

Semi-parametric Bayesian Inference of Accelerated Life Test Using Dirichlet Process
Mixture Model

A dissertation presented to
the faculty of
the Russ College of Engineering and Technology of Ohio University

In partial fulfillment
of the requirements for the degree
Doctor of Philosophy

Xi Liu

December 2015

© 2015 Xi Liu. All Rights Reserved.

This dissertation titled
Semi-parametric Bayesian Inference of Accelerated Life Test Using Dirichlet Process
Mixture Model

by

XI LIU

has been approved for
the Department of Industrial and Systems Engineering
and the Russ College of Engineering and Technology by

Tao Yuan

Associate Professor of Industrial and Systems Engineering

Dennis Irwin

Dean, Russ College of Engineering and Technology

ABSTRACT

LIU, XI, Ph.D., December 2015, Mechanical and Systems Engineering

Semi-parametric Bayesian Inference of Accelerated Life Test Using Dirichlet Process

Mixture Model (100 pp.)

Director of Dissertation: Tao Yuan

Accelerated life testing (ALT) is commonly used to estimate the reliability of highly reliable products. This dissertation develops statistical models to predict useful life of nano devices with data collected under constant-stress ALT and step-stress ALT. As an example of nano devices, nc-MoO_x embedded ZrHfO high-*k* dielectric thin film is studied with respect to its physical properties, failure mechanisms, and long-term stability. The devices used for ALT and reliability prediction demonstration have identical structure with this nc-MoO_x embedded device.

This research develops a semi-parametric Bayesian method to analyze ALT. The model assumes a log-linear lifetime-stress relationship, without assuming any parametric form of the failure-time distribution. The Dirichlet Weibull mixture model is employed to model the failure-time distribution under a given stress level. The model is fitted with a simulation-based algorithm, which implements Gibbs sampling to analyze ALT data and predicts the failure-time distribution at a normal stress level.

Two practical examples related to the reliability of nanoelectronic devices are presented for constant-stress ALT, including one right-censored data and one complete data set. One right-censored practical example is demonstrated for simple step-stress

ALT. All three examples illustrate the capability of the proposed methodology to provide accurate prediction of the failure-time distribution at a normal stress level.

DEDICATION

*This dissertation is dedicated to my dear father, Jianzhen Liu, and to my mother, Jinping
Deng.*

ACKNOWLEDGMENTS

Over the past five years of doctoral journey, I have received numerous encouragement and support from a great number of persons and institutions. My work cannot be completed without their support.

First and foremost, I would like to express my deepest appreciation to my advisor, Dr. Tao Yuan, for his continuous support and guidance throughout my graduate studies. I would never have been able to finish this dissertation without his persistent help and deep knowledge. I also would like to thank Dr. Yue Kuo, who has offered me the great opportunity to do tests in Thin Film Nano & Microelectronics Research Laboratory, Texas A&M University, for his support on this dissertation and for guiding me to the semiconductor devices' field.

Next, I would like to thank Dr. Andrew Snow, Dr. David Koonce and Dr. Diana Schwercha, for serving my committee members and offering many insightful comments and suggestions to my research.

I am deeply grateful to Dr. Chia-Han Yang and Dr. Chi-Chou Lin for fabricating samples, teaching me the device concepts and mentoring me on testing and measurements. Thanks Minghao Zhu for the XPS and XRD analyzes. Thanks Dr. Saleem Z. Ramadan and Dr. Wen Luo for supplying me data sets as my practical examples.

Lastly, I would like to thank my family, for their continuous support, inspiration and love.

This dissertation was partially supported by National Science Foundation project CMMI-0926420 and CMMI-0926379.

TABLE OF CONTENTS

	Page
Abstract.....	3
Dedication.....	5
Acknowledgments.....	6
List of Tables	9
List of Figures	10
Chapter 1: Introduction.....	11
1.1 Motivation and Objective.....	11
1.2 Reliability and Related Functions	12
1.3 Accelerated Life Testing.....	14
1.4 Acceleration Model.....	16
1.5 Step-Stress Accelerated Life Testing	17
1.6 Commonly Used Estimation Method.....	19
1.6.1 Parametric method.....	20
1.6.2 Nonparametric Bayesian method	22
1.6.3 Dirichlet process mixture model	24
1.7 Nanocrystals Embedded High- k Device.....	27
1.8 Dissertation Overview.....	28
Chapter 2: Literature Review.....	29
2.1 Inference of Constant-Stress Accelerated Life Testing.....	29
2.1.1 Parametric estimation of constant-stress ALT	29
2.1.2 Semi-parametric estimation of constant-stress ALT	33
2.2 Inference of Step-Stress Accelerated Life Testing.....	36
2.2.1 Parametric estimation of step-stress accelerated life testing.....	36
2.2.2 Semi-parametric estimation of step-stress accelerated life testing.....	40
2.3 Dirichlet Process Mixture Model	41
Chapter 3: Problem Statement	44
3.1 Notations	44
3.2 Assumptions	45

3.3	Problem Description.....	45
3.3.1	Investigation of metal oxide nanodots-embedded ZrHfO high-k film.....	46
3.3.2	Development of simulation-based algorithm	46
3.3.3	Comparison between parametric Weibull log-linear ALT model and DP Weibull mixture ALT model.....	46
Chapter 4: Memory Functions of Molybdenum Oxide Nanodots Embedded ZrHfO High- k^\dagger		47
4.1	Fabrication of Nanocrystals Embedded ZrHfO High- k Device	47
4.2	Physical Properties of Nanocrystals Embedded ZrHfO high- k Device	48
4.3	Charge Trapping and Detrapping Mechanisms.....	49
4.4	Charge Retention Capability	53
4.5	Conclusions	54
Chapter 5: Bayesian Analysis For Accelerated Life tests Using Dirichlet Process Weibull Mixture Model [†]		56
5.1	Methodologies.....	56
5.1.1	Dirichlet process Weibull mixture ALT model.....	56
5.1.2	Simulation-based model fitting	60
5.2	Illustrative Examples.....	69
5.2.1	Complete data set example.....	70
5.2.2	Right-censored data set example.....	74
5.3	Conclusions	76
Chapter 6: Bayesian Analysis For Simple Step-Stress Accelerated Life Testing		78
6.1	Methodologies.....	78
6.2	Illustrative Examples.....	81
Chapter 7: Conclusions.....		84
7.1	Memory Functions of MoO _x Nanodots Embedded ZrHfO High- k	84
7.2	Prediction of CDF of Failure-Time Distribution at Normal Stress Level.....	84
7.3	Future Research.....	85
References.....		87

LIST OF TABLES

	Page
Table 5.1: Example 1: times-to-breakdown of MOS capacitors tested at four electrical field stresses	70
Table 5.2: Example 2: times-to-breakdown of nanocrystals-embedded high- k memories tested at four voltage stresses	75
Table 6.1: Times-to-breakdown of nanocrystals-embedded high- k memories under simple SSALT	81

LIST OF FIGURES

	Page
Figure 1.1: The reliability function, CDF and PDF of the Weibull distribution with $\alpha = 1.5$ and $\lambda = 7$	13
Figure 1.2: Accelerated life testing concept	15
Figure 1.3: Step-stress accelerated life testing.....	18
Figure 1.4: Cross-sectional view of nanocrystals embedded ZrHfO capacitor	28
Figure 4.1: (a) XPS O 1s peak, and (b) XRD pattern of the MoO _x embedded ZrHfO sample	49
Figure 4.2: $C-V$ hysteresis curves of the nc-MoO _x embedded ZrHfO capacitor measured at 1MHz. The inset is the hysteresis curve of the control sample.....	51
Figure 4.3: $J-V$ curve of nc-MoO _x embedded capacitor V_g swept from -8 V to +8 V. The inset is the $J-V$ curve of control sample	52
Figure 4.4: Retention property of holes trapped in the nc-MoO _x embedded capacitor. The inset shows the extrapolation of the curve to 10 years projection.....	54
Figure 5.1: Example 1: predicted failure-time CDF at the normal stress $x_0 = 7.1$ MV/cm	72
Figure 5.2: Example 1: estimation of failure-time CDF at 7.9 MV/cm.....	74
Figure 5.3: Example 2: predicted failure-time CDF at the normal stress $x_0 = 7.1$ V	76
Figure 6.1: Trace plots of α^* , β , and λ^*	82
Figure 6.2: Predicted failure-time CDF at the normal stress $x_0 = 7.5$ V	83

CHAPTER 1: INTRODUCTION

1.1 Motivation and Objective

For the purpose of increasing competitive advantages, decreasing costs, and satisfying increasing customer expectations, manufacturers strive to design and produce highly reliable products. Consequently, the studies of reliability data analysis have been promoted. Quantitative methods to predict and assess product reliability are required to improve reliability of existing products and ensure the high reliability of new products [1]. The data obtained from life tests are commonly supplied to statistical method for reliability estimation [2]. Conventional lifetime data analysis assumes the failure time following certain parametric distributions such as exponential, Weibull, or log-normal and estimates the parameters of the distributions [3]. However, in some situations it may be difficult to choose the appropriate parametric model. For example, modern complex products may involve multiple failure modes and therefore simple lifetime distributions are not adequate to describe their failure mechanisms. For some new technologies such as nanotechnology, the failure mechanisms have not been well understood. Therefore, more flexible nonparametric methods are needed. In recent years, the development of Markov chain Monte Carlo (MCMC) methods for simulation-based implementation and analysis assures the feasibility of nonparametric data analysis [4].

This dissertation aims to assess the reliability under normal stress levels based on ALT using Bayesian models involving Dirichlet process mixture Models (DPMM). Both constant-stress ALT and step-stress ALT are studied in this work. The rest of this chapter

provides a brief background introduction of reliability and related functions, ALT, acceleration models, step-stress ALT, and commonly used estimation methods.

1.2 Reliability and Related Functions

The reliability is the probability that a system or component can perform its function under operation conditions for a specified period of time. It can be mathematically presented as:

$$R(t) = \Pr\{T \geq t\}, \quad (1.1)$$

where T is a random variable denoting the time to failure of a non-repairable device or the time to the first failure of a repairable device, and $R(t)$ is called the *reliability function* of the failure-time distribution.

Denote $F(t)$ as the *cumulative distribution function* (CDF) of the failure-time distribution, which is the probability that the lifetime is less than t , i.e.,

$$F(t) = 1 - R(t) = \Pr\{T < t\}. \quad (1.2)$$

If the CDF is a continuous function of t , a third function defined by the derivative of $F(t)$ is used to describe the shape of the failure-time distribution. It is called the *probability density function* (PDF), and can be expressed as:

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt}, \quad (1.3)$$

where $f(t)$ is a non-negative function.

Taking Weibull failure-time distribution as an example, its PDF, CDF, and reliability function can be written as:

$$f(t) = \frac{\alpha}{\lambda} t^{\alpha-1} \exp\left(-\frac{t^\alpha}{\lambda}\right), \quad (1.4)$$

$$F(t) = \int_0^t \frac{\alpha}{\lambda} t^{\alpha-1} \exp\left(-\frac{t^\alpha}{\lambda}\right) dt = 1 - e^{-t^\alpha/\lambda}, \quad (1.5)$$

and
$$R(t) = 1 - F(t) = e^{-t^\alpha/\lambda}. \quad (1.6)$$

where α and λ are the shape and scale parameters, respectively.

Figure 1.1 graphically shows these three functions when the time to failure follows the Weibull distribution with the shape parameter $\alpha = 1.5$ and the scale parameter $\lambda = 7$.

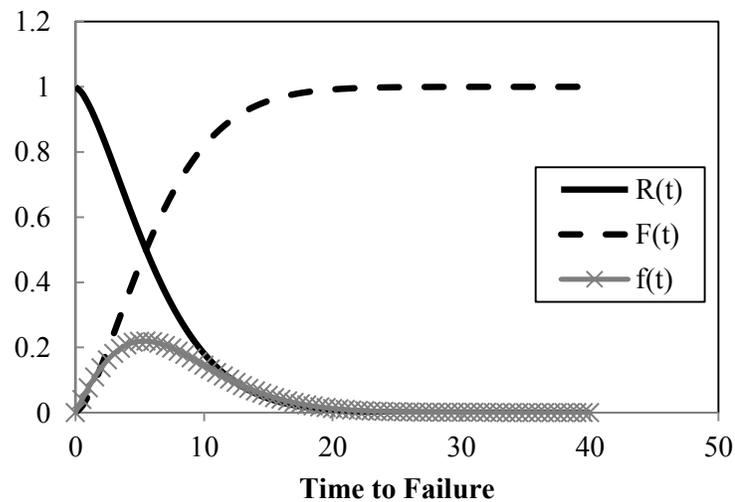


Figure 1.1. The reliability function, CDF and PDF of the Weibull distribution with $\alpha = 1.5$ and $\lambda = 7$.

Another important function, the *failure rate* or *hazard rate function*, is used to provide an instantaneous failure rate, which can be denoted as:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t \Pr\{T > t\}} = \frac{f(t)}{R(t)}. \quad (1.7)$$

When $h(t)$ is an increasing, constant, or decreasing function, the failure rate can be described as increasing (IFR), constant (CFR), or decreasing (DFR), respectively [5].

The mean time to failure (MTTF) and the median time to failure are commonly used as measures of the center of a failure-time distribution. The MTTF is the average length of time until failure, and can be defined as the expected value of T , i.e.,

$$\text{MTTF} = E(T) = \int_0^{\infty} t f(t) dt = \int_0^{\infty} R(t) dt. \quad (1.8)$$

For a given value of $p \in [0, 1]$, if t_p satisfies

$$F(t_p) = p, \quad (1.9)$$

then t_p is called the p th percentile of the lifetime distribution, which means $100p\%$ of the failures occur before time t_p . If $p=0.5$, then $t_{0.5}$ is the median time to failure. When the distribution is highly skewed, the median is preferentially used as the measure of the center location of a distribution. The median, $t_{0.5}$, as well as the lower and upper 25% percentiles, $t_{0.25}$ and $t_{0.75}$, respectively, are important characteristics of a lifetime distribution.

1.3 Accelerated Life Testing

Reliability life testing is carried out to obtain failure information for the purpose of quantifying reliability [5]. Highly reliable products, such as electronic devices, usually have long lifetimes. Therefore, only very few failures can be obtained within reasonable time period under actual operating condition and it is difficult to obtain adequate failure-time data for statistical methods. One approach to solve this problem is to use ALT, in which the units are placed under higher than operational stress conditions to speed up the failure occurrence [2]. Classical stress includes voltage, current, humidity, temperature,

pressure, cycling rate or load [6]. Then the failure-time data collected at elevated stress levels are analyzed and extrapolated to predict the reliability at the normal stress level through an acceleration model. ALT is based on the fundamental principles that the unit under test will have the same failure mechanisms in a short time at a high stress level as it exhibits in a longer time at a lower stress level [7]. Figure 1.2 shows the basic concept of an ALT estimation which uses lifetime data collected from a four-level single-stress test to estimate mean life under normal stress. The variable x_i denotes the stress level with x_0 representing the normal stress. According to Dasgupta and Pecht [8], there are four categories of failure mechanisms: stress-strength, damage-endurance, challenge-response, and tolerance-requirement. The ALT is the most appropriate with the damage-endurance failure and some cases of tolerance-requirement failure [7].

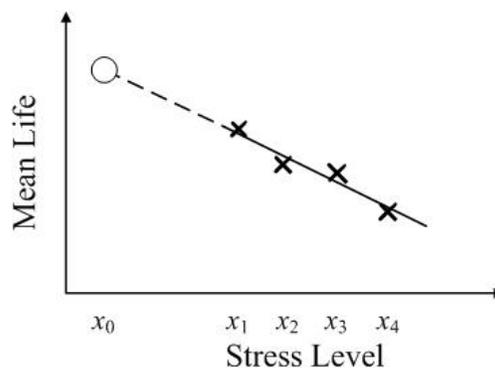


Figure 1.2. Accelerated life testing concept.

The constant-stress ALT and the step-stress ALT are the two typical types of accelerated life testing. In constant-stress ALT each single unit is placed only under one

higher than normal stress level, while the step-stress ALT (SSALT) allows several stress levels.

1.4 Acceleration Model

One difficulty of ALT data analysis is how to predict the reliabilities of units at the normal stress level from the failure-time data under higher stress levels. Basically, a functional relationship called the acceleration function is applied to describe the relationship between the lifetimes and the stress conditions [6]. The log-linear model which assumes the log-linear relation for the lifetime is widely used, it can be expressed as:

$$\ln \theta = a + b_1 x_1 + \dots + b_m x_m, \quad (1.10)$$

where θ denotes the stress related characteristic, and x_1, \dots, x_m denote the stress factors (or proper transformation of them). The log-linear model is used because it is simple and can be transformed from many other relationships.

For example, when the temperature primarily contributes to the failures, the Arrhenius model is commonly used:

$$r = A \exp \left\{ -\frac{E_a}{KT} \right\}, \quad (1.11)$$

where r is the reaction or process rate, A is constant, E_a represents the activation energy in electron volts, T is the absolute temperature($^{\circ}$ K), and K is the Boltzmann's constant, a known physical constant equals to 8.617×10^{-5} (eV/ $^{\circ}$ K) [9]. This equation can be transformed to the log-linear model as $\ln r = \ln A - (E_a/K) (1/T)$ with $x=1/T$.

The generalized Eyring model is applicable for failures related to two types of stresses, one thermal and one nonthermal stress. The simplest form of Eyring model can be presented as:

$$r = AT^\alpha \exp\left\{-\frac{E_a}{KT}\right\} \exp\left\{\left(B + \frac{C}{T}\right)s\right\}, \quad (1.12)$$

where r is the process rate, A , α , B , and C are constants, E_a is the activation energy, T is the absolute temperature(°K), K is the Boltzmann's constant, and s is the second stress [10]. When $\alpha=0$ and $C=0$, this equation can be transformed to the log-linear model as $\ln r = \ln A - (E_a / K)(1/T) + Bs$ with $x_1=1/T$ and $x_2=s$.

The E -model can be used to study the time-dependent dielectric breakdown of gate dielectric thin films where the stress is the electrical field. The E -model is expressed as:

$$\theta = \tau_L \exp[G_L(E_{bd}-E)], \quad (1.13)$$

where τ_L and G_L are unknown, temperature-dependent constants, E is the applied electrical field, and E_{bd} is the field above which breakdown occurs immediately [11]. The E -model can also be expressed as a log-linear function $\ln\theta = \ln\tau_L + G_L(E_{bd}-E)$.

In this study, the log-linear acceleration model will be used and the details will be introduced in section 5.1.1.

1.5 Step-Stress Accelerated Life Testing

Step-stress accelerated life testing (SSALT) can further ensure enough failures within reasonable time period by testing the units through more than one level of stress. In the SSALT, the stress levels change (usually increase) at pre-specified times (*time-step*

stress ALT) or after pre-specified numbers of failures (*failure-step stress ALT*). A test with only one change of stress is called a *simple step-stress ALT*, while a test with several stress changes is called a *multiple step-stress ALT* [6]. The number of failures in the time-step stress ALT is random at each level of stress. The time duration of each level of stress in the failure-step stress ALT is also random. If the test ends at a pre-specified time, it is called type-I censoring; if the test ends when a pre-specified number of failures has achieved, it is called type-II censoring. Figure 1.3 shows a multi-SSALT with four stresses and type-I censoring. The test ends at a pre-specified time t_c . All units that have not failed by t_c are censored. The x_i 's and τ_i 's are the stress levels and stress changing times, respectively.

In order to analyze SSALT data, a model describing the effect of changing stress is needed. Nelson [12] proposed the *cumulative exposure model*, which assumes that “the remaining life of specimens depends only on the current cumulative fraction failed and current stress- regardless of how the fraction accumulated.”

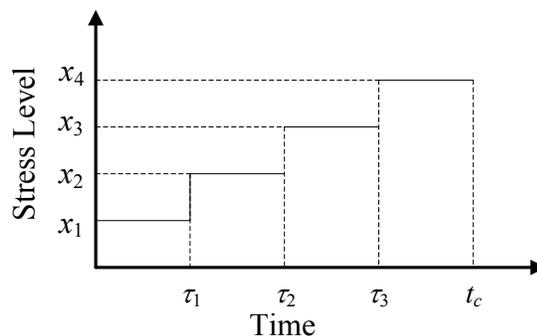


Figure 1.3. Step-stress accelerated life testing.

Denote F_i as the CDF of the failure-time distribution under stress x_i , F_i under a step-stress pattern expressed by the cumulative exposure model can be written as:

$$F_0(t) = \begin{cases} F_1(t), & 0 \leq t \leq \tau_1, \\ F_2(t - \tau_1 + u_1), & \tau_1 \leq t \leq \tau_2, \\ F_3(t - \tau_2 + u_2), & \tau_2 \leq t \leq \tau_3, \\ \dots & \\ \dots & \\ F_m(t), & \tau_m \leq t \leq \infty, \end{cases} \quad (1.14)$$

where τ_i is the time of changing the stress from the i th stress level to the $(i+1)$ th stress level, $F_i(t)$ is the CDF under the i th stress level, and u_i is the solution of

$$F_{i+1}(u_i) = F_i(\tau_i - \tau_{i-1} + u_{i-1}).$$

This dissertation studies the simple SSALT with two stress levels and assumes the failure-time distribution under each stress level following a Dirichlet process mixture model with Weibull kernel. The details of inference will be given in Chapter 6.

1.6 Commonly Used Estimation Method

The commonly used data estimation method for ALT analysis can be classified as parametrical and non-parametrical, depending on if there are parameters assumed in the model. Two widely applied parametric estimation groups are Maximum Likelihood Estimation and parametrical Bayesian estimation. While commonly used nonparametric estimation methods for ALT analysis are empirical method and nonparametrical Bayesian Inference. In addition, the Dirichlet Process Mixture Model is a popular nonparametric Bayesian estimation method frequently cited in literature.

1.6.1 Parametric method

The parametric method is one approach to performing ALT model inference. The parametric inference assumes the lifetime distribution under each stress level comes from the same parametric family and is preassigned with a theoretical distribution such as Weibull, exponential, or log-normal distribution. After that an acceleration model is chosen and the parameters are estimated. The Maximum Likelihood Estimation (MLE) is one of the most widely used parametric estimation method to estimate the model parameters given the sample data. Given n failure-time data $\mathbf{t} = (t_1 \dots t_n)$ collected from the test, the point estimator of parameters are obtained by maximizing the likelihood function $L(\Theta|\mathbf{t})$:

$$L(\Theta|\mathbf{t}) = \prod_{i=1}^n L_i(\Theta, t_i), \quad (1.15)$$

where L_i is the likelihood contribution of the i th observation and Θ is the parameter vector to be estimated. For the complete data set, $L_i = f(t_i|\Theta)$, and for rightly censored data, $L_i = R(t^*)$, where t^* is the censored time. The logarithm of likelihood function (*log-likelihood function*) instead of the likelihood function is usually maximized for computational convenience. In general, the maximum likelihood estimator is obtained by solving the following sets of equations:

$$\begin{aligned} \frac{\partial}{\partial \theta_1} \ln L(\theta_1, \mathbf{t}) &= 0, \\ \vdots & \\ \frac{\partial}{\partial \theta_k} \ln L(\theta_k, \mathbf{t}) &= 0, \end{aligned} \quad (1.16)$$

where k is the number of parameters, and θ is the estimated parameters .

The MLE method assumes unknown parameters as fixed. Therefore, in order to obtain precise estimation, a large sample size and accurate model assumptions are required. Another parametric approach to conducting ALT model inference, Bayesian inference, has been applied. Bayesian inference assumes the parameters are random and describes the uncertainties by a joint prior distribution. The prior distribution is formulated before data collection and is based on the historical data or experts' opinions. The main advantage of Bayesian inference is the ability of combining the collected data with any related information available for reliability analysis, which can relax the sample size limitation. The Bayesian data analysis estimates parameters using the posterior distribution $f(\Theta|\mathbf{t})$, which is obtained by incorporating its prior distribution $f(\Theta)$ and the likelihood function of data $L(\Theta)$. The prior distribution is updated after the data is collected according to the Bayes' theorem:

$$f(\Theta|\mathbf{t}) = \frac{L(\Theta|\mathbf{t})f(\Theta)}{f(\mathbf{t})}, \quad (1.17)$$

where

$$f(\mathbf{t}) = \int L(\Theta|\mathbf{t})f(\Theta), \quad (1.18)$$

which is called the preposterior marginal distribution of \mathbf{t} .

In many practical problems which are complex and involve more than one parameter, multiple levels of integration are necessary in Bayesian inference. Mostly these integrations are analytically intractable and therefore numerical methods are used instead. For example, the Markov chain Monte Carlo (MCMC) simulation has been widely used for numerical integrations. Generally, it "simulates a Markov chain in such a

way that the stationary distribution of the chain is the posterior distribution of the parameters,” and then uses the simulated data to compute Bayes estimation [13]. The Gibbs sampling, which is usually applied when it is difficult to directly sample from multivariate probability distribution is a type of MCMC simulation that is particularly useful in high dimensional problems. For example, when samples from $f(\theta_1, \theta_2|\mathbf{t})$ are needed, and it is difficult to sample directly from their marginal distributions $f(\theta_1|\mathbf{t})$ and $f(\theta_2|\mathbf{t})$, their conditional distributions $f(\theta_1|\mathbf{t}, \theta_2)$ and $f(\theta_2|\mathbf{t}, \theta_1)$ are sampled instead in each iteration. When the number of iterations is sufficiently large, the samples obtained from conditional distributions can be regarded as simulated observations sampled from their marginal distributions. In this study, the Gibbs sampling will be applied in the semi-parametric Bayesian methods for simulating posterior distributions.

Generally, the parametric reliability analysis is performed by fitting the lifetime data with a suitable parametric model. Usually, a failure distribution with parameter(s) Θ is assigned and then Θ is estimated based on the observed data. Some commonly used failure-time distributions are Weibull, exponential, or log-normal probability distribution. In the accelerated life testing, an acceleration relationship is also selected and then the parameters are estimated.

1.6.2 Nonparametric Bayesian method

The accuracy of conventional parametric estimation is based on the parametric assumptions that are assigned to the data, that is, the particular parametric family of distributions assumed. However, the failure mechanism of some products may be unknown and may involve multiple modes or steps which are impossible to model using

a simple lifetime distribution. Therefore, more flexible methods-nonparametric methods have been developed.

One class of nonparametric methods is empirical, which have no restrictive assumptions on the lifetime distributions and derive the reliability properties such as PDF and CDF directly from the data. Some commonly used empirical methods include Kaplan-Meier estimator and Median rank, which are shown in equations (1.19) and (1.20), respectively.

$$\hat{R}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{1}{n_j}\right), \quad (1.19)$$

where t_j is the ordered failure times and n_j is the number remaining at risk just prior to the j th failure.

$$\hat{F}(t_i) = \frac{i - 0.3}{n + 0.4}, \quad (1.20)$$

where i is the i th ordered failure and n is the sample size.

The nonparametric Bayesian inference is another group of nonparametric methods which has been proposed to estimate a probability distribution. The nonparametric Bayesian methods in the practical use are actually probability models with infinitely many parameters on function spaces [14]. In ALT analysis, the failure-time data under each stress level is not suggested by any standard model. Therefore it is distribution-free. Generally, a prior distribution on the class of all distribution functions is placed and the posterior distribution on the class of all distribution functions is obtained from data. The prior distributions for the underlying distribution functions constitute a stochastic process.

There are many nonparametric Bayesian methods for different applications, including Gaussian process (GP), spline models and DPMM, etc. [14]. These methods are widely used in statistical inference problems, such as density estimation, regression, and clustering. In this study, the DPMM is proposed to be used to estimate the ALT data.

Although the nonparametric methods are more flexible, when both parametric and nonparametric methods are applicable for a problem, the parametric method is preferred because of its efficiency and computational convenience [3].

1.6.3 Dirichlet process mixture model

The Dirichlet Process (DP) is “by far the most popular nonparametric model in the literature” [14]. The DP prior which was formally developed by Ferguson [15] is the first prior defined for spaces of distribution functions [16]. It fulfilled two desirable properties of prior distribution for nonparametric problems: large support of the prior distribution and analytically manageable posterior distributions [15].

The foundation of Dirichlet process is the Dirichlet distribution. Defining the probabilities of n discrete space $\chi = \{\chi_1, \dots, \chi_n\}$ are $\Theta = \{\theta_1, \dots, \theta_n\}$, i.e. $p(X = \chi_i) = \theta_i$. Then the PDF of Dirichlet distribution can be defined as:

$$p(\Theta | y_1, \dots, y_m) = \frac{1}{B(\mathbf{y})} \prod_{i=1}^n \theta_i^{y_i-1}, \quad (1.21)$$

where $B(\mathbf{y})$ is the normalizing constant expressed in terms of gamma function:

$$B(\mathbf{y}) = \frac{\prod_{i=1}^n \Gamma(y_i)}{\Gamma(\sum_{i=1}^n y_i)}. \quad (1.22)$$

Denote $\alpha = \sum_i y_i$ as the concentration parameter of χ , and $\mathbf{m} = \{m_1, \dots, m_n\} = x_i/\alpha$ as the base measure, the Dirichlet distribution can be expressed as:

$$p(\Theta | \alpha, M) = \frac{\Gamma(\alpha)}{\prod_{i=1}^n \Gamma(\alpha m_i)} \prod_{i=1}^n \theta_i^{\alpha m_i - 1}. \quad (1.23)$$

The concentration parameter shows how much the probability would be concentrated around \mathbf{m} . When $n=2$, the Dirichlet distribution reduces to a Beta distribution.

The Dirichlet process can be regarded as the extension of the Dirichlet distribution to continuous spaces, which includes two parameters: a positive scalar parameter M and a probability base measure G_0 . The base distribution G_0 is where the nonparametric distributions are centered and usually represents the prior belief [17]. The concentration parameter M indicates the degree of concentration of the distribution around G_0 . The greater the M , the more samples from Dirichlet process are concentrated around G_0 [18]. Therefore, the distribution function G with a DP prior can be written as [19]:

$$G \sim DP(MG_0). \quad (1.24)$$

It is worth noting that although the Dirichlet process is defined over a continuous space, it is still discrete, as it consists of countably infinite point probability mass.

Understanding the format of a mixture model is a necessary step before understanding DPMM. Given the observation of n independent random variables v_1, v_2, \dots, v_n , generated from a population with unknown PDF $k(v)$, a parametric mixture model with k components can be written as [20]:

$$f(v) = \sum_{j=1}^k \pi_j k_j(v|\boldsymbol{\theta}_j), \quad (1.25)$$

where $k(v|\boldsymbol{\theta})$ is a parametric kernel with parameter vector $\boldsymbol{\theta}$, and the mixing proportions $0 < \pi_j < 1$ satisfy $\sum \pi_j = 1$. Define a latent allocation variable z_i as the group to which the observation v_i belongs and z_i is supposed to be drawn independently from the distribution [21]:

$$p(z_i = j) = \pi_j, \quad j = 1, \dots, k. \quad (1.26)$$

The hierarchical form of the parametric mixture model can be written as:

$$\begin{aligned} v_i | z_i &\sim k(v_i | \boldsymbol{\theta}_{z_i}), \quad i = 1, 2, 3, \dots, n, \\ \boldsymbol{\theta}_j &\sim G_0(\boldsymbol{\theta}), \quad j = 1, 2, 3, \dots, k, \\ z_i &\sim \text{Multinomial}(\boldsymbol{\pi}), \quad i = 1, 2, 3, \dots, n, \\ \boldsymbol{\pi} | \mu &\sim \text{Dirichlet}(\mu/k, \dots, \mu/k), \end{aligned} \quad (1.27)$$

where μ is concentration parameter of Dirichlet distribution. In this model, all $\boldsymbol{\theta}_j$'s are assumed to come from a common distribution $G_0(\boldsymbol{\theta})$ and the mixing proportion $\boldsymbol{\pi}$ is assigned as a Dirichlet prior. When $k \rightarrow \infty$, the Dirichlet distribution becomes the Dirichlet process, and the parametric mixture model becomes the Dirichlet process mixture model given by:

$$\begin{aligned} v_i | \boldsymbol{\theta}_i &\sim k(v_i | \boldsymbol{\theta}_i), \quad i = 1, 2, 3, \dots, n, \\ \boldsymbol{\theta}_i &\sim G(\boldsymbol{\theta}), \quad i = 1, 2, 3, \dots, n, \\ G(\boldsymbol{\theta}) &\sim DP(\mu G_0). \end{aligned} \quad (1.28)$$

In the DPMM, the parameter vector $\boldsymbol{\theta}_i$'s are assumed to be from a common distribution function $G(\boldsymbol{\theta})$ with a DP prior. The unknown PDF $f(t)$ modeled by the DPMM can be expressed as:

$$f(v) = \int k(v|\boldsymbol{\theta})dG(\boldsymbol{\theta}). \quad (1.29)$$

The behavior of the DPMM is affected by the choice of the kernel distribution k . Many common distributions have been applied in the DPMM, such as DP Gaussian mixture model (DPGMM), DP exponential mixture model (DPEMM), and DP Weibull mixture model (DPWMM). In this study, the DPWMM is applied for ALT analysis.

1.7 Nanocrystals Embedded High- k Device

The ALT data which are analyzed in this study are the failure times of nanocrystals embedded high- k devices collected at Thin Film Nano & Microelectronics Research Laboratory, Texas A&M University.

When the metal–oxide–semiconductor field-effect transistor (MOSFET) is scaled down to the nano scale to satisfy the requirements of technology development, the thickness of the silicon dioxide (SiO₂) gate dielectric layer has to be reduced drastically. This degrades the device performance and reliability [22]. One effective solution is to use a high dielectric constant (high- k) film to replace SiO₂. The high- k films have also been used in memory devices [23]. The conventional poly-Si floating-gate nonvolatile memory (NVM) device contains a continuous poly-Si layer in the SiO₂ gate dielectric as the charge-trapping medium. Therefore, a single leakage path in the tunneling oxide may quickly drain all the charges. The nanocrystals embedded dielectric structure can solve this problem because the nanodots are isolated from each other by the surrounding dielectric materials. Therefore, a single leakage path can only drain charges stored in one or a few dots locally [24]. Various conductive and semiconductive materials, such as Si, ITO, ZnO, and MoO_x have been prepared into the nanocrystalline form and embedded in

high- k films as electron- or hole-trapping media [23], [25]–[27]. Figure 1.4 shows a cross-sectional view of a general single-layer nanocrystals embedded ZrHfO capacitor. The reliability of this kind of device has not been well studied, which is important to the practical applications. The time-dependent dielectric breakdown refers to damage-endurance failures [7] and therefore ALT can be applied for obtaining the failure data of dielectric devices. In this study, the failure-time data of this kind of device are applied in the DPMM for reliability analysis.

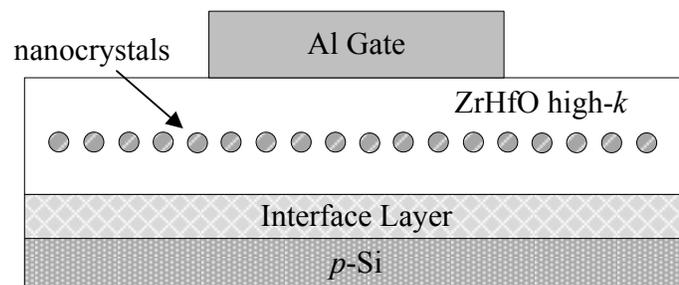


Figure 1.4. Cross-sectional view of nanocrystals embedded ZrHfO capacitor.

1.8 Dissertation Overview

The balance of this dissertation is organized as the following: Chapter 2 reviews previous research on inference of accelerated life testing and Dirichlet process mixture model; Chapter 3 describes the notations and assumptions, as well as the problem solved in this study; Chapter 4 introduces the nano-crystals embedded sample investigated in the research; Chapter 5 studies the constant-stress ALT; Chapter 6 investigates the step-stress ALT; and finally, Chapter 7 concludes the dissertation.

CHAPTER 2: LITERATURE REVIEW

This chapter reviews previous research on ALT inference, including constant-stress ALT and step-stress ALT. The review of Dirichlet process mixture model is also summarized.

2.1 Inference of Constant-Stress Accelerated Life Testing

A traditional parametric approach to develop the inference on constant-stress ALT requires two assumptions. Firstly, the failure-time distribution under each stress level is assumed to come from the same parametric distribution family. For example, it is typical to assume that a failure-time distribution is from a location-scale family, such as exponential, Weibull, log-normal or normal. Secondly, an acceleration relationship called the “time transformation function” such as the Arrhenius and Eyring law is assumed to relate the parameters of the distributions under various stress levels. A semi-parametric ALT analysis usually relaxes one of these two assumptions and has been applied widely in literature.

2.1.1 Parametric estimation of constant-stress ALT

The ALT assuming various time transformation functions and failure-time distributions have been analyzed with different parametric estimation methods in literature.

Singpurwalla et al. [28], Kahn [29] and Barbosa and Louzada-Neto [30] used Least Squares Estimation for the ALT. Singpurwalla et al. [28] handled the censored data using Eyring model with normal failure-time distribution. The authors defined a linear model for the parameters in Eyring model and estimated the parameters using Least

Square Estimation. Kahn [29] followed the approach of Singpurwalla et al. [28] and used the inverse power law with an exponential, rather than normal failure-time distribution. Barbosa and Louzada-Neto [30] estimated the mean lifetime of the units under working conditions. The censored data was handled. The authors assumed a Weibull distribution for lifetime, and a log-linear relationship as the time transformation function. The point estimation of parameters was obtained with Iteratively Reweighted Least Squares Algorithm and the interval estimation was obtained with MLE.

Whitman [31], Abdel-Ghaly et al. [32], Watkins [33], [34], Glaser [35], Hirose [36], and Newby [37] applied ML method to estimate the parameters. Whitman [31] assumed the failure-time distribution to be log-normal and used the Arrhenius model for the median time to failure as the time transformation function. The author estimated the parameters in the Arrhenius model and the median time to failure at a certain stress, and also provided their confidence intervals. Both complete data set and data with censoring have been modeled. Abdel-Ghaly et al. [32] estimated the data with type-II censoring, assuming a 3-parameter Pareto failure-time distribution and inverse power law as the time transformation function. The authors predicted the value of the shape parameter as well as the reliability function at a mission time under operating condition. Watkins [33], [34] fitted Weibull distribution to ALT data. Watkins [33] assumed power-law model and dealt with a complete data set. The common shape parameter of Weibull distribution and parameters of power-law model were estimated. The ML estimators were obtained by Newton-Raphson iterative method. Watkins [34] specified a log-linear relationship to describe the scale parameter of Weibull distribution. The author fitted data with

censoring and estimated parameters with Newton's method. Glaser [35] estimated a Weibull ALT model. Both scale and shape parameters were expressed as functions of the testing environment, while in most cases, the shape parameter was assumed to be constant. Besides complete data and censored data, this paper also handled grouped data by introducing a status indicator. Hirose [36] developed a Weibull inverse power law including a threshold stress considering type-I censoring. The authors considered both scenarios of common shape parameter and stress-related shape parameters. Newby [37] introduced a generalized treatment of AFT model using general shape, scale and location parameter families of distributions. The Arrhenius model was assumed in the paper.

Mazzuchi and Soyer [38] and Mazzuchi [39] did Bayesian inference for the ALT model with power-law transformation function. Mazzuchi and Soyer [38] assumed an exponential lifetime distribution. The author rewrote the power-law function by taking logarithm and then set up a linear model. The Linear Bayesian Approach was used to solve the model. Mazzuchi [39] presented a Bayesian procedure for inference assuming Weibull failure-time distribution. The procedure was based on the General Linear Model set up by "linearizing" the time transformation function and then it employed Linear Bayesian Approach to produce computable results.

Bai and Chung [40] proposed both constant-stress and progressive-stress ALT based on Weibull lifetime distribution and inverse power law. Both MLE and Bayesian methods were used to estimate the parameters. A Monte Carlo study was carried out to investigate the behavior of estimators.

Biernat et al. [41] dropped the assumption that the failure-time distribution at different stress levels was from the same family distribution. Instead, the authors assumed the failure-time distribution might be governed by different forms of distribution. The authors fitted each CDF of failure-time distribution with empirical CDF and computed quantiles of each CDF. They then performed regression on each quantiles based on the inverse power law. The estimated parameters of inverse power law function were obtained from regression, and the CDF in use condition was a step function based on the computed quantiles. This method can only produce a discrete CDF.

Kim and Bai [42], AL-Hussaini and Abdel-Hamid [43], [44] developed mixture models for data with more than one failure mode. Kim and Bai [42] analyzed ALT data under two failure modes by assuming the log lifetime followed mixture of two location-scale distributions and each location parameter had a linear relation with the stress. The ML estimates of the distribution parameters and the mixing proportions were obtained by the expectation and maximization algorithm. AL-Hussaini and Abdel-Hamid [43], [44] assumed that the type-II censoring lifetimes under various failure modes were distributed according to a finite mixture model, in which each failure mode was represented by a nonnegative and continuous function. In both papers, the power-law relationship applied to the mixture of two Weibull components was presented. AL-Hussaini and Abdel-Hamid [43] estimated the parameters, reliability and hazard rate functions using the Bayesian method, while AL-Hussaini and Abdel-Hamid [44] applied the MLE method. Moreover, AL-Hussaini and Abdel-Hamid [44] used mixtures of two exponentials, Rayleigh and Weibull components models as illustrative examples.

2.1.2 Semi-parametric estimation of constant-stress ALT

A commonly used semi-parametric estimation for ALT in literatures drops the first assumption of parametric form of lifetime distribution and retains a parametric acceleration function. For example, Shaked et al. [45], Shaked and Singpurwalla [46], Bai and Lee [47], and Basu and Ebrahimi [48] all followed this approach. Shaked et al. [45] assumed a parametric relationship between any two stress levels and estimated the MTTF under normal stress based on the empirical distribution function. Shaked and Singpurwalla [46] assumed inverse power law as time transformation function and also estimated the MTTF and CDF under use condition based on the empirical CDF. Moreover, the authors tested whether the estimated CDF followed a member of a specified family of CDF's using Kolmogorov-Smirnov statistic and obtained the confidence bounds for CDF based on the test result. Bai and Lee [47] assumed inverse power law and used empirical estimator to estimate the ALT under intermittent inspection, in which the test units were only inspected at specified points of time. Since the empirical distribution functions are discrete, it may be difficult to handle censoring data with empirical estimators. Basu and Ebrahimi [48] extended the work of Shaked and Singpurwalla[46] to include the censoring data by introducing the scale model.

Some literatures drop the second assumption- parametric acceleration model and retain the assumption that the failure-time distribution at each condition is from a family of distribution. Dorp and Mazzuchi [49], [50] proposed a model for general ALT, including regular ALT, constant-stress ALT, step-stress ALT, and profile-stress ALT. The lifetime distribution at each stress level was assumed to be exponential in Dorp and

Mazzuchi [49] and Weibull in Dorp and Mazzuchi [50]. The authors did not assume any parametric time transformation function. Instead, the multivariate Ordered Dirichlet distribution was used as prior information to define a multivariate prior distribution for the scale parameter at various stress levels and the common shape parameter. This approach required information of a specified quantile on the mission time reliability at the operating stress level to infer the use stress life parameters estimates. Louis [51] assumed the Weibull family and developed a scale-change model by introducing a scale change parameter to parameterize the difference between two survival distributions. Then the efficient score statistic was used to estimate the scale change parameter. Schmoyer [52] analyzed ALT with two-level single-stress, including an accelerated stress and a normal stress. The author proposed a general model of the form $\Pr(t; x) = F(g(x)h(t))$, where $\Pr(t; x)$ denoted the probability of failure by time t at stress level x , F was a CDF on $[0, \infty)$, g and h were nonnegative and nondecreasing, g had S -shaped curvature and $g(0)=0$. Either F or h was assumed to be known. If F was known and $F(u) = 1 - e^{-u}$, then this was a proportional hazards model; if h was known and $h(t) = t$, it was a accelerated failure time model. Based on these assumptions, the author developed confidence bounds for low-stress long-time probabilities and quantiles. Because no acceleration function was assumed, neither approach developed by Louis [51] nor Schmoyer [52] can predict the failure-time distribution under use condition.

Proschan and Singpurwalla [53] and Maciejewski [54] relaxed both assumptions. Proschan and Singpurwalla [53] divided the test time to intervals and assumed the probability $(p_{j,i})$ of failure of a unit in time interval i under stress j was Beta distributed

and estimated the failure rate with weighted average failure rate. The estimated $p_{j,i}$ was obtained with Bayesian estimation and the weight parameters were obtained with Least Square Estimation. Maciejewski [54] first fit the CDF at each stress level to a parametric family of functions with empirical CDF, then derived quantiles of the levels and their dispersions for each CDF. The author defined a function which could be derived from each quantile level and used this function to calculate the estimate of each quantile level in use condition. Then fit the CDF with quantiles in use condition.

Some more recent literatures assumed a linear or log-linear model between failure time and covariate, resulting in a more generic acceleration model. No parametric distribution function was assumed and regression was carried out. The regression model was defined as $(\ln)Y_i = \beta'X_i + \varepsilon_i, i=1, \dots, n$, where Y was the failure time, \mathbf{X} was a $p \times 1$ vector of covariates, β was a $p \times 1$ vector of unknown regression parameters, and ε_i 's were s-independent and had an unspecified distribution function. To estimate the regression coefficient β , Lin and Geyer [55] developed computational methods to implement rank regression procedures using simulated annealing. The unknown error function was estimated using the Kaplan-Meier estimator. Komarek and Lesaffre [56] developed a Bayesian linear regression model with paired doubly interval-censored data. The bivariate error distribution was assumed as a finite mixture of bivariate normal densities. The Bayesian approach with the Markov chain Monte Carlo (MCMC) methodology was used for inference. Komarek and Lesaffre [57] analyzed the multivariate doubly-interval-censored data considering clustering. The univariate densities of random errors and random effects were modeled as penalized Gaussian

mixture with an overspecified number of mixture components. The Bayesian approach with MCMC sampling was used to estimate the model parameters. Argiento et al. [58] examined the accelerated failure-time model for univariate data with right censoring. The error distribution was represented as a nonparametric hierarchical mixture of Weibull distribution on both shape and rate parameters, and the mixing measure was a priori distributed as the generalized gamma measure. Other linear or log-linear regression model involved with Dirichlet process mixtures will be reviewed in section 2.3.

2.2 Inference of Step-Stress Accelerated Life Testing

Besides the two traditional assumptions in constant-stress ALT, a classical approach to analyze a SSALT requires an additional assumption to relate the distribution under step stresses to the distribution under constant stress. Major models used in literatures include the Tampered Random Variable (TRV) model [59], the Cumulative Exposure (CE) model [12], and the Tampered Failure Rate (TFR) model [60].

2.2.1 *Parametric estimation of step-stress accelerated life testing*

The Cumulative Exposure model was first proposed by Nelson [12] and has been widely accepted and used in SSALT analysis. The CE model assumes the remaining life of the units depends only on the cumulative exposure the units have experienced, without memory on how this exposure was accumulated. Nelson [12] presented the inference using ML estimators for Weibull failure data under the inverse power law. The data of time to breakdown of electrical insulations was fitted as illustration. Xiong [61] presented the inference of parameters in an simple SSALT with type-II censoring, assuming CE model and exponential lifetime distribution with a mean that was a log-linear function of

stress. The confidence intervals were constructed using a pivotal quantity. Xiong and Ji [62] studied a similar problem with type-I censoring. Xiong and Milliken [63] constructed the model with the same assumptions to estimate the parameters in log-linear function and further predicted the lifetime under design stress as well as that during a future SSALT. The failure-step stress ALT was conducted in this paper. Zhang and Geng [64] constructed the analysis for both constant-stress and step-stress ALT by applying a Weibull lifetime distribution with a linear Arrhenius lifetime-stress relationship. The Least Square Method was used to estimate the Weibull parameters and a self-designed software was programmed to predict the life under use condition. Abdel-Hamid and AL-Hussaini [65] applied an exponentiated distribution, with a scale parameter which was a log-linear function of the stress and hold the CE model. Special attention was paid to an exponentiated exponential distribution. The ML estimates of parameters under consideration were obtained based on type-I censoring. Wang [66] derived the confidence intervals for the exponential SSALT model under progressive type-II censoring which employed the removal of surviving units at time of failure. The mean life was assumed to be a log-linear function of stress and the MLE was applied. Yin and Sheng [67] derived the lifetime distribution under progressive stress ALT, in which the stress was proportional to time. The lifetime distribution was assumed to follow an exponential or a Weibull distribution with inverse power law and the parameters were estimated using MLE. Gouno [68], and Lee and Pan [69]–[71] all assumed exponential lifetime distribution at each individual stress level and used failure rate to describe PDF and reliability under step-stress. This resulted in the same form with that derived from CE

model. Gouno [68] presented a practical method to analyze temperature SSALT data with type-II censoring considering an Arrhenius model. Both Least Squares and MLE were used to estimate the parameters of an Arrhenius model and failure rate under use condition. Lee and Pan [69], [70] presented Bayesian inference model for SSALT with type-II censoring. The mean lifetime at each stress was assumed to be a log-linear function of stress level. Lee and Pan [69] constructed the model for simple SSALT and derived the Bayesian inference with conjugate prior. Lee and Pan [70] constructed the model for multi-SSALT and used MCMC technique to deal with nonconjugate prior. Lee and Pan [71] analyzed multi-stress ALT with right censoring assuming a Generalized Linear Model (GLM) and log-linear relationship. Both the MLE and Bayesian estimation were used to estimate the GLM parameters.

Tang et al. [72] modified the CE model to analyze ALT with failure-free life (FFL), which is the age of a product below which no failure should occur. The FFL was characterized by a location parameter in the distribution. The authors proposed a Linear Cumulative Exposure Model (LCEM) which assumed the fractional exposure was linearly accumulated. The 3-parameter Weibull distribution with location and scale parameters expressed as inverse power law relationship with stress was used to illustrate the estimation procedure, and the MLE was used.

Khamis and Higgins [73] proposed the KH model as an alternative to the Weibull CE model in SSALT. The proposed model was based on a time transformation of the exponential CE model. The new model was as flexible as the Weibull CE model for fitting data while easier to obtain the ML estimates of the parameters.

When an SSALT only has two steps, i.e., simple SSALT, the time transformation function is not necessary and many researchers use original parameters at each stress level directly. For example, Balakrishnan et al.[74], [75], Kateri and Balakrishnan [76], Balakrishnan and Xie [77], Balakrishnan and Han[78], and Han and Balakrishnan [79] all used original parameters of different assumed distributions at each stress level as estimates and applied the MLE method. Moreover, Balakrishnan and Xie [77] considered type-II hybrid censoring scheme, in which the test was terminated at time T if the r th failure occurred before time T ; otherwise, the test was terminated as soon as the r th failure occurred. This type of censoring scheme ensures the test can obtain at least r failures. Balakrishnan and Han [78] and Han and Balakrishnan [79] considered the simple SSALT with two fatal causes for the failure and assumed different risk factors were independent and exponentially distributed. Balakrishnan and Han [78] analyzed type-II censored data and Han and Balakrishnan [79] analyzed type-I censored data. Balakrishnan et al. [80] developed the model for multi-SSALT without assuming any time transformation function. The authors assumed an exponential lifetime distribution and developed the order restricted MLE under right censored sampling situations.

DeGroot and Groel[59] introduced the Partially Accelerated Life Test (PALT) in which if a unit survived to a specified time at design stress, it was switched to a higher level of stress. The authors modeled the effect of switching the stress by multiplying the remaining lifetime of the unit by some unknown factor called the tampering coefficient and the model was called Tampered Random Variable (TRV) model. DeGroot and Groel [59] assumed the lifetime under use conditions is exponential and applied Bayesian

estimation. Madi [81], Abdel-Ghaly et al. [82], and Wang et al. [83] applied TRV model to analyze PALT considering different censoring schemes. Madi [81] proposed an empirical Bayes approach to pool data from several groups of units that were tested at different instances to estimate the parameters. Abdel-Ghaly et al. [82] and Wang et al. [83] assumed a Weibull lifetime distribution and used MLE to estimate the distribution parameters and tampering coefficient. Abdel-Ghaly et al. [82] considered type-I and type-II censoring, and Wang et al. [83] considered multiply censored data.

Battacharyya and Soejoeti [60] modified the TRV model by assuming that the effect of changing the stress was to multiply the failure rate function over the remaining life and the modified model was called Tampered Failure Rate (TFR) model. The authors assumed a Weibull lifetime distribution and estimated the parameters with MLE. An extension to fully SSALT was derived with the application of log-linear life-stress function. Wang and Fei [84] applied the TFR model to the progressive stress ALT, assuming a Weibull-time failure distribution with the scale parameter satisfying inverse power law. The parameters were estimated using MLE.

Zhao and Elsayed [85] proposed a general SSALT model based on the acceleration models and produced some commonly used lifetime-stress relationship and their acceleration factors. The MLE method was utilized to solve for the Weibull and lognormal lifetime distributions.

2.2.2 Semi-parametric estimation of step-stress accelerated life testing

Most semi-parametric estimation inference for SSALT is obtained by dropping the lifetime distribution assumption, while holding the parametric time transformation

function and the CE model. Lin and Fei[86], Bai and Chun [87], and Tyoskin and Krivolapov [88] all used this approach. The lifetime properties under use conditions were obtained based on transformed failure times and some empirical estimators. Shaked and Singpurwalla [89] developed a model which assumed the distribution of the lifetime was a function of the total accumulated V^α , where V was the stress and α was an unknown constant. The new model unified and generalized the TRV and CE models. Hu et al. [90] extended the work of Schmoyer [52] to a simple SSALT and obtained the upper confidence bounds for cumulative failure probability under use conditions.

Dorp et al. [91] assumed the failure times at each stress level were exponentially distributed and dropped the time-transformation function. The model was developed for SSALT considering linear ramping stress. The multivariate Ordered Dirichlet distribution was used as prior information to define a multivariate prior distribution for the failure rates at various stress levels. Bayes point estimates, as well as probability statement lifetime parameters under use conditions were developed.

2.3 Dirichlet Process Mixture Model

The DP was first formally developed by Ferguson [15] as a random probability model to define priors for spaces of distribution functions. It consists of countably infinite point probability masses. In order to relax the restriction of discreteness, the DPMM was developed and has been widely applied in the area of nonparametric Bayesian data analysis.

The DP Gaussian mixture model using a normal kernel has been extensively applied, e.g., in density estimation [92]–[95], curve fitting [96], and regression [97]–

[100]. Some other classical kernels have also been applied. Kottas [101] proposed a DPMM with a Beta distribution to estimate density and intensity. West et al. [102] introduced the DP Gamma mixture model for hierarchical linear regression and density estimation. Mukhopadhyay and Gelfand [103] developed the Generalized Linear Models with DP mixture of binomial and Poisson kernels. Carota and Parmigiani [104] developed the DP Poisson mixture model for regression problems in which the response variable is a count.

Some researchers have applied DPMM in survival data analysis. Kottas and Gelfand [97] introduced both semi-parametric and nonparametric Bayesian modelling approaches for the error distributions of median regression. Both models were based on DP normal mixture models. The censored survival data was handled in the paper. Gelfand and Kottas [98] demonstrated a computational approach to obtain the entire posterior distribution for nonparametric Bayesian inference with DPMM. The application of comparison of survival times from different populations under fairly heavy censoring was illustrated. Gelfand and Kottas [99] extended the nonparametric Bayesian modeling approach with DPMM in Gelfand and Kottas [97] to median residual life distribution. The DP normal mixture model was applied to survival data. Kottas [19] applied the DP Weibull mixture model to censored survival data. The mixes were applied over both the shape and scale parameters of the Weibull kernel.

Kuo and Mallick [18], Hanson [105], and Ghosh and Ghosal [106] developed a DP mixture model with log-linear acceleration relationship for constant-stress ALT analysis. Kuo and Mallick [18] proposed two hierarchical models to estimate error

distribution of log-linear model, which was shown in equation (1.14) as $\ln T_i = -x_i\beta + W_i$.

The first model applied the DP mixture model for the distribution of $V_i = \exp(W_i)$ and the other modeled the distribution of W_i . The DP mixture model with normal and log-normal kernel was used. Hanson [105] modeled the distribution of V_i by the DP mixture model with gamma kernel and Ghosh and Ghosal [106] applied the DP Weibull mixture with a fixed shape parameter.

CHAPTER 3: PROBLEM STATEMENT

This chapter first describes the notations and basic assumptions applied in the dissertation. Then the primary objective as well as its sub-tasks is presented to illustrate the problem that will be solved.

3.1 Notations

x_i : stress or a proper transformation of the stress

α, λ : shape and scale parameter of Weibull distribution

β : regression coefficient in the log-linear lifetime-stress relationship

t : failure time

t_c : censoring time

δ : censoring indicator

μ : precision parameter of the Dirichlet process

π : mixing proportions of the parametric mixture process

θ : parameter vector of the parametric kernel

$F(\cdot), f(\cdot)$: cumulative distribution function (CDF) and probability density function (PDF)

$G(\theta)$: random distribution function for θ

$G_0(\theta)$: base distribution of the Dirichlet process

$R(t)$: reliability function of failure time, $R(t)=1-F(t)$

$K(\cdot), k(\cdot)$: CDF and PDF of the parametric kernel

v : $v = t \exp(x\beta)$

τ : stress changing time in a simple SSALT

3.2 Assumptions

This dissertation makes the following basic assumptions:

- (1) n identical and independent units are placed in the test.
- (2) The relationship of failure time t_i under different stress levels is log-linear, which can be expressed as $\log t_i = -x_i\beta + w_i$, $i=1, 2, \dots, n$, where w_i is error term.
- (3) Each v_i is from an independent Weibull kernel $k(v_i|\alpha_i, \lambda_i)$.
- (4) A cumulative exposure model is assumed for simple SSALT.
- (5) For constant-stress ALT, three accelerated stresses are used in the test. The breakdown time was observed by the jump of leakage current. Each unit under test was tested individually and broke down independently. A stress lower than these three accelerated stress is assumed to be normal stress level.
- (6) For simple SSALT, a unit is first tested under the lower stress level x_L . If the unit has not failed by a pre-specified time τ , the stress level is increased to x_H at the changing time τ , and the test is continued until failure or the censoring time t_c (i.e., type-I censoring). The lower stress x_L is assumed to be normal stress level.

3.3 Problem Description

In this dissertation, the sample size n , the testing stress levels, and stress changing time τ for simple SSALT are pre-specified. The primary objective is to predict CDF of lifetime distribution under normal stress level from experimental data. Three experimental datasets, including one complete dataset and two right censored datasets are used to illustrate the applicability of the proposed methodology. In order to fulfill this objective, the following sub-tasks are performed.

3.3.1 Investigation of metal oxide nanodots-embedded ZrHfO high-k film

The experimental datasets used in this dissertation were collected at the Thin Film Nano & Microelectronics Research Laboratory at Texas A&M University. In order to further understand the device, this dissertation investigates the fabrication process, physical properties, and memory functions of the device tested for life time prediction.

3.3.2 Development of simulation-based algorithm

The DP Weibull mixture ALT model is a high dimensional problem with multiple parameters. Gibbs sampling is a popular algorithm to fit DP mixture model. This dissertation develops Gibbs sampling algorithm to formulate posterior inference on parameters, as well as failure-time CDF.

3.3.3 Comparison between parametric Weibull log-linear ALT model and DP Weibull mixture ALT model

A parametric ALT analysis is also performed for the purpose of comparison. It is assumed the failure time at a given stress level comes from a Weibull distribution, and the relationship between lifetime and stress level is log-linear. This Weibull log-linear ALT model is estimated with standard MLE method, and is implemented using Minitab. Then, the CDF of failure-time distribution under normal stress level predicted with parametric ALT model and the proposed DP mixture model, as well as empirical CDF are plotted and visual compared.

CHAPTER 4: MEMORY FUNCTIONS OF MOLYBDENUM OXIDE NANODOTS EMBEDDED ZRHFO HIGH- K^\dagger

This chapter introduces the device tested in the dissertation, including the fabrication process, its physical characteristics, charge trapping and detrapping mechanisms, as well as memory properties.

4.1 Fabrication of Nanocrystals Embedded ZrHfO High- k Device

The ZrHfO (tunnel oxide)/ MoO_x/ZrHfO (control oxide) gate dielectric stack was deposited on the HF precleaned p -type (10^{15}cm^{-3}) Si (100) wafer in one pump down without breaking the vacuum. The tunnel and control ZrHfO layers were sputtered from a composite Zr/Hf (12/88 wt %) target in an Ar/O₂ (1:1) mixture at 5 mTorr and 60 W for 2 and 10 min, separately. The MoO_x film was sputtered from the Mo target in Ar/O₂ (1:1) at 5 mTorr and 100 W for 15 s. After the gate dielectric stack was accomplished, the post deposition annealing (PDA) step was carried out by rapid thermal annealing (RTA) at 800°C in the pure N₂ ambient for 1 min. An aluminum (Al) film was sputter deposited, lithography patterned, and wet etched into gate electrodes. The Al film was also deposited on the backside of the wafer for ohmic contact. The final MOS capacitor was annealed at 300 °C under H₂/N₂ (10/90) for 5 min. The control sample, i.e., containing only the ZrHfO gate dielectric without the embedded nc-MoO_x layer, was also prepared and characterized for comparison. The capacitor's capacitance-voltage (C - V) and current-voltage (I - V) characteristics were measured with an Agilent 4284A LCR meter

[†]Reprinted with permission from Xi Liu, Chia-Han Yang, Yue Kuo, and Tao Yuan, *Electrochemical and Solid-State Letters*, 2012, vol. 15, issue 6, H192 (2012). Copyright 2012, The Electrochemical Society.

and an Agilent 4155C semiconductor parameter analyzer, respectively. The high- k stack was analyzed with X-ray photoelectron spectroscopy (XPS) for chemical bond states and X-ray diffraction (XRD) for the crystallinity.

4.2 Physical Properties of Nanocrystals Embedded ZrHfO high- k Device

Figure 4.1(a) shows the XPS O1s peak of the MoO_x embedded ZrHfO high- k stack prepared in this study. It has been deconvoluted into 4 sub peaks. The peak with the binding energy (BE) of 530.3 eV is MoO₃. Peaks with BE 529.6 eV and 531.3 eV are related to HfO₂ and ZrO₂. The peak with BE 531.65 eV is probably contributed by the Al₂O₃ from the gate electrode. A small Mo 3d_{5/2} peak at 227.3 eV BE was detected, which is the Mo^{δ+} element [107]. The nanodots are crystalline MoO₃ detected by X-ray diffraction (XRD), as shown in Fig. 4.1(b). The crystal size is about 28 nm, determined from the peak location and full width at half maximum using the Scherrer equation [108]. Previously, it was demonstrated that under the same process condition, discrete nc-ITO nanodots were formed in the ZrHfO film [109]. However, it is not clear if the embedded nc-MoO_x film was composed of discrete dots. In addition, the density of the nanodots in the dielectric layer is important to the charge retention and reliability. In order to obtain these data, the sample has to be analyzed with the high resolution TEM, which is under study now. The equivalent oxide thickness of the control sample and the nc-MoO_x embedded sample are 7.8 nm and 8.5 nm, respectively, calculated from the $C-V$ curve.

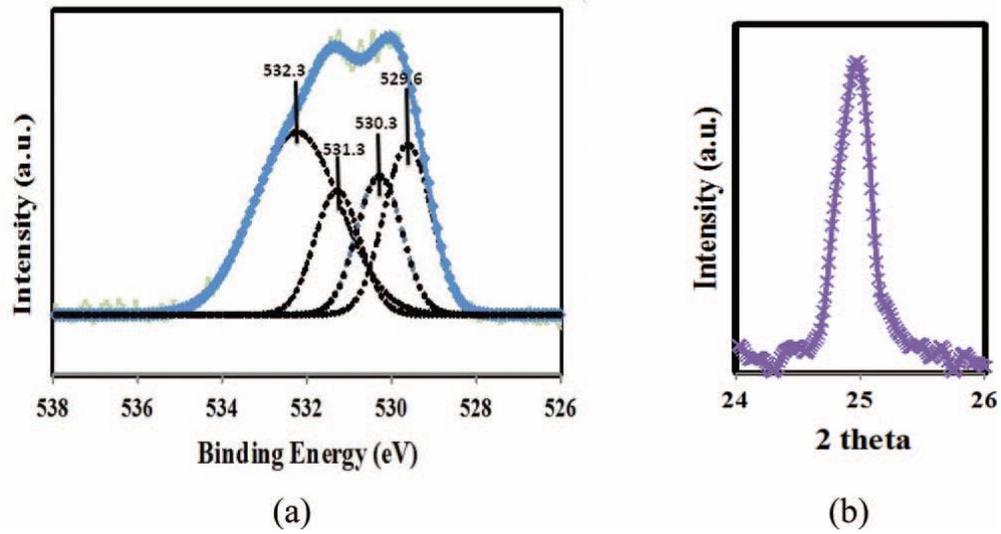


Figure 4.1. (a) XPS O 1s peak, and (b) XRD pattern of the MoO_x embedded ZrHfO sample

4.3 Charge Trapping and Detrapping Mechanisms

The charge trapping characteristics of the capacitor can be examined from the C - V hysteresis curves. Figure 4.2 shows the C - V hysteresis curves of the capacitor with the nc-MoO_x embedded ZrHfO dielectric measured at 1 MHz. The gate was stressed with a voltage (V_g) from negative to positive, i.e., the forward direction, and then back to negative, i.e., backward direction, in three ranges, i.e., -3V to 3V to -3V (± 3 V), -6V to 6V to -6V (± 6 V), and -8V to 8V to -8V (± 8 V), separately. The memory window can be defined as flat band voltage difference (ΔV_{FB}) between the V_{FB} of the forward curve and that of the backward curve. The control sample was prepared and compared with the nc-MoO_x embedded sample for defect formation consideration. The charge trapping density (Q) of the capacitor can be estimated from the following equation [110]:

$$Q = \frac{C_{FB} \times \Delta V_{FB}}{q}, \quad (4.1)$$

where C_{FB} is the flatband capacitance and q is the electron charge.

For the control sample, since the C - V hysteresis is very small, i.e., $\Delta V_{FB} = 0.06$ V, the ZrHfO film has negligible charge trapping capability. For the nc-MoO_x embedded sample, the C - V hysteresis phenomenon is more pronounced. The charge trapping densities Q_{ot} 's are 2.21×10^{10} cm⁻², 1.45×10^{11} cm⁻², and 7.13×10^{11} cm⁻² for the V_g sweep ranges of ± 3 V, ± 6 V, and ± 8 V, separately. Therefore, the charges are trapped to the nc-MoO_x site. The charge trapping capability is related to the amount of charges supplied to the high- k stack. It is worth to note that the magnitude of the V_{FB} of the forward C - V curve is highly dependent on the starting V_g , i.e., -0.49 V, -0.62 V, and -1.15 V for starting $V_g = -3$ V, -6V, and -8V, separately. The V_{FB} of the forward C - V curve of the fresh nc-MoO_x embedded sample is -0.46 V. Compared with the control sample, holes are trapped to the nc-MoO_x stack are a density of 6.62×10^{10} cm⁻², 3.87×10^{11} cm⁻², and 1.59×10^{12} cm⁻² with the starting V_g of -3V, -6V, and -8V, separately. Therefore, nc-MoO_x is effective in trapping holes, which is similar to the nc-RuO or nc-ITO case [111]–[113]. On the other hand, the V_{FB} of the backward curve is more negative with the increase of the sweeping voltages, i.e., -0.48 V at ± 3 V, -0.56 V at ± 6 V and -0.84 V at ± 8 V. Therefore, holes trapped in the forward sweeping are not completely erased in the back sweeping direction.

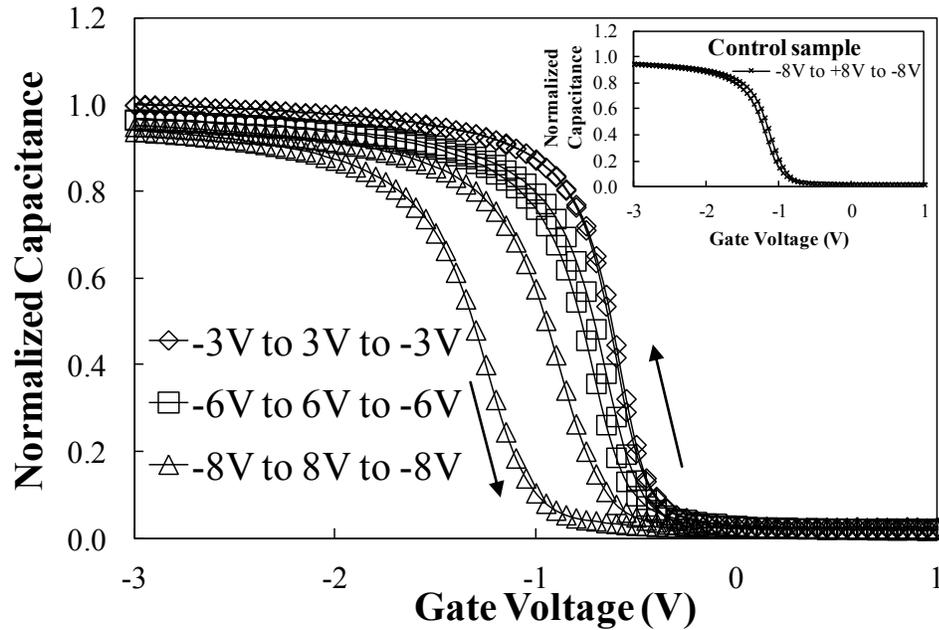


Figure 4.2 C - V hysteresis curves of the nc-MoO_x embedded ZrHfO capacitor measured at 1 MHz. The inset is the hysteresis curve of the control sample.

To further study the hole-trapping mechanism, the leakage current density-voltage (J - V) curves of the nc-MoO_x embedded sample and the control sample were measured from -8 V to +8 V, as shown in Figure 4.3. The polarity of the leakage current is defined as positive when the current flows toward the substrate and negative when the current flows toward the gate. Compared with the control sample, the nc-MoO_x embedded sample has a larger leakage current and the J - V curve is less smooth. There are several bumps in the J - V curve of nc-MoO_x embedded sample. First, a very small bump appears at point A, which is near the V_{FB} of the corresponding C - V curve in Fig.4.2 at $V_g = -1.15$ V. The current changes its polarity at this point because of the release of loosely trapped holes, which are probably located at the nc-MoO_x/ZrHfO interface[111], [113]. Second, a

slightly more obvious jump of the current is observed at point B near $V_g = 0$ V, which is due to the further release of a large number of remaining trapped holes from the change of the V_g polarity [111], [113]. Third, there are two obvious peaks at points C and D near $V_g = 1$ V. These are the negative differential resistance (NDR) peaks commonly observed in floating gate memory rapping devices [25], [114]. This is caused by the Coulomb blockade effect, i.e., the nc-MoO_x site is saturated with the trapped charges [23], [115], [116]. When the V_g is further increased, e.g., beyond 2 V, an inversion layer is fully established and the leakage current increases drastically with the increase of V_g .

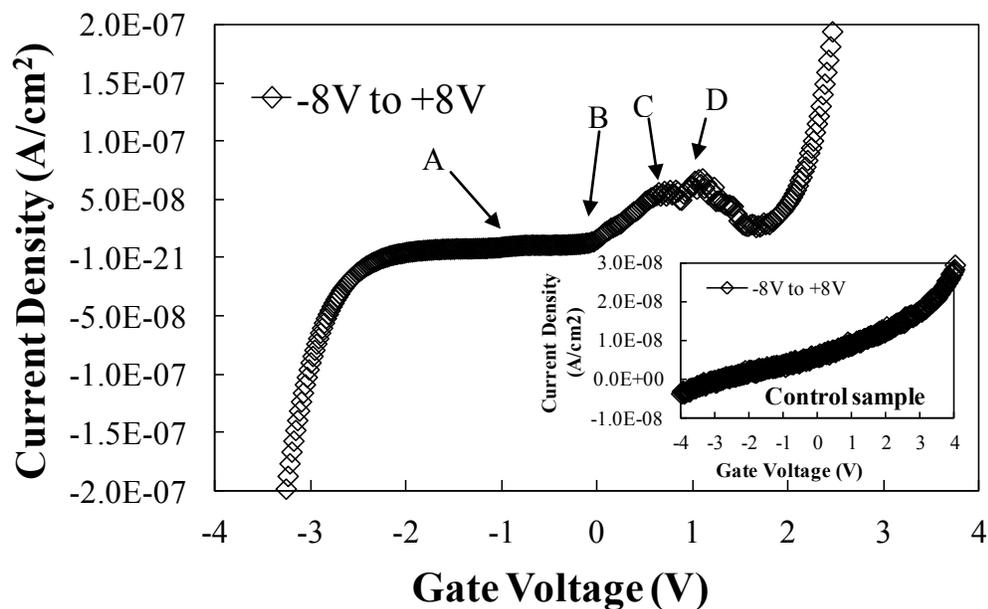


Figure 4.3 J - V curve of nc-MoO_x embedded capacitor V_g swept from -8 V to +8 V. The inset is the J - V curve of control sample.

4.4 Charge Retention Capability

The charge retention capability of the capacitor could be determined by measuring the percentage of charge remained in the device after releasing the stress voltage for a period of time. The following equation was used to calculate the percentage of charge from the V_{FB} shift over a period of time [113]:

$$\text{Charge remaining (\%)} = \frac{V_{FB}(t) - V_{FB}(\text{fresh})}{V_{FB}(\text{stress}) - V_{FB}(\text{fresh})}, \quad (4.2)$$

where $V_{FB}(\text{stress})$ is the V_{FB} immediately after the V_g stress, $V_{FB}(\text{fresh})$ is the V_{FB} before the V_g stress, and $V_{FB}(t)$ is the V_{FB} after releasing the stress for time t . After releasing the write stress, the C - V curve was measured over a small V_g range i.e., -2 V to +1 V, to determine the $V_{FB}(t)$. Figure 4.4 shows the charge retention curve of the nc-MoO_x embedded capacitor after being stress at $V_g = -8$ V for 10 seconds. The total measurement time was 10 hrs and the V_{FB} was determined every 1,800 s. First, a quick loss of nearly 20% of trapped holes occurred within 1,800 s after releasing the stress V_g , which is due to the detrap of the loosely-trapped holes from the nc-MoO_x/ZrHfO site [111], [113]. Then, the strongly-trapped holes were gradually released, e.g., totally 9% loss from 1,800s to 36,000s. The phenomenon of the 2-step release of trapped holes have been observed in the nc-RuO and nc-ITO embedded capacitors [111], [113]. The exact location of the charge trapping site can be clarified using the frequency dispersion method [26], [117]. The inset of figure 4.4 shows the extrapolation of the curve to 10 year period. About 54% of trapped holes remained in the nc-MoO_x embedded sample after releasing the stress V_g for 10 years, which is the desirable characteristic for the nonvolatile memory device.

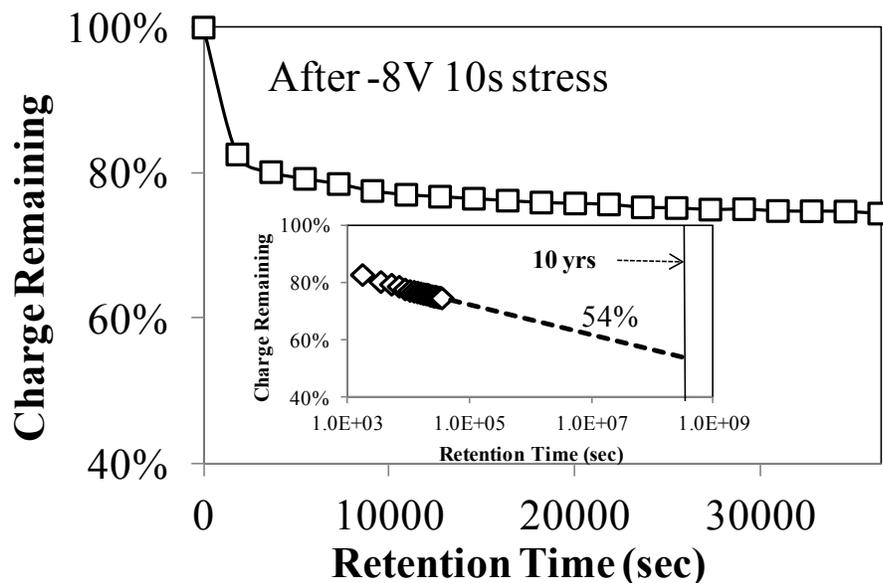


Figure 4.4 Retention property of holes trapped in the nc-MoO_x embedded capacitor. The inset shows the extrapolation of the curve to 10 years projection.

4.5 Conclusions

Memory functions of the nc-MoO_x embedded ZrHfO high-*k* MOS capacitor have been studied. The *C-V* hysteresis and *J-V* data show that holes were trapped to the nc-MoO_x site of the high-*k* stack. The charge trapping capability is affected by the supply of the charges from the wafer substrate. Holes trapped from the negative gate voltage stress could not be completely erased by a positive gate voltage with the same magnitude. The *J-V* curve of the nc-MoO_x embedded sample confirmed that those loosely trapped holes were released easily. The Coulomb blockade effect was observed under the electron trapping condition. The charge retention study shows that about 20% of the trapped holes were loosely trapping and more than half of those originally trapped holes remained in the device after 10 years. In principle, the nc-MoO_x embedded ZrHfO high-*k* capacitor is

a unique memory device of which the operation is based on hole trapping and detrapping. For the giga level application, in addition to the enlargement of the storage capacity through the optimization of the structure and operation parameters, detailed charge trapping site and reliability issues need to be investigated.

CHAPTER 5: BAYESIAN ANALYSIS FOR ACCELERATED LIFE TESTS USING
DIRICHLET PROCESS WEIBULL MIXTURE MODEL [†]

This chapter develops the Dirichlet process Weibull mixture model for constant-stress ALT. The model is employed to predict failure-time distribution at a given stress level. A simulation-based model fitting algorithm that implements Gibbs sampling is developed to analyze complete and right-censored ALT data and to predict the failure-time distribution at the normal stress level. Two practical examples related to the reliability of nanoelectronic devices are presented. The results have demonstrated that the proposed methodology is capable of providing accurate prediction of the failure-time distribution at the normal stress level without assuming any restrictive parametric failure-time distribution.

5.1 Methodologies

5.1.1 Dirichlet process Weibull mixture ALT model

Section 1.6.3 introduced the general form of Dirichlet process mixture model. In this section, a Dirichlet process mixture model with the Weibull kernel is developed to model failure-time distribution at a given stress level, and a log-linear regression model is assumed to describe the relationship of failure time under various stress levels.

Assume n items are tested in ALT and let $\mathbf{d} = \{(t_i, x_i, \delta_i), i = 1, 2, \dots, n\}$ denote the ALT data, where the i th unit is tested at the stress x_i . Note that x_i may be a transformation observation, and δ_i equals zero when t_i is a right-censored observation. Kuo and

[†]©2014 IEEE. Reprinted, with permission, from Tao Yuan, Xi Liu, Saleem Z. Ramadan, and Yue Kuo, Bayesian Analysis for Accelerated Life Tests Using a Dirichlet Process Weibull Mixture Model, IEEE Transactions on Reliability, vol. 63, No.1, March 2014

Mallick [18] and Ghost and Ghosal [106] considered the following semiparametric linear regression model for ALT

$$\ln t_i = -x_i\beta + w_i, i = 1, 2, \dots, n. \quad (5.1)$$

where x_i is the i th stress level, β is the coefficient of stress level, and w_i is the error term. If the error terms $w_i, i = 1, 2, \dots, n$ are assumed to be s -independent and identically distributed (*i.i.d.*) from the smallest extreme value distribution, or equivalently, if $v_i = \exp(w_i), i = 1, 2, \dots, n$ are *i.i.d.* Weibull random variables, the model given by Eq. (5.1) is the widely used Weibull ALT model with a log-linear lifetime-stress relationship [1], [118]. Kuo and Mallick [18] modeled the distribution of v_i by the DP mixture model with the normal kernel and the lognormal kernel. Ghost and Ghosal [106] modeled the distribution of v_i by the DP mixture model with the Weibull kernel with a fixed shape parameter.

This study extends the work of Kuo and Mallick [18] and Ghost and Ghosal [106], and uses the DP mixture model with the Weibull kernel, where the shape parameter is not fixed, to model the distribution of v_i . Kottas [19] pointed out that using Weibull kernel has the computational advantage over the normal and lognormal kernels when dealing with censoring because the Weibull kernel has a closed form CDF. In addition, mixing on both the shape and scale parameters of the Weibull kernel can result in a flexible mixture that can model a wide range of distributional shapes [19].

Assuming that $v_i, i = 1, 2, \dots, n$, are *iid* from the following PDF

$$f(v_i|G) = \int k(v_i|\alpha_i, \lambda_i)G(d\alpha_i, d\lambda_i), \quad (5.2)$$

where $k(v_i|\alpha_i, \lambda_i)$ is the PDF of the Weibull kernel. The PDF and CDF of the Weibull kernel are given by

$$k(v|\alpha, \lambda) = \lambda^{-1} \alpha v^{\alpha-1} \exp(-\lambda^{-1} v \alpha), \quad (5.3)$$

and

$$K(v|\alpha, \lambda) = 1 - \exp(-\lambda^{-1} v \alpha), \quad (5.4)$$

respectively, where $\alpha > 0$ is the shape parameter, and $\lambda > 0$ is the scale parameter. From $f(v_i|G)$ given by Eq. (5.2), and according to the log-linear relationship function, $t_i = v_i \exp(-x_i \beta)$, the PDF of t_i can be derived through the transformation of random variables:

$$f(t_i|\beta, G) = f(v_i|G) \left| \frac{dv_i}{dt_i} \right| = \int \exp(x_i \beta) k(t_i \exp(x_i \beta)|\alpha_i, \lambda_i) G(d\alpha_i, d\lambda_i). \quad (5.5)$$

The CDF of t_i then can be derived as:

$$\begin{aligned} F(t_i|\beta, G) &= \int_0^{t_i} \int \exp(x_i \beta) k(s \exp(x_i \beta)|\alpha_i, \lambda_i) G(d\alpha_i, d\lambda_i) ds \\ &= \int \int_0^{t_i} \exp(x_i \beta) k(s \exp(x_i \beta)|\alpha_i, \lambda_i) ds G(d\alpha_i, d\lambda_i) \\ &= \int K(t_i \exp(x_i \beta)|\alpha_i, \lambda_i) G(d\alpha_i, d\lambda_i). \end{aligned} \quad (5.6)$$

The base distribution G_0 can be considered as prior guess on α and λ , and in this dissertation the following G_0 is adopted:

$$G_0(\alpha, \lambda) = \text{Uniform}(\alpha|0, \phi) \times \text{inverse-Gamma}(\lambda|d, \gamma). \quad (5.7)$$

This base distribution can offer both computational convenience and modeling flexibility [19]. The inverse-Gamma distribution is the conditional conjugate prior for the Weibull scale parameter λ when the shape parameter is known. There is no nature conjugate prior for the Weibull shape parameter. $d=2$ is set so that the inverse-Gamma distribution has an infinite variance, which can convey the lack of prior knowledge.

In summary, the DP Weibull mixture ALT model can be written in the following hierarchical form:

$$\begin{aligned}
t_i | \alpha_i, \lambda_i, \beta &\sim \begin{cases} \exp(x_i \beta) k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) & \text{if } \delta_i = 1, \\ 1 - K(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) & \text{if } \delta_i = 0, \end{cases} & i = 1, 2, \dots, n, \\
(\alpha_i, \lambda_i) &\sim G(\alpha, \lambda), i = 1, \dots, n, \\
G(\alpha, \lambda) &\sim DP(\mu, G_0(\alpha, \lambda)), \\
\beta &\sim f(\beta | a_\beta, b_\beta), \\
\phi &\sim f(\phi | a_\phi, b_\phi), \\
\gamma &\sim f(\gamma | a_\gamma, b_\gamma), \\
\mu &\sim f(\mu | a_\mu, b_\mu),
\end{aligned} \tag{5.8}$$

where $f(\beta | a_\beta, b_\beta)$, $f(\phi | a_\phi, b_\phi)$, $f(\gamma | a_\gamma, b_\gamma)$, and $f(\mu | a_\mu, b_\mu)$ denote the prior distributions for β , ϕ , γ , and μ , respectively. δ_i denotes the indicator function. $\delta_i=1$ indicates exact failure time and $\delta_i=0$ indicates the censored observation. The following prior distributions are assumed:

$$\begin{aligned}
\beta &\sim \text{Normal}(a_\beta, b_\beta), \\
\phi &\sim \text{Pareto}(a_\phi, b_\phi), \\
\gamma &\sim \text{Gamma}(a_\gamma, b_\gamma), \\
\mu &\sim \text{Gamma}(a_\mu, b_\mu).
\end{aligned}$$

The Pareto distribution and Gamma distribution are the conjugate prior distributions for ϕ and γ , respectively. Using the Gamma prior for the precision parameter μ of the DP prior can result in a very attractive computational convenience [93]. A normal prior is assumed for β because β may be negative, especially when x_i is a transformation of the stress.

The proposed DP Weibull mixture ALT model handles censoring differently from the work of Kuo and Mallick [18] and Ghost and Ghosal [106]. The previous two studies

used an imputation method to replace the censored observations with simulated failure times. This dissertation uses the reliability function of t_i if t_i is a censored observation in the first stage of the hierarchical model (5.8). This is intuitively appealing because it is consistent with the general approach to deal with censoring in parametric data analysis, in which the likelihood contribution from a right-censored observation is the reliability function.

5.1.2 Simulation-based model fitting

This section presents a simulation-based algorithm for fitting the DP Weibull mixture ALT model. Gibbs sampling has become the standard algorithm for fitting the DP mixture models as a useful tool for high dimensional problems with many parameters. Each iteration of the Gibbs sampling cycles through the unknown parameters, sampling a value of one parameter conditioning on the latest values of all the other parameters. When the number of iterations is large enough, the sample drawn on one parameter can be regarded as simulated values from its marginal posterior distribution. Sample statistics can then be used to formulate posterior inference on that parameter [119].

A key feature of the DP mixture models is the discreteness of $G(\alpha, \lambda)$ under the DP assumption, inducing clustering of (α_i, λ_i) 's [15], [93]. Denote n^* as the number of clusters in (α_i, λ_i) , $i = 1, 2, \dots, n$, and $(\alpha_j^*, \lambda_j^*)$, $j = 1, 2, \dots, n^*$, as the distinct clusters. The vector of indicators $\mathbf{c} = \{c_1, c_2, \dots, c_n\}$ is introduced to indicate the clustering configuration. $c_i = j$ when $(\alpha_i, \lambda_i) = (\alpha_j^*, \lambda_j^*)$, indicating that the i th observation belongs to the j th cluster. Let n_j denote the number of members in the j th cluster.

This dissertation develops a Gibbs sampling algorithm to fit the model (5.8). Each iteration of the algorithm consists of the following steps:

(a) Draw (α_i, λ_i) from the conditional posterior distribution denoted by:

$$f(\alpha_i, \lambda_i | \{(\alpha_{i'}, \lambda_{i'}, c_{i'}), i' \neq i\}, \phi, \gamma, \mu, \beta, \mathbf{d}), \text{ for } i = 1, \dots, n;$$

(b) Adjust the cluster locations $(\alpha_j^*, \lambda_j^*)$ by sampling from conditional posterior

$$\text{distribution: } f(\alpha_j^*, \lambda_j^* | \mathbf{c}, \phi, \gamma, \mu, \beta, \mathbf{d}), \text{ for } j = 1, \dots, n^*;$$

(c) Update ϕ , γ , and μ by sampling from their conditional posterior distributions;

(d) Sample β from its conditional posterior distribution;

(e) Sample $F(t|\beta, G, x_0)$, the failure-time CDF at the normal stress level x_0 .

Details of these five steps will be described in the remainder of this section.

A. Step (a): Update (α_i, λ_i)

Step (a) draws new values of (α_i, λ_i) and updates the clustering location indicator c_i for each observation t_i . The new values of (α_i, λ_i) can either be one of the $(\alpha_{i'}, \lambda_{i'})$ $i' \neq i$, or could be new values drawn from G_0 . Denote n^{*-} as the number of clusters in $\{(\alpha_{i'}, \lambda_{i'}), i' \neq i\}$, i.e., the number of clusters when (α_i, λ_i) is removed from $\{(\alpha_1, \lambda_1), (\alpha_2, \lambda_2), \dots, (\alpha_n, \lambda_n)\}$, and let $(\alpha_j^{*-}, \lambda_j^{*-}), j = 1, \dots, n^{*-}$, denote the distinct clusters in $\{(\alpha_{i'}, \lambda_{i'}), i' \neq i\}$.

Also denote n_j^- as the number of members in the cluster $(\alpha_j^{*-}, \lambda_j^{*-})$, for $j = 1, \dots, n^{*-}$.

If t_i is an exact failure-time observation (i.e., if $\delta_i = 1$), the conditional posterior distribution in Step (a) has the following mixture form:

$$\begin{aligned}
& f(\alpha_i, \lambda_i | \{(\alpha_{i'}, \lambda_{i'}, c_{i'}), i' \neq i\}, \phi, \gamma, \mu, \beta, t_i, x_i) \\
&= \frac{q_0^o h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) + \sum_{j=1}^{n^*} n_j^- q_j^o \Delta_{(\alpha_j^*, \lambda_j^*)}}{q_0^o + \sum_{j=1}^{n^*} n_j^- q_j^o}, \tag{5.9}
\end{aligned}$$

where $\Delta_{(\alpha_j^*, \lambda_j^*)}$ denotes a point mass at $(\alpha_j^*, \lambda_j^*)$, and

$$q_j^o = k(t_i \exp(x_i \beta) | \alpha_j^*, \lambda_j^*) \exp(x_i \beta) \propto k(t_i \exp(x_i \beta) | \alpha_j^*, \lambda_j^*), \tag{5.10}$$

and

$$\begin{aligned}
q_0^o &= \mu \int \exp(x_i \beta) k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) G_0(\alpha_i, \lambda_i) d\alpha_i d\lambda_i \\
&\propto \mu \int_0^\phi \int_0^\infty k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) G_0(\alpha_i, \lambda_i) d\alpha_i d\lambda_i \\
&= \mu \int_0^\phi \int_0^\infty \lambda_i^{-1} \alpha_i [t_i \exp(x_i \beta)]^{\alpha_i - 1} \exp(-\lambda_i^{-1} [t_i \exp(x_i \beta)]^{\alpha_i}) \frac{1}{\phi} \frac{\gamma^d}{\Gamma(d)} \lambda_i^{-d-1} \exp(-\frac{\gamma}{\lambda_i}) d\alpha_i d\lambda_i \\
&= \frac{d\mu\gamma^d}{\phi} \int_0^\phi \frac{\alpha_i [t_i \exp(x_i \beta)]^{\alpha_i - 1}}{(\gamma + [t_i \exp(x_i \beta)]^{\alpha_i})^{d+1}} d\alpha_i, \tag{5.11}
\end{aligned}$$

which can be easily computed numerically. The superscript o is used to indicate that t_i is an exact failure time observation. Therefore, (α_i, λ_i) equals $(\alpha_j^*, \lambda_j^*)$ with the probability

of $n_j^- q_j^o / (q_0^o + \sum_{j=1}^{n^*} n_j^- q_j^o)$, and with the probability of $q_0^o / (q_0^o + \sum_{j=1}^{n^*} n_j^- q_j^o)$ are new values

drawn from $h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i)$. Herein,

$$\begin{aligned}
h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) &\propto k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) G_0(\alpha_i, \lambda_i) \\
&\propto k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) \frac{1}{\phi} \frac{\gamma^d}{\Gamma(d)} \lambda_i^{-d-1} \exp(-\frac{\gamma}{\lambda_i}) \mathbf{I}_{\{\alpha_i \in (0, \phi)\}}, \tag{5.12}
\end{aligned}$$

where \mathbf{I}_\emptyset is an indicator function. $h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i)$ can be expressed as:

$$h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) = f(\alpha_i | \phi, \gamma, \mu, \beta, t_i, x_i) f(\lambda_i | \alpha_i, \phi, \gamma, \mu, \beta, t_i, x_i),$$

where

$$\begin{aligned} f(\alpha_i | \phi, \gamma, \mu, \beta, t_i, x_i) &= \int_0^\infty h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) d\lambda_i \\ &\propto \frac{\alpha_i [t_i \exp(x_i \beta)]^{\alpha_i - 1}}{(\gamma + [t_i \exp(x_i \beta)]^{\alpha_i})^{d+1}} \mathbf{I}_{\{\alpha_i \in (0, \phi)\}}, \end{aligned} \quad (5.13)$$

and

$$\begin{aligned} f(\lambda_i | \alpha_i, \phi, \gamma, \mu, \beta, t_i, x_i) &\propto \lambda_i^{-d-2} \exp\left(-\frac{\gamma + [\mu_i \exp(x_i \beta)]^{\alpha_i}}{\lambda_i}\right) \\ &\sim \text{inverse-Gamma}(d+1, \gamma + [t_i \exp(x_i \beta)]^{\alpha_i}). \end{aligned} \quad (5.14)$$

In order to draw new values of (α_i, λ_i) from $h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i)$, first a value for α_i is sampled from $f(\alpha_i | \phi, \gamma, \mu, \beta, t_i, x_i)$ given by equation (5.13) by discretization, and then a value for λ_i is drawn from the inverse-Gamma distribution given by equation (5.14).

If t_i is a right-censored observation, (i.e., if $\delta_i = 0$), the conditional posterior distribution $f(\alpha_i, \lambda_i | \{(\alpha_{i'}, \lambda_{i'}, c_{i'}), i' \neq i\}, \phi, \gamma, \mu, \beta, t_i, x_i)$, can be derived in a similar way by replacing $k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i)$ with $1 - K(t_i \exp(x_i \beta) | \alpha_i, \lambda_i)$. Because the Weibull kernel has the closed form CDF, censoring can be easily handled in the computation as

$$\begin{aligned} &f(\alpha_i, \lambda_i | \{(\alpha_{i'}, \lambda_{i'}, c_{i'}), i' \neq i\}, \phi, \gamma, \mu, \beta, t_i, x_i) \\ &= \frac{q_0^c h^c(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) + \sum_{j=1}^{n^*} n_j^- q_j^c \Delta_{(\alpha_j^*, \lambda_j^-)}}{q_0^c + \sum_{j=1}^{n^*} n_j^- q_j^c}, \end{aligned} \quad (5.15)$$

where the superscript c is used to indicate that t_i is a right-censored observation,

$$q_j^c = 1 - K(t_i \exp(x_i \beta) | \alpha_j^{*-}, \lambda_j^{*-}),$$

$$\begin{aligned} q_0^c &= \mu \int_0^\phi \int_0^\infty [1 - K(t_i \exp(x_i \beta) | \alpha_i, \lambda_i)] G_0(\alpha_i, \lambda_i) d\alpha_i d\lambda_i \\ &= \frac{\mu \gamma^d}{\phi} \int_0^\phi \frac{1}{(\gamma + [t_i \exp(x_i \beta)]^{\alpha_i})^{d+1}} d\alpha_i, \end{aligned} \quad (5.16)$$

and

$$h^o(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) \propto [1 - K(t_i \exp(x_i \beta) | \alpha_i, \lambda_i)] G_0(\alpha_i, \lambda_i).$$

Express $h^c(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i)$ as:

$$h^c(\alpha_i, \lambda_i | \phi, \gamma, \mu, \beta, t_i, x_i) = f(\alpha_i | \phi, \gamma, \mu, \beta, t_i, x_i) f(\lambda_i | \alpha_i, \phi, \gamma, \mu, \beta, t_i, x_i),$$

it can be shown that

$$f(\alpha_i | \phi, \gamma, \mu, \beta, t_i, x_i) \propto \frac{1}{(\gamma + [t_i \exp(x_i \beta)]^{\alpha_i})^{d+1}} \mathbf{I}_{\{\alpha_i \in (0, \phi)\}}, \quad (5.17)$$

and

$$f(\lambda_i | \alpha_i, \phi, \gamma, \mu, \beta, t_i, x_i) \sim \text{inverse-Gamma}(d, \gamma + [t_i \exp(x_i \beta)]^{\alpha_i}). \quad (5.18)$$

After (α_i, λ_i) are sampled from the conditional posterior distribution given by equation

(5.9) or (5.15), the configuration indicator c_i and the cluster locations $(\alpha_j^*, \lambda_j^*)$ are updated

accordingly.

B. Step (b): update $(\alpha_j^, \lambda_j^*)$*

Once Step (a) is completed for all n observations, the clustering configuration vector \mathbf{c} and the cluster locations $(\alpha_j^*, \lambda_j^*)$, $j = 1, 2, \dots, n^*$ have been updated. Step (b)

adjusts the clustering locations conditioning on \mathbf{c} , ϕ , γ , μ , β , and \mathbf{d} . For each cluster location α_j^*, λ_j^* , let \mathbf{t}_j^o and \mathbf{t}_j^c denote, respectively, the set of exact failure time observations and the set of right censored observations in the j th cluster, i.e., $\mathbf{t}_j^o \equiv \{t_i : c_i = j \text{ and } \delta_i = 1\}$ and $\mathbf{t}_j^c \equiv \{t_i : c_i = j \text{ and } \delta_i = 0\}$. In addition, let $\mathbf{t}_j = \mathbf{t}_j^o \cup \mathbf{t}_j^c \equiv \{t_i : c_i = j\}$ denote all the observations belonging to the j th cluster. The conditional posterior distribution for $(\alpha_j^*, \lambda_j^*)$ is given by:

$$f(\alpha_j^*, \lambda_j^* | \phi, \gamma, \beta, \mathbf{c}, \mathbf{d}) \propto G_0(\alpha_j^*, \lambda_j^* | \phi, \gamma) \times \prod_{t_i \in \mathbf{t}_j^o} k(t_i \exp(x_i \beta) | \alpha_j^*, \lambda_j^*) \times \prod_{t_i \in \mathbf{t}_j^c} [1 - K(t_i \exp(x_i \beta) | \alpha_j^*, \lambda_j^*)]. \quad (5.19)$$

To draw the values for $(\alpha_j^*, \lambda_j^*)$, first draw a value for λ_j^* from $f(\lambda_j^* | \alpha_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d})$, and then draw a value for α_j^* from $f(\alpha_j^* | \lambda_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d})$. It can be shown that

$$\begin{aligned} f(\lambda_j^* | \alpha_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d}) &\propto (\lambda_j^*)^{-d-1} \exp\left(-\frac{\gamma}{\lambda_j^*}\right) \times \prod_{t_i \in \mathbf{t}_j^o} (\lambda_j^*)^{-1} \exp(-(\lambda_j^*)^{-1} (t_i \exp(x_i \beta))^{\alpha_j^*}) \\ &\quad \times \prod_{t_i \in \mathbf{t}_j^c} \exp(-(\lambda_j^*)^{-1} (t_i \exp(x_i \beta))^{\alpha_j^*}). \\ &\propto (\lambda_j^*)^{-d-n_j^o-1} \exp\left[-(\lambda_j^*)^{-1} \left(\gamma + \sum_{t_i \in \mathbf{t}_j} (t_i \exp(x_i \beta))^{\alpha_j^*}\right)\right] \\ &\sim \text{inverse-Gamma}(d + n_j^o, \gamma + \sum_{t_i \in \mathbf{t}_j} (t_i \exp(x_i \beta))^{\alpha_j^*}), \end{aligned} \quad (5.20)$$

where n_j^o denotes the number of exact failure time observations in the j th cluster, and

$$\begin{aligned}
f(\alpha_j^* | \lambda_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d}) &\propto (\alpha_j^*)^{n_j^o} \times \mathbf{I}_{\{\alpha_j^* \in (0, \phi)\}} \times \prod_{t_i \in \mathbf{t}_j^o} (t_i \exp(x_i \beta))^{\alpha_j^*} \\
&\times \prod_{t_i \in \mathbf{t}_j} \exp(-(\lambda_j^*)^{-1} (t_i \exp(x_i \beta))^{\alpha_j^*}),
\end{aligned} \tag{5.21}$$

which is a non-standard density function. The data augmentation method discussed by Damien et al. [120] can be implemented here to sample a value from $f(\alpha_j^* | \lambda_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d})$. Introduce auxiliary variables $\mathbf{u}_1 \equiv \{u_{1,i} : t_i \in \mathbf{t}_j^o\}$, and $\mathbf{u}_2 \equiv \{u_{2,i} : t_i \in \mathbf{t}_j\}$ so that the joint density can be expressed as:

$$\begin{aligned}
f(\alpha_j^*, \mathbf{u}_1, \mathbf{u}_2 | \lambda_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d}) &\propto (\alpha_j^*)^{n_j^o} \mathbf{I}_{\{\alpha_j^* \in (0, \phi)\}} \times \prod_{t_i \in \mathbf{t}_j^o} \mathbf{I}_{\{0 < u_{1,i} < (t_i \exp(x_i \beta))^{\alpha_j^*}\}} \\
&\times \prod_{t_i \in \mathbf{t}_j} \mathbf{I}_{\{0 < u_{2,i} < \exp(-(\lambda_j^*)^{-1} (t_i \exp(x_i \beta))^{\alpha_j^*})\}}.
\end{aligned} \tag{5.22}$$

The Gibbs sampling can be extended to first draw $u_{1,i}$ uniformly on the interval $(0, (t_i \exp(x_i \beta))^{\alpha_j^*})$, and $u_{2,i}$ uniformly on the interval $(0, \exp(-(\lambda_j^*)^{-1} (t_i \exp(x_i \beta))^{\alpha_j^*}))$, and then draw α_j^* from

$$f(\alpha_j^* | \mathbf{u}_1, \mathbf{u}_2, \lambda_j^*, \phi, \gamma, \mu, \beta, \mathbf{c}, \mathbf{d}) \propto (\alpha_j^*)^{n_j^o}, \tag{5.23}$$

restricted to the interval determined by

$$(0, \phi) \cap \{u_{1,i} < (t_i \exp(x_i \beta))^{\alpha_j^*}, \forall t_i \in \mathbf{t}_j^o\} \cap \{u_{2,i} < \exp(-(\lambda_j^*)^{-1} (t_i \exp(x_i \beta))^{\alpha_j^*}), \forall t_i \in \mathbf{t}_j\}.$$

This can be easily done using the inverse-CDF method.

C. Step(c): update ϕ , γ , and μ

Step (c) draw values for ϕ , γ , and μ from their conditional posterior distributions. Escobar and West [93] developed an augmentation method for sampling precision parameter μ from its conditional posterior distribution when the Gamma prior is assumed

for μ . An auxiliary variable u is introduced such that $u \mid \mu, \mathbf{c}, \mathbf{d} \sim \text{Beta}(\mu+1, n)$. Then the new value for μ is sampled from a mixture Gamma posterior distribution expressed as

$$p\text{Gamma}(a_\mu + n^*, b_\mu - \log u) + (1-p)\text{Gamma}(a_\mu + n^* - 1, b_\mu - \log u), \quad (5.24)$$

where $p = (a_\mu + n^* - 1) / (n(b_\mu - \log u) + a_\mu + n^* - 1)$.

The prior distribution for ϕ is pareto distribution, therefore according to Bayes Law the conditional posterior distribution for ϕ can be written as:

$$\begin{aligned} f(\phi \mid a_\phi, b_\phi, \alpha_j^*) &\propto \pi(\phi) f(\alpha_j^* \mid \phi) = \frac{a_\phi b_\phi^{a_\phi}}{\phi^{a_\phi+1+n^*}} \mathbf{I}_{\{\phi \geq b_\phi\}} \times \prod_{j=1}^{n^*} \frac{1}{\phi} \mathbf{I}_{\{\phi \geq \alpha_j^*\}} \propto \frac{1}{\phi^{a_\phi+1+n^*}} \mathbf{I}_{\{\phi \geq b_\phi\}} \mathbf{I}_{\{\phi \geq \alpha_j^*\}} \\ &\sim \text{Pareto}(\phi \mid a_\phi + n^*, \max(b_\phi, \max\{\alpha_j^*, j=1, 2, \dots, n^*\})). \end{aligned} \quad (5.25)$$

Similarly, given the prior of gamma distribution, the conditional posterior distribution for γ is:

$$\begin{aligned} f(\gamma \mid a_\gamma, b_\gamma, \lambda_j^*) &\propto \pi(\gamma) f(\lambda_j^* \mid \gamma) = \frac{\gamma^{a_\gamma-1} \exp(-b_\gamma \gamma) b_\gamma^{a_\gamma}}{\Gamma(a_\gamma)} \times \prod_{j=1}^{n^*} \frac{\gamma^d}{\Gamma(d)} \exp(-\frac{\gamma}{\lambda_j^*}) (\lambda_j^*)^{-(d+1)} \\ &\propto \gamma^{(a_\gamma+n^*d-1)} \exp(-(b_\gamma + \sum_{j=1}^{n^*} (\lambda_j^*)^{-1}) \gamma) \\ &\sim \text{Gamma}(a_\gamma + n^*d, b_\gamma + \sum_{j=1}^{n^*} (\lambda_j^*)^{-1}). \end{aligned} \quad (5.26)$$

D. Step (d): update β

Step (d) samples a new value for β from its conditional posterior distribution

$$\begin{aligned} f(\beta \mid \{(\alpha_i, \lambda_i), i=1, 2, \dots, n\}, \mathbf{d}) \\ &\propto f(\beta \mid a_\beta, b_\beta) \times \prod_{\delta_i=1} \exp(x_i \beta) k(t_i \exp(x_i \beta) \mid \alpha_i, \lambda_i) \times \prod_{\delta_i=0} [1 - K(t_i \exp(x_i \beta) \mid \alpha_i, \lambda_i)] \quad (5.27) \\ &\propto f(\beta \mid a_\beta, b_\beta) \times \prod_{\delta_i=1} \exp(\alpha_i x_i \beta) \times \prod_{i=1}^n \exp(-\lambda_i^{-1} t_i^{\alpha_i} \exp(\alpha_i x_i \beta)). \end{aligned}$$

The data augmentation method can be applied again to simulate a value from $f(\beta|\{(\alpha_i, \lambda_i), i=1,2,\dots,n\}, \mathbf{d})$. For each $i \in \{i: \delta_i=1\}$, draw an auxiliary variable $u_{1,i}$ uniformly on the interval $(0, \exp(\alpha_i x_i \beta))$. For each $i, i=1,2,\dots,n$, draw an auxiliary variable $u_{2,i}$ uniformly on the interval $(0, \exp(-\lambda_i^{-1} t_i^{\alpha_i} \exp(\alpha_i x_i \beta)))$. Then a value for β is sampled from $f(\beta|a_\beta, b_\beta)$ restricted to the interval

$$\left(\max_{\{i: \delta_i=1\}} \{\log u_{1,i} / \alpha_i x_i\}, \min_{\{i=1,2,\dots,n\}} \{(\alpha_i x_i)^{-1} \log(-\lambda_i^{-1} t_i^{\alpha_i} \log u_{2,i})\} \right).$$

It is routine to sample from the truncated normal distribution.

E. Step (e): update the failure-time distribution

One of the objectives of ALT analysis is to predict the failure-time distribution at the normal stress level x_0 . Kuo and Mallick [18] proposed two methods for predicting the failure-time CDF at x_0 . The first method simulates a sample of failure times at the normal stress level x_0 from the predictive density, and constructs the failure-time distribution at x_0 from the simulated failure times using some empirical estimators, such as the Kaplan-Meier estimator. The second method uses the functional form of the kernel for evaluation, and needs to combine draws from multiple chains of the Gibbs sampling. In this dissertation, the method proposed by Kottas [19] is extended. Once Steps (a)-(d) have completed in one iteration of Gibbs sampling, Step (e) draws values for the failure-time CDF at x_0 conditional on the values of $\{(\alpha_i, \lambda_i), i = 1, 2, \dots, n\}$, ϕ , γ , μ , and β obtained at that iteration. According to equation (5.6), the failure-time CDF at the normal stress level x_0 is given by:

$$F(t|\beta, G, x_0) = \int K(t \exp(x_0 \beta) | \alpha, \lambda) G(d\alpha, d\lambda). \quad (5.28)$$

The CDF given by equation (5.28) can be approximated by a finite mixture with a large number of mixing components, i.e., $F(t_m|\beta, G, x_0) \approx \sum_{l=1}^L \varpi_l K(t_m \exp(x_0 \beta) | \alpha_l', \lambda_l')$, for a prespecified grid of values $t_m, m = 1, 2, \dots, N$, over the support of $F(t|\beta, G, x_0)$, where L is the number of mixing components, and (α_l', λ_l') , $l = 1, 2, \dots, L$, are *i.i.d.* draws from the mixture $(\mu + n)^{-1} [\mu G_0(\alpha, \lambda | \phi, \gamma) + \sum_{i=1}^n \Delta_{\{\alpha_i, \lambda_i\}}]$. The weight coefficients $\varpi_l, l = 1, 2, \dots, L$ are simulated as follows. Introducing an auxiliary variable z_l , first simulate z_l *i.i.d.* from $\text{Beta}(1, \mu + n)$, and then compute ϖ_l according to $\varpi_1 = z_1, \varpi_l = z_l \prod_{s=1}^{l-1} z_s$, for $l = 1, 2, \dots, L-1$, and $\varpi_L = 1 - \sum_{j=1}^{L-1} \varpi_j$. The above procedure is based on the constructive definition of DP, which was discussed by Sethuraman [121]. $L = 2000$ is used. A sensitivity analysis has proven that 2000 is a reliable and conservative choice for L .

5.2 Illustrative Examples

This section uses two practical examples to illustrate the proposed ALT model and algorithm. Both examples use experimental data collected at the Thin Film Nano and Microelectronics Research Laboratory at Texas A&M University, College Station. The first example studies the reliability of a new mixed oxides high- k dielectric material for nanoelectronic applications [11]. The second example evaluates the reliability of a novel nanocrystals-embedded high- k nonvolatile memory device [113]. Metal-Oxide-Semiconductor (MOS) capacitors with the high- k dielectric film were subjected to accelerated electrical stresses.

For the purpose of comparison, a parametric ALT analysis is also performed. It is generally accepted that the breakdown of gate dielectrics belongs to the class of weakest-link problem of extreme value statistics, and thus is assumed to have the Weibull failure-time distribution at a given stress level [122]. As introduced in section 1.4, E -model is a widely used physical lifetime-stress relationship, which relates the time-to-breakdown to the electrical field stress. The E -model is actually a log-linear lifetime-stress relationship. Therefore, the Weibull distribution with the log-linear lifetime-stress relationship is frequently used as the parametric ALT model for studying the time-dependent dielectric breakdown. The standard MLE method is used to fit the Weibull log-linear ALT model and to predict the failure-time CDF at $x_0 = 7.1$ MV/cm.

Because of the absence of any prior knowledge, we choose noninformative priors to reflect our absence of prior knowledge. Apply the Gamma(1,0.001) prior for μ and γ , and the Normal(0, 10^6) prior for β . Both Gamma(1,0.001) and Normal(0, 10^6) are widely used diffuse priors in the sense of not favoring any value [123]. For ϕ , Pareto (1, 1) prior is used, which has an infinite variance to reflect large prior uncertainty due to absence of prior knowledge. The algorithm is coded in Matlab®. The Gibbs sampling algorithm runs for 10,000 iterations, and the first 5,000 iterations are discarded before data analysis. Convergence is verified by running multiple chains from diverse starting points, examining the trace plots, and monitoring the Gelman-Rubin statistics [119]. Posterior prediction on the failure-time CDF at x_0 is based on the posterior median.

Figure 5.1 shows the failure-time CDF at the normal stress level predicted by the proposed DP Weibull mixture ALT model and the Weibull log-linear ALT model. Using

the experimental data collected at $x_0 = 7.1$ MV/cm, an empirical CDF is computed from the median rank estimator, i.e.,

$$\hat{F}(t_i) = \frac{i - 0.3}{n + 0.4}, \quad i = 1, 2, \dots, n,$$

where $t_{(i)}$, $i = 1, 2, \dots, n$, denote the ordered failure times. The proposed DP Weibull mixture ALT model provides much better predictive result than the Weibull log-linear ALT model. The Weibull log-linear ALT model fails to accurately predict the failure-time distribution at the normal stress level. This may be caused by the involvement of multiple failure modes.

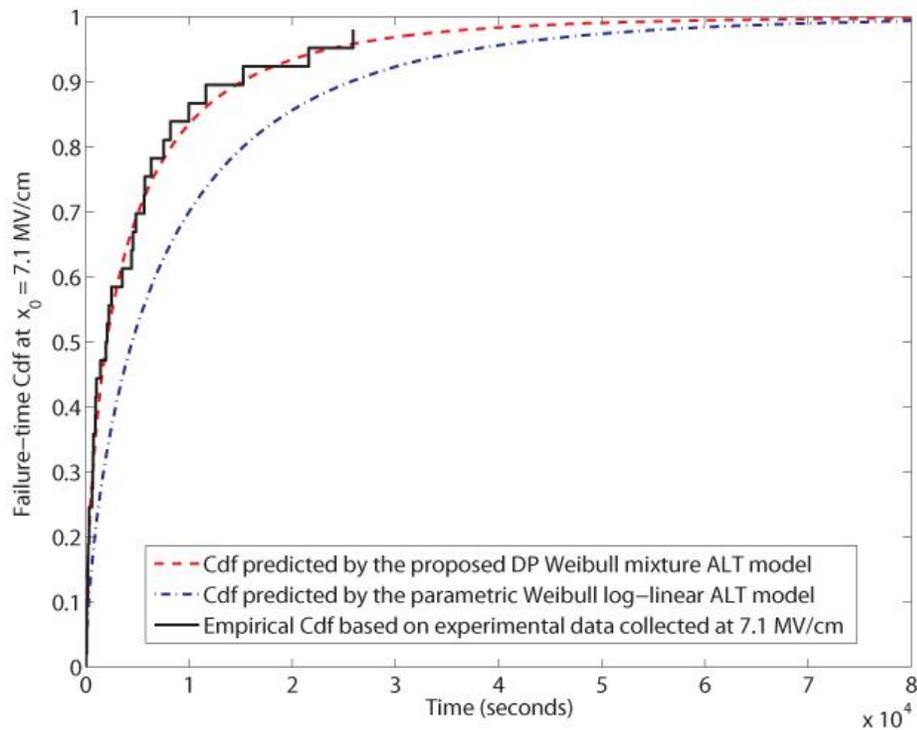


Figure 5.1. Example 1: predicted failure-time CDF at the normal stress $x_0 = 7.1$ MV/cm.

Figure 5.2 shows the Weibull fit to the data collected at 7.9 MV/cm and the empirical failure-time CDF at 7.9 MV/cm. The Weibull distribution does not fit the data adequately. Degraeve et al. [124] proposed a bimodal breakdown model with the following failure-time PDF

$$f(t|p, \alpha_i, \lambda_i, \alpha_a, \lambda_e) = p[f(t|\alpha_a, \lambda_e)R(t|\alpha_i, \lambda_i) + f(t|\alpha_i, \lambda_i)R(\alpha_a, \lambda_e)] + (1-p)f(t|\alpha_i, \lambda_i),$$

where the subscripts e and i denote, respectively, extrinsic breakdown and intrinsic breakdown, p is the fraction of devices with defects, and f and R are the Weibull PDF and the Weibull reliability function, respectively. For defect-free devices, the failures occur intrinsically. For devices with defects, the extrinsic and intrinsic failure modes are in competition with each other. Figure 5.2 indicates that the bimodal breakdown model fits the data better than the Weibull model. Because the Weibull distribution is not adequate to describe the failure-time distribution at a given stress level, the Weibull log-linear ALT model fails to accurately predict the failure-time distribution at the normal stress level. On the other hand, the DP Weibull mixture ALT model describes the failure-time distribution nonparametrically using the DP Weibull mixture model. This modeling flexibility leads to an improved prediction of the failure-time distribution at the normal stress level.

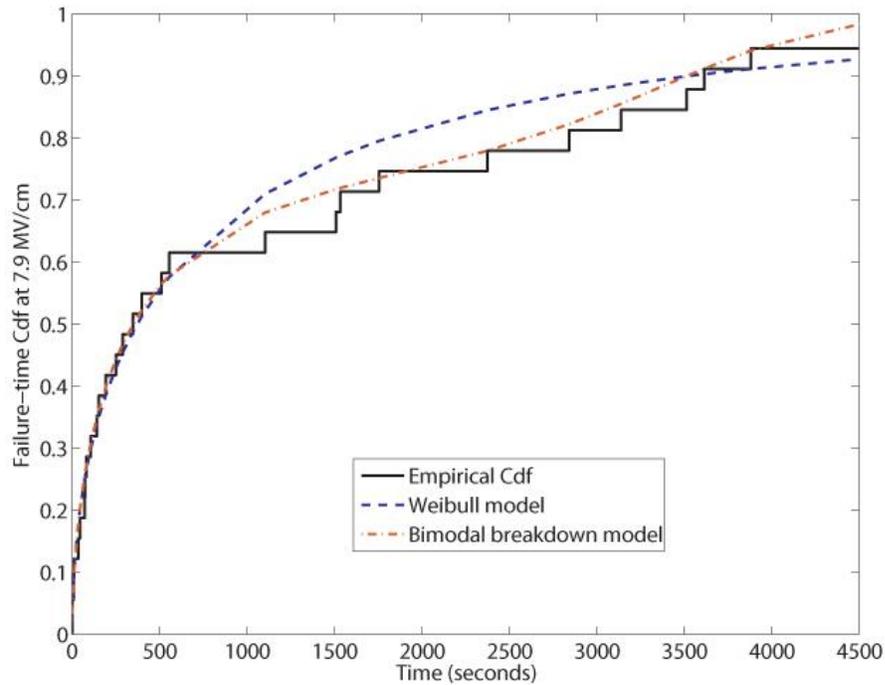


Figure 5.2. Example 1: estimation of failure-time CDF at 7.9 MV/cm

5.2.2 Right-censored data set example

Table 5.2 lists the times-to-breakdown of memories tested at four gate voltage stress levels. This data set consists of 10, 8, and 37 right-censored observations at 7.9, 7.5, and 7.1 V, respectively. Again, we assume the lowest stress level (i.e., 7.1 V) used in the experiment as the normal stress level, and use the data collected at 8.3, 7.9, and 7.5 V to predict the failure-time CDF at $x_0 = 7.1$ V.

Figure 5.3 shows the failure-time CDF at the normal stress level predicted by the DP Weibull mixture ALT model and the Weibull log-linear ALT model, and compares them with the empirical CDF computed by the median rank estimator. Because of the censoring, the empirical CDF is truncated at the censoring time of 600 seconds. For this

example, a visual comparison of the failure-time distributions predicted by the Weibull log-linear ALT model and the DP Weibull mixture ALT model shows that the two models provide very close results, and their predicted failure-time CDFs agree very well with the experimental data collected at $x_0 = 7.1$ V.

Table 5.2: Example 2: times-to-breakdown of nanocrystals-embedded high- k memories tested at four voltage stresses

Stress (volts)	Time-to-breakdown (seconds)									
8.3	2	5	5	6	6	6	10	11	15	18
	21	28	39	49	86					
7.9	5	8	8	12	15	26	29	39	45	69
	100	105	115	146	153	180 ⁺ (8 censored)				
7.5	5	29	31	31	33	39	46	46	54	66
	70	86	87	107	122	137	176	181	190	218
	225	259	277	334	356	371	443	480 ⁺ (10 censored)		
7.1	8	38	72	88	90	97	122	140	163	170
	188	198	199	223	232	256	257	265	318	371
	399	401	412	434	448	513	527	556	583	
	600 ⁺ (37 censored)									

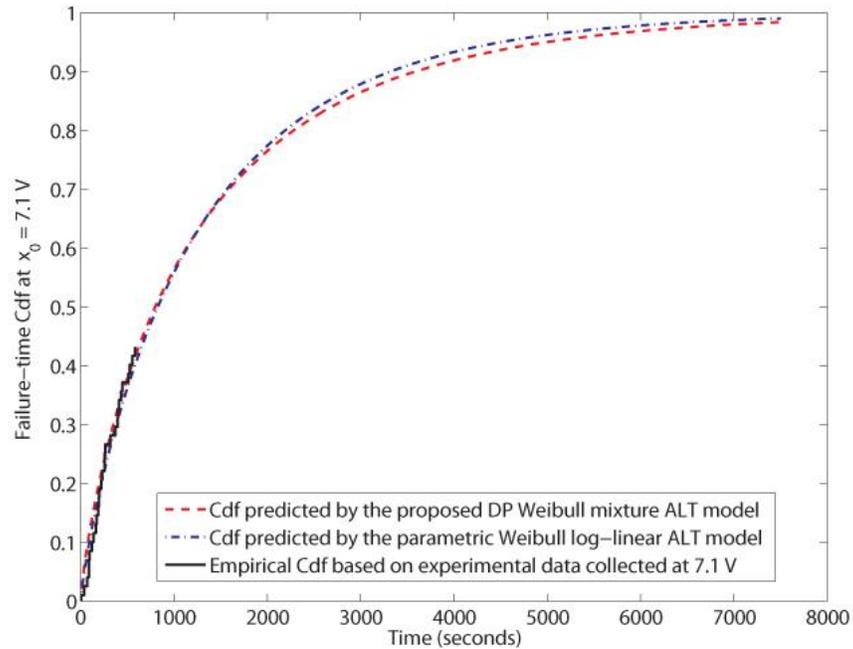


Figure 5.3: Example 2: predicted failure-time CDF at the normal stress $x_0 = 7.1$ V.

5.3 Conclusions

This study proposed the DP Weibull mixture ALT model and developed the Gibbs sampling algorithm for model fitting. This ALT model describes the failure-time distribution at a given stress level using the nonparametric DP mixture model with the Weibull kernel, which offers great modeling flexibility. Two practical examples related to the reliability of nanoelectronics have demonstrated that the proposed methodology is capable of providing accurate predication of the failure-time distribution at the normal stress level without assuming any restrictive parametric failure-time distribution.

The Gibbs sampling algorithm presented in Section 5.1.2, although cumbersome in appearance, is easy to implement. All the conditional posterior distributions involved in the algorithm are routine to sample. Of course, the proposed model and algorithm are

computationally more intense than the ML estimation of the Weibull log-linear ALT model. Another disadvantage of the nonparametric Bayesian analysis is that it may require larger sample sizes than the parametric analysis, especially when noninformative prior distributions are used.

This study considered only one stress variable. It is straightforward to extend the proposed methodology to consider multiple stress variables by allowing x_i and β in Eq. (5.1) to be vector-valued. This study considered right censoring. Other censoring mechanisms can be easily considered due to the fact that the Weibull kernel has the closed-form CDF and the flexibility and generality of the Gibbs sampling algorithm.

CHAPTER 6: BAYESIAN ANALYSIS FOR SIMPLE STEP-STRESS

ACCELERATED LIFE TESTING

This chapter develops Dirichlet process Weibull mixture model for simple step-stress ALT. The cumulative exposure model is applied to describe the effect of changing stress. The consistent model fitting algorithm with constant-stress ALT analysis is utilized. The results of practical example show this methodology is capable of accurately predicting the failure-time distribution at the normal stress level.

6.1 Methodologies

The SSALT model introduces cumulative exposure model to the constant-stress ALT model. The CDF and PDF of the failure-time distribution in a simple SSALT with log-linear acceleration model can be respectively expressed as:

$$F(t) = \begin{cases} F_1(t), & 0 < t < \tau, \\ F_2[\tau \exp(\beta(x_L - x_H)) + t - \tau], & \tau \leq t \leq \infty, \end{cases}$$

and,

$$f(t) = \begin{cases} f_1(t), & 0 < t < \tau, \\ f_2[\tau \exp(\beta(x_L - x_H)) + t - \tau], & \tau \leq t \leq \infty, \end{cases}$$

where τ is stress changing time, and x_L and x_H are lower and higher stress levels, respectively. Then, the hierarchical form of DP Weibull mixture model for SSALT can be written as:

$$\begin{aligned}
t_i | \alpha_i, \lambda_i, \beta &\sim \begin{cases} \exp(x_i \beta) k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i), & 0 < t_i \leq \tau, \\ \exp(x_i \beta) k([\tau \exp(\beta(x_L - x_H)) + t_i - \tau] \exp(x_i \beta) | \alpha_i, \lambda_i), & \tau < t_i, \text{ and } \delta_i = 1, \\ 1 - K([\tau \exp(\beta(x_L - x_H)) + t_i - \tau] \exp(x_i \beta) | \alpha_i, \lambda_i), & \delta_i = 0, \end{cases} \\
\alpha_i, \lambda_i &\sim G(\alpha, \lambda), i = 1, \dots, n, \\
G(\alpha, \lambda) &\sim DP(\mu, G_0(\alpha, \lambda)), \\
\beta, \phi, \gamma, \mu &\sim f(\beta | a_\beta, b_\beta) f(\phi | a_\phi, b_\phi) f(\gamma | a_\gamma, b_\gamma) f(\mu | a_\mu, b_\mu).
\end{aligned} \tag{6.1}$$

Again, assign normal prior for β , Pareto prior for ϕ , Gamma prior for γ and μ as following.

$$\begin{aligned}
\beta &\sim \text{Normal}(a_\beta, b_\beta), \\
\phi &\sim \text{Pareto}(a_\phi, b_\phi), \\
\gamma &\sim \text{Gamma}(a_\gamma, b_\gamma), \\
\mu &\sim \text{Gamma}(a_\mu, b_\mu).
\end{aligned}$$

The simulation algorithm of SSALT inference is similar to that of constant-stress ALT as following.

- (a) Draw (α_i, λ_i) from its conditional posterior distribution, and update clustering indicator for each t_i . For all $t_i > \tau$, replace t_i with t_i' , where

$$t_i' = t_i + \tau \exp(\beta(x_L - x_H)) - \tau;$$

- (b) Update cluster locations $(\alpha_j^*, \lambda_j^*)$ from $f(\alpha_j^*, \lambda_j^* | \mathbf{c}, \phi, \gamma, M, \beta, \mathbf{d})$, for $j = 1, \dots, n^*$, replace t_i with t_i' for all $t_i > \tau$,

- (c) Update ϕ , γ , and μ based on their conditional posterior distributions;

- (d) Sample value of β from its conditional posterior distribution using slice sampling.

For each iteration, samples a new value for β from its conditional posterior distribution

$$\begin{aligned}
& f(\beta | \{(\alpha_i, \lambda_i), i = 1, 2, \dots, n\}, \mathbf{d}) \\
& \propto f(\beta | a_\beta, b_\beta) \times \prod_{t_i=0}^{\tau} \exp(x_i \beta) k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) \\
& \quad \times \prod_{t_i > \tau}^{t_c} \exp(x_i \beta) k(t_i \exp(x_i \beta) | \alpha_i, \lambda_i) \\
& \quad \times \prod_{t_i > t_c} [1 - K(t_i \exp(x_i \beta) | \alpha_i, \lambda_i)] \\
& \propto f(\beta | a_\beta, b_\beta) \times \prod_{t_i=0}^{t_c} \exp(\alpha_i \beta x_i) \times \prod_{t_i=0}^{\tau} \exp(-\lambda_i^{-1} t_i^{\alpha_i} \exp(\alpha_i x_i \beta)) \times \prod_{t_i > \tau}^{t_c} t_i^{\alpha_i - 1} \\
& \quad \times \prod_{t_i > \tau} \exp(-\lambda_i^{-1} t_i^{\alpha_i} \exp(\alpha_i x_i \beta)) \\
& \propto f(\beta | a_\beta, b_\beta) \times \prod_{t_i=0}^{t_c} \exp(\alpha_i \beta x_i) \times \prod_{t_i=0}^n \exp(-\lambda_i^{-1} T_i^{\alpha_i} \exp(\alpha_i x_i \beta)), \tag{6.2}
\end{aligned}$$

where $T_i = \begin{cases} t_i, & 0 < t_i \leq \tau, \\ t_i + \tau \exp(\beta(x_L - x_H)) - \tau, & t_i > \tau. \end{cases}$ Again, apply data augmentation to

simulate a value for β from $f(\beta | \{(\alpha_i, \lambda_i), i = 1, 2, \dots, n\}, \mathbf{d})$. For each $i \in \{i: \delta_i = 1\}$, draw an

auxiliary variable $u_{1,i}$ uniformly on the interval $(0, \exp(\alpha_i x_i \beta))$. For each $i, i = 1, 2, \dots, n$,

draw an auxiliary variable $u_{2,i}$ uniformly on the interval $(0, \exp(-\lambda_i^{-1} T_i^{\alpha_i} \exp(\alpha_i x_i \beta)))$.

Then a value for β is sampled from the truncated normal distribution $f(\beta | a_\beta, b_\beta)$ with

restrict to the interval $\left(\max_{\{i: \delta_i = 1\}} \{\log u_{1,i} / \alpha_i x_i\}, \min_{\{i=1, 2, \dots, n\}} \{(\alpha_i x_i)^{-1} \log(-\lambda_i^{-1} T_i^{\alpha_i} \log u_{2,i})\} \right)$.

(a) Update the failure-time CDF at the normal use stress level.

6.2 Illustrative Examples

The practical example used in this section also use experimental data collected at the Thin Film Nano & Microelectronics Research Laboratory at Texas A&M University, College Station. Table 6.1 presents the times-to-breakdown observations of device tested at a simple SSALT. The low and high stresses applied in the test are 7.5V and 8.3 V, respectively. The stress changed at $\tau = 480$ s, and the tests ended at $t_c = 600$ s. The data set consists of 52 failure observations and 8 right-censored observations. Again, the lowest stress level (i.e., 7.5 V) used in the experiment is assumed as the normal stress level, and the data collected in the simple SSALT is used to predict the failure-time distribution at 7.5 V. This type of step-stress ALT is known as step-stress partially ALT, i.e., a unit is first tested under normal stress, if the unit does not fail by time τ , increase the stress level and test the unit until it fails or censors. The predicted distribution can be compared with the empirical CDF computed from the median rank estimator.

Table 6.1: Times-to-breakdown of nanocrystals-embedded high- k memories under simple SSALT.

Time-to-breakdown (seconds)									
2	3	4	5	6	6	7	8	11	12
14	15	22	22	25	31	31	33	33	35
36	38	55	55	64	73	91	103	147	148
170	187	213	221	263	277	278	299	305	381
443	449	489	491	495	497	504	508	520	566
569	575	600 ⁺ (8 censored)							

Noninformative priors are chosen to reflect absence of any prior knowledge. Assume Gamma(1, 0.001) prior for μ and γ , Pareto(1, 1) prior for ϕ , and Normal(0, 10^6) prior for β . Utilize Matlab® to code the algorithm, and run Gibbs sampling algorithm for 10,000 iterations, with the first 5,000 iterations discarded before data analysis. Figure 6.1 shows the trace plots for parameters α^* , β , and λ^* with three different start points. It is reasonable to conclude that the convergence has achieved since the three chains appear to mix well after 5,000 iterations.

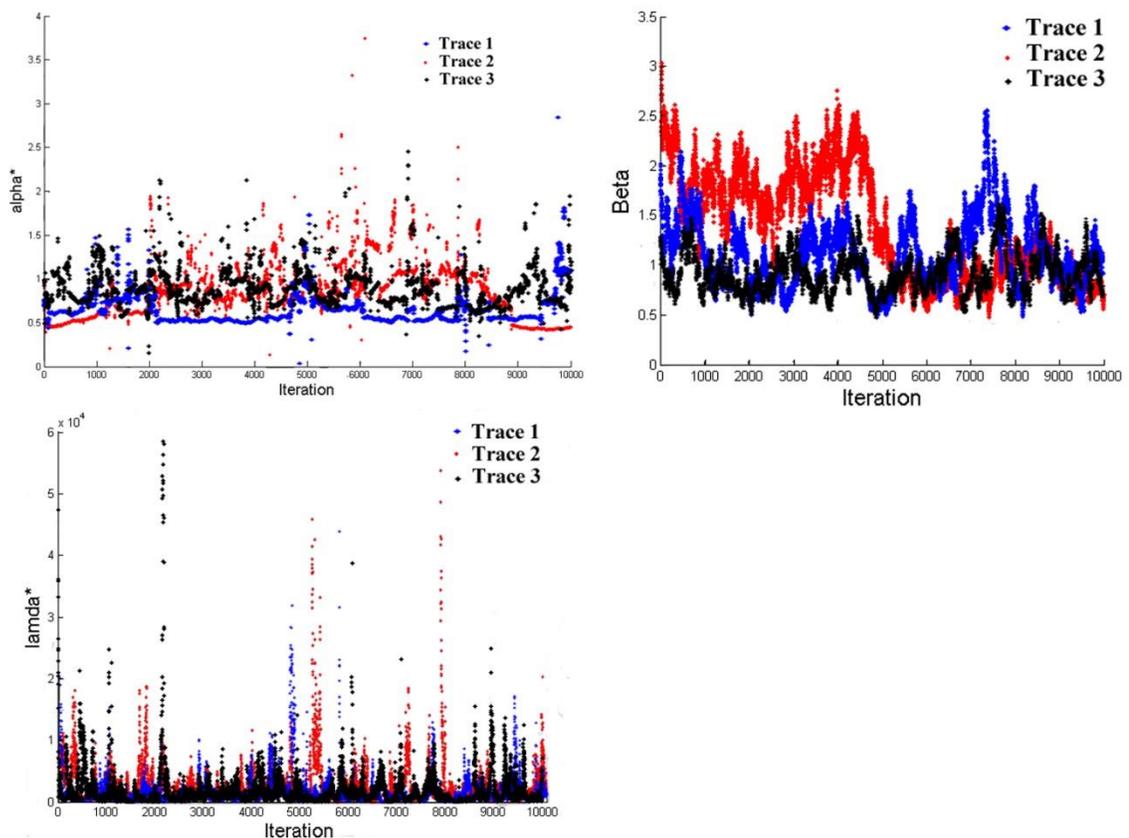


Figure 6.1 Trace plots of α^* , β , and λ^*

The posterior prediction on the failure-time CDF at 7.5V is based on the posterior median. The 95% posterior intervals are also plotted to demonstrate the accuracy of prediction. Figure 6.2 shows the failure-time CDF at 7.5 V predicted by DP Weibull mixture SSALT model, the parametric Weibull log-linear SSALT model, and empirical CDF computed from the median rank estimator. The empirical CDF is truncated at the censoring time of 600 seconds. DP Weibull mixture SSALT model and parametric model provide very similar results. In addition, the 95% posterior interval of DP Weibull mixture model covers all experimental data.

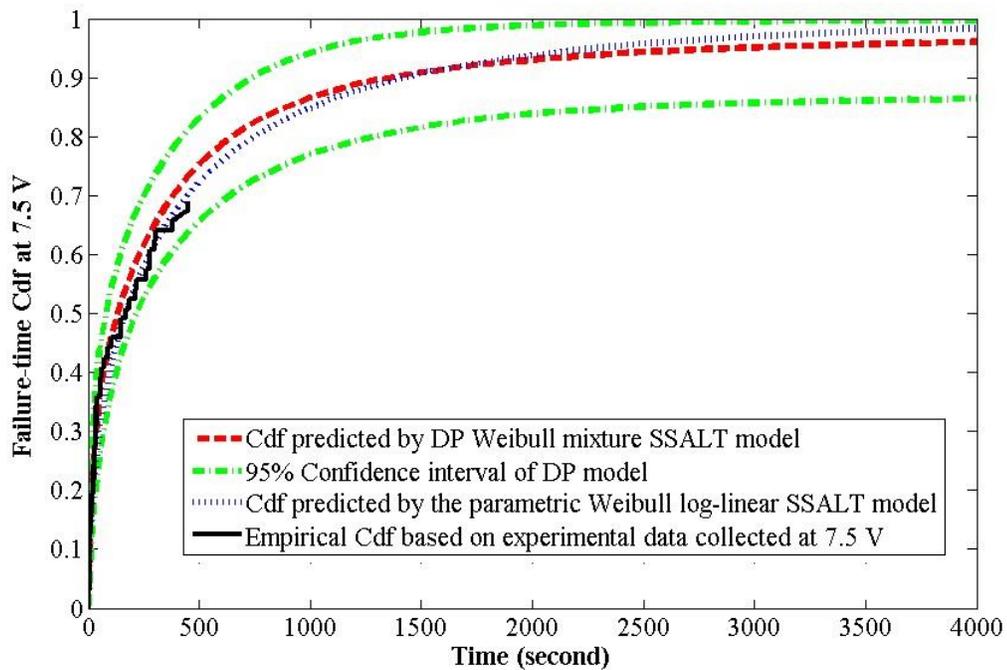


Figure 6.2 Predicted failure-time CDF at the use stress $x_0 = 7.5$ V

CHAPTER 7: CONCLUSIONS

This dissertation presents a semi-parametric Bayesian inference for accelerated life testing using Dirichlet process Weibull mixture model. A novel structure of nanoelectronics is investigated. Then the experimental datasets collected from similar structure of devices are applied in the semi-parametric ALT model proposed in this dissertation. The results demonstrate the capability of our model. In general, compare to parametric accelerated life testing model, the DPWM model is more applicable for emerging products, i.e., there exist uncertainties on the failure-time distribution of products, and very little prior knowledge on the model parameters are available.

7.1 Memory Functions of MoO_x Nanodots Embedded ZrHfO High-*k*

The memory properties of nc-MoO_x embedded ZrHfO high-*k* capacitor are shown to be contributed by hole trapping and detrapping. The *C-V* hysteresis and *J-V* curves detect that holes are trapped at the nc-MoO_x site. Those two curves also confirm that some loosely trapped holes are easily released. The retention study further demonstrates those loosely trapped hole (approximate 20% of total trapped holes) release quickly. More than 50% of trapped holes remain in the device after 10 years. This is desirable for nonvolatile memory devices.

7.2 Prediction of CDF of Failure-Time Distribution at Normal Stress Level

The constant-stress ALT model proposed in the dissertation assumes a log-linear relationship between lifetime and stress levels. The nonparametric DP Weibull mixture model is used to describe the failure-time distribution at a given stress to relax the limitation of lack of distribution information. The step-stress ALT model introduced the

CE model to describe the effect of changing stress. Matlab® is used to implement the simulation algorithm and plot the predicted CDF curve of failure-time distribution under use condition. The simulation has proved to be converged. The result is compared with the parametric Weibull log-linear ALT model. In complete dataset example, DPWM model yields closer curve to empirical CDF. For right censored dataset, both semiparametric and parametric models yield close curves.

7.3 Future Research

This dissertation develops a semiparametric method for ALT with one stress variable and it involves complete and type-I censored dataset. It can be generalized to include multiple stress variables and type-II censored data due to the flexibility and generality of Gibbs sampling algorithm. In addition, other distributions can be applied to Dirichlet process mixture model, such as exponential, normal or lognormal distribution, i.e., DPEM, DPNM, or DPLNM, respectively.

Another future research with DPWM can be degradation analysis. Instead of “hard failure”, the failures for some electronics are defined as “soft failure”. For example, failure time of light display devices is defined at the time when a device luminosity drops below 50% of its initial luminosity [125]. The failure of some laser products are defined as more than 20% output power loss. The analysis of performance degradation path, both under normal condition and accelerated stress condition, can further reduce lifetest time, and improve reliability inference. Degradation analysis can be performed as a nonparametric regression with DMWM model.

The cumulative exposure model is applied to deal with simple SSALT in this dissertation. Other models, for example, the Tampered Random Variable (TRV) model and the Tampered Failure Rate (TFR), can be applied and combined to DMWM model. Furthermore, simple SSALT modelling only deals with two stress levels. The algorithm involving multiple stress levels can be deduced by generalizing the cumulative exposure model with a log-linear function to multiple stresses.

REFERENCES

- [1] W. Q. Meeker and L. A. Escobar, *Statistical methods for reliability data*. New York: John Wiley & Sons, Inc., 1998.
- [2] N. R. Mann, et al., *Methods for statistical analysis of reliability and life data*. New York: John Wiley & Sons, Inc., 1974.
- [3] N. Cheng, “Bayesian nonparametric reliability analysis using Dirichlet process mixture model,” M.S. thesis, Dept. Ind. & Syst. Eng., Ohio University, Athens, OH, 2011.
- [4] R. Zhou, “Bayesian analysis of log-binomial models,” Ph.D. dissertation, Dept. Math. Sci., University of Cincinnati, Cincinnati, OH, 2005.
- [5] C. E. Ebeling, *An introduction to reliability and maintainability engineering*. Long Grove, IL: Waveland Press, Inc., 2005.
- [6] E. Gouno and N. Balakrishnan, “Step-stress accelerated life test,” in *Advances in Reliability*, N. Balakrishnan and C. Radhakrishna, Eds. New York: Elsevier North Holland, 2001, pp. 623–639.
- [7] L. W. Condra, *Reliability improvement with design of experiments*, 2nd ed. New York: Marcel Dekker, Inc., 2001.
- [8] A. Dasgupta and M. Pecht, “Material failure mechanisms and damage models,” *IEEE Trans. Reliab.*, vol. 40, no. 5, pp. 531–536, 1991.
- [9] F. Jensen, “Activation energies and the Arrhenius equation,” *Qual. Rel. Eng. Int.*, vol. 1, pp. 13–18, 1985.

- [10] P. A. Tobias and D. C. Trindade, *Applied reliability*. New York: Van Nostrand Reinhold Company, 1986, pp. 237.
- [11] W. Luo, "Reliability characterization and prediction of high- k dielectric thin film," Ph.D. dissertation, Dept. Ind. Eng. Texas A&M University, College Station, Texas, 2004.
- [12] W. Nelson, "Accelerated life testing-step-stress models and data analyses," *IEEE Trans. Rel.*, vol. 29, no. 2, pp. 103–108, Jun. 1980.
- [13] D. Ruppert, "Bayesian Data Analysis and MCMC," in *Statistics and Data Analysis for Financial Engineering*, D. Ruppert, Ed. New York: Springer New York, 2011, pp. 531–578.
- [14] P. Muller and F. A. Quintana, "Nonparametric Bayesian data analysis," *Stat. Sci.*, vol. 19, no. 1, pp. 95–110, Feb. 2004.
- [15] T. S. Ferguson, "A Bayesian analysis of some nonparametric problems," *Ann. Stat.*, vol. 1, no. 2, pp. 209–230, 1973.
- [16] A. Kottas, Class Lecture, Topic: "Dirichlet process prior." AMS 241, Faculty of Applied Mathematics and Statistics, University of California, Santa Cruz, Santa Cruz, CA, Fall. 2010.
- [17] D. Gorur and C. E. Rasmussen, "Dirichlet process Gaussian mixture models: choice of the base distribution," *J. Comput. Sci. Technol.*, vol. 25, no. 4, pp. 653–664, 2010.
- [18] L. Kuo and B. Mallick, "Bayesian semiparametric inference for the accelerated failure-time model," *Can. J. Stat.*, vol. 25, no. 4, pp. 457–472, Dec. 1997.

- [19] A. Kottas, "Nonparametric Bayesian survival analysis using mixtures of Weibull distributions," *J. Stat. Plan. Inference*, vol. 136, no. 3, pp. 578–596, Mar. 2006.
- [20] J. Diebolt and C. P. Robert, "Estimation of finite mixture distributions through Bayesian sampling," *J. Roy. Stat. Soc. Ser. B (Methodol.)*, vol. 56, no. 2, pp. 363–375, 1994.
- [21] S. Richardson and P. Green, "On Bayesian analysis of mixtures with an unknown number of components," *J. Roy. Stat. Soc. Ser. B (Methodol.)*, vol. 59, no. 4, pp. 731–792, 1997.
- [22] C. H. Yang, "Reliability Analysis of nanocrystal embedded high- k nonvolatile memories," Ph.D. dissertation, Dept. Ind. Eng., University of Tennessee, Knoxville, TN, 2011.
- [23] J. Lu, et al., "Nanocrystalline silicon embedded zirconium-doped hafnium oxide high- k memory device," *Jpn. J. Appl. Phys.*, vol. 45, no. 34, pp. L901–L903, 2006.
- [24] X. Liu, et al., "Memory functions of molybdenum oxide nanodots-embedded ZrHfO high- k ," *Electrochem. Solid-State Lett.*, vol. 15, no. 6, pp. H192–H194, 2012.
- [25] C. H. Yang, et al., "Temperature influence on nanocrystals embedded high- k nonvolatile C – V characteristics," *Electrochem. Solid-State Lett.*, vol. 14, no. 1, pp. H50–H52, 2011.
- [26] C. H. Lin and Y. Kuo, "Material and electrical properties of hole-trapping memory capacitors composed of nc-ITO embedded ZrHfO high- k films," *ECS Trans.*, vol. 35, no. 2, pp. 249–255, 2011.

- [27] X. Liu, et al., "Nanocrystalline MoO_x embedded ZrHfO high-*k* memories-charge trapping and retention characteristics," *ECS Trans.*, vol. 45, no. 6, pp. 203–209, Apr. 2012.
- [28] N. D. Singpurwalla, et al., "Inference from accelerated life tests using Eyring type re-parameterizations," *Nav. Res. Logist. Quart.*, vol. 22, pp. 289–296, 1975.
- [29] H. D. Kahn, "Least squares estimation for the inverse power law for accelerated life tests," *Appl. Stat.*, vol. 28, no. 1, pp. 40–46, 1979.
- [30] E. P. Barbosa and F. Louzada-Neto, "Analysis of accelerated life tests with Weibull failure distribution via generalized linear models," *Commun. Stat. Simul. Comput.*, vol. 23, no. 2, pp. 455–465, 1994.
- [31] C. S. Whitman, "Accelerated life test calculations using the method of maximum likelihood: an improvement over least squares," *Microelectron. Rel.*, vol. 43, no. 6, pp. 859–864, Jun. 2003.
- [32] A. A. Abdel-Ghaly, et al., "Estimation of the parameters of pareto distribution and the reliability function usin accelerated life testing with censoring," *Commun. Stat. Simul. Comput.*, vol. 27, no. 2, pp. 469–484, 1998.
- [33] A. J. Watkins, "On the analysis of accelerated life-testing experiments," *IEEE Trans. Rel.*, vol. 40, no. 1, pp. 98–101, 1991.
- [34] A. J. Watkins, "Review: Likelihood method for fitting weibull log-linear models to accelerated life-test data," *IEEE Trans. Rel.*, vol. 43, no. 3, pp. 361–365, 1994.
- [35] R. E. Glaser, "Estimation for a Weibull accelerated life testing model," *Nav. Res. Logist. Quart.*, vol. 31, no. 4, pp. 559–570, 1984.

- [36] H. Hirose, "Estimation of threshold stress in accelerated life-testing," *IEEE Trans. Rel.*, vol. 42, no. 4, pp. 650–657, 1993.
- [37] M. Newby, "Accelerated failure time models for reliability data analysis," *Rel. Eng. Syst. Saf.*, vol. 20, pp. 187–197, 1988.
- [38] T. A. Mazzuchi and R. Soyer, "Dynamic models for statistical inference from accelerated life tests," in *Proc. Annu. Rel. Maint. Symp.*, 1990, pp. 67–70.
- [39] T. A. Mazzuchi, et al., "Linear Bayesian inference for accelerated Weibull model," *Lifetime Data Anal.*, vol. 3, no. 1, pp. 63–75, Jan. 1997.
- [40] D. S. Bai and S. W. Chung, "An accelerated life test model with the inverse power law," *Rel. Eng. Syst. Saf.*, vol. 24, no. 3, pp. 223–230, Jan. 1989.
- [41] J. Biernat, et al., "Reliability considerations in accelerated life testing of electrical insulation with generalized life distribution function," *IEEE Trans. Power Syst.*, vol. 7, no. 2, pp. 656–664, 1992.
- [42] C. M. Kim and D. S. Bai, "Analyses of accelerated life test data under two failure modes," *Int. J. Rel. Qual. Saf. Eng.*, vol. 09, no. 02, pp. 111–125, Jun. 2002.
- [43] E. K. AL-Hussaini and A. H. Abdel-Hamid, "Bayesian estimation of the parameters, reliability and hazard rate functions of mixtures under accelerated life tests," *Commun. Stat. Simul. Comput.*, vol. 33, no. 4, pp. 963–982, Oct. 2004.
- [44] E. K. Al-Hussaini and A. H. Abdel-Hamid, "Accelerated life tests under finite mixture models," *J. Stat. Comput. Simul.*, vol. 76, no. 8, pp. 673–690, 2006.
- [45] M. Shaked, W. J. Zimmer, and C. A. Ball, "A nonparametric approach to accelerated life testing," *J. Am. Stat. Assoc.*, vol. 74, no. 367, pp. 694–699, 1979.

- [46] M. Shaked and N. D. Singpurwalla, "Nonparametric estimation and goodness-of-fit testing of hypotheses for distributions in accelerated life testing," *IEEE Trans. Rel.*, vol. R-31, no. 1, pp. 69–74, 1982.
- [47] D. S. Bai and N. Y. Lee, "Nonparametric estimation for accelerated life tests under intermittent inspection," *Rel. Eng. Syst. Saf.*, vol. 54, pp. 53–58, 1996.
- [48] A. P. Basu and N. Ebrahimi, "Nonparametric accelerated life testing," *IEEE Trans. Rel.*, vol. R-31, no. 5, pp. 432–435, Dec. 1982.
- [49] J. René Van Dorp and T. a. Mazzuchi, "A general Bayes exponential inference model for accelerated life testing," *J. Stat. Plan. Inference*, vol. 119, no. 1, pp. 55–74, Jan. 2004.
- [50] J. René Van Dorp and T. a. Mazzuchi, "A general Bayes Weibull inference model for accelerated life testing," *Rel. Eng. Syst. Saf.*, vol. 90, no. 2–3, pp. 140–147, Nov. 2005.
- [51] T. A. Louis, "Nonparametric analysis of an accelerated failure time model," *Biometrika*, vol. 68, no. 2, pp. 381–390, 1981.
- [52] R. L. Schmoyer, "Nonparametric analyses for two-level single-stress accelerated life tests," *Technometrics*, vol. 33, no. 2, pp. 175–186, 1991.
- [53] F. Proschan and N. D. Singpurwalla, "A new approach to inference from accelerated life tests," *IEEE Trans. Rel.*, vol. R-29, no. 2, pp. 98–102, 1980.
- [54] H. Maciejewski, "Accelerated life test data analysis with generalised life distribution function and with no aging model assumption," *Microelectron. Reliab.*, vol. 35, no. 7, pp. 1047–1051, 1995.

- [55] D. Y. Lin and C. J. Geyer, "Computational methods for semiparametric linear regression with censored data," *J. Comput. Graph. Stat.*, vol. 1, no. 1, pp. 77–90, 1992.
- [56] A. Komárek and E. Lesaffre, "Bayesian semi-parametric accelerated failure time model for paired doubly interval-censored data," *Stat. Modelling*, vol. 6, pp. 3–22, Mar. 2006.
- [57] A. Komárek and E. Lesaffre, "Bayesian accelerated failure time model with multivariate doubly interval-censored data and flexible distributional assumptions," *J. Am. Stat. Assoc.*, vol. 103, no. 482, pp. 523–533, Jun. 2008.
- [58] R. Argiento, et al., "A semiparametric Bayesian mixed-effects model for failure time data," *Proc. SCo 2009: Complex data Modeling and Computationally Intensive Statistical Methods for Estimation and Prediction*, pp. 17-22, 2009.
- [59] M. DeGroot and P. Goel, "Bayesian estimation and optimal designs in partially accelerated life testing," *Nav. Res. Logist. Quart.*, vol. 26, no. 2, pp. 223–235, 1979.
- [60] G. K. Bhattacharyya and Z. Soejoeti, "A tampered failure rate model for step-stress accelerated life test," *Commun. Stat. Theory Methods*, vol. 18, no. 5, pp. 1627–1643, 1989.
- [61] C. J. Xiong, "Inferences on a simple step-stress model with type-II censored exponential data," *IEEE Trans. Rel.*, vol. 47, no. 2, pp. 142–146, 1998.
- [62] C. J. Xiong and M. Ji, "Analysis of grouped and censored data from step-stress life test," *IEEE Trans. Rel.*, vol. 53, no. 1, pp. 22–28, Mar. 2004.

- [63] C. J. Xiong and G. A. Milliken, "Step-stress life-testing with random stress-change times for exponential data," *IEEE Trans. Rel.*, vol. 48, no. 2, pp. 141–148, 1999.
- [64] J. P. Zhang and X. M. Geng, "Constant-step stress accelerated life test of VFD under Weibull distribution case," *J. Zhejiang Univ. Sci.*, vol. 6A, no. 7, pp. 722–727, Jul. 2005.
- [65] A. H. Abdel-Hamid and E. K. AL-Hussaini, "Estimation in step-stress accelerated life tests for the exponentiated exponential distribution with type-I censoring," *Comput. Stat. Data Anal.*, vol. 53, no. 4, pp. 1328–1338, Feb. 2009.
- [66] B. X. Wang, "Interval estimation for exponential progressive type-II censored step-stress accelerated life-testing," *J. Stat. Plan. Inference*, vol. 140, no. 9, pp. 2706–2718, Sep. 2010.
- [67] X. Yin and B. Sheng, "Some aspects of accelerated life testing by progressive stress," *IEEE Trans. Rel.*, vol. R-36, no. 1, pp. 150–155, 1987.
- [68] E. Gouno, "An inference method for temperature step-stress accelerated life testing," *Qual. Rel. Eng. Int.*, vol. 17, pp. 11–18, 2001.
- [69] J. Lee and R. Pan, "Bayesian inference model for step-stress accelerated life testing with type-II censoring," in *Proc. Annu. Rel. Maint. Symp.*, pp. 91-96, 2008.
- [70] J. Lee and R. Pan, "Bayesian analysis of step-stress accelerated life test with exponential distribution," *Qual. Rel. Eng. Int.*, vol. 28, no. 3, pp. 353–361, Apr. 2012.
- [71] J. Lee and R. Pan, "Analyzing step-stress accelerated life testing data using generalized linear models," *IIE Trans.*, vol. 42, no. 8, pp. 589–598, 2010.

- [72] L. C. Tang, et al., "Analysis of step-stress accelerated-life-test data: a new approach," *IEEE Trans. Rel.*, vol. 45, no. 1, pp. 69–74, 1996.
- [73] I. H. Khamis and J. J. Higgins, "A new model for step-stress testing," *IEEE Trans. Rel.*, vol. 47, no. 2, pp. 131–134, 1998.
- [74] N. Balakrishnan, et al., "Point and interval estimation for a simple step-stress model with Type-II censoring," *J. Qual. Technol.*, vol. 39, no. 1, pp. 35–46, 2007.
- [75] N. Balakrishnan, et al, "Exact inference for a simple step-stress model from the exponential distribution under time constraint," *Ann. Inst. Stat. Math.*, vol. 61, no. 1, pp. 251–274, Jul. 2007.
- [76] M. Kateri and N. Balakrishnan, "Inference for a simple step-stress model with type-II censoring, and Weibull distributed lifetimes," *IEEE Trans. Rel.*, vol. 57, no. 4, pp. 616–626, Dec. 2008.
- [77] N. Balakrishnan and Q. Xie, "Exact inference for a simple step-stress model with type-II hybrid censored data from the exponential distribution," *J. Stat. Plan. Inference*, vol. 137, no. 8, pp. 2543–2563, Aug. 2007.
- [78] N. Balakrishnan and D. Han, "Exact inference for a simple step-stress model with competing risks for failure from exponential distribution under type-II censoring," *J. Stat. Plan. Inference*, vol. 138, no. 12, pp. 4172–4186, Dec. 2008.
- [79] D. Han and N. Balakrishnan, "Inference for a simple step-stress model with competing risks for failure from the exponential distribution under time constraint," *Comput. Stat. Data Anal.*, vol. 54, no. 9, pp. 2066–2081, Sep. 2010.

- [80] N. Balakrishnan, et al., "Order restricted inference for exponential step-stress models," *IEEE Trans. Rel.*, vol. 58, no. 1, pp. 132–142, Mar. 2009.
- [81] M. T. Madi, "Bayesian inference for partially accelerated life tests using Gibbs sampling," *Microelectron. Rel.*, vol. 37, no. 8, pp. 1165–1168, 1997.
- [82] A. A. Abdel-Ghaly, et al., "The maximum likelihood estimates in step partially accelerated life tests for the Weibull parameters in censored data," *Commun. Stat. Theory Methods*, vol. 31, no. 4, pp. 551–573, 2002.
- [83] F. K. Wang, et al., "Partially accelerated life tests for the Weibull distribution under multiply censored data," *Commun. Stat. Simul. Comput.*, vol. 41, no. 9, pp. 1667–1678, Oct. 2012.
- [84] R. H. Wang and H. Fei, "Statistical inference of Weibull distribution for tampered failure rate model in progressive stress accelerated life testing," *J. Syst. Sci. Complex.*, vol. 17, no. 2, pp. 237–243, 2004.
- [85] W. Zhao and E. a. Elsayed, "A general accelerated life model for step-stress testing," *IIE Trans.*, vol. 37, no. 11, pp. 1059–1069, Nov. 2005.
- [86] Z. N. Lin and H. Fei, "A nonparametric approach to progressive stress accelerated life testing," *IEEE Trans. Rel.*, vol. 40, no. 2, pp. 173–176, 1991.
- [87] D. S. Bai and Y. R. Chun, "Nonparametric inferences for ramp stress tests under random censoring," *Rel. Eng. Syst. Saf.*, vol. 41, pp. 217–223, 1993.
- [88] O. I. Tyoskin and S. Y. Krivolapov, "Nonparametric model for step-stress accelerated life testing," *IEEE Trans. Rel.*, vol. 45, no. 2, pp. 346–350, 1996.

- [89] M. Shaked and N. Singpurwalla, “Inference for step-stress accelerated life tests,” *J. Stat. Plan. Inference*, vol. 7, pp. 295–306, 1983.
- [90] C.-H. Hu, et al., “Step-stress accelerated life tests: a proportional hazards-based non-parametric model,” *IIE Trans.*, vol. 44, no. 9, pp. 754–764, 2012.
- [91] J. Dorp Rene van, et al., “A Bayes approach to step-stress accelerated life testing,” *IEEE Trans. Rel.*, vol. 45, no. 3, pp. 491–498, 1996.
- [92] A. Y. Lo, “On a class of Bayesian nonparametric estimates: I. density estimates,” *Ann. Stat.*, vol. 12, no. 1, pp. 351–357, 1984.
- [93] M. Escobar and M. West, “Bayesian density estimation and inference using mixtures,” *J. Am. Stat. Assoc.*, vol. 90, no. 430, pp. 577–588, 1995.
- [94] M. Gasparini, “Bayesian density estimation via Dirichlet density processes,” *J. Nonparametr. Stat.*, vol. 6, pp. 355–366, 1996.
- [95] S. Ghosal, et al., “Posterior consistency of Dirichlet mixtures in density estimation,” *Ann. Stat.*, vol. 27, no. 1, pp. 143–158, 1999.
- [96] P. Müller, et al., “Bayesian curve fitting using multivariate normal mixtures,” *Biometrika*, vol. 83, no. 1, pp. 67–79, 1996.
- [97] A. Kottas and A. E. Gelfand, “Bayesian semiparametric median regression modeling,” *J. Am. Stat. Assoc.*, vol. 96, no. 456, pp. 1458–1468, 2001.
- [98] A. E. Gelfand and A. Kottas, “A computational approach for full nonparametric Bayesian inference under Dirichlet process mixture models,” *J. Comput. Graph. Stat.*, vol. 11, no. 2, pp. 289–305, 2002.

- [99] A. E. Gelfand and A. Kottas, “Bayesian semiparametric regression for median residual life,” *Scand. J. Stat.*, vol. 30, no. 4, pp. 651–665, Dec. 2003.
- [100] L. A. Hannah, et al., “Dirichlet process mixtures of generalized linear models,” *J. Mach. Learn. Res.*, vol. 12, pp. 1923–1953, 2011.
- [101] A. Kottas, “Dirichlet process mixtures of beta distributions, with applications to density and intensity estimation,” *Proc. Work. Learn. with Nonparametric Bayesian Methods*, 2006.
- [102] M. West, et al., “Hierarchical priors and mixture models, with application in regression and density estimation,” in *Aspects of Uncertainty: A Tribute to D.V. Lindley*, A. F. M. Smith and P. Freeman, Eds. New York: Wiley, 1994, pp. 363–386.
- [103] S. Mukhopadhyay and A. E. Gelfand, “Dirichlet process mixed generalized linear models,” *J. Amer. Stat. Assoc.*, vol. 92, no. 438, pp. 633–639, 1997.
- [104] C. Carota and G. Parmigiani, “Semiparametric regression for count data,” *Biometrika*, vol. 89, no. 2, pp. 265–281, 2002.
- [105] T. E. Hanson, “Modeling censored lifetime data using a mixture of gammas baseline,” *Bayesian Anal.*, vol. 1, no. 3, pp. 575–594, 2006.
- [106] S. Ghosh and S. Ghosal, “Semiparametric accelerated failure time models for censored data,” *Bayesian Stat. its Appl.*, pp. 213–229, 2006.
- [107] A. Khademi, et al., “Growth and field emission study of molybdenum oxide nanostars,” *J. Phys. Chem. C*, vol. 113, no. 44, pp. 19298–19304, Nov. 2009.

- [108] L. V. Azaroff, *Elements of X-Ray Crystallography*. New York: McGraw-Hill, 1986.
- [109] C. H. Lin and Y. Kuo, "Nonvolatile memories with dual-layer nanocrystalline ZnO embedded Zr-Doped HfO₂ high-*k* dielectric," *Electrochem. Solid-State Lett.*, vol. 13, no. 3, pp. H83, 2010.
- [110] Y. Kuo, "Status review of nanocrystals embedded high-*k* nonvolatile memories," *ECS Trans.*, vol. 35, no. 3, pp. 13–31, 2011.
- [111] C.H. Lin and Y. Kuo, "Nanocrystalline ruthenium oxide embedded zirconium-doped hafnium oxide high-*k* nonvolatile memories," *J. Appl. Phys.*, vol. 110, no. 2, pp. 024101, 2011.
- [112] C. H. Lin and Y. Kuo, "Embedding of nanocrystalline ruthenium oxide in zirconium-doped hafnium oxide high-*k* film for nonvolatile memories," *ECS Trans.*, vol. 13, no. 1, pp. 465–470, 2008.
- [113] C. H. Lin and Y. Kuo, "Single- and dual-layer nanocrystalline indium tin oxide embedded ZrHfO high-*k* films for nonvolatile memories – material and electrical properties," *J. Electrochem. Soc.*, vol. 158, no. 8, pp. H756, 2011.
- [114] T. Maeda, et al., "Electrical properties of Si nanocrystals embedded in an ultrathin oxide," *Nanotechnology*, vol. 10, no. 2, pp. 127–131, Jun. 1999.
- [115] J. Lu, et al., "Nonvolatile memories based on nanocrystalline zinc oxide embedded zirconium-doped hafnium oxide thin films," *ECS Trans.*, vol. 11, no. 4, pp. 509–518, 2007.

- [116] D. N. Kouvatsos, et al., "Charging effects in silicon nanocrystals within SiO₂ layers, fabricated by chemical vapor deposition, oxidation, and annealing," *Appl. Phys. Lett.*, vol. 82, no. 3, pp. 397, 2003.
- [117] C. H. Lin and Y. Kuo, "Charge trapping sites in nc-RuO embedded ZrHfO high-*k* nonvolatile memories," *MRS Proc.*, vol. 1250, 2010.
- [118] W. Nelson, *Accelerated testing: statistical models, test plans, and data analysis*. Hoboken: Wiley, 1990.
- [119] A. Gelman, et al., *Bayesian data analysis*, 2nd ed. New York: Chapman & Hall/CRC, 2004.
- [120] P. Damien, et al., "Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables," *J. Roy. Stat. Soc. Ser. B*, vol. 61, pp. 331–344, 1999.
- [121] J. Sethuraman, "A constructive definition of Dirichlet priors," *Statistica Sinica.*, vol. 4, pp. 639–650, 1994.
- [122] E. Y. Wu, et al., "Ultra-thin oxide reliability for ULSI applications," *Semicond. Sci. Technol.*, vol. 15, no. 5, pp. 425–435, May 2000.
- [123] P. Congdon, *Applied Bayesian modelling*. England: Wiley, 2003.
- [124] R. Degraeve, et al., "A new statistical model for fitting oxide breakdown distributions at different field conditions," *Microelectron. Rel.*, vol. 36, no. 11, pp. 1651–1654, 1996.
- [125] S. J. Bae, et al., "Degradation analysis of nano-contamination in plasma display panels," *IEEE Trans. Rel.*, vol. 57, no. 2, pp. 222–229, Jun. 2008.



OHIO
UNIVERSITY

Thesis and Dissertation Services