\bigg|_{r=F_1(y_1)} = 0.$$

From (A.5), (A.6), and (A.2) it follows that

$$\mathrm{E}\left(\frac{\partial^2 \log c\Big(F_1(Y_1), F_2(Y_2)\Big)}{\partial\boldsymbol{\beta}_l\boldsymbol{\beta}_l^T}\right) = \mathrm{E}\left(\frac{c_l^2\Big(F_1(Y_1), F_2(Y_2)\Big)}{c^2\Big(F_1(Y_1), F_2(Y_2)\Big)} \frac{\partial F_l(Y_l)}{\partial\boldsymbol{\beta}_l}\left(\frac{\partial F_l(Y_l)}{\partial\boldsymbol{\beta}_l}\right)^T\right)$$
$$- \int_0^\infty \int_0^\infty \frac{c_l^2\Big(F_1(y_1), F_2(y_2)\Big)}{c\Big(F_1(y_1), F_2(y_2)\Big)} \frac{\partial F_l(y_l)}{\partial\boldsymbol{\beta}_l}\left(\frac{\partial F_l(y_l)}{\partial\boldsymbol{\beta}_l}\right)^T f_1(y_1)\, f_2(y_2)\,\mathrm{d}y_1\mathrm{d}y_2, \tag{A.7}$$

for $l = 1, 2$.

Next we show that

$$\mathrm{E}\left(\frac{c_{12}\Big(F_1(Y_1), F_2(Y_2)\Big)}{c\Big(F_1(Y_1), F_2(Y_2)\Big)} \frac{\partial F_1(Y_1)}{\partial\boldsymbol{\beta}_1}\left(\frac{\partial F_2(Y_2)}{\partial\boldsymbol{\beta}_2}\right)^T\right) =) \tag{A.8}$$
$$\int_0^\infty \int_0^\infty c\Big(F_1(y_1), F_2(y_2)\Big) \frac{\partial f_1(y_1)}{\partial\boldsymbol{\beta}_1}\left(\frac{\partial f_2(y_2)}{\partial\boldsymbol{\beta}_2}\right)^T \mathrm{d}y_1\mathrm{d}y_2.$$

The expectaion on the l.h.s. of (A.8) equals

$$\int_0^\infty \int_0^\infty c_{12}\Big(F_1(y_1), F_2(y_2)\Big) \frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} \left(\frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2}\right)^T f_1(y_1)\, f_2(y_2)\, \mathrm{d}y_1 \mathrm{d}y_2.$$

Writing

$$c_{12}\Big(F_1(y_1), F_2(y_2)\Big) \frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} f_1(y_1) = \frac{\partial}{\partial \boldsymbol{\beta}_1}\Big[c_2\Big(F_1(y_1), F_2(y_2)\Big) f_1(y_1)\Big]$$
$$- c_2\Big(F_1(y_1), F_2(y_2)\Big) \frac{\partial f_1(y_1)}{\partial \boldsymbol{\beta}_1},$$

the last double integral rewrites as

$$\int_0^\infty \left\{ \int_0^\infty \Big[\frac{\partial}{\partial \boldsymbol{\beta}_1}\Big[c_2\Big(F_1(y_1), F_2(y_2)\Big) f_1(y_1)\Big] - c_2\Big(F_1(y_1), F_2(y_2)\Big)\frac{\partial f_1(y_1)}{\partial \boldsymbol{\beta}_1}\Big] \mathrm{d}y_1 \right\}$$
$$\left(\frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2}\right)^T f_2(y_2)\, \mathrm{d}y_2.$$

Now for any fixed y_2,

$$\int_0^\infty \frac{\partial}{\partial \boldsymbol{\beta}_1}\Big[c_2\Big(F_1(y_1), F_2(y_2)\Big) f_1(y_1)\Big] \mathrm{d}y_1 = \frac{\partial}{\partial \boldsymbol{\beta}_1} \int_0^1 c_2\Big(r, F_2(y_2)\Big) \mathrm{d}r = \mathbf{0},$$

since the last integral does not depend on $\boldsymbol{\beta}_1$. We have obtained that the expectation on the l.h.s. of (A.8) is equal to

$$-\int_0^\infty \int_0^\infty c_2\Big(F_1(y_1), F_2(y_2)\Big) \frac{\partial f_1(y_1)}{\partial \boldsymbol{\beta}_1} \left(\frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2}\right)^T f_2(y_2)\, \mathrm{d}y_1 \mathrm{d}y_2$$
$$= -\int_0^\infty \frac{\partial f_1(y_1)}{\partial \boldsymbol{\beta}_1} \left\{ \int_0^\infty c_2\Big(F_1(y_1), F_2(y_2)\Big) \left(\frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2}\right)^T f_2(y_2)\, \mathrm{d}y_2 \right\} \mathrm{d}y_1.$$

Writing

$$c_2\Big(F_1(y_1), F_2(y_2)\Big) \left(\frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2}\right)^T f_2(y_2) =$$
$$\left(\frac{\partial}{\partial \boldsymbol{\beta}_2}\Big[c\Big(F_1(y_1), F_2(y_2)\Big) f_2(y_2)\Big]\right)^T - c\Big(F_1(y_1), F_2(y_2)\Big) \left(\frac{\partial f_2(y_2)}{\partial \boldsymbol{\beta}_2}\right)^T,$$

and observing that for any fixed y_1

$$\int_0^\infty \frac{\partial}{\partial \boldsymbol{\beta}_2}\Big[c\Big(F_1(y_1), F_2(y_2)\Big) f_2(y_2)\Big] \mathrm{d}y_2 = \frac{\partial}{\partial \boldsymbol{\beta}_2} \int_0^\infty c\Big(F_1(y_1), s\Big) \mathrm{d}s = \mathbf{0},$$

we get (A.8). From (A.4) and (A.8) we get

$$
\mathrm{E}\left(\frac{\partial^2 \log c\big(F_1(Y_1), F_2(Y_2)\big)}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}_2}\right) = \tag{A.9}
$$

$$
\int_0^\infty \int_0^\infty \Bigg[c\big(F_1(y_1), F_2(y_2)\big) \frac{\partial f_1(y_1)}{\partial \boldsymbol{\beta}_1} \frac{\partial f_2(y_2)}{\partial \boldsymbol{\beta}_2}
$$

$$
- \frac{c_1\big(F_1(y_1), F_2(y_2)\big) c_2\big(F_1(y_1), F_2(y_2)\big)}{c\big(F_1(y_1), F_2(y_2)\big)} \frac{\partial F_1(y_1)}{\partial \boldsymbol{\beta}_1} \frac{\partial F_2(y_2)}{\partial \boldsymbol{\beta}_2} f_1(y_1) f_2(y_2) \Bigg] \mathrm{d}y_1 \mathrm{d}y_2.
$$

Observing that

$$
\frac{\partial F_l(y_l)}{\partial \boldsymbol{\beta}_l} = \frac{\partial F_l(y_l)}{\partial \gamma_l} \frac{\partial \gamma_l}{\partial \boldsymbol{\beta}_l}, \quad \frac{\partial f_l(y_l)}{\partial \boldsymbol{\beta}_l} = \frac{\partial f_l(y_l)}{\partial \gamma_l} \frac{\partial \gamma_l}{\partial \boldsymbol{\beta}_l}, \tag{A.10}
$$

$$
\text{and} \quad \frac{\partial \gamma_l}{\partial \boldsymbol{\beta}_l} = \gamma_l(x_l) \big(1, x_l\big)^T, \quad l = 1, 2, \tag{A.11}
$$

formulas (6.25) in Subsection 6.3.2 follow from (A.7) and (A.9).

The derivatives $\frac{\partial f_l(y_l)}{\partial \gamma_l}$ and $\frac{\partial F_l(y_l)}{\partial \gamma_l}$ are given by

$$
\frac{\partial}{\partial \gamma_l} f_l(y_l) = \frac{\delta_l \Big(\log(y_l) - \log(\nu_l) - \psi(\kappa_l) \Big)}{\Gamma(\kappa_l) \nu_l^{\kappa_l}}, \tag{A.12}
$$

where $\kappa_l = \gamma_l(x_l) \Delta$ and $\delta_l = \exp\Big(-y_l/\nu_l\Big) y_l^{\kappa_l - 1} \Delta$, and

$$
\frac{\partial}{\partial \gamma_l} F_l(y_l) = \frac{\partial}{\partial \gamma_l} \frac{\tilde{\Gamma}(\kappa_l, y_l/\nu_l)}{\Gamma(\kappa_l)} = \Delta\Bigg(- \Gamma(\kappa_l)(y_l/\nu_l)^{\kappa_l}
$$

$$
{}_2\tilde{F}_2(\kappa_l, \kappa_l; \kappa_l + 1, \kappa_l + 1; -y_l/\nu_l) - \psi(\kappa_l) \frac{\tilde{\Gamma}(\kappa_l, y_l/\nu_l)}{\Gamma(\kappa_l)} \tag{A.13}
$$

$$
- \exp(y_l/\nu_l) \frac{\Gamma(\kappa_l, y_l/\nu_l, 0)}{\Gamma(\kappa_l)} \Bigg),
$$

such that $\psi(\kappa) = \frac{\partial}{\partial \kappa} \ln(\Gamma(\kappa))$ indicates the digamma function, $\Gamma(s, z, 0) = \Gamma(s, z) - \Gamma(s)$, $\tilde{\Gamma}(\kappa_l, y_l/\nu_l)$ refers to the lower incomplete Gamma function, and ${}_2\tilde{F}_2$ denotes the regularized hypergeometric function which is a slight modification of the generalized hypergeometric function ${}_2F_2(\kappa, \kappa; \kappa + 1, \kappa + 1; -y/\nu)$ and given by

$$
{}_2\tilde{F}_2(\kappa, \kappa; \kappa + 1, \kappa + 1; -y/\nu) = \frac{1 + \sum_{k=1}^\infty \left(\frac{\kappa}{\kappa+k}\right)^2 \frac{(-y/\nu)^k}{k!}}{\Gamma(\kappa + 1)^2}.
$$

Bibliography

Abdel-Hameed, M. (1975). A gamma wear process. *IEEE transactions on Reliability*, **24**,2, 152–153.

Adegbola, A. and Yuan, X.-X. (2019). A Multivariate Gamma Process for Dependent Degradation Modelling and Life Phenomena. *13th International Conference on Applications of Statistics and Probability in Civil Engineering(ICASP13), Seoul, South Korea, May 26-30, 2019*.

AghaKouchak, A., Bárdossy, A., and Habib, E. (2010). Copula-based uncertainty modelling: application to multisensor precipitation estimates. *Hydrological Processes*, **24**,15, 2111–2124.

Atkinson, A., Donev, A., and Tobias, R. (2007). *Optimum experimental designs, with SAS*. **34**, Oxford University Press.

Azrulhisham, E. A., Mohamad, W. M. W., and Hamid, H. F. A. (2013). *Inverse Power Law Model for Operative Life Estimation of Carbon Steel Stub Axle*. Tech. rep. SAE Technical Paper.

Balakrishnan, N. and Qin, C. (2019). Nonparametric evaluation of the first passage time of degradation processes. *Applied Stochastic Models in Business and Industry*, **35**,3, 571–590.

Boulanger, M. and Escobar, L. A. (1994). Experimental design for a class of accelerated degradation tests. *Technometrics*, **36**,3, 260–272.

Carey, M. and Tortorella, M (1988). Analysis of degradation data applied to MOS devices. In: *6th International Conference on Reliability and Maintainability, Strasbourg, France*.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics*, 586–602.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**,1, 141–151.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B*, **34**,2, 187–202.

Debusho, L. K. and Haines, L. M. (2008). V- and D-optimal population designs for the simple linear regression model with a random intercept term. *Journal of Statistical Planning and Inference*, **138**,4, 1116–1130.

Deng, A., Chen, X., Zhang, C., and Wang, Y. (2007). A comprehensive review of accelerated degradation testing. *Acta Armamentarii*, **28**,8, 1002–1007.

Denman, N., McGree, J. M., Eccleston, J. A., and Duffull, S. B. (2011). Design of experiments for bivariate binary responses modelled by Copula functions. *Computational statistics & data analysis*, **55**,4, 1509–1520.

Dette, H., Pepelyshev, A., and Holland-Letz, T. (2010). Optimal designs for random effect models with correlated errors with applications in population pharmacokinetics. *The Annals of Applied Statistics*, **4**,3, 1430–1450.

Dette, H., Pepelyshev, A., and Zhigljavsky, A. (2008). Improving updating rules in multiplicative algorithms for computing D-optimal designs. *Computational Statistics & Data Analysis*, **53**,2, 312–320.

Dorfleitner, G. and Klein, T. (1999). Rounding with multiplier methods: An efficient algorithm and applications in statistics. *Statistical Papers*, **40**,2, 143–157.

Dror, H. A. and Steinberg, D. M. (2006). Robust experimental design for multivariate generalized linear models. *Technometrics*, **48**,4, 520–529.

Duan, F. and Wang, G. (2018). Optimal step-stress accelerated degradation test plans for inverse Gaussian process based on proportional degradation rate model. *Journal of Statistical Computation and Simulation*, **88**,2, 305–328.

Duan, F. and Wang, G. (2019). Optimal design for constant-stress accelerated degradation test based on gamma process. *Communications in Statistics-Theory and Methods*, **48**,9, 2229–2253.

Elfving, G. (1952). Optimum Allocation in Linear Regression Theory. *The Annals of Mathematical Statistics*, **23**,2, 255–262.

Elsayed, E. A. (2012). *Reliability engineering*. **88**, John Wiley & Sons.

Entholzner, M., Benda, N., Schmelter, T., and Schwabe, R. (2005). A note on designs for estimating population parameters. *Biometrical Letters – Listy Biometryczne*, **42**, 25–41.

Fang, G., Pan, R., and Hong, Y. (2020). Copula-based reliability analysis of degrading systems with dependent failures. *Reliability Engineering & System Safety*, **193**, 106618.

Filipiak, K., Markiewicz, A., and Szczepańska, A. (2009). Optimal designs under a multivariate linear model with additional nuisance parameters. *Statistical Papers*, **50**,4, 761–778.

Ford, I., Torsney, B, and Wu, C. J. (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *Journal of the Royal Statistical Society, Series B*, **54**,2, 569–583.

Frank, M. J. (1979). On the simultaneous associativity ofF (x, y) andx+y- F (x, y). *Aequationes mathematicae*, **19**,1, 194–226.

Gebraeel, N. Z. and Lawley, M. A. (2008). A neural network degradation model for computing and updating residual life distributions. *IEEE Transactions on Automation Science and Engineering*, **5**,1, 154–163.

Gertsbackh Kordonskiy, e. a. (1969). Model of Failure (English Translation From the Russian Version). *Springer-Verlag*, 172–173.

Graßhoff, U., Doebler, A., Holling, H., and Schwabe, R. (2012). Optimal design for linear regression models in the presence of heteroscedasticity caused by random coefficients. *Journal of Statistical Planning and Inference*, **142**,5, 1108–1113.

Guan, Q., Tang, Y., and Xu, A. (2016). Objective Bayesian analysis accelerated degradation test based on Wiener process models. *Applied Mathematical Modelling*, **40**,4, 2743–2755.

Gumbel, E. J. (1960). Distributions des valeurs extremes en plusiers dimensions. *Publ. Inst. Statist. Univ. Paris*, **9**, 171–173.

Guo, J. and Li, Z. (2017). Bivariate Gamma Processes for Modeling Lithium Ion Battery Aging Mechanism. In: *IIE Annual Conference. Proceedings*. Institute of Industrial and Systems Engineers (IISE), 139–144.

Haghighi, F. and Bae, S. J. (2015). Reliability Estimation from Linear Degradation and Failure Time Data With Competing Risks Under a Step-Stress Accelerated Degradation Test. *IEEE Transactions on Reliability*, **64**,3, 960–971.

Haghighi, F. (2014). Accelerated test planning with independent competing risks and concave degradation path. *International Journal of Performability Engineering*, **10**,1, 15–22.

Hong, L. and Ye, Z. (2017). When is acceleration unnecessary in a degradation test? *Statistica Sinica*, 1461–1483.

Jiang, R and Jardine, A. K. (2008). Health state evaluation of an item: A general framework and graphical representation. *Reliability Engineering & System Safety*, **93**,1, 89–99.

Jin, G., Matthews, D. E., and Zhou, Z. (2013). A Bayesian framework for on-line degradation assessment and residual life prediction of secondary batteries inspacecraft. *Reliability Engineering & System Safety*, **113**, 7–20.

Kahneman, D. and Tversky, A. (1986). Rational choice and the framing of decisions. *Journal of business*, **59**,4, 251–278.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, **12**, 363–366.

Kiefer, J. and Wolfowitz, J. (1964). Optimum extrapolation and interpolation designs, I. *Annals of the Institute of Statistical Mathematics*, **16**,1, 79–108.

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society, Series B*, **21**,2, 272–304.

Koucký, M. (2003). Exact reliability formula and bounds for general k-out-of-n systems. *Reliability Engineering & System Safety*, **82**,2, 229–231.

Krantz, S. G. and Parks, H. R. (2012). *The Implicit Function Theorem: History, Theory, and Applications*. Springer.

Li, H., Pan, D., and Chen, C. P. (2014). Reliability modeling and life estimation using an expectation maximization based Wiener degradation model for momentum wheels. *IEEE transactions on cybernetics*, **45**,5, 969–977.

Li, Q. and Kececioglu, D. B. (2006). Design of an optimal plan for an accelerated degradation test: a case study. *International Journal of Quality & Reliability Management*, **23**,4, 426–440.

Lim, H. (2015). Optimum Accelerated Degradation Tests for the Gamma Degradation Process Case under the Constraint of Total Cost. *Entropy*, **17**,5, 2556–2572. ISSN: 1099-4300.

Lim, H., Kim, Y. S., Bae, S. J., and Sung, S.-I. (2019). Partial accelerated degradation test plans for Wiener degradation processes. *Quality Technology & Quantitative Management*, **16**,1, 67–81.

Limon, S., Yadav, O. P., and Liao, H. (2017). A literature review on planning and analysis of accelerated testing for reliability assessment. *Quality and Reliability Engineering International*, **33**,8, 2361–2383.

Liu, L., Li, X.-Y., Zio, E., Kang, R., and Jiang, T.-M. (2017). Model uncertainty in accelerated degradation testing analysis. *IEEE Transactions on Reliability*, **66**,3, 603–615.

Liu, Z., Ma, X., Yang, J., and Zhao, Y. (2014). Reliability modeling for systems with multiple degradation processes using inverse Gaussian process and copulas. *Mathematical Problems in Engineering*, **2014**,

Lu, C. J. and Meeker, W. O. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, **35**,2, 161–174.

Luo, H., Baker, N., Iannuzzo, F., and Blaabjerg, F. (2017). Die degradation effect on aging rate in accelerated cycling tests of SiC power MOSFET modules. *Microelectronics Reliability*, **76**, 415–419.

Markiewicz, A and Szczepańska, A (2007). Optimal designs in multivariate linear models. *Statistics & Probability Letters*, **77**,4, 426–430.

Meeker, W. Q., Escobar, L. A., and Lu, C. J. (1998). Accelerated degradation tests: modeling and analysis. *Technometrics*, **40**,2, 89–99.

Mireh, S., Khodadadi, A., and Haghighi, F. (2019). Copula-based reliability analysis of gamma degradation process and Weibull failure time. *International Journal of Quality & Reliability Management*, **36**,5, 654–668.

Mukhopadhyay, S. and Khuri, A. (2008). Comparison of designs for multivariate generalized linear models. *Journal of Statistical Planning and Inference*, **138**,1, 169–183.

Müller, W. G. (2007). *Collecting Spatial Data : Optimum Design of Experiments for Random Fields. 3rd rev. and extended ed.* Springer.

Näther, W. (1985). *Effective Observation of Random Fields.* Teubner.

Nelsen, R. B. (2007). *An introduction to copulas.* Springer Science & Business Media.

Nelson, W. B. (2005). A bibliography of accelerated test plans. *IEEE Transactions on Reliability*, **54**,2, 194–197.

Nelson, W. B. (2009). *Accelerated testing: statistical models, test plans, and data analysis.* John Wiley & Sons.

Nelson, W. (1990). Accelerated life testing: statistical models, data analysis and test plans. *Accelerated life testing: statistical models data analysis and test plants.*

Palayangoda, L. K., Ng, H. K. T., and Butler, R. W. (2020). Improved techniques for parametric and nonparametric evaluations of the first-passage time for degradation processes. *Applied Stochastic Models in Business and Industry*, **36**,4, 730–753.

Pan, Z. and Balakrishnan*, N (2010). Multiple-steps step-stress accelerated degradation modeling based on Wiener and gamma processes. *Communications in Statistics-Simulation and Computation*, **39**,7, 1384–1402.

Pan, Z., Balakrishnan*, N., and Sun, Q. (2011). Bivariate Constant-Stress Accelerated Degradation Model and Inference. *Communications in Statistics - Simulation and Computation*, **40**,2, 247–257.

Pan, Z. and Balakrishnan, N. (2011). Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. *Reliability Engineering & System Safety*, **96**,8, 949–957.

Pan, Z. and Sun, Q. (2014). Optimal Design for Step-Stress Accelerated Degradation Test with Multiple Performance Characteristics Based on Gamma Processes. *Communications in Statistics - Simulation and Computation*, **43**,2, 298–314.

Pan, Z., Balakrishnan, N., Sun, Q., and Zhou, J. (2013). Bivariate degradation analysis of products based on Wiener processes and copulas. *Journal of Statistical Computation and Simulation*, **83**,7, 1316–1329.

Pang, Z., Si, X., Hu, C., Zhang, J., and Pei, H. (2020). A review on modeling and analysis of accelerated degradation data for reliability assessment. *Microelectronics Reliability*, **107**, 113602.

Pascual, F. (2007). Accelerated Life Test Planning With Independent Weibull Competing Risks With Known Shape Parameter. *IEEE Transactions on Reliability*, **56**,1, 85–93.

Peng, C.-Y. (2015). Inverse Gaussian processes with random effects and explanatory variables for degradation data. *Technometrics*, **57**,1, 100–111.

Peng, C.-Y. and Tseng, S.-T. (2010). Progressive-stress accelerated degradation test for highly-reliable products. *IEEE Transactions on Reliability*, **59**,1, 30–37.

Peng, W., Li, Y.-F., Yang, Y.-J., Zhu, S.-P., and Huang, H.-Z. (2016a). Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas. *IEEE Transactions on Reliability*, **65**,2, 624–639.

Peng, W., Li, Y.-F., Mi, J., Yu, L., and Huang, H.-Z. (2016b). Reliability of complex systems under dynamic conditions: A Bayesian multivariate degradation perspective. *Reliability Engineering & System Safety*, **153**, 75–87.

Perrone, E. and Müller, W. G. (2016). Optimal designs for copula models. *Statistics*, **50**,4, 917–929.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Sahm, M. and Schwabe, R. (2001). A note on optimal bounded designs. In: *Optimum Design 2000*. Kluwer, 131–140.

Schmelter, R. S. and Schwabe, R. (2008). On optimal designs in random intercept models. *Tatra Mt. Math. Publ*, **39**, 145–153.

Schmelter, T. (2007). The Optimality of Single-group Designs for Certain Mixed Models. *Metrika*, **65**,2, 183–193.

Schmelter, T., Benda, N., and Schwabe, R. (2007). Some curiosities in optimal designs for random slopes. In: *mODa 8 - Advances in Model-Oriented Design and Analysis*. Physica, 189–195.

Schwabe, R. (1996a). *Optimum Designs for Multi-factor Models*. Springer.

Schwabe, R. and Schmelter, T. (2008). On optimal designs in random intercept models. *Tatra Mt. Math. Publ*, **39**, 145–153.

Schwabe, R. (1996b). Optimal design of experiments for multivariate re-sponse in two-factor linear models. In: *Multidimensional Statistical Analysis and Theory of Random Matrices. Proceedings of the Sixth Eugene Lukacs Symposium, Bowling Green, OH. VSP, Utrecht*, 235–242.

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.

Shat, H. (2021). Optimal Design of Stress Levels in Accelerated Degradation Testing for Multivariate Linear Degradation Models. *arXiv preprint arXiv:2106.09379*.

Shat, H. and Gaffke, N. (2021). Optimal Accelerated Degradation Testing Based on Bivariate Gamma Process with Dependent Components. *arXiv preprint arXiv:2106.13540*.

Shat, H. and Schwabe, R. (2019). Experimental Designs for Accelerated Degradation Tests Based on Gamma Process Models. *arXiv preprint arXiv:1912.04202*.

Shat, H. and Schwabe, R. (2021). Experimental Designs for Accelerated Degradation Tests Based on Linear Mixed Effects Models. *arXiv preprint arXiv:2102.09446.*

Shen, L., Sun, D., Ye, Z., and Zhao, X. (2017). Inference on an adaptive accelerated life test with application to smart-grid data-acquisition-devices. *Journal of Quality Technology*, **49**,3, 191–212.

Shi, Y. and Meeker, W. Q. (2014). "Planning Accelerated Destructive Degradation Tests with Competing Risks". In: *Statistical Models and Methods for Reliability and Survival Analysis.* Wiley. Chap. 22, 335–356. ISBN: 9781118826805.

Shiomi, H and Yanagisawa, T (1979). On distribution parameter during accelerated life test for a carbon film resistor. *Bulletin of the Electrotechnical Laboratory*, **43**, 330–345.

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.

Silvey, S. (1980). *Optimal design: an introduction to the theory for parameter estimation.* **1**, Chapman and Hall, London.

Silvey, S., Titterington, D., and Torsney, B (1978). An algorithm for optimal designs on a design space. *Communications in Statistics-Theory and Methods*, **7**,14, 1379–1389.

Sklar, M (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, **8**, 229–231.

Srivastava, P. W. and Gupta, T. (2017). Optimum modified ramp-stress ALT plan with competing causes of failure. *International Journal of Quality & Reliability Management.*

Sun, F., Wang, N., Li, X., and Cheng, Y. (2018). A time-varying copula-based prognostics method for bivariate accelerated degradation testing. *Journal of Intelligent & Fuzzy Systems*, **34**,6, 3707–3718.

Sun, Q., Zhou, J., Zhong, Z., Zhao, J., and Duan, X. (2004). Gauss-Poisson joint distribution model for degradation failure. *IEEE transactions on plasma science*, **32**,5, 1864–1868.

Torsney, B. (1983). A moment inequality and monotonicity of an algorithm. In: *Semi-infinite programming and applications.* Springer, 249–260.

Torsney, B. and Martín-Martín, R. (2009). Multiplicative algorithms for computing optimum designs. *Journal of Statistical Planning and Inference*, **139**,12, 3947–3961.

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.-R., Sung, W.-Y., Lio, Y., Chang, S. I., and Lu, J.-C. (2015). Optimal two-variable accelerated degradation test plan for gamma degradation processes. *IEEE Transactions on Reliability*, **65**,1, 459–468.

Wang, F.-K. and Chu, T.-P. (2012). Lifetime predictions of LED-based light bars by accelerated degradation test. *Microelectronics Reliability*, **52**,7, 1332–1336.

Wang, H., Xu, T., and Mi, Q. (2015). Lifetime prediction based on Gamma processes from accelerated degradation data. *Chinese Journal of Aeronautics*, **28**,1, 172–179.

Wang, H., Wang, G.-j., and Duan, F.-j. (2016). Planning of step-stress accelerated degradation test based on the inverse Gaussian process. *Reliability Engineering & System Safety*, **154**, 97–105.

Wang, X., Balakrishnan, N., Guo, B., and Jiang, P. (2015a). Residual life estimation based on bivariate non-stationary gamma degradation process. *Journal of Statistical Computation and Simulation*, **85**,2, 405–421.

Wang, Y. and Pham, H. (2011). Modeling the dependent competing risks with multiple degradation processes and random shock using time-varying copulas. *IEEE Transactions on Reliability*, **61**,1, 13–22.

Wang, Y., Zhang, C., Zhang, S., Chen, X., and Tan, Y. (2015b). Optimal design of constant stress accelerated degradation test plan with multiple stresses and multiple degradation measures. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, **229**,1, 83–93.

Wang, Z.-b., Zhai, G.-f., Huang, X.-y., and Yi, D.-x. (2013). Combination forecasting method for storage reliability parameters of aerospace relays based on grey-artificial neural networks. *International Journal of Innovative Computing, Information and Control*, **9**,9, 3807–3816.

Weaver, B. P. and Meeker, W. Q. (2013). Methods for planning repeated measures accelerated degradation tests. *Applied Stochastic Models in Business and Industry*, **30**,6, 658–671.

Wen, D., Wu, Z., Tang, Y., Li, M., and Qiang, Z. (2018). Accelerated degradation of sulfamethazine in water by VUV/UV photo-Fenton process: Impact of sulfamethazine concentration on reaction mechanism. *Journal of hazardous materials*, **344**, 1181–1187.

Whitmore, G. and Schenkelberg, F. (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime data analysis*, **3**,1, 27–45.

Wu, C.-F. and Wynn, H. P. (1978). The convergence of general step-length algorithms for regular optimum design criteria. *The Annals of Statistics*, 1273–1285.

Wu, J., Deng, C., Shao, X., and Xie, S. (2009). A reliability assessment method based on support vector machines for CNC equipment. *Science in China Series E: Technological Sciences*, **52**,7, 1849–1857.

Ye, Z.-S. and Chen, N. (2014). The inverse Gaussian process as a degradation model. *Technometrics*, **56**,3, 302–311.

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.

Ye, Z.-S., Chen, L.-P., Tang, L. C., and Xie, M. (2014). Accelerated degradation test planning using the inverse Gaussian process. *IEEE Transactions on Reliability*, **63**,3, 750–763.

Yu, Y. et al. (2010). Monotonic convergence of a general algorithm for computing optimal designs. *The Annals of Statistics*, **38**,3, 1593–1606.

Zhai, Q., Chen, P., Hong, L., and Shen, L. (2018). A random-effects Wiener degradation model based on accelerated failure time. *Reliability Engineering & System Safety*, **180**, 94–103.

Zhang, X. (2013). "Experiment Design and Reliability Analysis of Accelerated Degradation Test". PhD thesis. University of Cincinnati.

Zhao, W. and Elsayed, E. A. (2005). A general accelerated life model for step-stress testing. *Iie Transactions*, **37**,11, 1059–1069.

Zhou, J., Pan, Z., and Sun, Q. (2010). Bivariate degradation modeling based on gamma process. In: *Proceedings of the World Congress on Engineering.* **3**,

Zyskind, G (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Annals of Mathematical Statistics*, **38**,4, 1092–1109.

List of Abbreviations

RDT	Reliability Demonstration Test
ALT	Accelerated Life Test
ADT	Accelerated Degradation Test
ADDT	Accelerated Destructive Degradation Test
CSADT	Constant Stress Accelerated Degradation Test
SSADT	Step Stress Accelerated Degradation Test
PSADT	Progressive Stress Accelerated Degradation Test
BMA	Bayesian Model Averaging
LMEM	Linear Mixed Effects Model
IG	Inverse Gaussian
GP	Gamma Process
RUL	Remaining Useful Life
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimator
GLM	Generalized Linear Model

List of Symbols

Y	univariate random variable
y	univariate observation (value of Y)
$\boldsymbol{Y}$	multivariate random variable
$\boldsymbol{y}$	multivariate observations vector (value of $\boldsymbol{Y}$)
t	time variable
$\boldsymbol{t}$	time plan/vector of observation times
Δ	time interval
T	failure time random variable
F_T	failure time distribution function
y_0, z_0	predetermined failure threshold
R_T	survival probability function
n	number of units
k	number of measurement times
$C(u_1, ..., u_n)$	copula function
$\varkappa$	copula dependence parameter
Z_t	stochastic process
$B(t)$	standard Brownian motion
$\gamma(x)$	rate function of gamma process at x
ν	scale parameter of gamma process
$\mathrm{var}(Y)$	variance of Y
Σ	variance covariance matrix of random effect parameters
ς	vector of variance covariance parameters
$\mathbf{V}$	overall variance covariance matrix
ε	random error
$\boldsymbol{f}$	vector of regression functions
f	component of $\boldsymbol{f}$
$\boldsymbol{\beta}$	marginal vector of model parameters
p	dimension of model parameter $\boldsymbol{\beta}$
$\boldsymbol{\theta}$	overall vector of model parameters
q	dimension of model parameter $\boldsymbol{\theta}$
$\ell(\boldsymbol{\theta})$	log likelihood function

$\mathcal{X}$	experimental (design) region
x	experimental setting, $x \in \mathcal{X}$
$\mathbb{R}^r$	r-dimensional set of real numbers
ξ	design
m	support size of ξ
w	weight of ξ at x
$\xi_{\mathbf{x}}$	one-point design of $\mathbf{x}$
ξ^*	optimal design
$\mathbf{x}^*$	design point of ξ^*
w^*	optimal weight of ξ^* at $\mathbf{x}^*$
$\boldsymbol{I}$	identity matrix
$\mathbf{1}$	vector of ones
$\mathbf{0}$	vector of zeros
$\mathbf{M}_{\boldsymbol{\theta}}(\mathbf{x})$	Fisher information matrix at one point $\mathbf{x}$
$\mathbf{M}_{\boldsymbol{\theta}}(\xi)$	Fisher information matrix of design ξ
$\boldsymbol{\Phi}$	optimality criterion function
$\boldsymbol{\Psi}(\xi, \xi_{\mathbf{x}})$	directional derivative of $\boldsymbol{\Phi}$ at ξ in the direction of $\xi_{\mathbf{x}}$
$\mathrm{tr}(\boldsymbol{A})$	trace of matrix $\boldsymbol{A}$
$\det(\boldsymbol{A})$	determinant of matrix $\boldsymbol{A}$
$\lambda(\xi, \boldsymbol{\theta})$	eigenvalue of $\mathbf{M}_{\boldsymbol{\theta}}(\xi)$
$\mathbf{F}$	design matrix
$\hat{\boldsymbol{\theta}}$	parameter estimators
$\otimes$	Kronecker product, product of designs
ρ	exponent/ correlation
$\mathrm{eff}(\xi; \boldsymbol{\beta})$	efficiency of design ξ at $\boldsymbol{\beta}$

List of Figures

147

List of Tables