Degradation modeling based on a time-dependent Ornstein-Uhlenbeck process and prognosis of system failures

Yingjun Deng

Thèse
de doctorat
de l'UTT

Yingjun DENG

Degradation Modeling Based on a Time-dependent Ornstein-Uhlenbeck Process and Prognosis of System Failures

Spécialité :
Optimisation et Sûreté des Systèmes

2015TROY0004

Année 2015

utt
université de technologie
Troyes

THESE

pour l'obtention du grade de

DOCTEUR de l'UNIVERSITE DE TECHNOLOGIE DE TROYES

Spécialité : OPTIMISATION ET SURETE DES SYSTEMES

présentée et soutenue par

Yingjun DENG

le 24 février 2015

Degradation Modeling Based on a Time-dependent Ornstein-Uhlenbeck Process and Prognosis of System Failures

JURY

M. I. NIKIFOROV	PROFESSEUR DES UNIVERSITES	Président
Mme A. BARROS	PROFESSEUR DES UNIVERSITES	Directrice de thèse
Mme A. GÉGOUT-PETIT	PROFESSEUR DES UNIVERSITES	Rapporteur
M. A. GRALL	PROFESSEUR DES UNIVERSITES	Directeur de thèse
M. M. PANDEY	PROFESSOR	Rapporteur
M. C. PAROISSIN	MAITRE DE CONFERENCES	Examinateur
M. M. ROUSSIGNOL	PROFESSEUR EMERITE	Examinateur

Acknowledgments

The initial motivation to bridge between mathematics and engineering problems drives me to the interesting topic proposed by my supervisors Prof. Anne Barros (currently at NTNU, Norway) and Prof. Antoine Grall. And here I appreciate them a lot for giving me the opportunity to do this PhD at Institut Charles Delaunay, UMR CNRS 6279 & Université de Technologie de Troyes (UTT), also for their inspiration, patience and support in last 3 years. The financial support from China Scholarship Council is appreciated.

It is my honor that Prof. Anne Gégout-Petit and Prof. Mahesh Pandey have accepted to review this thesis. I am also glad that Prof. Michel Roussignol, Prof. Igor Nikiforov, and Dr. Christian Paroissin have accepted to be my jury committee members.

During my PhD period, I have been benefitted a lot from frequent discussions with Prof. Michel Roussignol at Université de Paris-Est Marne-la-Vallée (UPEM). His insight and rigorousness in mathematical research influence me a lot, without whom I would be lost in problems and mistakes. I am grateful for those outside experience with Prof. Damien Lamberton for stochastic analysis at UPEM, with Dr. Hans Van der Weide for first passage problems at TUDelft, Netherlands, and also in Séminaire Bachelier Paris at Institut Henri Poincaré, Paris VI for financial mathematics.

Many thanks are due to my sincere colleagues at UTT, especially to Igor Nikiforov, Mitra Fouladirad, Edith Grall-Maës, Estelle Deloux, Malika Kharouf, Nicholas Lefebvre, Yann Dijoux, Yves Langeron, Tuan Huynh, Elias Khoury and Khanh Le Son. Sincere assistance on administrative issues from Marie-José Rousselet, Veronique Banse and Bernadette Andre, makes my life much easier in France. And it is really a fantastic experience to work in the same office during last 3 years with Kim Anh Nguyen, Danh Ngoc Nguyen and Houda Ghamlouch.

Technical Discussions with my friends help me a lot, especially those with Hao Jiang, Jiange Li, and Chengfa Wu. Also my Chinese friends in Troyes make me feel like at home, especially with Yugang Li, Xiaowei Lv, Hongchang Han, Hui Shang, Fei Zhu, Heping Li, Yaofu Cao etc..

Not accompanying with them in such a long period, I am really in debt to my mother, father, sister and my girlfriend Hanbing. Their understanding, encourage and support are essential for me to complete this thesis.

Yingjun Deng

Abstract

This thesis consists of four parts : stochastic degradation modeling, prognosis of system failures, failure level estimation and maintenance optimization. These parts are connected and dedicated to three core issues of system failures : description, prediction and prevention.

The first part about stochastic degradation modeling proposes a degradation model based on a time-dependent Ornstein-Uhlenbeck (OU) process. Such a process utilizes the inspection records to establish the dynamic description of degradation process. The time-dependent OU process is proved superior by its statistical properties on controllable mean, variance and correlation. Its mean-reverting property can be introduced to interpret temporary correlated fluctuations from an overall degrading trend in degradation records. Corresponding parameter estimation is proposed based on maximum likelihood estimation method. A case study is performed to test the model's fitting goodness based on a degradation data-set of a passive component in power plants.

The second part about prognosis of system failures is discussed further based on the time-dependent OU process. And the first passage time is introduced as the system failure time, to a pre-set failure level. Later how to estimate the failure time is discussed based on two kinds of views : partial differential equation and integral equation. These two views lead to various estimation techniques from different concentrations, and they can be classified into 3 categories : analytical approximations, numerical algorithms and Monte-Carlo simulation methods. Simulation tests are done to calculate first passage density based on proposed methods.

The third part about failure level estimation proposes some techniques to estimate failure levels based on inverse first passage problems. In previous literature the failure level is generally treated as physical barriers or experts' opinions, based on which failure prognosis from first passage failures can hardly fit existing failure records. Therefore the effort in this part is paid to make up the gap between failure records and inspection records under the definition of first passage failure based on inverse first passage problems. When the lifetime distribution is given or estimated from failure records, we emphasizes on numerically reproducing the failure level under which the first passage time of the given stochastic process can have the same lifetime distribution with the given lifetime distribution.

The fourth part about maintenance optimization investigates how to optimize maintenance policies based on the time-dependent OU process. Based on monitored system conditions and prognosis of system failures, condition-based maintenance is adopted to introduce preventive maintenance such that the balance can be achieved between operation costs and disastrous results caused by system failures. In this part, corresponding maintenance optimization problems are discussed based on the time-dependent OU process and the hypothesis of continuously monitored system. Due to the unexplicit expression for prognosis of system failures, classical heuristic optimization procedures cannot be fulfilled. Therefore approximate first passage density is introduced to fulfill the maintenance optimization.

Résumé

Cette thèse est organisé en quatre parties :

1. la modélisation stochastique de la dégradation,
2. le pronostic de l'instant de défaillance du système,
3. l'estimation du niveau dégradation associé à la défaillance,
4. l'optimisation de la maintenance.

Ces différentes parties sont liées entre elles et tentent de décrire, prévoir et prévenir la défaillance du système.

Dans la première partie, un modèle stochastique de la dégradation s'appuyant sur un processus d'Ornstein-Uhlenbeck (OU) dépendant du temps et sur l'exploitation conjointe des donnés d'inspection est proposé. Les qualités de ce modèle sont démontrées au travers de ses propriétés statistiques qui permettent d'ajuster de manière indépendante la moyenne, la variance et la corrélation. Une propriété de "convergence" vers la moyenne est ensuite exploitée pour interpréter la corrélation temporelle des fluctuations autour d'une tendance globale de dégradation. Puis, s'appuyant sur une technique de maximisation de la vraisemblance, une méthode d'estimation des paramètres de ce modèle est proposée. Enfin, un cas d'application portant sur l'étude de la dégradation d'un composant passif de central électrique est traité.

La deuxième partie de la thèse est consacrée au pronostic de l'instant de défaillance du système en s'appuyant sur un processus OU dépendant du temps. Cet instant de défaillance est défini comme le premier temps d'atteinte d'un état de dégradation, critique i.e. d'un état de santé inacceptable.. L'estimation de cet instant de défaillance est abordé selon deux approches : *i)* équations aux dérivés partielles, *ii)* équations intégrales. Ces approches conduisent à différentes techniques d'estimation qui peuvent être classées selon le schéma suivant :

- les techniques d'approximation analytique,
- les techniques d'approximation numérique,
- les techniques de simulation de Monté-Carlo.

Des essais numériques destinés au calcul de la densité de l'instant de défaillance et permettant la confrontation de ces différentes techniques, concluent cette seconde partie de la thèse.

L'estimation du niveau dégradation associé à la défaillance, que l'on appelle dans la suite par commodité niveau de défaillance, est l'objet de la troisième partie du document. Classiquement, ce niveau de défaillance est déterminé sur la base de caractéristiques physiques ou d'avis d'experts. Cependant ce niveau de défaillance "théorique" n'est pas toujours cohérent avec les données associées à des défaillances réelles. L'accent est donc mis sur la réduction de cet écart. Pour ce faire, la loi de la durée de vie est supposée connue ou, tout au moins, estimée sur la base de données de défaillances. Le niveau de défaillance peut alors être déterminé de telle sorte que le processus stochastique de dégradation considéré conduise à une distribution du premier temps d'atteinte du niveau de défaillance qui corresponde à la densité estimée.

La quatrième partie est dédiée à l'optimisation de la maintenance lorsque le processus de dégradation considéré est un processus OU dépendant du temps. S'appuyant sur les données de surveillance continue du système et sur le pronostic de l'instant de défaillance, un compromis peut être trouvé entre les coûts de maintenance préventive et ceux associés à une

défaillance du système. Dans ce contexte, la formulation non explicite du pronostic de l'instant de défaillance ne permet pas d'exploiter les techniques d'optimisation classiques. Une approximation de la densité de l'instant de défaillance est donc proposée pour mener à bien cette étape d'optimisation.

Table des matières

Abbreviations and Notations

Abbreviations

pdf	Probability density function
cdf	Cumulative distribution function
MTTF	Mean time to failure
OU	Ornstein-Uhlenbeck
FPT	First passage time
RFPT	Randomized first passage time
RUL	Residual useful lifetime
PHM	Prognostics and health management
VIE	Volterra integral equation
PDE	Partial differential equation
SDE	Stochastic differential equation
MC	Monte-Carlo
AIC	Akaike information criterion
LPT	Last passage time
IFPT	Inverse first passage time
CBM	Condition-based maintenance
argmax	Argument of the maximum
argmin	Argument of the minimum
$\mathbf{IPA}_k$	Indicator of type k for prognosis assessment

Notations

$X_t^{y,s}$	The process X_t with an initial value y at time s.
$L(t)$	An enough smooth, upper and time-dependent failure level.
$\wedge$	The smaller one between two values.
$\vee$	The larger one between two values.
$\tau_{y,s}$	Conditional first passage time of the process to a given boundary based on the observation of y at time s
$a(t)$	Mean-reversion coefficient in $dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t$
$b(t)$	Drift coefficient in $dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t$
$\sigma(t)$	Diffusion coefficient in $dX_t = (a(t)x_t + b(t))dt + \sigma(t)dB_t$
$c(t)$	$\sigma^2(t)/2$
$\alpha(t,s)$	$-\int_s^t a(u)du$
$\beta(t,s)$	$-\int_s^t b(u)e^{\alpha(u,s)}du$
$\gamma(t,s)$	$\int_s^t c(u)e^{2\alpha(u,s)}du$
$\delta(x)$	Dirac measure centered at 0
$p(x,t\|y,s)$	Transition probability density function from $X_s = y$ to $X_t = x$, $t \geq s$
$F(x,t\|y,s)$	Transition probability from $X_s = y$ to $X_t = x$, $t \geq s$
$u(x,t\|y,s)$	Joint probability of $P(X_t \leq x, \tau_{y,s} > t\|X_s = y)$, $t \geq s$

$w(x,t\|y,s)$	$\frac{\partial u(x,t\|y,s)}{\partial x}$
$\rho_{x,y}$	Correlative coefficient of random variables x and y
$\mathrm{cov}(x,y)$	Covariance function of random variables x and y
$\mathrm{var}(x)$	Variance function of the random variable x
$\mathbb{E}(x)$	Expectation of the random variable x
$\inf A$	The infimum of the set A
$\sup A$	The supremum of the set A
$\sim \mathcal{N}(\mu,\sigma^2)$	Subject to normal distribution with mean μ and variance σ^2

General Introduction

Background

The core issue discussed in this thesis is system failures. And the target is to utilize observed information in industrial systems to describe, predict and prevent system failures. This is a general topic discussed in reliability engineering [60], prognosis and system health management (PHM) [70] and condition-based maintenance [16, 34]. To the extent of the author's understanding, reliability engineering concerns more on the descriptive research of the system failures, PHM concerns more on the prediction techniques for the failures, and condition-based maintenance concerns more on the prevention of the system failures.

The above understanding induces the organization of discussions in this thesis, the same three issues for system failures will be reproduced. However we emphasize in this thesis on the prognosis of system failures, and the title of this thesis is given based on this consideration.

This work is done in a model-based view to suppose the system degradation process is determined by a justified model [1]. Therefore the system health state can be predicted based on this model. Moreover supposing the system failure is described by the first passage failure related to the system state, the prediction of system failures can be done based on the prediction of system states. To prevent the system failures, the preventive maintenance is introduced based on maintenance optimization problems. The role of first passage failures in a model-based view is presented in Figure 1 [1].

1. Technical issues such as failure level, first passage failure, preventive maintenance, maintenance optimization will be explained later in corresponding chapters.

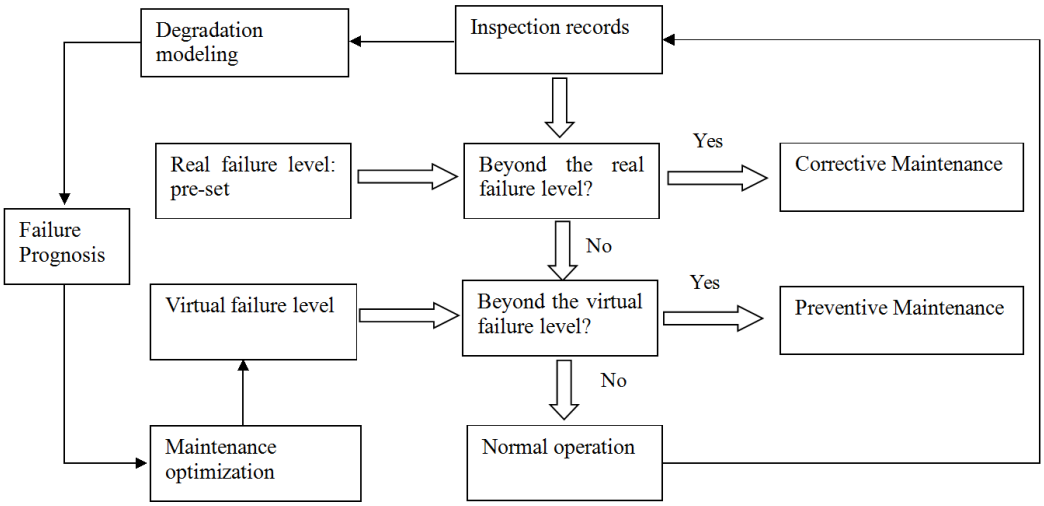


FIGURE 1 – System failures as first passage failures : description, prediction and prevention.

Thesis Structure

This thesis consists of four parts : stochastic degradation modeling, prognosis of system failures, failure level estimation and maintenance optimization. As stated before, these parts are connected and dedicated to three core issues of system failures : description, prediction and prevention.

The first part about stochastic degradation modeling proposes a degradation model based on a time-dependent Ornstein-Uhlenbeck (OU) process. Such a process utilizes the inspection records to establish the dynamic description of degradation process. The time-dependent OU process is proved good by its statistical properties on controllable mean, variance and correlation. Its mean-reverting property can be introduced to interpret temporary correlated fluctuations from an overall degrading trend in degradation records. Corresponding parameter estimation is proposed based on maximum likelihood estimation method. A case study is performed to test the model's fitting goodness based on a degradation data-set of a passive component in power plants.

The second part about prognosis of system failures is discussed further based on the time-dependent OU process. And the first passage time is introduced as the system failure time, to a pre-set failure level. Later how to estimate the failure time is discussed based on two kinds of views : partial differential equation and integral equation. These two views lead to various estimation techniques from different concentrations, and they can be classified into 3 categories : analytical approximations, numerical algorithms and Monte-Carlo simulation methods. Case studies are done to calculate first passage density based on proposed methods.

The third part about failure level estimation proposes some techniques to estimate failure levels based on inverse first passage problems. In previous literature the failure level is generally treated as physical barriers or experts' opinions, based on which failure prognosis from first passage failures can hardly fit existing failure records. Therefore the effort in this part is paid to make up the gap between failure records and inspection records under the definition of first passage failure based on inverse first passage problem. And suppose the lifetime distribution is given or estimated from failure records, we emphasizes on numerically reproducing the failure level under which the first passage time of the given stochastic process can have the same lifetime distribution with the given lifetime distribution.

The fourth part about maintenance optimization investigates how to optimize maintenance policies based on time-dependent OU process. Based on monitored system conditions and prognosis of system failures, condition-based maintenance can be adopted to introduce preventive maintenance such that the balance can be achieved between operation costs and disastrous results caused by system failures. In this part, corresponding maintenance optimization problems are discussed based on the time-dependent OU process and the hypothesis of continuously monitored system. Due to the unexplicit expression for prognosis of system failures, classical heuristic optimization procedures cannot be fulfilled. Therefore approximate first passage density is used to fulfill the maintenance optimization.

Main Contributions

Compared to previous literature, this thesis is new from the following aspects :

- A time-dependent OU process is introduced into the field of stochastic degradation modeling with applications to prognosis and health management, and maintenance optimization.
- Prognosis of system failures based on first passage problems is considered from a

technical view. Analytical approximations, numerical algorithms and a piecewise quasi-linear Monte-Carlo method are proposed and fulfilled.

- The gap between failure records and failure prognosis based on first passage failures is investigated from a data-analysis view, where the inverse first passage problem is introduced to do failure level estimation.
- Maintenance optimization for a continuously monitored system is presented for the time-dependent OU process.

Chapitre 1

Degradation Modeling Based on a Time-Dependent Ornstein-Uhlenbeck Process

This chapter is written mainly based on the submitted manuscript [20].

In this chapter, a time-dependent Ornstein-Uhlenbeck (OU) process is introduced to stochastic degradation modeling. The time-dependent OU process is proposed from its statistical properties on controllable mean, variance and correlation. Its mean-reverting property can be introduced to interpret temporary correlated fluctuations from an overall degrading trend in degradation records. Corresponding parameter estimation is proposed based on the maximum likelihood estimation method. Several simulation tests are done to test the model's fitting goodness based on a degradation data-set of a passive component in power plants.

The whole chapter is organized as follows. In Section 1.1, we will review several existing stochastic process models and describe the considered system. In Section 1.3, we will introduce the statistical properties for the time-dependent OU process, its mean-reverting property will also be emphasized. In Section 1.4, a maximum likelihood estimation for statistical inference will be introduced. In Section 1.5, a case study will be considered based on several information criterions to compare the fitting goodness of the OU process and a linear-diffusion model, where the degradation data-set comes from a passive component in power plants. At the end, several conclusions and perspectives are presented in Section 1.6.

1.1 Introduction

In this thesis, we concentrate on deteriorating systems subject to gradual degradation processes such as fatigue crack, corrosion, erosion etc.. Generally speaking, inspection data related to this kind of gradual degradation phenomena show many different recorded paths for the same type of system in the same environment or test parameters. The variations from one path to another are usually considered uncertainties due to measurement errors, unit-to-unit difference and internal deterioration mechanisms [68]. Purely deterministic physical models can not always explain and represent properly such uncertainties.

Hence, when degradation mechanisms are not revealed or too complex, the physical modeling may be inadequate and the degradation modeling is therefore a regression or filtering problem for chosen models. In such cases, the flexibility of the degradation models is of importance as it is connected to the model's fitting-goodness to degradation records. In this context, we propose to model the system deteriorating state by a continuous time stochastic process which is called degradation process. A realistic hypothesis is adopted that temporary fluctuations in degradation process can be tolerated, while the degradation process should have an overall degrading trend.

For commonly used processes such as Gamma process [84] and drifted Brownian motion [66], the mean tendency or the drift can be chosen quite freely but the variance strongly depends on the considered process. Following this idea, the modeling stage considered in this thesis is started from controlling statistical properties of the degradation records and possibly taking into account expert opinions. Specifically the choice of both a mean degradation tendency and a variance function will be considered. The possibility to build a model with a free choice for the mean degradation tendency and for the variance function is the main contribution of this chapter and justifies the use of a time-dependent OU process.

From the overall degrading trend, we can consider stochastic processes devoted to the modeling of degradation as a pure monotone phenomenon where increasing Lévy processes and jump processes are most widely used including Gamma process [84], inverse Gaussian process [88] etc. This comes from the hypothesis that the damage of system is caused by small shocks which accumulate without any "self-repair" in systems [75]. Also the monotone property results in an explicit formula to describe the failure time for a constant failure level when first passage failure is considered [84, 88]. This advantage makes later estimation easier on failure time, residual useful lifetime and mean time to failure (MTTF) etc.

However, the monotone process cannot explain well on fluctuations existing in some degradation records. To describe such fluctuations, a first intuitive idea is that fluctuations come from external stochastic impacts such as measurement errors, unit-to-unit difference, noisy observations etc.. This idea induces the introduction of external stochastic impacts for existing models. For instance, Khanh considered introducing independent Gaussian noises for Gamma process model [77]. Blain treated the fluctuations in the stress crack corrosion as external noises [8]. And the stochastic filtering models for Wiener process [68], Markov-switching models [65] and references therein. A good overview on statistical data-driven RUL estimation is given by Si et al. in [67].

Especially, a good applicable tool would be to use random coefficient regression models or mixed effect models [94]. However, as stated in the review [67], these models give the pdf of the RUL in some simple cases or under the assumption of monotonic degradation paths (then applying this kind of methods may generate a conservative estimation). This approximation is adopted in [94], the degradation model is supposed to be monotone whereas it is not (assumption A2 therein). Another drawback of these kind of models is that they do not take into account the temporal variability of the RUL [51].

A second intuitive idea is to investigate those fluctuations from the internal system uncertainties which may be due to noises such as non-homogeneous material, random loading and thermal effects in fatigue crack modeling [75]. In these noise-excited dynamic systems, white noise is commonly accepted and introduced to induce correlated Gaussian fluctuations in degradation records. Correspondingly diffusion processes are therefore introduced for degradation modeling. For example, to describe crack growth, physical modeling based on deterministic crack models [52] leads to the consideration of fluctuations caused by random residual stress [73, 76] etc.

These diffusion degradation models are established based on specific physical laws such that they can hardly be applied to a general degradation process where degradation mechanisms are not revealed. Moreover when boundary-crossing failures are considered, corresponding boundary-crossing problems are difficult for these general diffusion processes. Therefore previously most of diffusion models adopted in the degradation modeling are still related to Brownian motion [54, 45], where the first passage failure can be described explicitly by inverse Gaussian distribution. However the linearity of Brownian motion restricts its application and

to increase the model's fitting-goodness, Si et al. [66], Tseng and Peng [82] have considered a process with nonlinear drift and linear diffusion whose RUL estimation are based on an approximate expression of first passage density of Gauss processes [25]. This model is also used by Le Son et al. to test the PHM 2008 data challenge data in [78]. A more general model named Gauss-Markov process [23] attracts also researchers' attention recently, in Peng and Tseng [55] corresponding RUL estimation is considered based on numerical solutions to an integral equation [23].

Under the framework of Gauss-Markov process, the average performance of the degradation process can be described well as in Si et al. [66, 55]. However in these Brownian-like processes, uncertainties are introduced with little control such that the flexibility to adjust the model's volatility is absent. This chapter is therefore triggered to introduce a dynamic model which balances between the model's complexity and the efficiency on RUL estimation. As is shown later, a time-dependent Ornstein-Uhlenbeck process will be proposed to control the statistical quantities of models on mean, variance and correlation, where the conditional first passage failure to a time-dependent boundary is adopted for RUL estimation. Moreover, its mean-reverting property is interpreted as a self-repair mechanism to compensate uncertainties introduced in this time-dependent OU process. This interesting property is emphasized from a basic fact : when we talk about a degradation process, it should have an overall degrading trend if this is a real degradation. Even fluctuations exist in degradation records, these fluctuations should be temporary and system's degradation will tend to move to a monotone average degradation performance in long-term period.

Summarize all the above and for mathematical simplification, the degradation process or system health state considered in this thesis is supposed to have the following properties :

- The system deteriorates gradually without the risk of large shocks, such that the degradation process can be supposed to be continuous.
- The degradation records are disturbed by noises from an underlying degradation process, which has a long-term monotone trend.
- Noises in degradation records are gaussian and they come from internal system mechanisms, which can accumulate and influence later records.
- The system is with the memory-less property, i.e. the current system state determines the future evolution of the system.

These hypotheses are not restrictive. They are rather common and the foundation for many literature related to diffusion processes, see [68, 66, 65] and references therein.

1.2 The Overview on Stochastic Degradation Processes

In this section, several stochastic processes used in stochastic degradation modeling will be introduced. The discussion concerns :

- Homogeneous Gamma process, which belongs to pure jump processes. This leads to the introduction of nonhomogeneous Gamma process.
- Brownian motion, which belongs to continuous processes. This leads to the introduction of stochastic differential equations.

Our contribution in this thesis related to passage problems is limited in the framework of stochastic differential equations. And it should be noticed that a key characteristic to distinguish the stochastic degradation modeling with other stochastic modeling areas is failure-oriented modeling. Stochastic degradation modeling is discussed in the field of reliability engineering, and it serves as a first step for reliability analysis, prognostics and health ma-

nagement, and condition-based maintenance. So the fitting goodness to existing inspection records is not the only indicator to choose models, how to connect this modeling work with the failure's description and prognosis attracts more attention. The emphasis on the computability of failure prognosis corresponding to the modeling work should be put into the first place during modeling.

Throughout this thesis, the system failure is considered as a first passage time, that is to say, given a pre-set failure level, the system is considered as failed when the system degradation is beyond the failure level. This definition is utilizable and widely accepted in engineering problems, but also it provides the possibility to consider the system failure in a mathematical way combining with the stochastic modeling work. Actually, if the stochastic degradation process X_t has been established, it would be convenient to consider the failure time as the random time $\tau = \inf_{t>0}\{X_t \geq L(t)\}$ where $L(t)$ is the given failure level.

So summarize all the above, the introduction presented in this section emphasizes on three points : the definition of the stochastic process, the transition pdf leading to model's estimation and the possibility of solving first passage problems.

1.2.1 Gamma Process

Homogenous Gamma Processes

The statement follows the one in [84, 63]. Homogeneous Gamma process is one of the most widely accepted degradation models in the field of reliability engineering. And the Gamma process can describe well degradation phenomenons caused by cumulative small shocks, such as the crack growth due to shocks in the trackway [63].

Recall for a Gamma random variable $Ga(v, u)$, for $v, u \in \mathbb{R}^+$ the corresponding Gamma distribution function is given by

$$Ga(x|v,u) = \frac{u^v}{\Gamma(v)} x^{v-1} \exp(-ux) I_{[0,+\infty)}(x), \tag{1.2.1}$$

where $I_A(x)$ is the indicator function, and $\Gamma(a) = \int_0^{+\infty} z^{a-1} e^{-z} dz$ is the Gamma function for $a > 0$. u is called scale parameter and v shape parameter.

When the shape parameter in the Gamma distribution is a linear function $v \cdot t$ with $v > 0$, the homogeneous Gamma process is derived as follows

Definition 1.2.1. A continuous-time stochastic process $\{X_t, t \geq 0\}$ is called homogeneous Gamma process if for the shape function $v \cdot t$ and scale parameter $u > 0$,

1. $X_0 = 0$, almost surely,
2. $X_t - X_s \sim Ga(v \cdot (t-s), u), \forall t > s \geq 0$,
3. X_t has independent and stationary increments.

Therefore the transition pdf, that is to say, the pdf of the distribution of X_t conditioned on $X_s = y$ is given by

$$p(x, t|y, s) = Ga(x - y|v \cdot (t-s), u). \tag{1.2.2}$$

Noticing the right part is only related to $t - s$, therefore $p(x, t|y, s) = p(x, t-s|y, 0)$, so X_t is homogeneous.

Some properties of the homogeneous Gamma process are of interest :

1. $\mathbb{E}(X_t) = \frac{v}{u}t$ and $\text{var}(X_t) = \frac{v}{u^2}t$ are all linear functions. Therefore homogeneous Gamma process can model those degradation process with linear tendency.
2. $\frac{\mathbb{E}(X_t)}{\text{var}(X_t)}$ is a constant. This strong constraint limits its application where inspection records don't show this appearance.
3. X_t is a pure jump, Markov process. So this memory-less property provides a much easier situation when problems are considered.
4. X_t is with nondecreasing trajectories. This property fits most of degradation processes, as in general the engineering system is considered to degrade all the time without maintenance.

Nonhomogeneous Gamma Processes

However, the linear tendency cannot fit all degradation in consideration, and therefore in some cases homogeneous Gamma process is not preferred. So it is natural to introduce a non-homogeneous Gamma process as follows.

Let $v(t)$ be a non-decreasing, right-continuous, real-valued function for $t \geq 0$, with $v(0) = 0$. Then it comes to the definition of non-homogeneous Gamma process [84].

Definition 1.2.2. A continuous-time stochastic process $\{X_t, t \geq 0\}$ is called Gamma process if for the shape function $v(t) > 0$ and scale parameter $u > 0$,

1. $X_0 = 0$, almost surely,
2. $X_t - X_s \sim Ga(v(t) - v(s), u), \forall t > s \geq 0$,
3. X_t has independent increments.

Therefore the transition pdf is given by

$$p(x, t|y, s) = Ga(x - y|v(t) - v(s), u), \tag{1.2.3}$$

Noticing the right part cannot be presented with a function of $t - s$, therefore $p(x, t|y, s) \neq p(x, t - s|y, 0)$, so X_t is nonhomogeneous.

Some properties of the nonhomogeneous Gamma process are of interest :

1. $\mathbb{E}(X_t) = \frac{v(t)}{u}$ and $\text{var}(X_t) = \frac{v(t)}{u^2}$ are all ajustable functions. Therefore nonhomogeneous Gamma process can model those degradation process with nonlinear tendency by choosing proper shape function $v(t)$.
2. $\frac{\mathbb{E}(X_t)}{\text{var}(X_t)}$ is a constant. This strong constraint limits its application where inspection records don't show this appearance.
3. X_t is a pure jump, Markov process. So this memory-less property provides a much easier situation when problems are considered.
4. X_t is with nondecreasing trajectories. So like homogeneous Gamma process, it suppose that the engineering system is considered to degrade all the time without maintenance.

Failure time for Increasing Processes

As stated before, when the degradation process has been established, the system failure time can be modeled by a first passage time to a preset failure level. So deriving a description of such a failure time could be of interest in the field of reliability engineering. Furthermore

it is noticed that for an increasing process X_t with initial value x_0 (e.g. Gamma process), for a constant boundary L, its first passage time $\tau := \inf_{t \geq t_0}\{X_t \geq L\}$ satisfies :

$$P(\tau > t) = P(X_t < L) = \int_{x_0}^{L} p(x, t|x_0, t_0)dx, \tag{1.2.4}$$

where $p(x, t|y, s)$ is the transition pdf of X_t if it is existing.

This property provides an explicit expression on the statistical description of the failure time for increasing processes. Actually the monotone property promises that the process cannot go back to its previous state, so there is no problem of "first passage", and the current state of the process determines whether it is beyond or not the crossing boundary for the future.

1.2.2 Brownian Motion

Gamma processes and other increasing processes are adopted widely in the field of reliability engineering to model degradation processes. However some cases cannot be explained clearly by increasing processes, when degradation records are with fluctuations. This may be due to internal mechanisms such as crack closure [75], or external mechanisms such as measurement errors. For the external fluctuations, we can still model the degradation process by increasing processes with the help of filtering. However for internal fluctuation mechanisms, it needs more tools. So the Brownian motion is proposed to model a Gaussian noise in the system.

The Brownian motion is also called Wiener process and it is defined as follows, which can describe linearly increasing and fluctuating degradation records.

Definition 1.2.3. A stochastic process $\{B_t, t \geq 0\}$ is called a (standard) Brownian motion, if

1. $B_0 = 0$,
2. $\{B_t, t \geq 0\}$ has independent and stationary increments,
3. $B_t \sim \mathcal{N}(0, t), \forall t > 0$.

Then the transition pdf for Brownian motion B_t is given by

$$p(x, t|y, s) = \frac{1}{2\pi\sigma^2(t-s)} \exp(-\frac{(x-y)^2}{2\sigma^2(t-s)}) \tag{1.2.5}$$

Some properties of the Brownian motion are of interest :

1. $\mathbb{E}(B_t) = 0$ and $\text{var}(B_t) = t$, so its mean is zero, and its variance is a linear function.
2. B_t is not a jump process, but it is still a Markov process. So this memory-less property provides a much easier situation when problems are considered.
3. B_t is with fluctuating trajectories.
4. Brownian motion is almost surely continuous, unlike Gamma processes, it is without jumps. So it suppose there is no huge shocks in the systems such that the system state can be continuous.

Brownian motion can model well the mean-zero, linear-variance Gaussian noises in the degradation process, so it is natural to propose a drifted Brownian motion X_t to model degradation processes as follows

$$X_t = x_0 + \mu t + \sigma B_t, \tag{1.2.6}$$

where $x_0 \in \mathbb{R}, \mu > 0, \sigma > 0$. With such an expression, the engineering system is not always degrading, but sometimes it can automatically have some self-repair.

Therefore for the drifted Brownian motion X_t, it has a mean $\mathbb{E}(X_t) = x_0 + \mu t$ and other properties remain the same with Brownian motion. It should be noticed that similar to Gamma process, $\frac{\mathbb{E}(X_t)}{\text{var } X_t}$ is a constant for the drifted Brownian motion if $x_0 = 0$.

To describe the system failure, we consider a constant failure level L for $X_t = x_0\mu t + \sigma B_t$, then the failure density function is given by the inverse Gaussian distribution :

$$g(t) = \frac{L - x_0}{\sqrt{2\pi\sigma^2 t^3}} \exp(-\frac{(L - x_0 - \mu t)^2}{2\sigma^2 t}). \tag{1.2.7}$$

Then the reliability function or the survival function is given by $\int_t^{+\infty} g(s)ds$.

This explicit expression also simplifies the analysis to apply such a process in engineering problems. However it is also remarked that for the Brownian motion, except for a few cases of failure levels, the first passage failure distribution cannot be solved explicitly.

Gauss-Markov Process

The Brownian motion is with a linearly increasing variance, and in some cases, this modeling could put over-estimated uncertainties into the degradation process. So the Brownian motion can be extended to Gauss-Markov process to control such a variance, it is defined as follows :

Definition 1.2.4. A stochastic process $\{X_t, t \geq 0\}$ is called Gauss-Markov process, if for a non-zero function $h(t)$ and non-decreasing function $f(t)$,

1. $X_0 = 0$,
2. $\{X_t, t \geq 0\}$ has independent increments,
3. $X_t = h(t)B(f(t)), \forall t > 0$, $B(*)$ is a standard Brownian motion.

Some properties of the Gauss-Markov process are of interest :

1. $\mathbb{E}(X_t) = 0$ and $\text{var}(X_t) = h(t)^2 f(t)$, so its mean is zero, and its variance can be adjusted by choosing appropriate functions h, f for applications.
2. X_t is a Markov process. So this memory-less property provides a much easier situation when problems are considered.
3. X_t is with fluctuating trajectories.
4. Gauss-Markov process is almost surely continuous, unlike Gamma processes, it is without jumps. So it suppose there are no huge shocks in the systems such that the system state can be continuous.

Gauss-Markov process can model well nonlinear-variance, zero-mean Gaussian noises, so it is natural to propose a drifted Gauss-Markov X_t to model degradation processes as follows

$$X_t = n(t) + h(t)B(f(t)). \tag{1.2.8}$$

With such an expression, the engineering system is not always degrading, but sometimes it can automatically have some self-repair. And $\mathbb{E}(X_t) = n(t)$, but other properties remain the same with Gauss-Markov process.

For the Gauss-Markov process, it is hard to get an explicit expression for the failure description based on the first passage failure definition, although it can flexibly model degradation processes. So this is a challenge, but also a current research emphasis in the field of reliability engineering [66, 68].

1.2.3 Itô's Stochastic Differential Equation

Although we presented many models to model degradation processes such as Gamma processes and drifted Brownian motion etc., one default for those modeling work is that these processes are proposed based on more intuitive consideration. That is to say, we just consider simple models to describe generally the tendency of the degradation models. It is hard to find the micro understanding of these processes to describe a system mechanism.

Therefore to consider more complex modeling work, and especially to combine the modeling work with existing physical laws, it is natural to introduce stochastic differential equations (SDEs) based on Brownian motion. SDEs first appear in the beginning of 20th century due to some research on statistical mechanics, and is strictly formulated by Itô in 1950s. It can be accepted from the following Langevin-type equation [30, 42] :

$$\frac{dX_t}{dt} = \tilde{m}(X_t, t) + \sigma(X_t, t)\xi_t, t \geq 0, \tag{1.2.9}$$

where ξ_t is white noise, i.e. $\mathbb{E}(X_t) = 0, \mathbb{E}(X_t X_s) = \delta(t-s)$, where δ is the Dirac measure centered at 0 (e.g. p39, [30]). This system can be viewed as the original ordinary differential system influenced by the noise $\sigma(X_t, t)\xi_t$.

Consider an informally transformed expression of white noise by Brownian motion $\xi_t = \frac{dB_t}{dt}$, we can get the general expression of SDE. However this physical illustration leads to the Stratonovich SDE (e.g. p.111, [30]), corresponding to the following Itô's SDE :

$$dX_t = m(X_t, t)dt + \sigma(X_t, t)dB_t, \tag{1.2.10}$$

where $m(X_t, t) = \tilde{m}(X_t, t) - \frac{1}{2}\sigma_x(X_t, t)\sigma(X_t, t)$, B_t is a standard Brownian motion. $m, \sigma : R \times [0, +\infty) \to R$ are properly defined functions and let X_0 be the initial value independent of B_t.

The solution of (1.2.10) is a Markov process with continuous trajectories a.s.. And its transition pdf $p(x, t|y, s)$ when $t > s$ satisfies the following Fokker-Planck equation :

$$\frac{\partial p(x,t|y,s)}{\partial t} = -\frac{\partial (m(x,t))p(x,t|y,s)}{\partial x} + \frac{1}{2}\frac{\partial^2 \sigma^2(x,t)p(x,t|y,s)}{\partial x^2}. \tag{1.2.11}$$

Several properties of SDEs are of interest :

1. SDEs can be derived directly from existing physical models by introducing Gaussian white noises, such that many classical results and empirical research can be extended to a stochastic version. In the field of reliability engineering, the effort has been paid to extend the work of crack modeling, e.g. the Paris' Law to the SDE models [74, 75, 73].
2. By introducing stochastic influences, SDEs are with much more flexibility in the modeling work than the deterministic models, and it could be expected to have more accurate fitting results than deterministic models.

3. It is an almost surely continuous, Markov process. This memory-less property provides a simplified discussion.

But because of the complexity of SDEs, only a few cases can be solved explicitly with also explicit transition pdf. Therefore our concentration in this thesis will be put on the following time-dependent Ornstein-Uhlenbeck (OU) process

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0. \tag{1.2.12}$$

It will be presented later that the OU process has an explicit solution and explicit transition pdf.

Moreover it is not an easy work to estimate the failure time based on SDE models. And this problem based on first passage definition is the main topic discussed in this thesis.

1.3 Ornstein-Uhlenbeck Process

1.3.1 Main Properties

General Process

As stated before, the OU process is one of few cases which can be solved explicitly with also explicit probability laws in stochastic differential equations. Therefore to apply such a process into the field of reliability engineering could be interesting to balance between the modeling and the computability. Also based on proposed system hypotheses, it will be presented that a time-dependent Ornstein-Uhlenbeck process could be a good candidate for degradation modeling. OU process is widely applied in fields such as financial market, to describe physical dynamics of systems which stabilize at its equilibrium point. It is given by

$$dX_t = (aX_t + b)dt + \sigma dB_t, \tag{1.3.1}$$

where a, b, σ are constants, B_t is the standard Brownian motion.

However it is inconvenient to be applied directly in the degradation modeling as the mean of OU process is a constant which disobeys the overall degrading trend for degradation data. Two modifications to compensate this inconvenience include adding a drift as in [47] and considering time-dependent coefficients as in [2]. As a consequence the following Itô's stochastic differential equation will be considered throughout to define stochastic processes for degradation modeling [92] :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0. \tag{1.3.2}$$

where a, b and σ are enough smooth functions and B_t is standard Brownian motion. The initial value X_0 is a random variable independent of B_t (including the deterministic case where the distribution degrades to Dirac measure).

Furthermore, following notations are considered throughout this thesis :

$$\begin{aligned} \alpha(t,s) &= -\int_s^t a(u)du, \\ \beta(t,s) &= -\int_s^t b(u)e^{\alpha(u,s)}du, \\ \gamma(t,s) &= \int_s^t c(u)e^{2\alpha(u,s)}du, \end{aligned} \tag{1.3.3}$$

where $c(t) = \frac{\sigma^2(t)}{2} > 0$.

Equation (1.3.2) is too general to be connected with the degradation modeling in a straight way but several interesting properties can be derived.

- An explicit form of X_t can be derived from Equation (1.3.2). It is given under notations in (1.3.3) :

$$X_t = e^{-\alpha(t,0)}\left[X_0 - \beta(t,0) + \int_0^t \sigma(s)e^{\alpha(s,0)}dB_s\right]. \tag{1.3.4}$$

- The mean of the process is given by :

$$\mathbb{E}(X_t) = e^{-\alpha(t,0)}\left(\mathbb{E}(X_0) - \beta(t,0)\right). \tag{1.3.5}$$

- The process' covariance is :

$$\operatorname{cov}(X_t, X_s) = e^{-(\alpha(t,0)+\alpha(s,0))}\left\{\operatorname{var}(X_0) + \int_0^{t\wedge s} \sigma^2(u)e^{2\alpha(u,0)}du\right\}, \tag{1.3.6}$$

  which means correspondingly that the variance can be written as :

$$\operatorname{var}(X_t) = e^{-2\alpha(t,0)}\left\{\operatorname{var}(X_0) + \int_0^t \sigma^2(u)e^{2\alpha(u,0)}du\right\}. \tag{1.3.7}$$

- Moreover, the process' correlation coefficient is :

$$\rho_{X_t,X_s} = \frac{\operatorname{cov}(X_t, X_s)}{\sqrt{\operatorname{var}(X_t)\operatorname{var}(X_s)}} = \frac{\sqrt{\operatorname{var}(X_0) + \int_0^s \sigma^2(u)e^{2\alpha(u,0)}du}}{\sqrt{\operatorname{var}(X_0) + \int_0^t \sigma^2(u)e^{2\alpha(u,0)}du}}, t \geq s \tag{1.3.8}$$

What should be mentioned more is that the OU process is Markov, with continuous trajectories (almost surely).

Preserving Mean and Variance

Let us recall that one of the aims of this chapter is to propose and consider a process X_t such that

$$\mathbb{E}(X_t) = m(t) \text{ and } \operatorname{var}(X_t) = v(t). \tag{1.3.9}$$

where $m \in C^1[0, +\infty)$ and $v \in C^1[0, +\infty)$ are functions chosen from statistical analysis or from expert opinion with $v(t) \geq 0$, $\forall t \geq 0$.

When $v(0) > 0$, from Equations (1.3.5) and (1.3.7) it can be shown that (1.3.9) is obtained by defining functions a and b in (1.3.2) as :

$$a(t) = \frac{v'(t) - \sigma^2(t)}{2v(t)}, \tag{1.3.10}$$

$$b(t) = m'(t) - a(t)m(t), \tag{1.3.11}$$

where the pending $\sigma(t)$ is introduced to adjust further the covariance of the process. The case of a deterministic initial value i.e. with $v(0) = 0$ has to be considered carefully where Equation (1.3.10) may not be applied. A specific case will be considered later where $v(t)$ is chosen as an exponential function.

A Special Time-dependent OU process

From now on and from (1.3.11), the following special time-dependent OU process X_t is considered in this chapter instead of the general process (1.3.2). It is described by the following Itô's stochastic differential equation :

$$dX_t = (a(t)X_t + m'(t) - a(t)m(t))dt + \sigma(t)dB_t, \quad t \geq 0, \tag{1.3.12}$$

where X_0 is a random variable with $\mathbb{E}(X_0) = m(0), \mathrm{var}(X_0) = v(0)$. Remark here $a(t)$ can be replaced by expressions in (1.3.10) when $v(0) > 0$ such that Equation (1.3.9) holds.

The process X_t is non-stationary and time-inhomogeneous and its explicit expression can be formally derived from (1.3.4). It comes :

$$X_t = e^{-\alpha(t,0)}\Big(X_0 - \beta(t,0) + \int_0^t \sigma(u)e^{\alpha(u,0)}dB_u\Big), \tag{1.3.13}$$

with updated notations

$$\alpha(t,s) = -\int_s^t a(u)du \text{ and } \beta(t,s) = m(s) - m(t)e^{\alpha(t,s)}.$$

It can be derived naturally from (1.3.13) that $\mathbb{E}(X_t) = m(t)$.

Some particular cases of this process can be considered specifically.

- The limit case where $a(t) = 0$ and $\sigma(t) = \sigma$ is a constant was studied by Si et al. [66]. In this case $X_t = X_0 + (m(t) - m(0)) + \sigma B_t$ is drifted Brownian motion. Moreover, we have $\mathbb{E}[X_t] = m(t), \mathrm{var}(X_t) = v_0 + \sigma^2 t$.
- When $a(t)$ is a negative function, several additional technical conditions on $a(t)$ and $\sigma(t)$ can be proposed :

$$\alpha(+\infty,0) = -\int_0^{+\infty} a(u)du = +\infty, \quad \gamma(+\infty,0) = \frac{1}{2}\int_0^{+\infty} \sigma^2(u)e^{2\alpha(u,0)}du = +\infty. \tag{1.3.14}$$

Corresponding physical meaning of such conditions will be revealed later by consideration of mean-reversion / self-repairing mechanism.
When the condition (1.3.14) is satisfied, from l'Hôpital's rule we could find that :

$$\lim_{t\to+\infty} \mathrm{var}(X_t) = \lim_{t\to+\infty} \frac{v(0) + \int_0^t e^{2\alpha(u,0)}\sigma^2(u)du}{e^{2\alpha(t,0)}} = \lim_{t\to+\infty} -\frac{\sigma^2(t)}{2a(t)}. \tag{1.3.15}$$

Also the asymptotic correlation coefficient is obtained when $h > 0$ is fixed :

$$\begin{aligned}\lim_{s\to+\infty} \rho_{X_{s+h},X_s} &= \frac{\mathrm{cov}(X_{s+h},X_s)}{\sqrt{\mathrm{var}(X_{s+h})\,\mathrm{var}(X_s)}} = \lim_{s\to+\infty}\sqrt{\frac{\int_0^s e^{2\alpha(u,0)}\sigma^2(u)du}{\int_0^{s+h} e^{2\alpha(u,0)}\sigma^2(u)du}} \\ &= \lim_{s\to+\infty}\frac{\sigma(s)}{\sigma(s+h)}\exp(\int_s^{s+h} a(u)du).\end{aligned} \tag{1.3.16}$$

Therefore the mean of the process (1.3.12) is expressed by $m(t)$. The variance and the correlation coefficient of the process when time is large can be fully determined by the function $\sigma(t)$ and the function $a(t)$. If the variance and the covariance are given by a statistical analysis of the data-set, the model can be roughly established without further calculation.

- The case $a(t) = -r$ being a constant ($r > 0$) leads to some particular asymptotic properties. From (1.3.15), it can be derived that :

$$\lim_{t\to+\infty} \text{var}(X_t) = \lim_{t\to+\infty} \frac{\sigma^2(t)}{2r}. \tag{1.3.17}$$

The asymptotic coefficient of correlation is obtained from (1.3.16) when $h > 0$ is fixed :

$$\lim_{s\to+\infty} \rho_{X_{s+h},X_s} = \lim_{s\to+\infty} \frac{\sigma(s)}{\sigma(s+h)} \exp(-rh). \tag{1.3.18}$$

Therefore, if $\sigma(t) \to \sigma$ when $t \to +\infty$, then we know the correlation coefficient when time is large is almost $\exp(-rh)$.

- Special attention can be paid to model the case when a constant variance (or almost a constant) is observed ($v(t) \sim v$). In such a case we can propose $\sigma(t) = \sigma$ as a constant from (1.3.17) where r is also constant. As a consequence the process is proposed as in [18, 17] :

$$dX_t = (-rX_t + m'(t) + rm(t))dt + \sigma dB_t, t \geq 0. \tag{1.3.19}$$

The solution of Equation (1.3.19) based on X_0 can be given from (1.3.13) by :

$$X_t = m(t) + (X_0 - m(0)) \exp(-rt) + \sigma \int_0^t \exp(-r(t-u))dB_u. \tag{1.3.20}$$

And for $t \geq s$ it comes to

$$\begin{aligned} E(X_t) = m(t), \text{cov}(X_t, X_s) &= \frac{\sigma^2}{2r} \exp(-r(t+s))(\exp(2rs) + \frac{2rv(0) - \sigma^2}{\sigma^2}), \\ \text{var}(X_t) &= \frac{\sigma^2}{2r} \exp(-2rt) \times (\exp(2rt) + \frac{2rv(0) - \sigma^2}{\sigma^2}). \end{aligned} \tag{1.3.21}$$

From Equation (1.3.21), when $v(0) = \frac{\sigma^2}{2r}$, then the variance of X_t is a constant $v(0)$. Therefore it can be a substitute a random-effect model $X_t = m(t) + \xi, \xi \sim N(0, v_0)$. It reserves consideration of fixed-variance uncertainty while it also includes the interaction of uncertainties at different times.
Moreover when X_0 is arbitrary, the asymptotic distribution of X_t is Gaussian with mean $m(t)$. And the long-term performance of (1.3.20) is derived from (1.3.15) and (1.3.16) :

$$\lim_{t\to+\infty} \text{var}(X_t) = \frac{\sigma^2}{2r} \text{ and } \lim_{s\to+\infty} \rho_{X_{h+s},X_s} = \exp(-rh), \quad h \geq 0. \tag{1.3.22}$$

From Equation (1.3.22), the asymptotic variance is a constant $v = \frac{\sigma^2}{2r}$, and the correlation coefficient is described by r which can be adjusted in the modeling.

- Recall further Equations (1.3.9), to consider the case when $X_0 = m(0)$ is deterministic, it is supposed here $v(t) = \nu(1 - e^{-ut})$, $u, \nu > 0$ and $m(t)$ are given. Then the following process can be proposed to describe the given mean and variance $m(t)$, $v(t)$:

$$dX_t = (-\frac{u}{2}X_t + m'(t) + \frac{u}{2}m(t))dt + \sqrt{u\nu}dB_t, t \geq 0. \tag{1.3.23}$$

These special cases are with different statistical properties. This can help to establish the model, with a preliminary comprehension of the fitting data. And also two specific cases i.e. non-linear drift, linear-diffusion processes [66], and Equation (1.3.19) will be applied to the case study later.

Illustrative Examples on Controlling Statistical Properties

The discussion in this subsection is to present that controlling the process' mean and variance is not enough in the modeling work. Actually if we consider a simple model as in (1.3.19), then it is found similar with the following random effect model :

$$Y_t = m(t) + \xi,\ \xi \sim N(0, v(0)), \tag{1.3.24}$$

with $\mathbb{E}(Y_t) = m(t), \text{var}(Y_t) = v(0), \rho(Y_t, Y_s) = 1$.

Compared to the random effect model (1.3.24), the OU process (1.3.19) (with $v(0) = \frac{\sigma^2}{2r}$) is with one more control on the correlation of the process. And the question comes : is it useful to consider such a flexility on the correlation of the process? This question will be answered by illustrative examples in the following.

Suppose the process X_t is given as in (1.3.19), with the same mean $m(t) = 5.6338738((t+1)^{0.5964851} - 1) + 1.7922389$ and the same variance $v(t) = 11.86187$ as a constant. And the correlation coefficient $\rho_{X_t,X_s} = \exp(a(t-s))$ only depends on a.

The simulation tests here is to fix the mean $m(t)$ and variance $v(t)$, meanwhile we change σ and a to adjust the process' correlation coefficients. Therefore figures are produced to present simulated trajectories for processes with different correlation coefficients, see Figure 1.1.

From the figures, the necessity to introduce one more adjustable parameter on the correlation can be concluded :

- When the correlation is stronger, less fluctuations are introduced into the process.
- When the correlation is weaker, more fluctuations are introduced into the process.

1.3.2 Time-Change and Gauss-Markov Process

The following lemma states a time-change of Itô's stochastic integral, see Proposition 7.6 in [80], which leads to an expression of Gauss-Markov process for OU processes with a deterministic start.

Proposition 1.3.1. *[80] If $f \in C[0,T]$, then the process defined by $Z_t = \int_0^t f(s)dB_s, t \in [0,T]$ is a mean zero Gaussian process with independent increments and with covariance function*

$$Cov(Z_s, Z_t) = \int_0^{s \wedge t} f^2(u)du. \tag{1.3.25}$$

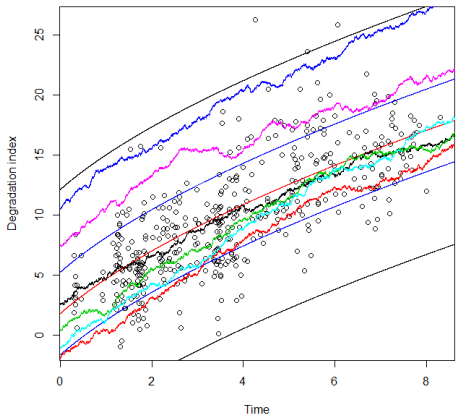
Proof. see Proposition 7.6 by Steele in [80]. □

Moreover, Z_t can be treated as a Brownian motion under the time-change $Z_t = B_{\rho(t)}$ with $\rho(t) := \int_0^t f^2(u)du$. Consider the general OU process Y_t in (1.3.2), when initial condition is considered to be a constant y at time $s \geq 0$, the solution is expressed from (1.3.4) by :

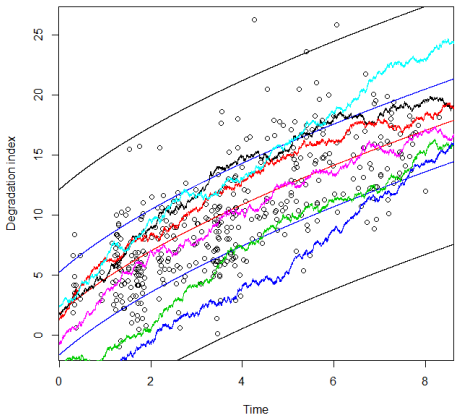
$$Y_t := Y_t^{y,s} = e^{-\alpha(t,s)}\left[y - \beta(t,s) + \int_s^t \sigma(u)e^{\alpha(u,s)}dB_u\right], \quad t \geq s. \tag{1.3.26}$$

Moreover recall in Definition 1.2.4, a Gauss-Markov process is expressed by

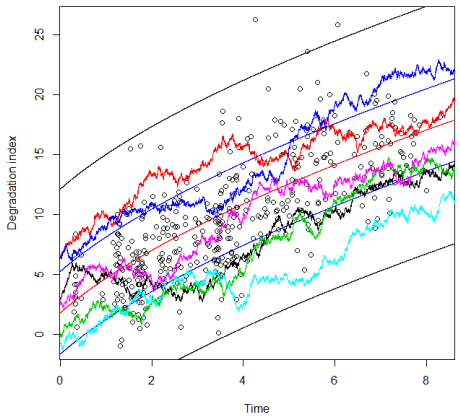
$$X_t = h(t)B_{f(t)}, \tag{1.3.27}$$



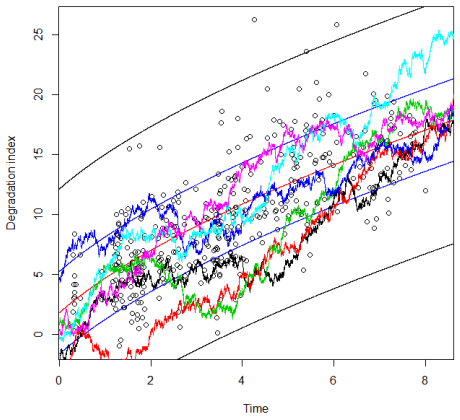
(a) Simulated trajectories of M_{OUR} with $a = -0.06683215$.



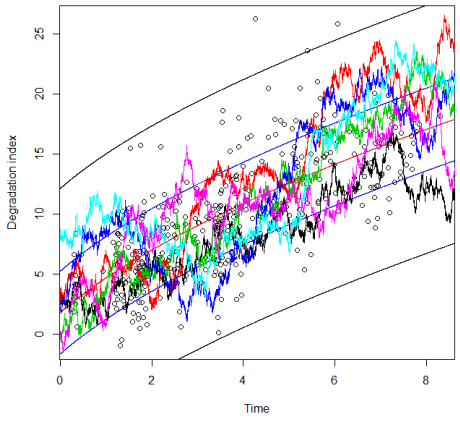
(b) Simulated trajectories of M_{OUR} with $a = -0.09451493$.



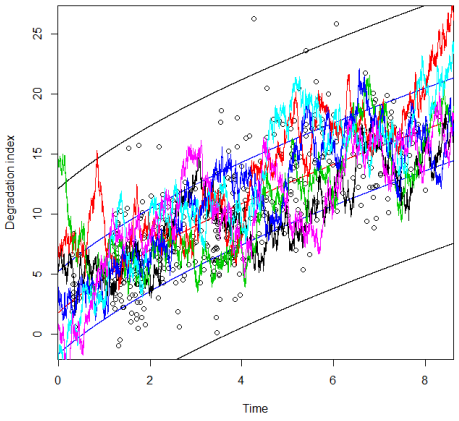
(c) Simulated trajectories of M_{OUR} with $a = -0.1056709$.



(d) Simulated trajectories of M_{OUR} with $a = -0.2113418$.



(e) Simulated trajectories of M_{OUR} with $a = -0.6340254$.



(f) Simulated trajectories of M_{OUR} with $a = -1.268051$.

FIGURE 1.1 – Trajectories of M_{OUR} with the same mean $m(t) = 5.6338738((t+1)^{0.5964851} - 1) + 1.7922389$ and variance 11.86187, the correlation coefficient $\exp(a(t-s))$ varies.

where $h(t), f(t)$ are properly defined deterministic functions, and B_* is a standard Brownian motion. Then from (1.3.26) and Proposition 1.3.1, if the initial condition is given by a constant y at time s, we know under the time-change $2\gamma(t,s) := \int_s^t \sigma^2(u)\exp(2\alpha(u,s))du$, $Y_t = (y - \beta(t,s))e^{-\alpha(t,s)} + e^{-\alpha(t,s)}B_{2\gamma(t,s)}$. So under notations in (1.3.3), we have expressed Y_t as a drifted Gauss-Markov process $Y_t = n(t) + h(t)B(f(t))$, where $n(t) = (y - \beta(t,s))e^{-\alpha(t,s)}$, $h(t) = e^{-\alpha(t,s)}$, $f(t) = 2\gamma(t,s)$, $t \geq s$.

Remark : the time-dependent OU process (1.3.2) is a Gauss-Markov process only if the initial value Y_0 is deterministic.

This time-change technique can lead to the transition pdf for Y_t. As $\sigma(u)e^{\alpha(u,s)}$ is deterministic and continuous, so the stochastic integral $Z_t := \int_s^t \sigma(u)e^{\alpha(u,s)}dB_u, t \geq s$ is also a mean zero Gaussian process from Proposition 1.3.1. Moreover we know $Z_t \sim \mathcal{N}(0, 2\gamma(t,s))$, and $\mathcal{N}(\mu, \sigma^2)$ represents the normal distribution with mean μ and variance σ^2.

Therefore from (1.3.26), the transition pdf of Y_t, $p(x,t|y,s)$, $t \geq s$ can be derived with $Z_t e^{-\alpha(t,s)} = Y_t + (y - \beta(t,s))e^{-\alpha(t,s)}$. This leads to

$$p(x,t|y,s) = \frac{1}{\sqrt{4\pi\gamma(t,s)e^{-2\alpha(t,s)}}}\exp(-\frac{(x + (\beta(t,s) - y)e^{-\alpha(t,s)})^2}{4\gamma(t,s)e^{-2\alpha(t,s)}}). \tag{1.3.28}$$

The transition pdf also satisfies a Fokker-Planck equation, which can also lead to (1.3.28). Related details are included in Appendix A in this thesis.

1.3.3 Mean-Reversion Properties

In the previous section, a time-dependent OU process is proposed in (1.3.12). It is shown that the process is exactly determined by the knowledge of the functions $m(t)$, $v(t)$ and $\sigma(t)$ which describe respectively its mean, its variance and the way to adjust its long-term correlation relationship. From a statistical view, this process can be considered relevant for observations fluctuating in multiple ways. Furthermore the time-dependent OU process exhibits a mean-reverting property which makes the expectation of X_t tend to drift toward its long-term mean over time. This property is discussed in this section.

Assume that the degradation process has been perfectly observed at time s with $X_s = y$. From equation (1.3.13), the mean $\mathbb{E}(X_t^{y,s})$, $t \geq s$ can be written as :

$$\mathbb{E}(X_t^{y,s}) = m(t) + (y - m(s))\exp(\int_s^t a(u)du). \tag{1.3.29}$$

When $a(t)$ is negative and satisfies Equation (1.3.14), the deviation $(y - m(s))$ from the mean function is weighted by $\exp(\int_s^t a(u)du)$ which tends to 0 as t tends to infinity. Therefore, the influence of a monitored fluctuation on the long-term average performance of the system decreases in time. Figure 1.2 proposes an illustration of the mean-reversion property by comparison of an OU process which exhibits the property and a drifted Brownian motion which doesn't. The two processes are defined respectively by $X_t = \int_0^t(-X_u + u)du + \int_0^t 4dB_u$ and $Y_t = t + 2B_t$ i.e. in a such way that starting from the origin, they have the same mean function $m(t) = t$ ($E(X_t) = E(Y_t) = t$). The stars represent observed values every 0.5 unit of time up to time 1.5. Blue and red curves are possible paths starting from the last observed degradation level which is quite far from the mean value. The blue curves are trajectories produced for the OU process the red curves are trajectories of the drifted Brownian motion. The corresponding conditional expectations are drawn by a solid light blue and red curve

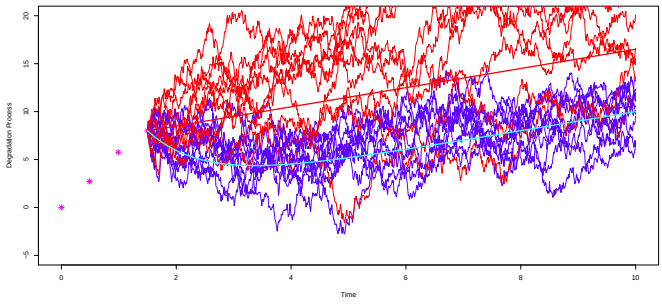


FIGURE 1.2 – Comparison of OU process and Brownian motion for mean reverting property illustration : OU process in blue, Brownian motion in red. The stars represents degradation levels observed at the inspection times starting from the origin. Later blue and red curves are possible trajectories for degradation prognosis based on the last inspection for respectively OU process and Brownian motion with the same mean $m(t)$.

respectively. As described by equation (1.3.29) the temporary deviation vanishes for the OU process, and its conditional mean function tends to the original mean function. However Brownian motion doesn't have the same property.

The way fluctuations accumulate is important in degradation models. For instance, simple addition of gaussian white noises leads to linear-diffusion models as in Si et al. [66] where models' uncertainties increase rapidly as time moves on. The mean-reverting property provides a localizing mechanism to control uncertainties in stochastic models such that uncertainties can be adjusted.

Moreover, the mean-reverting property can also be supported by some cases from a physical modeling view as in the two-state crack growth model proposed by Sobczyk and Spencer [75]. Actually in the stochastic process (1.3.12), $m(t)$ can be treated as the "correct" degradation mechanism, and the mean-reversion can be introduced as a self-repair mechanism to describe the recovery ability of the system to its "correct" state spontaneously. Therefore $a(t)$ can represent the recovery rate for the accompanying self-repairing mechanism, and $\sigma(t)$ for the fluctuating influence of surrounding environment. The total uncertainties in the time-dependent OU process therefore are determined by the combination between the self-repair and environmental influences.

This general description of fluctuations in degradation processes from a self-repairing mechanism deserves its reasonability for fatigue cracks, due to revealed mechanisms e.g. crack-closure. It could be adopted in more general degradation processes related to fatigue cracks. Also rather than an actual self repairing of the considered system this hypothesis refers to the fact that the internal intrinsic randomness of the degradation phenomenon can not cause its instability. Hence it is possible to stabilize fluctuations around the average trend

More specifically, the system dynamic state x_t can be mainly determined by a continuous dynamic system, for example Paris' law for the crack growth [52], where $m(t)$ is supposed to be the "correct" model :

$$dx_t = m'(t)dt. \tag{1.3.30}$$

To consider the fluctuations in degradation records, suppose during the period $[t, t+dt)$, the fluctuation can be described by a Gaussian variable which is given in the form of $\sigma(t)dB_t$,

where $dB_t \sim N(0, dt)$, then the model (1.3.30) turns to the following one :

$$dx_t = m'(t)dt + \sigma(t)dB_t. \tag{1.3.31}$$

The time-dependent strength of noises $\sigma(t)$ can be treated as a constant. This case was adopted by Si et al. [66], Tseng and Peng [82]. And Equation (1.3.31) is not bad to describe the local performance of system states, however this addition of noises leads to a rapidly increasing variance such that the long-term system state may be with over-estimated uncertainties. Therefore the mean-reversion can be proposed to adjust Equation (1.3.31).

We here avoid a strict exploration on this phenomenon, but think the self-repairing mechanism acts like damping on a permanently vibrating system : the further the system state is away from the underlying mechanism, the stronger the system will react to return to its underlying mechanism. Therefore, an item $a(t)$ is introduced to represent strength of mean-reversion, and $a(t)(x_t - m(t))$ represents the correcting power from abnormal states such that :

$$dx_t = (a(t)(x_t - m(t)) + m'(t))dt + \sigma(t)dB_t. \tag{1.3.32}$$

It should also be noticed that the general OU process includes a wide range of degradation processes e.g. the drifted Brownian motion, which is without mean-reverting property. So the mean-reversion is not a strong technical constraint for the modeling, but just a way to avoid unwanted properties induced by the process, e.g. the rapidly-increasing variance of the Brownian motion.

Summarize all the above, the modeling work is started by macro performance of degradation phenomenons, and fluctuations are introduced by the co-effect of exciting noises and self-repairing mechanisms. Under this framework, the condition (1.3.14) is proposed to promise the mean-reversion of X_t..

1.4 Parameter Estimation

In this section, the problem of statistical inference is considered for OU process. The first step consists in choosing some parametric functions for $m(t)$, $v(t)$ and $\sigma(t)$. In the sequel the vector of useful parameters will be denoted θ and the corresponding functions respectively $m(t|\theta)$, $v(t|\theta)$ and $\sigma(t|\theta)$. These functions depend on the case study and can be chosen from a predetermined list by considering descriptive statistics from the dataset or expert opinion.

Let $X_{t;\theta}$ be the stochastic process described by (1.3.12). It has been stated that its transition pdf $p(x, t|x_s, s; \theta)$, $t > s \geq 0$ can be derived from the framework of stochastic analysis based on properties of Itô's stochastic integral [80]. It can be shown directly that the analytical expression of $p(x, t|y, s; \theta)$ is given by (1.3.28) :

$$p(x, t|y, s; \theta) = \frac{e^{\alpha(t,s|\theta)}}{\sqrt{4\pi\gamma(t, s|\theta)}} \exp\left(-\frac{(xe^{\alpha(t,s|\theta)} + \beta(t, s|\theta) - x_s)^2}{4\gamma(t, s|\theta)}\right), \tag{1.4.1}$$

where $\alpha(t, s|\theta), \beta(t, s|\theta), \gamma(t, s|\theta)$ with $c(u|\theta) := \frac{\sigma(u|\theta)^2}{2}$are given by :

$$\begin{aligned}
\alpha(t, s|\theta) &= -\int_s^t a(u|\theta)du, \\
\beta(t, s|\theta) &= m(s|\theta) - m(t|\theta)e^{\alpha(t,s|\theta)}, \\
\gamma(t, s|\theta) &= \int_s^t c(u|\theta)e^{2\alpha(u,s|\theta)}du.
\end{aligned} \tag{1.4.2}$$

Let us further suppose that the degradation records have been obtained from n independent components and that m_i is the number of records collected on the ith component ($1 \leq i \leq n$). Furthermore (x_{ij}, t_{ij}) stands for the record at the jth time on the ith component with $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m_i\}$.

In the general case X_0 is subject to a distribution function $F_{0,\theta}$ and the corresponding transition pdf based on random initial condition $p(x, t|X_0, 0; \theta)$ can be written as :

$$p(x, t|X_0, 0; \theta) = \int_{-\infty}^{+\infty} p(x, t|u, 0) dF_{0,\theta}(u) \tag{1.4.3}$$

If X_0 is supposed to be deterministic, its probability law becomes a Dirac measure with $t_{i0} = 0$ and $x_{i0} = x_0$ for all $i \in \{1, \ldots, n\}$.

The adopted stochastic process (1.3.12) is a Markov process. As a consequence the conditional probability density of every record based on the previous records in the same trajectory can be simplified by Markov property to the transition pdf based on the closest record i.e. :

$$p(x_{ij}, t_{ij}|\{(x_{iz}, t_{iz})\}_{z=0}^{j-1}; \theta) = p(x_{ij}, t_{ij}|x_{i(j-1)}, t_{i(j-1)}; \theta).$$

Then the log-likelihood function for the ith trajectory can be written as follows :

$$\log L_i(\theta) = \sum_{j=1}^{m_i-1} \log\Big(p\big(x_{i(j+1)}, t_{i(j+1)}|x_{ij}, t_{ij}; \theta\big)\Big) + \log(p(x_{i1}, t_{i1}|X_0, 0; \theta), \tag{1.4.4}$$

where the transition pdf to $(x_{i(j+1)}, t_{i(j+1)})$ from (x_{ij}, t_{ij}) can be given directly from (1.4.1) and (1.4.3).

Noticing that different trajectories are independent with each other, therefore the log-likelihood function for the whole data set is as follows :

$$\log L(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{m_i-1} \log\Big(p\big(x_{i(j+1)}, t_{i(j+1)}|x_{ij}, t_{ij}; \theta\big)\Big) + \sum_{i=1}^{n} \log(p(x_{i1}, t_{i1}|X_0, 0; \theta)). \tag{1.4.5}$$

The maximum likelihood estimate $\theta^* = \operatorname{argmax}_\theta \log L(\theta)$ of the parameters can be obtained by maximization of the log-likelihood function (1.4.5). As it is generally not easy to calculate analytically, numerical algorithms can be used e.g. Nelder-Mead method. However to reach the global optimum, a good initial estimate is crucial.

An initial estimate can be obtained by using the least square estimation method for $m(t|\theta)$, $a(t|\theta)$ and $\sigma(t|\theta)$ respectively based on their statistical meaning. Suppose $\theta = (\theta_1, \theta_2, \theta_3)$ where θ_1, θ_2 and θ_3 denote independent parameters such that $m(t|\theta) = m(t|\theta_1)$, $a(t|\theta) = a(t|\theta_2)$ and $\sigma(t|\theta) = \sigma(t|\theta_3)$. The optimal value for θ_1 can be estimated from the least square estimation problem :

$$\operatorname*{argmin}_{\theta_1} \sum_{i=1}^{n} \sum_{j=0}^{m_i-1} (m(t_{ij}|\theta_1) - x_{ij})^2. \tag{1.4.6}$$

The above discretization introduces new errors into the estimation. However it should be noticed the least-square estimation is designed for an initial estimate for the maximum likelihood estimation. So even it is with errors, a rough estimate near the real estimate is enough for this two-stage estimation.

The second step considers the Euler discretized equation [36] as an approximation for (1.3.12) :

$$X_{t+\Delta t} = X_t + \Big(a(t|\theta_2)\big(X_t - m(t)\big) + m'(t)\Big)\Delta t + \sigma(t|\theta_3)\Delta B_t, \tag{1.4.7}$$

where $\Delta B_t \sim \mathcal{N}(0, \Delta t)$.

Parameters in $a(t|\theta_2)$ can therefore be estimated from minimizing the following objective function :

$$\underset{\theta_2}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=0}^{m_i-1} \Big(x_{i(j+1)} - x_{ij} - \Big(a(t_{ij}|\theta_2)\big(x_{ij} - m(t_{ij})\big) + m'(t_{ij})\Big)\big(t_{i(j+1)} - t_{ij}\big)\Big)^2. \tag{1.4.8}$$

Furthermore, parameters θ_3 in the diffusion part $\sigma(t|\theta_3)$ can be found by fitting the standard deviation of residuals.

The parameters estimation for the initial distribution $F_{0,\theta}$ can be done according to the values of $\mathbb{E}(X_0)$ and $\mathrm{var}(X_0)$ hence according to the values of $m(0)$, $a(0)$ and $\sigma(0)$. In such a case $\theta_1, \theta_2, \theta_3$ can be estimated by ignoring the items involving $\{x_{i0}\}_{i=1}^n$ in (1.4.6) and (1.4.8).

1.5 Case Study

1.5.1 Degradation Data Set

In the following a real degradation data-set is considered which is composed of 415 records of degradation level of a passive component in a power plant for 159 independent equipments. Figure 1.3 expresses the degradation records and recorded trajectories on different equipments. Single points in Figure 1.3 mean that only one record is recorded on that equipment.

On average, only about 3 records exist for each component. Furthermore, if we consider one trajectory carefully, there exists a large fluctuation. This datas-set could be considered sparse, noisy, and non-smooth. However with records on different components, the sparse data-set can still show statistical laws, which is the foundation for degradation modeling. This is exactly the same consideration for the crack-growth data used in [94].

Moreover we don't observe an obvious tendency of increasing variance in this case, such that this hypothesis of rapidly increasing variance is not preferred in the modeling. This is why Drifted Brownian motion is not adopted to describe this data-set. But a simple random-effect model with fixed variance, e.g. $m(t) + \xi$, $\xi \sim \mathcal{N}(0, c)$, can hardly explain the temporal variability in the data-set [51].

Throughout this section n denotes the number of monitored equipments i.e. the number of recorded degradation path, m_i denotes the number of inspection after time $t = 0$ for equipment number i i.e. the number of points on the i-th trajectory. As previously stated, t_{ij} and x_{ij} respectively stand for the jth recording time and measured degradation level on component i. It is suggested from expert opinion that all the equipments have the same initial condition, which means that $t_{i0} = 0$ and $x_{i0} = m_0$ for all $i \in \{1, \ldots, n\}$. However initial uncertainties due to various reasons, such as initial degradation, will also be considered and compared in the following as in [17].

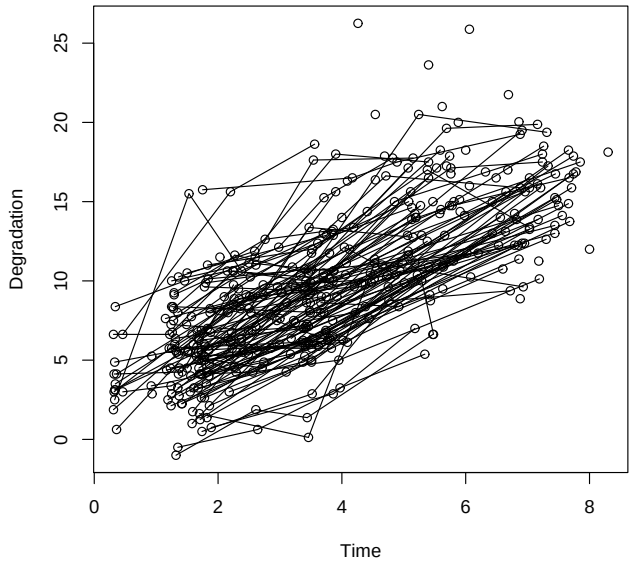


FIGURE 1.3 – Recorded Degradation Trajectories

1.5.2 Evaluation Criterions

As the inspection times are not periodic and the proposed stochastic process is not time homogeneous, the corresponding increments of degradation are not identically distributed. Usual properties of classical goodness of fit criterions are failing. In this section, three different criterions are proposed in order to assess the relevance of the considered models for a specific data-set and for prognosis purpose.

Akaike Information Criterion

Akaike information criterion (AIC) is a widely used information criterion to evaluate the fitting goodness of models [3]. The value which is noted as AIC thereafter is given by :

$$AIC = 2k - 2logL(\theta^*), \tag{1.5.1}$$

where k is the number of unknown parameters in the model, and L and θ^* are defined as in (1.4.5). In such a criterion, not only the maximization of the likelihood function is considered, but also a punishment of introducing too complex models is introduced such that the problem of over-fitting can be avoided to a certain extent. A better model should have a smaller AIC value.

Indicators for Prognosis Assessment

The mathematical model considered for degradation modeling may be used for residual lifetime prediction. As a stochastic process it is intrinsically designed to take into account the variability of the measured degradation index which is characterized by the probability law of the degradation level at a given future time. In order to make the best use of the data which

are available by degradation path, a class of indicators for prognosis assessment is proposed in [4]. In this class, the Indicator for Prognosis Assessment of type k is denoted as IPA_k and defined as follows :

$$\text{IPA}_k(\varepsilon) = \frac{1}{S_k}\sum_{i=1}^{n}\sum_{j=0}^{m_i-1}\sum_{s=j+1}^{\min(j+k;m_i)} P\{|x(t_{is}) - x_{is}| \leq \varepsilon / x(t_{ij}) = x_{ij}\}, \tag{1.5.2}$$

$$\text{with } S_k = \sum_{i=1}^{n}\{\mathbf{1}_{k\leq m_i+1}\left(k(m_i-k) + \frac{k(k+1)}{2}\right) + \mathbf{1}_{k>m_i+1}\frac{m_i(m_i+1)}{2}\}. \tag{1.5.3}$$

To obtain a normalized value, S_k is equal to the total number of terms considered in the definition of IPA_k. The aim of this indicator is to reflect the accuracy of the model prediction. Each point of each trajectory is successively considered as the last inspection data and the gap between the model prediction and the available data is considered on the next k future records on the same trajectory, if available. If less than k future inspection times are available they all are considered. The probability that the model prediction stays within a range of width 2ε around the next degradation records given the last inspection data is considered for prognosis assessment. It can be written as :

$$P\{|x(t_{is}) - x_{is}| \leq \varepsilon / x(t_{ij}) = x_{ij}\} = \int_{x_{is}-\varepsilon}^{x_{is}+\epsilon} p(\tau, t_{is}|x_{ij}, t_{ij})d\tau, \quad s > i, \tag{1.5.4}$$

where the transition pdf $p(x,t|y,s)$ is given by (1.4.1). Unfortunately the expression given by equation (1.5.4) cannot be expressed analytically in general. Therefore a numerical integral scheme is considered for the above integral expression. Given ε, the interval $[x_{is}-\epsilon, x_{is}+\epsilon]$ is discretized in P parts with a step-size $\Delta := \frac{2\epsilon}{P}$. By denoting $z_{ij}^k = x_{ij} - \epsilon + k\Delta$, $k = 0, \ldots, P$ the calculation for (1.5.4) is based on compound trapezoid formula :

$$\begin{aligned} P\{|x(t_{is}) - x_{is}| \leq \varepsilon / x(t_{ij}) = x_{ij}\} \approx \frac{\Delta}{2}\Big\{&p(z_{is}^0, t_{is}|x_{ij}, t_{ij}) \\ &+ \sum_{k=1}^{P-1} p(z_{is}^k, t_{is}|x_{ij}, t_{ij}) + p(z_{is}^P, t_{is}|x_{ij}, t_{ij})\Big\}. \end{aligned} \tag{1.5.5}$$

Under the proposed Indicator for Prognosis Assessment, a better model should have a larger value. Nevertheless a good capacity of prediction with a unimodal peaky probability density function $p(x,t|y,s)$ leads to a higher value of $\text{IPA}_k(\varepsilon)$ than if the pdf is spread out. Hence it depends on the data variability.

In this chapter, the Prognosis Assessment of type 1 ($\text{IPA}_1(\varepsilon)$) and ∞ $\text{IPA}_\infty(\varepsilon)$ are considered. For the first one, only the record following the point considered as the inspection data is taken into account. For the second one, all the end of the trajectory is considered.

1.5.3 Comparison on Fitting Results on Different Models

Compared Models

This section proposes a comparison between the time-dependent OU process and the linear-diffusion model to test the validity of introducing self-repairing mechanism / mean-reversion for cumulating fluctuations. The models considered here are the time-dependent

OU process given in equation (1.3.19), and the linear-diffusion model in Si et al. [66]. These models are introduced respectively in Tables 1.1, 1.2 and 1.3. The first 2 models are supposed to start from an unobservable constant m_0, and the third one is supposed to start from an unobservable gaussian variable. For short in the following, we will note the linear-diffusion model as M_{LD}, the model of time-dependent OU process as M_{OU} and the model of time-dependent OU process with initial uncertainties as M_{OUR}.

These models are expected to have same average performance, before statistical inference. And the only different hypothesis is the uncertainty introduced in M_{LD}, M_{OU} and M_{OUR}. Here the variance of M_{LD} is supposed to linearly increase, the variance of M_{OU} is supposed to increase but to converge to a constant and the variance of M_{OUR} is a constant.

The mean function m(t) can be chosen freely, not limited to the shape in this paper. The expression $m(t) = \alpha((t+1)^{\beta} - 1) + m_0$ is proposed due to some preliminary experts' opinions on the fitting data-set. Unfortunately because of confidentiality reasons this point cannot be developed in the paper. It is also a very general polynomial function with plenty of flexibility. Choosing a parametric function is a common choice for the modeling work, e.g. the mean is pre-determined as a quadratic function in [94].

TABLE 1.1 – General description of M_{LD}

parameters	$\alpha, \beta, \sigma, m_0$
expression	$Y_t = m(t) + \sigma B_t$
mean	$m(t) = \alpha((t+1)^{\beta} - 1) + m_0$
variance	$\sigma^2 t$
covariance $(t \geq s)$	$\sigma^2 s$
initial value	m_0
nb. of parameters	4

TABLE 1.2 – General description of M_{OU}

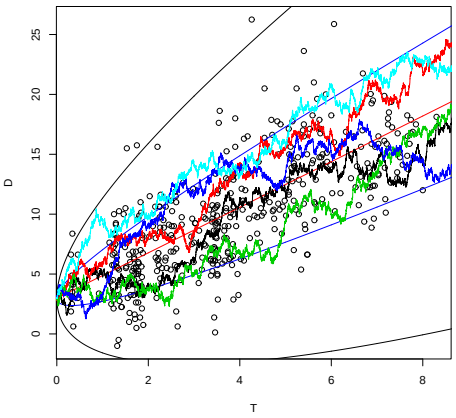
parameters	$\alpha, \beta, a, \sigma, m_0$
expression	$dX_t = (a(X_t - m(t)) + m'(t))dt + \sigma dB_t$
mean	$m(t) = \alpha((t+1)^{\beta} - 1) + m_0$
variance	$\frac{\sigma^2}{-2a}(1 - \exp(2at))$
covariance $(t \geq s)$	$\frac{\sigma^2}{-2a}\exp(a(t+s))(\exp(-2as) - 1)$
initial value	m_0
nb. of parameters	5

TABLE 1.3 – General description of M_{OUR}

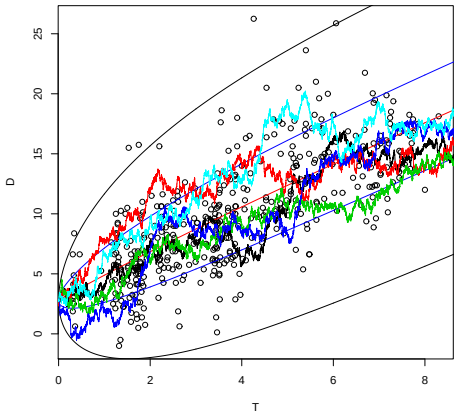
parameters	$\alpha, \beta, a, \sigma, m_0$
expression	$dX_t = (a(X_t - m(t)) + m'(t))dt + \sigma dB_t$
mean	$m(t) = \alpha((t+1)^{\beta} - 1) + m_0$
variance	$\frac{\sigma^2}{-2a}$
covariance $(t \geq s)$	$\frac{\sigma^2}{-2a}\exp(a(t-s))$
initial value	$\mathcal{N}(m_0, -\frac{\sigma^2}{2a})$
nb. of parameters	5

Fitting Results

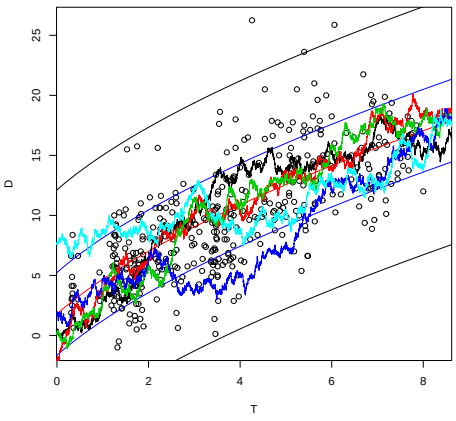
Based on the method of maximum likelihood estimation as stated in section 1.4, the estimated parameters listed in Table 1.4 have been found with help of Nelder-Mead algorithm in R software. The simulated trajectories then are produced based on Euler-Maruyama scheme with the step-size 0.01 [36], see Figure 1.4(a), Figure 1.4(b) and Figure 1.4(c) respectively. In these pictures, some auxiliary curves are drawn to show statistical properties of corresponding processes : the red curve for the mean, the blue curves for the mean $\pm$ standard variance, and the black curves for the mean $\pm$ 3 times of standard variance.



(a) Simulated trajectories of M_{LD}.



(b) Simulated trajectories of M_{OU}.



(c) Simulated trajectories of M_{OUR}.

FIGURE 1.4 – Comparison on trajectories produced by M_{LD}, M_{OU} and M_{OUR}.

Fitting goodness

The values of AIC, IPA_1 and IPA_∞ from (1.5.1) and (1.5.2) for M_{LD}, M_{OU} and M_{OUR} respectively are listed in Table 1.5. Under the information criterions AIC, IPA_1 and IPA_∞, M_{OU} and M_{OUR} have a slightly better fitting result compared to M_{LD}.

TABLE 1.4 – Fitting results

parameters	$M_{\mathbf{LD}}$	$M_{\mathbf{OU}}$	$M_{\mathbf{OUR}}$
α	1.873542	2.4402845	5.6338738
β	1.005893	0.8892020	0.5964851
σ	2.152958	2.4640884	2.2391552
a	–	-0.1806708	-0.2113418
m_0	2.988090	2.8074561	1.7922389
X_0	2.988090	2.8074561	$\mathcal{N}(1.7922389, 11.86187)$
variance	4.635228t	16.8033(1-exp(-0.3613416t))	11.86187

However, it is interesting to see that M_{OU} and M_{OUR} don't have uniform fitting goodness order under criterions AIC, IPA_1 and IPA_∞. It is recalled that the criterion of AIC concerns more on describing existing records, while IPA_1 and IPA_∞ are proposed to predict system state based on current observation. Although the difference here is not huge, it can be distinguished that M_{OU} has a better ability to describe data while M_{OUR} has a better ability for the prognosis assessment.

TABLE 1.5 – Fitting goodness

criterion [1]	$M_{\mathbf{LD}}$	$M_{\mathbf{OU}}$	$M_{\mathbf{OUR}}$
AIC	2048.64	2021.044	2025.84
IPA_1	0.314464	0.3255434	0.3368724
IPA_∞	0.3333623	0.3524595	0.3687379

1.6 Conclusions and Perspectives

In this chapter, a time-dependent OU process is introduced to describe a degradation process. The time-dependent OU process is proposed from its statistical properties on controllable mean, variance and correlation. Its mean-reverting property can be introduced to interpret temporary correlated fluctuations from an overall degrading trend in degradation records. Corresponding parameter estimation is proposed based on maximum likelihood estimation method. Several simulation tests are done to test the model's fitting goodness based on a degradation data-set of a passive component in power plants.

The characteristics to distinguish stochastic degradation modeling with other stochastic modeling areas e.g. financial mathematics are :

- Long-term trend. As emphasized in this chapter, when degradation modeling is discussed, the possible models should show a long-term trend to final failure. Therefore the stochastic degradation modeling refers more to physical laws and empirical research rather than purely mathematical analysis.
- Mean-reversion. Some self-repairing mechanisms e.g. crack closure show the possibility of recovery of physical degradation phenomenons, which reveals a micro comprehension of stable macro appearance of slowly degradation phenomenons. This is one of the reasons why the fluctuations is introduced in degradation modeling. But to control such fluctuations to reach the overall degrading trend should be considered meanw-

1. In the calculation of IPA_1 and IPA_∞, the tolerance level $\epsilon = 0.25$ where the step-size is 0.01 for the numerical integral.

hile. In this chapter, the mean-reversion of the time-dependent OU process could be interesting and reasonable to fulfill this task.

- Failure-oriented modeling. Stochastic degradation modeling is discussed in the field of reliability engineering, and it serves as a first step for reliability analysis, prognostics and health management, and condition-based maintenance. So the fitting goodness to existing inspection records is not the only indicator to choose models, how to connect this modeling work with the failure's description and prognosis attracts more attention.

Meanwhile, the mean-reversion in stochastic degradation modeling implies the possibility to consider 2-stage models furthermore as in [75] and also Section 1.3.3. That is to say, the real degradation process is treated as a should-be processes given by physical laws. To make up the gap between this should-be process and observed data, the co-effect of self-repair and exciting-noise leads to a modified model.

Also the mean-reversion could be a starting point to introduce Levy-driven OU processes into stochastic degradation modeling. This is triggered by considering the risk of sudden change of system states or jumps in mathematical descriptions. However this is from the modeling view, it would be difficult to consider prognosis of system failures based on such processes and involve advanced mathematical tools.

Chapitre 2

Prognosis of Systems Failures via First Passage Problems

Considering the time-dependent OU process introduced in the last chapter as the stochastic degradation process, prognosis of system failures is considered in this chapter based on first passage times. This chapter contains several results in [20, 18, 17], and it emphasizes on estimating the probability laws for the first passage time (FPT) of the time-dependent OU process to a given failure level. Different methods are proposed, and their derivation is tried to be given in a self-contained form in this chapter. This effort is not a direct reproduction of existing ideas, but an interesting connection through Fortet's equation [32, 57].

The whole chapter is organized as follows. In Section 2.1, first passage time is introduced to describe the failure time. In Section 2.2, it is presented that the first passage problems can be connected to the process itself, which forms a foundation for Section 2.3 and Section 2.4. In Section 2.3, the first passage problem is considered from the view of partial differential equation, and related results are presented. In Section 2.4, the first passage passage problem is considered from the view of integral equation, where computable results are presented. In Section 2.5, simulation tests are done to calculate first passage density based on proposed methods. Conclusions and perspectives are stated in Section 2.6.

2.1 Introduction

2.1.1 Background

The aim in this chapter is to estimate the system's residual useful lifetime (RUL). RUL is often used as a decision indicator in the field of reliability engineering and prognostics and health management (PHM). It is commonly defined as the "residual useful lifetime of a system at time t given all available information up to time t".

To fulfill the above task, the core issue is to determine when the system failure occurs. Various ways exist to describe a system failure. In our research, we concentrate on deterioration phenomenons, eg. corrosion, erosion, crack etc., in which deterioration cumulates until it reaches the material strength. So we will consider first passage failure throughout this chapter : given a failure level [1] of deterioration, a system failure occurs when accumulating deterioration exceeds the failure level.

1. In other literature, similar expressions exist corresponding to different emphases, such as alarm threshold, safety limit, safety barrier, crossing boundary, hitting level, passage level etc. We don't distinguish these different names, and also in this thesis, it is treated differently due to different application. *Failure level* is used when the applied problems are discussed, *crossing boundary* is used when mathematical issues are discussed.

Such first passage problems attract plenty of attention with various applications to serve as decision rules. And in the field of reliability engineering, when increasing processes are considered for stochastic degradation processes, it would be trivial to derive the distribution of FPTs to a constant failure level due to the monotone property. That's why FPT is not discussed frequently before in reliability as most of considered degradation processes belong to the group of monotone processes, e.g. Gamma process [84], Poisson process, inverse Gaussian process [93]. However from a mathematical view, there are still few results for FPTs of increasing processes to time-dependent boundaries.

Recently, due to the consideration of fluctuations in degradation processes, diffusion processes are discussed more and more. However when FPTs for diffusion processes are discussed, it would be more challenging as stochastic processes may return to the safe area after its first passage. Therefore the most celebrated result is inverse Gaussian distribution for the first passage density of Brownian motion with a linear boundary, which luckily can be solved analytically. Other issues are discussed more to develop efficient numerical algorithms, and these views can generally categorized to 2 ideas : integral equation [57] and Fokker-Planck Equation [43, 53]. In this chapter, these 2 basic ideas will be reviewed and reproduced for the OU process. Such an effort aims to reveal a connection between these 2 different ideas through Fortet's equation [32, 57]. An overview on first passage problems for diffusion processes can refer to [24, 83].

These two ideas lead to different methods to estimate the first passage density or first passage distribution, which can be categorized into three classes : analytical approximation, numerical algorithms and Monte-Carlo simulation. These methods are designed for different targets, with different advantages and disadvantages.

The consideration in this chapter not only contributes to the ordinary first passage problem, but also to the conditional first passage problem. That is due to the main concentration in the field of PHM : when the system state is updated, the prognosis of system failures based on current observation should also be updated. This treatment is not trivial as the adopted degradation process is nonhomogeneous, non-stationary and without independent increments, such that the results of ordinary first passage problems cannot be extended directly to the conditional first passage problems.

2.1.2 System Description

In this chapter, supposing the degradation process is described by a stochastic process $\{X_t, t \geq t_0\}$ defined by (1.3.2) on a complete probability space $(\Omega, \mathfrak{F}, \mathbf{P})$, the following time-dependent Ornstein-Uhlenbeck process X_t will be considered throughout this chapter :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq t_0, \tag{2.1.1}$$

where $a(t), b(t), \sigma(t) \in C^1(R^+)$, $\sigma(t)$ is positive, B_t is a standard Brownian motion. $X(t_0) = x_0$ is a random variable independent of B_t with a pdf f_0 (if it exists) and distribution function F_0. When x_0 is deterministic, $F_0(du) = \delta_{x_0}(du)$, $\delta_{x_0}(*)$ as the Dirac measure concentrated at x_0.

It is derived from Equation (2.1.1) that

$$X_t := X_t^{x_0,t_0} = e^{-\alpha(t,t_0)}\left[x_0 - \beta(t,t_0) + \int_{t_0}^{t} \sigma(u)e^{\alpha(u,t_0)}dB_u\right], \quad t \geq t_0, \tag{2.1.2}$$

where α, β, γ are defined in (1.3.3).

Furthermore given $t > s$, the transition pdf $p(x,t|y,s)$ of X_t, satisfies the following Fokker-Planck equation (FPE), see Appendix A.1 :

$$\frac{\partial p(x,t|y,s)}{\partial t} = (-\frac{\partial}{\partial x})\left[(a(t)x + b(t))p(x,t|y,s)\right] + c(t)\frac{\partial^2}{\partial x}p(x,t|y,s), \tag{2.1.3}$$

where $c(t) := \frac{\sigma^2(t)}{2}, p(\pm\infty,t|y,s) = 0, p(x,s|y,s) = f_y(x)$, f_y as the pdf of y when it exists.

The transition pdf $p(x,t|y,s), t \geq s$ of X_t can be solved directly when y is deterministic :

$$p(x,t|y,s) = \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\exp(-\frac{(xe^{\alpha(t,s)} + \beta(t,s) - y)^2}{4\gamma(t,s)}), \tag{2.1.4}$$

where α, β, γ are defined in (1.3.3). This expression is derived by a time-change in Section 1.3.2, or from Fokker-Planck Equation (2.1.3) directly by a time-space change as in Appendix A.2 or by a Lie-algebraic method as in [91].

When y is a random variable with distribution F_y, $p(x,t|y,s)$ is also defined based on the expectation :

$$p(x,t|y,s) = \int_{-\infty}^{+\infty} p(x,t|u,s)dF_y(u). \tag{2.1.5}$$

In such a way, the uniform notation for transition pdf $p(x,t|y,s)$ is given no matter y is deterministic or not. And in following sections, except for special attention, these 2 cases will not be distinguished.

2.1.3 First Passage Time

When the degradation process has been represented by a stochastic process X_t, first passage failure is translated to classical first passage time (FPT) problem for the process X_t in a mathematical view.

The first passage time $\tau_{y,s}$ when X_t reaches $L(t)$ based on the observation (y,s) i.e. with $X_s = y$, as a constant, is given by

$$\tau_{y,s} := \inf_{t \geq s}\left\{t|X_t^{y,s} \geq L(t)\right\}, \tag{2.1.6}$$

where $L(t)$ is enough smooth, upper, pre-set and time-dependent failure level. In such a definition, when the observation $y \geq L(s)$, $\tau_{y,s} = s$. And throughout this thesis, we suppose conditions are satisfied such that $P(\tau_{y,s} < \infty) = 1$, i.e. the process will reach the boundary at a finite time almost surely.

The residual useful lifetime based on current observation (y,s) is naturally defined by $\text{RUL}_s := \tau_{y,s} - s$, which is essentially determined by the conditional first passage time based on the current observation. Correspondingly, the conditional mean time to failure (MTTF_s) is defined by the expectation of the failure time based on current observation (y,s), i.e. $E(\tau_{y,s}) - s$.

The definition of RUL is different with the one in [66] (see p.54 therein.). In [66], the authors consider the RUL definition from a conditional view. That is to say, the RUL is defined as $E(\tau_{X_0,0} - s|\tau_{X_0,0} > s)$ for the current time s. It is conditioned on the event that the system doesn't fail before the current time.

However in general we can observe the current system state y at time s. Based on the Markov property of the adopted OU process, the future state of the system is determined

only by the current system state. Therefore the future failure is not conditioned on previous events. Therefore in this paper the RUL_s is proposed to be $\tau_{y,s} - s$. We don't have a preference for these two definitions. Actually the definition in [66] could be useful when the real system state cannot be observed or can only partially observed. This could be another discussion beyond the consideration in this paper.

Moreover due to the lack of perfect observation, e.g. due to undetectable small cracks [17], when the initial value x_0 is treated as a random variable with distribution function F_0, the first passage time τ_{x_0,t_0} can be considered as a randomized first passage time (RFPT) as described in [38, 48, 83]. Its distribution function is calculated for $t \geq t_0$ from (2.1.6) :

$$\mathbb{P}(\tau_{x_0,t_0} \leq t) = \int_{-\infty}^{L(t_0)} \mathbb{P}(\tau_{u,t_0} \leq t) dF_0(u) + \int_{L(t_0)}^{+\infty} dF_0(u). \tag{2.1.7}$$

Remark :

1. in the following, we only concentrate on the case when the initial value is centered on $[-\infty, L(t_0)]$. In the real case, this means we don't consider a component which is already in the failed state.
2. The tolerance of initial uncertainties for x_0 is not only due to imperfect observation, but also can process a simple case of random boundary for X_t. On one hand, for X_t satisfying $dX_t = m(t)dt + n(t)dB_t$, with $X_0 = x_0$. And suppose the boundary is given by $L(t) + \xi$, ξ is a random variable. On the other hand, if we consider $Y_t = X_t - \xi$ for the boundary $L(t)$ with the initial value $Y_0 = x_0 - \xi$, then the first passage time is the same with the one for X_t. In this simple case, the randomness in the boundary is converted to the initial uncertainties instead.

The core issues to describe the first passage time $\tau_{y,s}$ is to derive its pdf and cdf, which are discussed in following sections. And we note hereafter for $\tau_{y,s}$ in (2.1.6)

$$g(t|y,s) := \frac{\partial P(\tau_{y,s} \leq t | X_s = y)}{\partial t}, G(t|y,s) = P(\tau_{y,s} \leq t | X_s = y). \tag{2.1.8}$$

Moreover, we denote that

$$u(x,t|y,s) := P(X_t < x, \tau_{y,s} > t | X_s = y), w(x,t|y,s) = \frac{\partial u(x,t|y,s)}{\partial x}. \tag{2.1.9}$$

The denoted functions $g(t|y,s), G(t|y,s), w(x,t|y,s), u(x,t|y,s)$ are considered throughout this thesis. And actually following notations in (2.1.8) and (2.1.9), their connection holds for $t \geq s$:

$$\begin{aligned} P(\tau_{y,s} > t | X_s = y) &= 1 - G(t|y,s) = 1 - \int_s^t g(z|y,s) dz \\ &= P(X_t < L(t), \tau_{y,s} > t | X_s = y) = u(L(t),t|y,s) = \int_{-\infty}^{L(t)} w(z,t|y,s) dz. \end{aligned} \tag{2.1.10}$$

We should notice that Equation (2.1.10) holds because the event $\{\tau_{y,s} > t\}$is included in $\{X_t < L(t)\}$. Therefore the desired cdf $G(t|y,s)$ and pdf $g(t|y,s)$ for the first passage time can be converted to the value of $u(x,t|y,s)$ and also corresponding $w(x,t|y,s)$, vice versa. It is difficult to discuss the direct problems for the first passage density $g(t|y,s)$ and cdf

$G(t|y,s)$. And at a first glance it seems strange to introduce more complex issues $w(x,t|y,s)$ and $u(x,t|y,s)$ which not only depends on t, but also on x. This seems more difficult than the original problem. However in later sections we can see this complication can lead to the desired first passage density $g(t|y,s)$ by introducing new tools.

2.1.4 The Difficulty to Apply Results for Brownian motion

As is described in Section 1.3.2, the OU process can be treated as a Gauss-Markov process with a drift part, i.e. $n(t)+h(t)B(f(t))$. So at a first glance, it seems trivial to consider the analysis on the OU process, as many results for Brownian motion can be extended to the OU process under the time-change.

However, as the OU process considered in this thesis is time-dependent with nonlinear appearance, the time-change $f(t)$ given for the Gauss-Markov process may not have an explicit expression for its inverse. So when applying those results of Brownian motion, it would induce a big problem on the calculation. Even treated numerically, to find the inverse for a general function $f(t)$ is not efficient as this procedure could be involved in the calculation many times.

2.2 The Connection between First Passage density and the Process' Probability Laws

The discussion in this subsection contributes to the probability laws of the OU process is connected with the first passage problems by several integral equations, which are given in Lemma 2.2.1 and Lemma 2.2.2.

Fortet's Equation

The first connection is given by the following integral equation between the first passage density $g(t|x_0,t_0)$ and transition pdf of the process $p(x,t|y,s)$ in (2.1.4). Such an equation is called Fortet's equation in general [32, 57].

Lemma 2.2.1. *[32, 57] Suppose $L(t)\in C[t_0,+\infty)$, for X_t in (2.1.1) with $x_0<L(t_0), t\geq t_0$, the following equation holds :*

$$p(x,t|x_0,t_0)=\int_{t_0}^{t} g(s|x_0,t_0)p(x,t|L(s),s)ds, \quad \forall x\geq L(t). \tag{2.2.1}$$

Proof. First, we should notice that by strong Markov property, the following equation holds

$$\begin{aligned} P(X_t<x,\tau_{x_0,t_0}\leq t|x_0,t_0) &= \int_{t_0}^{t} p(X_t<x|\tau_{x_0,t_0}=s)g(s|x_0,t_0)ds \\ &= \int_{t_0}^{t} p(X_t<x|X_s=L(s))g(s|x_0,t_0)ds \\ &= \int_{t_0}^{t}\int_{-\infty}^{x} g(s|x_0,t_0)p(z,t|L(s),s)dsdz. \end{aligned} \tag{2.2.2}$$

Therefore we have :

$$u(x,t|x_0,t_0) = P(X_t < x, \tau_{x_0,t_0} > t|X_{t_0} = x_0) = P(X_t < x|X_{t_0} = x_0) - P(X_t < x, \tau_{x_0,t_0} \leq t|X_{t_0} = x_0)$$
$$= \int_{-\infty}^{x} p(z,t|x_0,t_0)dz - \int_{-\infty}^{x}\int_{t_0}^{t} g(s|x_0,t_0)p(z,t|L(s),s)dsdz, \tag{2.2.3}$$

From Lebesgue's dominated convergence theorem, differentiate (2.2.3) for x, we have

$$w(x,t|x_0,t_0) := \frac{\partial u(x,t|x_0,t_0)}{\partial x} = p(x,t|x_0,t_0) - \int_{t_0}^{t} g(s|x_0,t_0)p(x,t|L(s),s)ds. \tag{2.2.4}$$

Noticing $u(x,t|x_0,t_0) = u(L(t),t|x_0,t_0)$, $\forall x \geq L(t)$, so $w(x,t|x_0,t_0) = 0$, $\forall x > L(t)$:

$$0 = p(x,t|x_0,t_0) - \int_{t_0}^{t} g(s|x_0,t_0)p(x,t|L(s),s)ds. \tag{2.2.5}$$

When $x = L(t)$, from the continuity of $w(x,t|y,s)$, the right limit at $L(t)$ exists and

$$p(L(t),t|x_0,t_0) - \int_{t_0}^{t} g(s|x_0,t_0)p(L(t),t|L(s),s)ds = \lim_{x\to L(t)+}\{w(x,t|x_0,t_0)\} = 0. \tag{2.2.6}$$

Summarize all the above, it comes to (2.2.1). □

Remarks :

- Equation (2.2.1) can also be verified based on intuitive consideration. Actually as X_t is a diffusion process, from its continuity, strong Markov property, and the continuity of $L(t)$ it is promised that the state of the process is on the boundary at the first passage time. Therefore for any current observation $x \geq L(t)$ at time t, the process must cross the boundary $L(s)$ at a previous time $s \leq t$ [57], see Figure 2.1.
- Fortet's equation also reveals the Markov property, i.e. the conditional first passage time $\tau_{y,s}$ has the same probability law with the first passage time of of a new process $X_t^{y,s}$ to the boundary.
- $w(x,t|y,s)$ is not differentiable at $x = L(t)$ such that $w_x(L(t),t|y,s)$ doesn't exist. However it should be noticed that $w_x(L(t)-,t|y,s)$ and $w_x(L(t)+,t|y,s)$ exist.

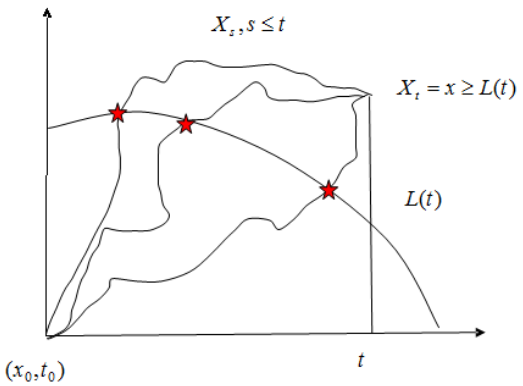


FIGURE 2.1 – An illustration of Fortet's equation.

The Master Equation

By integrating both sides, Lemma 2.2.1 on Fortet's equation leads to the following equation. This is called the master equation [57].

Lemma 2.2.2. *Suppose $L(t) \in C[t_0, +\infty)$, for X_t in (2.1.1) with $x_0 < L(t_0), t \geq t_0$, the following equation holds :*

$$1 - F(x,t|x_0,t_0) = \int_{t_0}^{t} g(s|x_0,t_0)(1 - F(x,t|L(s),s))ds, x \geq L(t) \tag{2.2.7}$$

where noting $\Phi(*)$ as the cdf of standard normal distribution,

$$F(x,t|y,s) := \int_{-\infty}^{x} p(z,t|y,s)dz = \Phi(\frac{xe^{\alpha(t,s)} + \beta(t,s) - y}{\sqrt{2\gamma(t,s)}}). \tag{2.2.8}$$

Proof. From Lemma 2.2.1, integrating (2.2.1) on $[x, +\infty)$ for $x \geq L(t)$, then from Fubini's theorem we get :

$$1 - F(x,t|x_0,t_0) = \int_{t_0}^{t} g(s|x_0,t_0)(1 - F(x,t|L(s),s))ds. \tag{2.2.9}$$

Furthermore, we have

$$\begin{aligned} F(x,t|y,s) &= \int_{-\infty}^{x} \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}} \exp(-\frac{(ze^{\alpha(t,s)} + \beta(t,s) - y)^2}{4\gamma(t,s)})dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{(ze^{\alpha(t,s)} + \beta(t,s) - y)^2}{4\gamma(t,s)})d\frac{ze^{\alpha(t,s)}}{\sqrt{2\gamma(t,s)}} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{xe^{\alpha(t,s)}+\beta(t,s)-y}{\sqrt{2\gamma(t,s)}}} \exp(-z^2/2)dz \\ &= \Phi(\frac{xe^{\alpha(t,s)} + \beta(t,s) - y}{\sqrt{2\gamma(t,s)}}). \end{aligned} \tag{2.2.10}$$

□

2.3 Initial-Boundary Value Problem for Fokker-Planck Equation

The Fokker-Planck equation is a partial differential equation which is strongly connected with the transition pdf of the OU process. And this section is written to explain the view of first passage problems based on the initial-boundary value problem for Fokker-Planck equation. The discussion in this section can be divided into several steps for the function $w(x,t|y,s)$ defined in (2.1.9) :

1. It is proved that the joint pdf $w(x,t|y,s)$ satisfies an initial-boundary value problem for Fokker-Planck Equation. This is done by using Fortet's equation which is a new way to consider the existing result. See Section 2.3.1.
2. The connection between $g(t|y,s)$ and $w(x,t|y,s)$ is explained based on Fokker-Planck Equation. See Section 2.3.2.

3. The initial-boundary value problem for Fokker-Planck Equation for $w(x,t|y,s)$ is solved by the method of images, which leads to explicit solutions under some special failure level. See Section 2.3.3.
4. Based on the derived result, analytical approximations of first passage density and distribution are derived for the general boundary. See Section 2.3.4.
5. Based on the method of images in Section 2.3.3, the linear programming approach is introduced to approximate the first passage density. This is a direct application of the method of images. See Section 2.3.5.
6. Based on a special boundary given in Section 2.3.3, it is considered to derive an explicit expression of first passage density $g(t|y,s)$ for a piecewise boundary where in each interval the boundary is of the given form. See Section 2.3.6.
7. With a preliminary result for mean-zero Gaussian processes, an approximation is derived for $g(t|y,s)$. This is found to be the same with the approximation derived in Section 2.3.4. See Section 2.3.7.

2.3.1 Derivation

It is a classical result that the first passage density is connected with the initial-boundary value problem of Fokker-Planck equation, see [5, 13, 33, 62]. However this connection is generally derived from a physical illustration to introduce an absorbing boundary for FPE. That is to say, releasing particles at the initial time, the first passage time is investigated by counting remaining particles when an absorbing boundary is proposed to remove particles crossed before current time [33, 62].

To the extent of the author's knowledge, a rigorous proof for the representation of FPT problem based on the PDE approach is given for the Brownian motion in [43]. However this question confuses the authors in more complex situations in [53]. In this section we will try to make clear such a problem in a simple situation from a new angle of Fortet's equation (2.2.1), where only one-dimensional process to one-sided boundary is considered. What should be mentioned more is that the analysis for Brownian motion in [43] is from an angle of constructing crossing boundaries from the explicit expression based on the method of images. Our starting point is from Fortet's equation when the crossing boundary is given.

The initial-boundary value problem for FPE is given by the following lemma for $w(x,t|y,s)$ defined in (2.1.9). The key idea is to use Equation (2.2.4) to derive corresponding calculation.

Lemma 2.3.1. *Denote $w(x,t) := w(x,t|y,s)$ as in (2.1.9), then $\forall t \geq s$ it satisfies :*

$$\frac{\partial w(x,t|y,s)}{\partial t} = -\frac{\partial}{\partial x}((a(t)x + b(t))w(x,t|y,s)) + c(t)\frac{\partial^2 w(x,t|y,s)}{\partial^2 x}, \tag{2.3.1}$$

with $w(x,t|y,s) = 0, \forall x \geq L(t)$, $w(-\infty,t|y,s) = 0$, $w(x,s|y,s) = f_y(x), x < L(s)$ and f_y is the pdf of y if it is existing, or the Dirac measure centered at y when y is deterministic.

Proof. First, it is well-known that $p(x,t|y,s)$ satisfies Fokker-Planck equation (2.1.3) with initial and boundary conditions $p(\pm\infty,t|y,s) = 0$, $p(x,s|y,s) = f_y(x)$, $t \geq s$.

Second, the initial and boundary conditions for $w(x,t|y,s)$ can be verified from Lemma 2.2.1.

- $\forall x \geq L(t), t \geq s$, from (2.2.1) and (2.2.4), it comes to $w(x,t|y,s) = 0$.

◦ From (2.2.4),

$$w(-\infty, t|y, s) = p(-\infty, t|y, s) - \int_s^t g(z|x_0, t_0)p(-\infty, t|L(z), z)dz = 0. \tag{2.3.2}$$

And also $\forall x < L(s)$,

$$w(x, s|y, s) = p(x, s|y, s) = f_y(x). \tag{2.3.3}$$

Third, when $x < L(t)$, from (2.2.4), further calculation contributes to :

$$\frac{\partial w(x, t|y, s)}{\partial t} = \frac{\partial p(x, t|y, s)}{\partial t} - \int_s^t g(z|y, s)\frac{\partial p(x, t|L(z), z)}{\partial t}dz - g(t|y, s)p(x, t|L(t), t). \tag{2.3.4}$$

Notice $p(x, t|L(t), t)$ is a Dirac measure centered at $L(t)$, where $p(x, t|L(t), t) = 0$ when $x < L(t)$. Therefore we know for $x < L(t)$,

$$\frac{\partial w(x, t|y, s)}{\partial t} = \frac{\partial p(x, t|y, s)}{\partial t} - \int_s^t g(z|y, s)p_t(x, t|L(z), z)dz, \tag{2.3.5}$$

Furthermore,

$$\frac{\partial w(x, t|y, s)}{\partial x} = \frac{\partial p(x, t|y, s)}{\partial x} - \int_s^t g(z|y, s)\frac{\partial p(x, t|L(z), z)}{\partial x}dz, \tag{2.3.6}$$

also

$$\frac{\partial^2 w(x, t|y, s)}{\partial x^2} = \frac{\partial^2 p(x, t|y, s)}{\partial x^2} - \int_s^t g(z|y, s)\frac{\partial^2 p(x, t|L(z), z)}{\partial x^2}dz. \tag{2.3.7}$$

From (2.3.5), (2.3.6), (2.3.7), and Fokker Planck equation (2.1.3), denote $w(x, t) := w(x, t|y, s), g(t) := g(t|y, s), p^1(x, t) := p(x, t|y, s), p^2(x, t) := p(x, t|L(s), s)$, we can derive :

$$\begin{aligned}
&-\frac{\partial}{\partial x}((a(t)x + b(t))w(x, t)) + c(t)\frac{\partial^2(w(x, t))}{\partial^2 x} = -a(t)w(x, t) - ((a(t)x + b(t))w_x(x, t) + c(t)w_{xx}(x, t)\\
&= -a(t)(p^1(x, t) - \int_s^t g(z)p^2(x, t)dz) - (a(t)x + b(t))(p^1_x(x, t) - \int_s^t g(z)p^2_x(x, t)dz)\\
&+ c(t)(p^1_{xx}(x, t) - \int_s^t g(z)p^2_{xx}(x, t)dz)\\
&= -a(t)p^1(x, t) - ((a(t)x + b(t))p^1_x(x, t) + c(t)p^1_{xx}(x, t)\\
&+ a(t)\int_s^t g(z)p^2(x, t)dz + (a(t)x + b(t))\int_s^t g(z)p^2_x(x, t)dz - c(t)\int_s^t g(z)p^2_{xx}(x, t)dz\\
&= p^1_t(x, t|y, s) - \int_s^t g(z)p^2_t(x, t|L(z), z)dz = \frac{\partial w(x, t)}{\partial t}.
\end{aligned} \tag{2.3.8}$$

Therefore from (2.3.8) and previously stated boundary condition, the proof is completed. □

From Lemma 2.3.1, a free boundary problem can be introduced for $u(x, t|y, s)$, which is proved well-posed in [12, 13].

Lemma 2.3.2. *Denote $u(x, t) := u(x, t|y, s)$, then $\forall t \geq s$ it satisfies :*

$$\frac{\partial u(x, t)}{\partial t} = -(a(t)x + b(t))\frac{\partial u(x, t)}{\partial x} + c(t)\frac{\partial^2 u(x, t)}{\partial^2 x}, \tag{2.3.9}$$

where $u(x, t|y, s) = u(L(t), t|y, s), \forall x \geq L(t)$, $u(-\infty, t|y, s) = 0$, $u(x, s|y, s) = F_y(x), x < L(s)$ and F_y is the distribution function of y.

Proof. As $u(x,t|y,s) = \int_{-\infty}^{x} w(z,t|y,s)dz$, and $w(x,t|y,s)$ satisfies (2.3.1). By integrating Equation (2.3.1) for both sides on $(-\infty, x)$, it is natural to get (2.3.9). □

As $u(L(t),t) = P(\tau_{y,s} > t)$ is unknown in direct problems, the free boundary can be proposed when the inverse problem is considered, i.e. $u(L(t),t)$ is known but the boundary is to be solved.

2.3.2 Connection with First Passage Density

In the last subsection, $w(x,t|y,s)$ is proved to satisfy a boundary problem for FPE. Its connection with first passage density will be explained as follows.

The connection between $w(x,t|y,s)$ and $\tau_{y,s}$ is explained by the first passage distribution as in (2.1.10). Following statements in [5, 13, 43, 62], another expression will be introduced by first passage density. From (2.1.10), the pdf $g(t|y,s)$ of $\tau_{y,s}$ satisfies :

$$-g(t|y,s) = \frac{\partial P(\tau_{y,s} > t)}{\partial t} = \frac{\partial P(X_t < L(t), \tau_{y,s} > t | X_s = y)}{\partial t} = \frac{\partial \int_{-\infty}^{L(t)} w(x,t|y,s)dx}{\partial t} \tag{2.3.10}$$

When $x < L(t)$, $w(x,t|y,s), w_t(x,t|y,s) \in C(R \times R^+)$. Then we have the following equation from analysis knowledge,

$$\frac{\partial(\int_{-\infty}^{L(t)} w(x,t|y,s)dx)}{\partial t} = \int_{-\infty}^{L(t)} \frac{\partial w(x,t|y,s)}{\partial t}dx + w(L(t),t|y,s)L'(t). \tag{2.3.11}$$

From the boundary condition of (2.3.1), we have

$$-g(t|y,s) = \frac{\partial \int_{-\infty}^{L(t)} w(x,t|y,s)dx}{\partial t} = \int_{-\infty}^{L(t)} \frac{\partial w(x,t|y,s)}{\partial t}dx. \tag{2.3.12}$$

It is of the following form from (2.3.1) :

$$-g(t|y,s) = \int_{-\infty}^{L(t)} (-\frac{\partial}{\partial x}[(a(t)x + b(t))w(x,t|y,s)] + c(t)\frac{\partial^2(w(x,t|y,s))}{\partial^2 x})dx. \tag{2.3.13}$$

Noticing corresponding boundary conditions $w(L(t),t|y,s) = 0, w(-\infty,t|y,s) = 0$, then

$$g(t|y,s) = -c(t)\frac{\partial(w(x,t|y,s))}{\partial x}|_{x=L(t)}. \tag{2.3.14}$$

Here the notation $f(x)|_{x=z}$ is introduced to represent the value of $f(x)$ at z. Summarizing all the above, a way to get FPT distribution (2.1.10) or first passage density (2.3.14) is to solve $w(x,t|y,s)$ from FPE (2.3.1).

2.3.3 Method of Images

As stated in the last section, the first passage density can be solved from $w(x,t|y,s)$, therefore $w(x,t|y,s)$ will be solved by the method of images in this subsection. The discussion in this subsection for $w(x,t|y,s)$ is divided into several steps as follows :

1. As the original boundary $L(t)$ is time-dependent, it results in the confusion to consider corresponding symmetric area in the method of images. So by introducing a variable-change process $Y_t = X_t - L(t)$, the original boundary $L(t)$ for X_t is converted to 0 for the changed process Y_t, which corresponds to a joint pdf $\tilde{w}(x,t|y_0,s) = w(x + L(t),t|y,s), y_0 = y - L(s)$.
2. Introducing the pending initial condition $U(r)$ the explicit expression for $\tilde{w}(x,t|y_0,s)$ is given in (2.3.17) by the method of images.
3. Under a quasi-linear boundary defined in (2.3.21), the explicit expression for the first passage density is derived in (2.3.22).
4. An explicit condition for the boundary is given whose corresponding first passage time is promised to be almost surely finite.
5. A special case of derived results is the inverse Gaussian distribution for Brownian motion, which is extended to the time-changed Brownian motion.

Variable-Change

In this subsection, following results in [11, 43, 96], the method of image is discussed for the boundary-value problem itself to derive possible explicit solutions, which requires solving a Fredholm integral equation of first kind.

By variable-change of $Y_t = X_t - L(t)$, denote $y_0 = y - L(s), \tau^Y_{y_0,s} := \inf_{t\geq s}\{Y_t \geq 0\}$, it is derived that

$$dY_t = \left[a(t)Y_t + b(t) + a(t)L(t) - L'(t)\right]dt + \sigma(t)dB_t, \quad t \geq s, Y_s = y_0. \tag{2.3.15}$$

Then it is obvious that $\tau^Y_{y_0,s} = \tau_{y,s}$. From Lemma 2.3.1, denote $\tilde{b}(t) = b(t)+a(t)L(t)-L'(t)$, $\tilde{w}(z,t|y_0,s) := \frac{\partial P(Y_t<z,\tau^Y_{y_0,s}>t|Y_s=y_0)}{\partial z}$ satisfies :

$$\frac{\partial \tilde{w}(z,t|y_0,s)}{\partial t} = -\frac{\partial}{\partial z}((a(t)z + \tilde{b}(t))\tilde{w}(z,t|y_0,s)) + c(t)\frac{\partial^2 \tilde{w}(z,t|y_0,s)}{\partial^2 z}, \tag{2.3.16}$$

with $\tilde{w}(z,t|y_0,s) = 0, \forall z \geq 0,\ \tilde{w}(-\infty,t|y_0,s) = 0,\ \tilde{w}(z,s|y_0,s) = f_{y_0}(z), z < 0$ and $f_{y_0}(z) = f_y(z + L(s))$ is the pdf of y_0.

From (2.1.4), (2.1.5) and the method of images, we know that a general solution to (2.3.16) is given by :

$$\begin{aligned}\tilde{w}(z,t|y_0,s) = \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\{&\int_{-\infty}^{0}\exp(-\frac{(ze^{\alpha(t,s)} + \tilde{\beta}(t,s) - r)^2}{4\gamma(t,s)})f_{y_0}(r)dr \\ &- \int_0^{+\infty}\exp(-\frac{(ze^{\alpha(t,s)} + \tilde{\beta}(t,s) - r)^2}{4\gamma(t,s)})U(r)dr\},\end{aligned} \tag{2.3.17}$$

where $U(r), r \in [0,+\infty)$ is a pending function to be solved. Moreover,

$$\tilde{\beta}(t,s) = -\int_s^t \tilde{b}(r)e^{\alpha(r,s)}dr = \beta(t,s) - L(s) + L(t)e^{\alpha(t,s)}. \tag{2.3.18}$$

It is noticed further that $\tilde{w}(x,t|y_0,s) = w(x + L(t),t|y,s), y_0 = y - L(s)$, such that

$$w(x,t|y,s) = p(x,t|y,s) - \int_0^{+\infty} p(x,t|L(s) + r,s)U(r)dr. \tag{2.3.19}$$

Noticing the boundary condition $\tilde{w}(0,t|y_0,s)=0, \forall t\geq 0$, then $U(r)$ can be solved from the following Fredholm integral equation of first kind :

$$h(t)=\int_0^{+\infty} K(r,t)U(r)dr, \quad \forall t\geq s, \tag{2.3.20}$$

where $K(r,t) = \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\exp(-\frac{(\tilde{\beta}(t,s)-r)^2}{4\gamma(t,s)}) = p(L(t),t|L(s)+r,s), h(t) = \int_{-\infty}^0 K(r,t)f_{y_0}(r)dr = p(L(t),t|y,s)$.

Equation (2.3.20) is generally hard to be solved, even numerically, so much attention is paid to (2.3.20) to search an explicit expression for $U(r)$ [43, 46, 11]. In the following we will derive a special boundary from (2.3.20), under which $w(x,t|y,s)$ can be solved explicitly.

Quasi-Linear Boundary

[46, 11, 79] contribute to establish the special failure level with several adjustable parameters. Such an idea is reproduced in the following proposition. However, the key idea here is not to directly search a special boundary, but to propose a boundary $L(t)$ based on the boundary $L(s)$ at s. As can be seen later, the treatment of variable-change saves plenty of calculation on the tangent method and the piecewise approximation of the original boundary.

Before the discussion, we first introduce the definition of *quasi-linear*.

Definition 2.3.3. A boundary $L(t)$ is called *quasi-linear*, if for a constant $C\in\mathbb{R}$,

$$L(t)=e^{-\alpha(t,s)}(L(s)-\beta(t,s)+C\gamma(t,s)), \tag{2.3.21}$$

where α,β,γ are defined in (1.3.3).

Remark : We call the boundary (2.3.21) as quasi-linear boundary because for Brownian motion, this boundary is simplified to a linear boundary. This will be shown later.

Then we have the following proposition.

Proposition 2.3.4. *For a boundary $L(t)$, suppose the process X_t is given as in (2.1.1) with the initial start (y,s) where y is a random variable with the pdf f_y defined on $(-\infty,L(s))$. If there exists a constant $C\in\mathbb{R}$ such that the boundary $L(t)$ satisfies (2.3.21), then the first passage density $g(t|y,s)$ satisfies :*

$$g(t|y,s)=\frac{c(t)e^{2\alpha(t,s)}}{\sqrt{4\pi\gamma^3(t,s)}}\int_{-\infty}^{L(s)}(L(s)-r)\exp(-\frac{(\beta(t,s)+L(t)e^{\alpha(t,s)}-r)^2}{4\gamma(t,s)})f_y(r)dr. \tag{2.3.22}$$

Moreover the distribution function $G(t|y,s):=P(\tau_{y,s}\leq t)$ satisfies :

$$G(t|y,s)=\int_{-\infty}^{L(s)}\left[\Phi(\frac{-L(s)-C\gamma(t,s)+r}{\sqrt{2\gamma(t,s)}})+e^{C(r-L(s))}\Phi(\frac{C\gamma(t,s)-L(s)+r}{\sqrt{2\gamma(t,s)}})\right]f_y(r)dr \tag{2.3.23}$$

where $\Phi(*)$ is the normal distribution function.

Proof. First, the following transform comes directly from the boundary condition (2.3.20)

$$\int_0^{+\infty}\exp(-\frac{(\tilde{\beta}(t,s)-r)^2}{4\gamma(t,s)})(U(r)-\exp(-r\frac{\tilde{\beta}(t,s)}{\gamma(t,s)})f_{y_0}(-r))dr=0. \tag{2.3.24}$$

Therefore the following necessary condition can be introduced to promise (2.3.20)

$$U(r) = \exp(-r\frac{\tilde{\beta}(t,s)}{\gamma(t,s)})f_{y_0}(-r), \quad \forall r \geq 0. \tag{2.3.25}$$

Noticing the left side of (2.3.25) depends only on r, there should be a constant $C \in R$ to promise

$$\frac{\tilde{\beta}(t,s)}{\gamma(t,s)} = C. \tag{2.3.26}$$

By calculating Equation (2.3.26) explicitly, it is found $L(t)$ is quasi-linear as in (2.3.21).

Second, under the condition (2.3.26), $U(r) = e^{-Cr}f_{y_0}(-r)$, that leads to

$$\begin{aligned}
\tilde{w}(z,t|y_0,s) &= \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\{\int_{-\infty}^{0}\exp(-\frac{(ze^{\alpha(t,s)}+\tilde{\beta}(t,s)-r)^2}{4\gamma(t,s)})f_{y_0}(r)dr \\
&- \int_{0}^{+\infty}\exp(-\frac{(ze^{\alpha(t,s)}+\tilde{\beta}(t,s)-r)^2}{4\gamma(t,s)})e^{-Cr}f_{y_0}(-r)dr\} \\
&= \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\{\int_{-\infty}^{0}\exp(-\frac{(ze^{\alpha(t,s)}+\tilde{\beta}(t,s)-r)^2}{4\gamma(t,s)})f_{y_0}(r)dr \\
&- \int_{0}^{+\infty}\exp(-\frac{(ze^{\alpha(t,s)}+\tilde{\beta}(t,s)+r)^2-4rze^{\alpha(t,s)}}{4\gamma(t,s)})f_{y_0}(-r)dr\},
\end{aligned} \tag{2.3.27}$$

Third, it should be noticed that $\tilde{w}(z,t|y_0,s) = w(z+L(t),t|y,s)$ such that

$$\begin{aligned}
&w(x,t|y,s) \\
&= \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\int_{-\infty}^{L(s)}(1-\exp(-\frac{(r-L(s))(x-L(t))e^{\alpha(t,s)}}{\gamma(t,s)}))\exp(-\frac{(xe^{\alpha(t,s)}+\beta(t,s)-r)^2}{4\gamma(t,s)})f_y(r)dr
\end{aligned} \tag{2.3.28}$$

This is also translated to the following form when y is defined on $(-\infty, L(s))$,

$$w(x,t|y,s) = \int_{-\infty}^{L(s)}\left[p(x,t|r,s) - e^{C(r-L(s))}p(x,t|2L(s)-r,s)\right]f_y(r)dr. \tag{2.3.29}$$

Moreover, from (2.1.10), (2.2.8) and the condition (2.3.21),

$$\begin{aligned}
&1-G(t|y,s) = \int_{-\infty}^{L(t)}w(x,t|y,s)dx \\
&= \int_{-\infty}^{L(s)}\left[F(L(t),t|r,s) - e^{C(r-L(s))}F(L(t),t|2L(s)-r,s)\right]f_y(r)dr \\
&= \int_{-\infty}^{L(s)}\left[\Phi(\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-r}{\sqrt{2\gamma(t,s)}}) - e^{C(r-L(s))}\Phi(\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-2L(s)+r}{\sqrt{2\gamma(t,s)}})\right]f_y(r)dr \\
&= \int_{-\infty}^{L(s)}\left[\Phi(\frac{L(s)+C\gamma(t,s)-r}{\sqrt{2\gamma(t,s)}}) - e^{C(r-L(s))}\Phi(\frac{C\gamma(t,s)-L(s)+r}{\sqrt{2\gamma(t,s)}})\right]f_y(r)dr.
\end{aligned} \tag{2.3.30}$$

Noticing that $1-\Phi(x) = \Phi(-x)$, it comes naturally to (2.3.23).

Fourth, from (2.3.28), $\frac{\partial w(x,t|y,s)}{\partial x}$ satisfies :

$$\frac{\partial w(x,t|y,s)}{\partial x}|_{x=L(t)} = \frac{e^{2\alpha(t,s)}}{\sqrt{4\pi\gamma^3(t,s)}}\int_{-\infty}^{L(s)}(r-L(s))\exp(-\frac{(\beta(t,s)+L(t)e^{\alpha(t,s)}-r)^2}{4\gamma(t,s)})f_y(r)dr. \tag{2.3.31}$$

By Equations (2.3.14) and (2.3.31), it is natural to achieve Equation (2.3.22). □

Remarks :

- the proposition can also be possibly derived by time-change based on inverse Gaussian distribution for Brownian motion.
- (2.3.28) provides an expression of conditional survival probability $P(\tau_{y,s} > t|X_t = x)$, which will be addressed later. And if we derive from the general method of images without the variable-change, it is hard to see clearly this connection.

Hitting the Boundary with a Finite Time

As stated at the beginning, it is supposed that the first passage time $\tau_{y,s}$ considered in this thesis is always almost surely finite without rigorous consideration. Here from (2.3.23), an explicit condition to promise $P(\tau_{y,s} < \infty) = 1$ for $y \in (-\infty, L(s))$ is given by the following corollary.

Corollary 2.3.5. *If there exists a constant $C < 0$, such that $L(t)$ satisfies*

$$L(t) \leq e^{-\alpha(t,s)}(L(s) - \beta(t,s) + C\gamma(t,s)), \tag{2.3.32}$$

then

$$P(\tau_{y,s} = \infty) = 0, \text{ or equivalently, } P(\tau_{y,s} < \infty) = 1. \tag{2.3.33}$$

Proof. It is denoted here $\tilde{L}(t) := e^{-\alpha(t,s)}(L(s) - \beta(t,s) + C\gamma(t,s))$, and $\tau_1 := \inf_{t\geq s}\{t|X_t \geq \tilde{L}(t)\}$. From (2.3.21), $\tilde{L}(t)$ is quasi-linear such that its first passage distribution is explicit.

Then from (2.3.32), the boundary $L(t) \leq \tilde{L}(t)$, it is natural to see the event of $\{\tau_{y,s} > t\}$ is included in the event of $\{\tau_1 > t\}$. Then from (2.3.23) with the hypothesis of $C < 0$, we have

$$\begin{aligned} 0 \leq P(\tau_{y,s} = \infty) &\leq P(\tau_1 = \infty) \\ &= \lim_{t\to\infty}\{1 - \Phi(\frac{-L(s) - C\gamma(t,s) + y}{\sqrt{2\gamma(t,s)}}) - e^{C(r-L(s))}\Phi(\frac{C\gamma(t,s) - L(s) + y}{\sqrt{2\gamma(t,s)}})\} = 0. \end{aligned} \tag{2.3.34}$$

□

Drifted Brownian Motion

As introduced in Section 1.2, the drifted Brownian motion is a widely used degradation process. It is of special interest although it is also a special OU process. Proposition 2.3.4 provides another view for the inverse Gaussian distribution for Brownian motion, actually we have

Corollary 2.3.6. *For the drifted Brownian motion $X_t = \mu t + \sigma B_t + x_0$ with a deterministic start x_0 at 0, its first passage density to a constant boundary $L > x_0$ satisfies inverse Gaussian distribution :*

$$g(t|x_0, 0) = \frac{L - x_0}{\sqrt{2\pi\sigma^2 t^3}} \exp(-\frac{(L - \mu t - x_0)^2}{2\sigma^2 t}). \tag{2.3.35}$$

Correspondingly, the first passage distribution is given by

$$G(t|x_0, 0) = \Phi(\frac{-L + \mu t + x_0}{\sigma\sqrt{t}}) + e^{\frac{2\mu}{\sigma^2}(L - x_0)}\Phi(\frac{-\mu t - L + x_0}{\sigma\sqrt{t}}), \tag{2.3.36}$$

where $\Phi(*)$ is the normal distribution function.

Proof. As $dX_t = \mu dt + \sigma dB_t$ with initial distribution $\delta(x - x_0)$ as Dirac measure centred at x_0, notations in (2.3.22) are modified to

$$\alpha(t, 0) = 0, \quad \beta(t, 0) = -\mu t, \quad \gamma(t, 0) = \frac{\sigma^2 t}{2}. \tag{2.3.37}$$

Substitute all the above to (2.3.22), and let $C = -\frac{2\mu}{\sigma^2}$ in (2.3.21), we come to (2.3.35).

Moreover, from (2.2.8), we know

$$\begin{aligned} 1 - G(t|x_0, 0) &= F(L(t), t|x_0, 0) - e^{C(x_0 - L(s))}F(L(t), t|2L(s) - x_0, 0) \\ &= \Phi(\frac{L - \mu t - x_0}{\sigma\sqrt{t}}) - e^{-\frac{2\mu}{\sigma^2}(x_0 - L)}\Phi(\frac{L - \mu t - 2L + x_0}{\sigma\sqrt{t}}). \end{aligned} \tag{2.3.38}$$

□

Time-Changed Brownian Motion

The drifted Brownian motion is with a linearly-increasing mean and variance. This limits its applications in degradation modeling where the nonlinear performance is presented. So here it is of interest to extend the simplicity of Brownian motion to a more general case, where its mean and variance can be adjusted. But at the same time, it could keep the simple calculation on the first passage density. This leads to the consideration of time-changed Brownian motion [49]. We here will also apply Proposition 2.3.4 to this process.

Definition 2.3.7. X_t is called a time-changed Brownian motion, if X_t is of the form

$$dX_t = \frac{C}{2}\sigma^2(t)dt + \sigma(t)dB_t, X_0 = x_0, t \geq 0 \tag{2.3.39}$$

where $C > 0$, and $\sigma(t) > 0$ is a continuous function.

Remark : The process X_t here can be essentially translated to the case of time-changed drifted Brownian motion $X_t = \frac{C}{2}\Lambda(t) + B_{\Lambda(t)}$ [49] based on the time-change $\Lambda(t) = 2\gamma(t, 0)$, see Section 1.3.2. However we would like to discuss this process in the framework of stochastic differential equations.

The notations in (1.3.3) for (2.3.39) are updated to

$$\alpha(t, s) = 0, \gamma(t, s) = \frac{1}{2}\int_s^t \sigma^2(z)dz, \beta(t, s) = -C \times \gamma(t, s), \tag{2.3.40}$$

It is solved explicitly

$$X_t = x_0 - \beta(t,0) + \int_0^t \sigma(s) dB_s. \tag{2.3.41}$$

Correspondingly its transition pdf is given by

$$p(x,t|y,s) = \frac{1}{\sqrt{4\pi\gamma(t,s)}} \exp(-\frac{(x-y-C\gamma(t,s))^2}{4\gamma(t,s)}), \tag{2.3.42}$$

such that the transition distribution is given by

$$F(x,t|y,s) = \Phi(\frac{x-y-C\gamma(t,s)}{\sqrt{2\gamma(t,s)}}). \tag{2.3.43}$$

Following Proposition 2.3.4, it is interesting to see that the first passage problem for the time-changed Brownian motion is also explicitly solved.

Corollary 2.3.8. *For the time-changed Brownian motion (2.3.39) with the observation of y at time s , its first passage density to a constant boundary $L > y$ satisfies the following equation :*

$$g(t|y,s) = \frac{\sigma^2(t)(L-y)}{2\gamma(t,s)} p(L,t|y,s). \tag{2.3.44}$$

Correspondingly, the first passage distribution is given by

$$G(t|y,s) = \Phi(\frac{-L+C\gamma(t,s)+y}{\sqrt{2\gamma(t,s)}}) + e^{-C(y-L)}\Phi(\frac{-C\gamma(t,s)-L+y}{\sqrt{2\gamma(t,s)}}), \tag{2.3.45}$$

where $\Phi(*)$ is the normal distribution function.

Proof. It is noticed that, for C in (2.3.39) and $L(s) = L$,

$$L(t) = L(s) - \beta(t,s) - C\gamma(t,s) = L(s) = L, \ \ for\ t \geq s. \tag{2.3.46}$$

Therefore L satisfies is quasi-linear boundary by (2.3.21), and from Proposition 2.3.4, the first passage density and distribution are given explicitly. □

Remark : The time-changed Brownian motion coincides with the drifted Brownian motion when $\sigma(t) = \nu$ as a constant.

2.3.4 The Tangent Approximation

In the last subsection, we have presented under a quasi-linear boundary, the first passage density can be given explicitly. In this subsection, it aims to approximate a general boundary by the quasi-linear boundary. And we concern the following points :

1. A tangent approximation of the first passage density and distribution is presented, this gives a good estimate when the time is near the current observation.
2. This tangent approximation can have a global accuracy under some conditions.

The tangent approximation for first passage problems of Brownian motion is to approximate the real boundary by the linear boundary. This is well studied in [43], and in this subsection we will try to reproduce the idea of tangent approximation for the OU process based on the quasi-linear boundary (2.3.21). The classical result is only for the first passage density of Brownian motion. Here we will also consider the first passage distribution for the OU process.

Recall that the discussion for the quasi-linear boundary $L(t)$ in (2.3.21) is given by

$$L(t) = e^{-\alpha(t,s)}(L(s) - \beta(t,s) + C\gamma(t,s)), \tag{2.3.47}$$

where $C \in \mathbb{R}$ and $L(s)$ is the initial boundary. Then the first passage density and distribution are given by (2.3.22) and (2.3.23) respectively.

For a general boundary $L(t)$ not in the form of (2.3.47), we want to approximate the real boundary by the quasi-linear one. This induces the tangent approximation. And the remaining statements are due to the approximation on $L(s)$ and C in (2.3.47).

Actually when $L(z) \in C^1[s,t]$ and s is near t,

$$\tilde{C}(t,s) := \frac{L'(t) - a(t)L(t) - b(t)}{c(t)e^{\alpha(t,s)}} \approx \frac{L(t)e^{\alpha(t,s)} - L(s) + \beta(t,s)}{\gamma(t,s)}, \tag{2.3.48}$$

such that for $z \in [s,t]$, we can approximate the original boundary $L(z)$ by the following one,

$$\tilde{L}(z) = L(t)e^{\alpha(t,z)} + \beta(t,z) - \tilde{C}(t,s)\gamma(t,z). \tag{2.3.49}$$

Therefore, an approximation for the first passage density and first passage distribution are given by the following propositions.

Proposition 2.3.9. *If for a constant T, the boundary $L(t) \in C^1[s,T]$, for $\tilde{C}, \tilde{L}$ defined in (2.3.48) and (2.3.49),*

$$\tilde{G}(t|y,s) := 1 - F(L(t),t|y,s) + e^{\tilde{C}(t,s)(y-\tilde{L}(s))}F(L(t),t|2\tilde{L}(s) - y, s) \tag{2.3.50}$$

Then we have

$$G(t|y,s) = \tilde{G}(t|y,s)(1 + O(\gamma(t,s))),\ t \to s. \tag{2.3.51}$$

Proof. First step : control the original boundary by quasi-linear boundaries

As $L(t) \in C^1[s,T]$, therefore $\tilde{C}(t,z)$ is continuous, such that $\tilde{C}(t,z) \le K/2$ for an enough large constant K in the given area, $[s,T] \times [s,T]$. This K depends only on the end-time T.

It is known from (2.3.49) when $z \in [s,t]$, by mean-value theorem,

$$\begin{aligned} |\tilde{L}(z) - L(z)| &= |L(t)e^{\alpha(t,z)} + \beta(t,z) - L(z) - \tilde{C}(t,s)\gamma(t,z)| \\ &= \gamma(t,z)|\tilde{C}(\xi,z) - \tilde{C}(t,s)|,\ \text{for some } \xi \in (z,t) \\ &\le K\gamma(t,z). \end{aligned} \tag{2.3.52}$$

Therefore we have

$$\tilde{L}(z) - K\gamma(t,z) \le L(z) \le \tilde{L}(z) + K\gamma(t,z),\ \ for\ z \in [s,t]. \tag{2.3.53}$$

Second step : control the original first passage distribution by two explicit expressions

Now consider t is enough close to s, such that $\tilde{L}(s)+K\gamma(t,s)-y>0$ due to $L(s)-y>0$. Denote $\tilde{L}_1(z):=\tilde{L}(z)-K\gamma(t,z)$, and $\tilde{L}_2(z):=\tilde{L}(z)+K\gamma(t,z)$, then we have $\tilde{L}_1(z)\leq L(z)\leq \tilde{L}_2(z)$, for $z\in[s,t]$. Denote the first passage times to these two boundaries respectively as $\tilde{\tau}_1,\tilde{\tau}_2$. It is noticed that $\tilde{L}_1(z),\tilde{L}_2(z)$ are also quasi-linear boundaries from Equation (2.3.49) such that for $z\in(s,t)$

$$\begin{aligned} &P(\tilde{\tau}_2\leq z)\leq G(z|y,s)\leq P(\tilde{\tau}_1\leq z),\\ &P(\tilde{\tau}_2\leq z)\leq \tilde{G}(z|y,s)\leq P(\tilde{\tau}_1\leq z). \end{aligned} \tag{2.3.54}$$

Here we have

$$\begin{aligned} 1-P(\tilde{\tau}_2\leq z)&=F(L(z),z|y,s)-e^{(\tilde{C}(z,s)+K)(y-\tilde{L}(s)-K\gamma(z,s))}F(L(z),z|2(\tilde{L}(s)+K\gamma(z,s))-y,s).\\ 1-P(\tilde{\tau}_1\leq z)&=F(L(z),z|y,s)-e^{(\tilde{C}(z,s)-K)(y-\tilde{L}(s)+K\gamma(z,s))}F(L(z),z|2(\tilde{L}(s)-K\gamma(z,s))-y,s). \end{aligned} \tag{2.3.55}$$

By the expression of $F(x,t|y,s)$ in (2.2.8), such that for t close to s, there exists a constant $N_{t,s}$ only depending on t and s,

$$P(\tilde{\tau}_1\leq z)-P(\tilde{\tau}_2\leq z)\leq N_{t,s}\gamma(z,s),\ \ for\ z\in[s,t]. \tag{2.3.56}$$

Therefore due to $|G(t|y,s)-\tilde{G}(t|y,s)|\leq P(\tilde{\tau}_1\leq t)-P(\tilde{\tau}_2\leq t)$, it leads directly to

$$G(t|y,s)=\tilde{G}(t|y,s)(1+O(\gamma(t,s))). \tag{2.3.57}$$

□

Remark : An illustration for constructing two quasi-linear boundaries to control the original boundary can be seen in Figure 2.2.

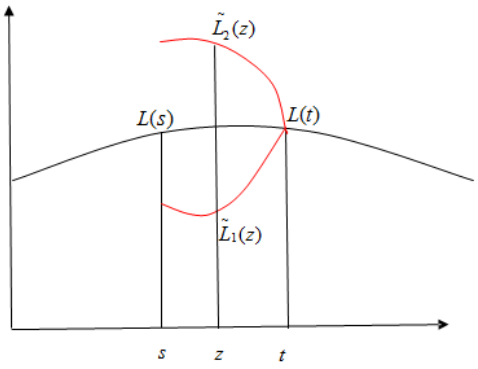


FIGURE 2.2 – Control the original boundary (black) by quasi-linear boundaries (red).

For the first passage density, the estimation requires much more efforts, however it is natural to derive the following result for the practical calculation.

Proposition 2.3.10. *Under some conditions, if for a constant T, the boundary $L(t)\in C^1[s,T]$, for $\tilde{C},\tilde{L}$ defined in (2.3.48) and (2.3.49),*

$$\tilde{g}(t|y,s)=\frac{c(t)e^{\alpha(t,s)}}{\gamma(t,s)}(\tilde{L}(s)-y)p(L(t),t|y,s). \tag{2.3.58}$$

Then

$$g(t|y,s) = \tilde{g}(t|y,s)(1+o(1)), t \to s. \tag{2.3.59}$$

Proof. The detailed proof is omitted here. However an indirect proof can be derived from the time-change of the OU process. Then the conclusion comes by using the preliminary result for Brownian motion [43]. □

Global accuracy

The tangent approximation converges when $t \to s$, however it also has a global accuracy under some conditions. And it seems strange at a first glance, because the tangent approximation is a local approximation. Also interestingly, it is found that this approximation is the same with the approximation by Durbin in [25, 27], which will be explained explicitly later.

This global accuracy is discussed in [43] for Brownian motion. In that monograph, the global accuracy is considered based on asymptotic analysis for given parametric boundaries solved from the method of images. However, for a general boundary, this treatment means we should construct a convergent boundary series satisfying the hypothesis of the method of images. This would be tedious and technical. Another error estimation of the tangent approximation is given based on an integral equation method in [31]. In this thesis, we will consider the global accuracy from a practical view, and the estimate of the global error would be missing. This rough consideration is adopted for Brownian motion before for the fitting of life tables [71] and also in residual useful life estimation [66].

2.3.5 Parametric Approximation

Now we go back to the original initial-boundary problem (2.3.1) for a general boundary $L(t)$. The discussion in this subsection concerns 2 points as follows :

1. By the method of images, it is proposed to consider an approximate function for the first passage density with several parameters. And this is later estimated by a linear programming approach.
2. a special case of 2-point approximation induces the explicit expression of first passage density for a special class of boundaries, which are called quasi-Daniels boundaries.

A Linear Programming Approach

As $w(x,t|y,s)$ satisfies an initial-boundary problem in (2.3.1), and it can be generally written by the method of images from (2.3.19)

$$w(x,t|y,s) = p(x,t|y,s) - \int_0^{+\infty} p(x,t|L(s)+r,s)U(r)dr, \tag{2.3.60}$$

where $U(r)$ can be obtained from the Fredholm integral equation (2.3.20). Such an equation can hardly be solved analytically. Even numerical solutions to a Fredholm integral equation can not be direct as such a problem is generally ill-posed where collocation methods could be useful [35]. However here we do not adopt a rigorous way to consider numerical problems, but consider a more heuristic way leading to linear programming problems [95].

Following the idea in [95], suppose mesh points $\{r_i\}_{i=1}^N$ are given for $(0, +\infty)$, and then an approximation for $U(r)$ is introduced by the point-measure :

$$U(r)dr = \sum_{i=1}^{N} \xi_i \delta_{L(s)+r_i}(dr), \tag{2.3.61}$$

where $\delta_y(*)$ is the Dirac measure centered at y, $\{\xi_i\}_{i=1}^N$ are nonnegative and pending to be solved. With such an approximation, correspondingly for $w(x,t|y,s)$

$$w(x,t|y,s) \approx p(x,t|y,s) - \sum_{i=1}^{N} \xi_i p(x,t|L(s)+r_i,s). \tag{2.3.62}$$

This leads to the explicit expression of first passage density from (2.3.14) :

$$\begin{aligned} g(t|y,s) &= -c(t)(p_x(x,t|y,s) - \sum_{i=1}^{N} \xi_i p_x(x,t|L(s)+r_i,s))|_{x=L(t)} \\ &\approx \frac{c(t)e^{2\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\{\int_{-\infty}^{L(s)} \frac{\beta(t,s)+L(t)e^{\alpha(t,s)}-r}{2\gamma(t,s)} \exp(-\frac{(\beta(t,s)+L(t)e^{\alpha(t,s)}-r)^2}{4\gamma(t,s)}) f_y(r)dr \\ &- \sum_{i=1}^{N} \xi_i \frac{\beta(t,s)+L(t)e^{\alpha(t,s)}-L(s)-r_i}{2\gamma(t,s)} \exp(-\frac{(L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)-r_i)^2}{4\gamma(t,s)})\} \end{aligned} \tag{2.3.63}$$

Correspondingly, the first passage distribution $G(t|y,s)$ is given

$$1 - G(t|y,s) = F(L(t),t|y,s) - \sum_{i=1}^{N} \xi_i F(L(t),t|L(s)+r_i,s)). \tag{2.3.64}$$

Here when a mesh $\{t_j\}_{j=1}^M$ is considered for $(0,T]$, $M \geq N$, the remaining task is to determine ξ_i such that the error term for the approximation at mesh points can be minimized. The error term $\epsilon(t)$ by approximating (2.3.20) can be given by

$$\epsilon(t) = h(t) - \sum_{i=1}^{N} \xi_i K(r_i,t) \tag{2.3.65}$$

When the initial value y at s is deterministic, then we have

$$h(t) = p(L(t),t|y,s), K(r,t) = p(L(t),t|L(s)+r,s). \tag{2.3.66}$$

As stated in [95], it is natural to derive a P-weighted linear programming problem which minimize the error term $\epsilon(t)$. Introducing a positive weight M-dimensional vectors P, $W = (\xi_i)_{N\times 1}$, we denote N-dimensional vectors $E = (\epsilon(t_i))_{M\times 1}$, $H = (h(t_i))_{M\times 1}$, $N \times M$ matrix $\tilde{K} = (K(r_i,t_j))_{N\times M}$.

Then the following linear programming problem gives the estimate of ξ_i.

$$\begin{aligned} &\text{Minimize} \quad Z = P^T E, \\ &\text{subject to :} \quad E + W^T \tilde{K} = H, \quad E \geq 0, \quad W \geq 0. \end{aligned} \tag{2.3.67}$$

Other parametric expressions based on different choices of $U(r)$ rather than the point measure can refer to those in [95].

Quasi-Daniels Boundary

A special case in such a parametric approximation is to consider a two-parameter approximation in (2.3.61) leading to a quasi-Daniels boundary [15]. This is given by the following proposition, and also we would like to propose $L(t)$ based on $L(s)$ rather than an explicit function.

Before the discussion, we first introduce the definition of *quasi-Daniels*.

Definition 2.3.11. A boundary $L(t) \in C[s,t]$ is called *quasi-Daniels*, if for $\xi_1, \xi_2 \in \mathbb{R}^+$ and $L(s) > y$,

$$L(t) = e^{-\alpha(t,s)}\{L(s) - \beta(t,s) + \frac{\gamma(t,s)}{y - L(s)} \log(\frac{\xi_1 + \sqrt{\xi_1^2 + 4\xi_2 \exp(-\frac{2(y-L(s))^2}{\gamma(t,s)})}}{2})\}. \quad (2.3.68)$$

where α, β, γ are defined in (1.3.3).

Remark : We call the boundary (2.3.21) quasi-linear boundary because for Brownian motion, this boundary is simplified to a linear boundary. This will be shown later.

Proposition 2.3.12. *For the process X_t in (2.1.1) with the current observation (y,s), when the boundary is quasi-Daniels as given in (2.3.68), then the corresponding first passage density is given by*

$$\begin{aligned} g(t|y,s) = \frac{c(t)e^{\alpha(t,s)}}{2\gamma(t,s)}\{&(\beta(t,s) + L(t)e^{\alpha(t,s)} - y)p(L(t),t|y,s) \\ &- \sum_{i=1}^{2} \xi_i(\beta(t,s) + L(t)e^{\alpha(t,s)} - L(s) - r_i)p(L(t),t|L(s)+r_i,s)\}, \end{aligned} \quad (2.3.69)$$

where $r_1 = L(s) - y, r_2 = 3(L(s) - y)$.

Proof. It is natural to see (2.3.69) from (2.3.63), for $r_1, r_2 > 0$

$$\tilde{U}(r) = \xi_1 \delta(r - r_1) + \xi_2 \delta(r - r_2). \quad (2.3.70)$$

Then we want to solve $L(t)$ explicitly based on (2.3.70) from the Fredholm equation of first kind (2.3.20) where y_0 and $\tilde{\beta}$ are defined, we have

$$\exp(-\frac{(\tilde{\beta}(t,s) - y_0)^2}{4\gamma(t,s)}) - \sum_{i=1}^{2} \xi_i \exp(-\frac{(\tilde{\beta}(t,s) - r_i)^2}{4\gamma(t,s)}) = 0, \quad (2.3.71)$$

This leads to

$$\exp(-\frac{y_0^2}{4\gamma(t,s)}) = \sum_{i=1}^{2} \xi_i \exp(-\frac{(r_i^2 - 2\tilde{\beta}(t,s)(r_i - y_0))}{4\gamma(t,s)}) \quad (2.3.72)$$

When it is chosen that $r_2 = -3y_0, r_1 = -y_0$ where $y_0 = y - L(s) < 0$ with initial value y at time s, then (2.3.72) becomes a quadratic equation for $z := \exp(y_0 \frac{\tilde{\beta}(t,s)}{\gamma(t,s)})$ such that

$$\xi_2 \exp(-\frac{2y_0^2}{\gamma(t,s)}) + \xi_1 z - z^2 = 0. \quad (2.3.73)$$

The solution is given by

$$z = \frac{\xi_1 + \sqrt{\xi_1^2 + 4\xi_2 \exp(-\frac{2y_0^2}{\gamma(t,s)})}}{2}, \tag{2.3.74}$$

where the negative solution is ignored as z is nonnegative. This is due to $\xi_2 \geq 0$.

Furthermore, recalling that $\tilde{\beta}(t,s) = \beta(t,s) - L(s) + L(t)e^{\alpha(t,s)}$, therefore an explicit expression for $L(t)$ is given by :

$$\begin{aligned} L(t) &= e^{-\alpha(t,s)}\{L(s) - \beta(t,s) + \frac{\gamma(t,s)}{y_0}\log(\frac{\xi_1 + \sqrt{\xi_1^2 + 4\xi_2 \exp(-\frac{2y_0^2}{\gamma(t,s)})}}{2})\} \\ &= e^{-\alpha(t,s)}\{L(s) - \beta(t,s) + \frac{\gamma(t,s)}{y - L(s)}\log(\frac{\xi_1 + \sqrt{\xi_1^2 + 4\xi_2 \exp(-\frac{2(y-L(s))^2}{\gamma(t,s)})}}{2})\}. \end{aligned} \tag{2.3.75}$$

□

Proposition 2.3.12 leads directly to the Daniels' boundary for Brownian motion [15]. This is given by the following corollary.

Corollary 2.3.13. *For the Brownian motion $B(t)$ with the initial start (y,s), when the boundary is given by*

$$L(t) = \{L(s) + \frac{t-s}{2(y - L(s))}\log(\frac{\xi_1 + \sqrt{\xi_1^2 + 4\xi_2 \exp(-\frac{4(y-L(s))^2}{t-s})}}{2})\} \tag{2.3.76}$$

where $\xi_1, \xi_2 \in \mathbb{R}^+$ and $L(s) > y$ are adjustable values, then the corresponding first passage density is given by

$$g(t|y,s) = \frac{1}{2(t-s)}\{(L(t) - y)p(L(t),t|y,s) - \sum_{i=1}^{2}\xi_i(L(t) - L(s) - r_i)p(L(t),t|L(s) + r_i, s)\}, \tag{2.3.77}$$

where $r_1 = L(s) - y, r_2 = 3(L(s) - y)$.

Proof. The Brownian motion B_t coincides with Equation (2.1.1) with $a(t) = 0, b(t) = 0, \sigma(t) = 1$, $c(t) = \frac{1}{2}$ and

$$\alpha(t,s) = 0, \quad \beta(t,s) = 0, \quad \gamma(t,s) = \frac{(t-s)}{2}. \tag{2.3.78}$$

Substitute all the above into Proposition 2.3.12, the result comes naturally. □

2.3.6 Piecewise Quasi-Linear Monte-Carlo Method

In Section 2.3.3, we have derived the explicit first passage density and distribution under the quasi-linear boundary. This result will be extended to a piecewise quasi-linear Monte-Carlo method for the first passage time to a general boundary in this subsection. The discussion in this subsection concerns two points :

1. For a piecewise quasi-linear boundary, the corresponding first passage distribution can be given explicitly.

2. For a general boundary, the result of piecewise quasi-linear boundary can be extended by approximating the general boundary by the piecewise boundary. Therefore a piecewise quasi-linear Monte-Carlo method is derived to calculate the first passage distribution.

It is known that for Brownian motion, the piecewise linear boundary can be proposed such that corresponding first passage density is expressed by a closed form [86, 97]. Such an idea will be reproduced here for the OU process, and previously such a generalization is also considered based on time-change [87]. However this is not utilizable for computation as the inverse of time-change is not explicit for the general OU process.

The piecewise boundary is considered based on the idea of conditional first passage time among independent intervals from the Markov property of the process, and this fact is given by the following lemma.

Lemma 2.3.14. *For a Markov process X_t, and a time $\vartheta \in (s,t)$, let $\tau_{y,s}$ denote the first passage time to $L(t)$ based on the current observation (y,s) and let $u(x,t|y,s) := P(X_t < x, \tau_{y,s} > t|X_s = y)$. We have*

$$u(x,t|y,s) = \int_{-\infty}^{L(\vartheta)} u(x,t|z,\vartheta)P(\tau_{y,s} > \vartheta|X_\vartheta = z)p(z,\vartheta|y,s)dz. \tag{2.3.79}$$

Proof. For a time $\vartheta \in (s,t)$, by the Markov property of X_t, we have

$$\begin{aligned}
&P(\tau_{y,s} > t, X_t < x|X_s = y) = P(\{X_\eta < L(\eta), s \le \eta \le t\}\&\{X_t < x\}|X_s = y)\\
&= \int_{-\infty}^{L(\vartheta)} P(\{X_\eta < L(\eta), s \le \eta \le \vartheta|X_\vartheta = z, X_s = y)\\
&\times P(\{X_\eta < L(\eta), \vartheta \le \eta \le t\}\&\{X_t < x\}|X_\vartheta = z, X_s = y)p(z,\vartheta|y,s)dz\\
&= \int_{-\infty}^{L(\vartheta)} P(\tau_{y,s} > \vartheta|X_\vartheta = z)P(\{X_\eta < L(\eta), \vartheta \le \eta \le t\}\&\{X_t < x\}|X_\vartheta = z)p(z,\vartheta|y,s)dz
\end{aligned} \tag{2.3.80}$$

Therefore Equation (2.3.79) comes naturally. □

Remark : There is no requirement for the continuity of $L(t)$ at the chosen time ϑ, therefore this lemma provides the possibility to investigate piecewise continuous boundary.

The lemma leads directly to the following corollary to state a view of $w(x,t|y,s)$.

Corollary 2.3.15. *For a time $\vartheta \in (s,t)$, $w(x,t|y,s) = \frac{\partial u(x,t|y,s)}{\partial x}$*

$$w(x,t|y,s) = \int_{-\infty}^{L(\vartheta)} w(x,t|z,\vartheta)P(\tau_{y,s} > \vartheta|X_\vartheta = z)p(z,\vartheta|y,s)dz. \tag{2.3.81}$$

Proof. Differentiate both sides of Equation (2.3.79) by x, then by Lebesgue's dominated convergence theorem, Equation (2.3.81) comes naturally. □

Iteratively using Equation (2.3.79), we achieve at the following proposition.

Proposition 2.3.16. *Suppose the time mesh is given $\{t_i\}_{i=0}^N$, with $t_0 = s, t_N = T$, and the boundary values at these mesh points $\{L(t_i)\}_{i=0}^N$ are also given. Moreover, the boundary $L(t)$ is chosen such that $L(t) = e^{-\alpha(t,t_i)}(L(t_i) - \beta(t,t_i) + \eta_i\gamma(t,t_i))$ for $t \in (t_i, t_{i+1}]$, with*

$\eta_i = (L(t_{i+1})e^{\alpha(t_{i+1},t_i)} + \beta(t_{i+1},t_i) - L(t_i))/\gamma(t_{i+1},t_i)$. *That is to say,* $L(t)$ *is quasi-linear in each interval* $(t_i, t_{i+1}]$.

Denote the function $\Upsilon_k(t_0,t_1,...,t_k,z_0,z_1,...,z_k)$, $k \geq 1$ *as*

$$\Upsilon_k(t_0,t_1,...,t_k,z_0,z_1,...,z_k) = \Pi_{i=0}^{k-1} I(z_{i+1} < L(t_{i+1}))(1-\exp(-\frac{(z_i - L(t_i))(z_{i+1} - L(t_{i+1}))e^{\alpha(t_{i+1},t_i)}}{\gamma(t_{i+1},t_i)})), \tag{2.3.82}$$

where $I(*)$ *is the indicator function.*

Then for any $t \in (t_k, t_{k+1}]$ *and* $x \leq L(t)$, *we have*

$$u(x,t|y,s) = \mathbb{E}\left[u(x,t|X_{t_k},t_k)\Upsilon_k(s,t_1,...,t_k,y,X_{t_1},...,X_{t_k})\right] \tag{2.3.83}$$

$$= \mathbb{E}\left[\int_{-\infty}^{x}(1-\exp(-\frac{(X_{t_k} - L(t_k))(\eta - L(t))e^{\alpha(t,t_k)}}{\gamma(t,t_k)}))p(\eta,t|X_{t_k},t_k)d\eta\Upsilon_k(s,t_1,...,t_k,y,X_{t_1},...,X_{t_k})\right]$$

Proof. First step : quasi-linear boundary

It is noticed that for an observation z at time ϑ related to the boundary $L(t)$, and $x \leq L(t)$

$$u(x,t|z,\vartheta) = \int_{-\infty}^{x} P(\tau_{y,s} > t|X_t = \eta)p(\eta,t|z,\vartheta)d\eta, \tag{2.3.84}$$

such that

$$w(x,t|z,\vartheta) = P(\tau_{y,s} > t|X_t = x)p(x,t|z,\vartheta). \tag{2.3.85}$$

Therefore when the crossing boundary in the interval $[\vartheta, t]$ satisfies the quasi-linear boundary given by (2.3.21), from (2.3.28), for $x \leq L(t)$ we know that

$$P(\tau_{z,\vartheta} > t|X_t = x) = 1 - \exp(-\frac{(z - L(\vartheta))(x - L(t))e^{\alpha(t,\vartheta)}}{\gamma(t,\vartheta)}). \tag{2.3.86}$$

Second step : piecewise quasi-linear boundary

Now recall X_t is given by (2.1.1), select the mesh $\{t_i\}_{i=1}^N$ with $t_0 = s, t_N = T$, and the boundary in each interval $(t_i, t_{i+1}]$ is chosen as in the proposition, such that the boundary is quasi-linear in each interval $(t_i, t_{i+1}]$.

For $t_0 = s, t_1 > t_0$, it is derived from (2.3.79) that for $t \in (t_k, t_{k+1}]$

$$u(x,t|y,s) = \int_{-\infty}^{L(t_1)} u(x,t|z_1,t_1)P(\tau_{y,s} > t_1|X_{t_1} = z_1)p(z_1,t_1|y,s)dz_1. \tag{2.3.87}$$

Iteratively, do the same transform from (2.3.79) for $u(x,t|z_1,t_1)$ for t_2, ...,t_k, finally we have

$$u(x,t|y,s) = \int_{-\infty}^{L(t_k)} \cdots \int_{-\infty}^{L(t_1)} u(x,t|z_k,t_k)\Pi_{i=0}^{k-1}\left[P(\tau_{z_i,t_i} > t_{i+1}|X_{t_{i+1}} = z_{i+1})p(z_{i+1},t_{i+1}|z_i,t_i)\right]dZ, \tag{2.3.88}$$

where $dZ := \Pi_{i=1}^{k} dz_i$, $z_0 = y$, $t_0 = s$.

Third step : introducing the function of Υ_k

Then from (2.3.86), for $z_{i+1} \in (-\infty, L(t_{i+1}))$, $z_i \in (-\infty, L(t_i))$, we have

$$P(\tau_{z_i,t_i} > t_i|X_{t_{i+1}} = z_{i+1}) = (1 - \exp(-\frac{(z_i - L(t_i))(z_{i+1} - L(t_{i+1}))e^{\alpha(t_{i+1},t_i)}}{\gamma(t_{i+1},t_i)})) \tag{2.3.89}$$

with $z_0 = y,\ t_0 = s$.

If we introduce the function Υ_k in (2.3.82), Equation (2.3.88) can be rewritten as

$$u(x,t|y,s) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{\infty} u(x,t|z_k,t_k)\Upsilon_k(t_0,t_1,...,t_k,z_0,z_1,...,z_k)\Pi_{i=0}^{k-1} p(z_{i+1},t_{i+1}|z_i,t_i)dZ, \tag{2.3.90}$$

where $dZ := \Pi_{i=1}^{k} dz_i,\ z_0 = y,\ t_0 = s$.

Fourth step : explicit expression of $u(x,t|y,s)$

It is further noticed that $\Pi_{i=0}^{k-1} p(z_{i+1},t_{i+1}|z_i,t_i)$ is the joint distribution of $(y, X_{t_1},...,X_{t_k})$ for the OU process. And a calculation for $u(x,t|z_k,t_k)$ from (2.3.84) and (2.3.86) leads to the following equation

$$u(x,t|y,s) = \tag{2.3.91}$$
$$\mathbb{E}\left[\int_{-\infty}^{x} (1 - \exp(-\frac{(X_{t_k} - L(t_k))(\eta - L(t))e^{\alpha(t,t_k)}}{\gamma(t,t_k)}))p(\eta,t|X_{t_k},t_k)d\eta\Upsilon_k(s,t_1,...,t_k,y,X_{t_1},...,X_{t_k})\right].$$

Equation (2.3.83) is derived directly from (2.3.91). □

Remark : It is noticed that there is no requirement for the global continuity of $L(t)$, but just the continuity in each small interval $(t_i, t_{i+1}]$. Here we continue the discussion in the proposition with the hypothesis that the boundary $L(t)$ should be Càdlàg, i.e. right continuous with left limits. Then we can extend naturally the proposition to investigate those piecewise continuous boundary, without any technical difficulty.

It is natural to derive the corresponding expression for $w(x,t|y,s)$ such that

Corollary 2.3.17. *Under the piecewise boundary proposed in Proposition 2.3.16, for* $t \in (t_k, t_{k+1}]$ *and* $x \le L(t)$

$$w(x,t|y,s) = \mathbb{E}\left[w(x,t|X_{t_k},t_k)\Upsilon_k(s,t_1,...,t_k,y,X_{t_1},...,X_{t_k})\right] \tag{2.3.92}$$
$$= \mathbb{E}\left[(1 - \exp(-\frac{(X_{t_k} - L(t_k))(x - L(t))e^{\alpha(t,t_k)}}{\gamma(t,t_k)}))p(x,t|X_{t_k},t_k)\Upsilon_k(s,t_1,...,t_k,y,X_{t_1},...,X_{t_k})\right].$$

Proof. Differentiating both sides of (2.3.83) with respect to x, by Lebesgue's dominated convergence theorem, Equation (2.3.92) comes naturally. □

Corollary 2.3.18. *Under the piecewise boundary proposed in Proposition 2.3.16, for* $t \in (t_k, t_{k+1}]$, *denote* $\bar{G}(t|y,s) = 1 - G(t|y,s)$ *as the survival function.*

$$\bar{G}(t|y,s) = \mathbb{E}\left[\Upsilon_k(s,t_1,...t_k,t,y,X_{t_1},...,X_{t_k},X_t)\right], \tag{2.3.93}$$

Proof. By Equation (2.3.83),

$$\bar{G}(t|y,s) = P(\tau_{y,s} > t) = u(L(t),t|y,s) = \mathbb{E}\left[u(L(t),t|X_{t_k},t_k)\Upsilon_k(s,t_1,...,t_k,y,X_{t_1},X_{t_2},...,X_{t_k})\right] \tag{2.3.94}$$

Further it is noticed from (2.3.86) that

$$u(L(t),t|z_k,t_k) = \int_{-\infty}^{+\infty} I(x < L(t))(1 - \exp(-\frac{(z_k - L(t_k))(x - L(t))e^{\alpha(t,t_k)}}{\gamma(t,t_k)}))p(x,t|z_k,t_k)dx. \tag{2.3.95}$$

From (2.3.90), then we achieve at the expression for $P(\tau_{y,s} > t)$ in (2.3.93). □

Piecewise Quasi-Linear Monte-Carlo Simulation

Following the same procedures in [86], we can extend the discussion by approximating a general boundary $L(t)$ based on the piecewise quasi-linear boundary discussed before. In this thesis we are interested to fulfill the calculation, rigorous consideration is on going, which would be fulfilled later.

Actually we approximate a general $L(t)$ by a piecewise quasi-linear boundary $\tilde{L}(t)$ as follows :

1. Select the mesh $\{t_i\}_{i=1}^N$ with $t_0 = s, t_N = T$, define $\Lambda_N = \max_{i=0,...,N-1}\{t_{i+1} - t_i\}$.
2. For $L(t_i)$ to be the real boundary at t_i, construct the piecewise boundary $\tilde{L}(t)$ between $(t_i, t_{i+1}]$ as $\tilde{L}(t) = e^{-\alpha(t,t_i)}(L(t_i) - \beta(t, t_i) + \eta_i\gamma(t, t_i)), t \in (t_i, t_{i+1}]$ with $\eta_i = (L(t_{i+1})e^{\alpha(t_{i+1},t_i)} + \beta(t_{i+1}, t_i) - L(t_i))/\gamma(t_{i+1}, t_i)$.

Then we can imagine when Λ_N is enough small, the constructed piecewise boundary is near the original boundary also. Therefore when Λ_N is small, the first passage distribution calculated from Equation (2.3.93) can approximate the real first passage distribution accurately. The construction can be directly seen in Figure 2.3.

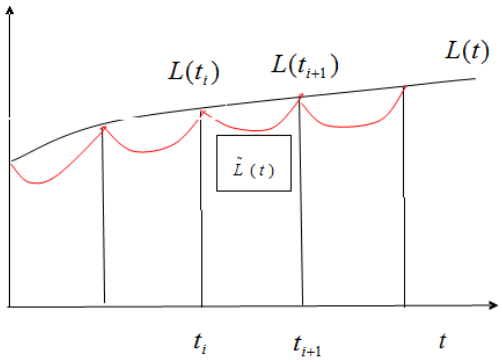


FIGURE 2.3 – Approximating the original boundary (black) by the piecewise quasi-linear boundary (red).

Then by adopting proper numerical algorithms for X_t, e.g. Euler-Maruyama scheme [36], Monte-Carlo simulation can be adopted to produce the estimate of $G(t|y, s)$ based on (2.3.93).

By Euler-Maruyama scheme [36] and the expression of X_t in (3.1.1), one trajectory can be produced for $\{X_i\}_{i=0}^N$, here X_i is the value of X_{t_i}

$$X_0 = x_0, \quad X_{i+1} = (a(t_i)X_i + b(t_i))(t_{i+1} - t_i) + \sigma(t_i)\Delta B_i, \quad i = 0, ..., N-1, \tag{2.3.96}$$

where $\Delta B_i \sim N(0, t_{i+1} - t_i)$.

For a mesh $\{t_i\}_{i=1}^N$, we produce J trajectories $\{z_i^j\}_{j=1}^J$ with J enough large, here z_i^j denote the value of X_{t_i} at the j-th trajectory. An approximate expression is achieved for $G(t|y, s)$ from (2.3.93) :

1. $G(t_0|y, s) = 0$, with $z_0 = y, t_0 = s$.
2. For $k \geq 1$,

$$G(t_k|y, s) \approx 1 - \frac{1}{J}\sum_{j=1}^{J} \Upsilon_k(s, t_1, ..., t_k, y, z_1^j, ..., , z_k^j). \tag{2.3.97}$$

Comparison with a Simple Monte-Carlo Method

We have developed a piecewise quasi-linear Monte-Carlo simulation method to calculate the first passage distribution, in this subsection it is aimed to show its advantages compared to a simple Monte-Carlo simulation method.

Intuitively, using Monte-Carlo simulation methods to produce trajectories of the given process, the first passage time can be statistically described by recording when those trajectories cross the boundary. Here the selected numerical scheme for MC simulation of X_t in (2.1.1) is the Euler-Maruyama scheme [36], which is described by the following iterative algorithm for a time mesh $\{t_i\}_{i=0}^N$:

$$X_{i+1} = X_i + (a(t_i)X_i + b(t_i))(t_{i+1} - t_i) + \sigma(t_i)\Delta B_i,\ \ i = 0, ..., N \tag{2.3.98}$$

where $\Delta B_i \sim \mathcal{N}(0, t_{i+1} - t_i)$.

Therefore, the trajectory of the process X_t can be produced based on (2.3.98). The boundary $L(t)$ is used based on its discretization $L_i = L(t_i)$. The first passage time for the jth simulated path reaching the boundary is τ_j. The MC algorithm aims to produce J trajectories of the deterioration process X_t to investigate statistical description of the corresponding first passage times :

step 1 $j = 0$;

step 2 while $j \leq J$, take $j = j + 1$, $X_0 = x_0$ or produce X_0 from the initial distribution function, and $i = 0$;

step 3 while $X_i < L_i$, calculate X_{i+1} from (2.3.98) and let $i = i + 1$;

step 4 if $X_i \geq L_i$, record t_{i-1} as the first passage time τ_j and go to step 2.

In such a case, we have a lower estimate for the first passage time, however this error can be controlled by reducing the step-size Δt. Later the first passage density can be estimated based on the values $\{\tau_j\}_{j=1}^J$ by methods such as kernel density estimation [64], which is included in R software as the density function to provide the fitted pdf. The first passage distribution can be given by the cumulative values of τ_j at different times directly.

Compared with the simple Monte-Carlo simulation method, the piecewise quasi-linear Monte-Carlo simulation method shows its advantages on the following aspects :

1. In the procedure of producing trajectories, the piecewise quasi-linear Monte-Carlo simulation doesn't need the calculation until the trajectory reaches the boundary. It can return the value of first passage distribution immediately when the trajectories at the current time are produced. This provides the possibility to do parallel computation when producing trajectories.
2. When the time mesh and the number of produced trajectories are fixed, the piecewise quasi-linear Monte-Carlo simulation method is expected to be more accurate than the simple Monte-Carlo simulation.

2.3.7 Durbin's Approximation

It concerns two points in this subsection :

1. An analytical approximation of first passage density is derived from the result of Durbin in [25] for mean-zero Gaussian processes.
2. This approximation of first passage density, interestingly is the same with the result derived from the tangent approximation given in (2.3.58).

For a mean-zero process Y_t with the boundary $L_y(t)$, we denote

$$e(t) := \lim_{z\to t-}(t-z)^{-1}E\left[I(\tau^Y_{y,s} > z)(L_y(z) - Y_z)|Y_t = L_y(t)\right], \tag{2.3.99}$$

$$d(t) := (2\pi \operatorname{var}(Y_t))^{-1/2}\exp(-\frac{1}{2}L_y^2(t)/\operatorname{var}(Y_t)), \tag{2.3.100}$$

where $\tau^Y_{y,s} := \inf_{t\geq s}\{t|Y_t \geq L_y(t)\}$ and $I(\tau^Y_{y,s} > z)$ is the indicator function which equals to 1 if the sample path doses not cross the boundary prior to time z and equals to 0 otherwise.

Lemma 2.3.19. *[25] Suppose the following conditions are satisfied :*

1. *The boundary $L_y(z)$ is continuous for $s \leq z \leq t$, and left-differentiable at t.*
2. *The covariance function $\operatorname{cov}(Y_r, Y_z)$ is positive definitive and has continuous first-order partial derivative on $(r, z) \in [0, t] \times [0, t]$, where appropriate left or right derivatives are taken at the end values of considered intervals.*
3. *The variance of the increment $Y_t - Y_z$ satisfies the condition*

$$\lim_{z\to t}\operatorname{var}(Y_t - Y_z) = \lambda_t, 0 < \lambda_t < \infty. \tag{2.3.101}$$

Then the first passage density $g(t|0, 0)$ of a mean-zero process Y_t to an upper, continuous boundary $L(t)$ satisfies

$$g(t|0, 0) = e(t)d(t). \tag{2.3.102}$$

Proof. see [25] □

Remark : $\lim_{z\to t}\operatorname{var}(Y_t - Y_z) = \lambda_t, 0 < \lambda_t < \infty$ is also equivalent to $\lim_{z\to t}(\frac{\partial \operatorname{cov}(Y_z,Y_t)}{\partial z} - \frac{\partial \operatorname{cov}(Y_z,Y_t)}{\partial t}) = \lambda_t \in (0, +\infty)$, see [25].

From Durbin's Result to the Tangent Approximation

It is hard to calculate $e(t)$ directly except for some special cases, an approximate expression is suggested by Durbin [25] by ignoring the indicator function $I(\tau^Y_{y,s} > z)$ such that

$$g(t|y, s) \approx \lim_{z\to t}\frac{1}{t-z}\mathbb{E}((L_y(z) - Y_z)|Y_t = L_y(t))d(t). \tag{2.3.103}$$

To simplify the calculation, we first consider the conditional expectation of mean-zero Gaussian processes. An important property for the mean-zero Gaussian process Y_t is essential to reach the approximation, and it is given as follows

Lemma 2.3.20. *[25] For a mean-zero Gaussian process, and $L_y(t) \in C^1[0, +\infty)$,*

$$\lim_{z\to t-}(t-z)^{-1}\mathbb{E}\left[(L_y(z) - Y_z)|Y_t = L_y(t)\right] = \frac{L_y(t)}{\operatorname{var}(Y_t)}\frac{\partial \operatorname{cov}(Y_t, Y_z)}{\partial z}|_{z=t-} - L'_y(t). \tag{2.3.104}$$

Proof. As Y_t is Gaussian and the conditional expectation $E(Y_z|Y_t)$ is the projection of Y_z on the space produced by Y_t,

$$\mathbb{E}\left[Y_z|Y_t\right] = \mathbb{E}(Y_z) + \frac{\operatorname{cov}(Y_t, Y_z)}{\operatorname{var}(Y_t)}(Y_t - \mathbb{E}(Y_t)). \tag{2.3.105}$$

Conditioned on $Y_t = L_y(t)$ and due to the mean-zero property of Y_t, Equation (2.3.105) leads to

$$\begin{aligned}
&\lim_{z\to t-}(t-z)^{-1}\mathbb{E}\left[(L_y(z)-Y_z)|Y_t = L_y(t)\right] \\
&= \lim_{z\to t-}(t-z)^{-1}\{\mathbb{E}\left[(L_y(z)-L_y(t))|Y_t = L_y(t)\right] + \mathbb{E}\left[(L_y(t)-Y_z)|Y_t = L_y(t)\right]\} \\
&= -L_y'(t) + \frac{L_y(t)}{\mathrm{var}(Y_t)}\frac{\partial\,\mathrm{cov}(Y_t,Y_z)}{\partial z}|_{z=t-}. \qquad (2.3.106)
\end{aligned}$$

□

Now we turn back to the OU process (2.1.1) to construct a mean-zero Gaussian process. It is direct to see X_t with a deterministic initial value y at time s can be expressed from (2.1.2) that

$$X_t := X_t^{y,s} = e^{-\alpha(t,s)}\left[y - \beta(t,s) + \int_s^t \sigma(u)e^{\alpha(u,s)}dB_u\right], \quad t \geq s. \qquad (2.3.107)$$

Corresponding statistical properties are derived as follows :

$$\begin{aligned}
&\mathbb{E}(X_t) = e^{-\alpha(t,s)}\left(y - \beta(t,s)\right), \mathrm{var}\,(X_t) = e^{-2\alpha(t,s)}\left\{\int_s^t \sigma^2(u)e^{2\alpha(u,s)}du\right\}, \\
&\mathrm{cov}(X_t,X_z) = e^{-(\alpha(t,s)+\alpha(z,s))}\left\{\int_s^{t\wedge z}\sigma^2(u)e^{2\alpha(u,s)}du\right\}, \quad t,z \geq s. \qquad (2.3.108)
\end{aligned}$$

It is known that X_t is Gaussian and with almost surely continuous trajectories, however to fit the mean-zero hypothesis, the following process Y_t with initial value $y_0 := 0$ at s will be discussed in this subsection

$$Y_t := X_t - \mathbb{E}(X_t), \quad t \geq s. \qquad (2.3.109)$$

With Equations in (2.3.108), we have

$$\mathbb{E}(Y_t) = 0, \mathrm{var}(Y_t) = \mathrm{var}(X_t), \mathrm{cov}(Y_t,Y_z) = \mathrm{cov}(X_t,X_z), t,z \geq s. \qquad (2.3.110)$$

The original FPT $\tau_{y,s}$ of X_t to the boundary $L(t)$ is equal to the first passage time τ of Y_t to the boundary

$$L_y(t) := L(t) - \mathbb{E}(X_t). \qquad (2.3.111)$$

Therefore we achieve at an approximate expression for the first passage density for (2.1.1) to an upper boundary $L(t)$ based on Lemmas 2.3.19 and 2.3.20.

Proposition 2.3.21. *Suppose the following conditions are satisfied : the process X_t given by (2.1.1)*

1. *The boundary $L(z)$ is continuous for $s \leq z \leq t$, and left-differentiable at t.*
2. $$\lim_{z\to t}\left(\frac{\partial\,\mathrm{cov}(X_z,X_t)}{\partial z} - \frac{\partial\,\mathrm{cov}(X_z,X_t)}{\partial t}\right) = \lambda_t \in (0,+\infty) \qquad (2.3.112)$$

An approximate expression for the first passage density to the boundary $L(t) \in L^1$ is given by

$$g(t|y,s) \approx p(L(t),t|y,s)\left[a(t)\mathbb{E}(X_t)+b(t)-L'(t)+(L(t)-\mathbb{E}(X_t))(\frac{\sigma^2(t)}{\text{var}(X_t)}+a(t))\right]. \tag{2.3.113}$$

Proof. First, it is noticed that $g(t|y,s) = f_\tau(t)$, where $f_\tau(t)$ is the first passage density of Y_t defined in (2.3.109) to an upper boundary $L_y(t)$ defined in (2.3.109). Also $d(t) = p(L(t),t|y,s)$. And with (2.3.108), $\mathbb{E}(Y_t) = 0$, $\text{var}(Y_t) = \text{var}(X_t)$, $\text{cov}(Y_t, Y_z) = \text{cov}(X_t, X_z)$, $t, z \geq s$. Therefore $\text{cov}(X_t, X_z)$ is differentiable and with (2.3.112), the conditions in Lemma 2.3.19 are satisfied.

Second, as Y_t is a continuous, mean-zero Gaussian process, $f_\tau(t) = e(t)p(L(t),t|y,s)$ from Lemma 2.3.19, with $e(t)$ defined in (2.3.99).

Third, Lemma 2.3.20 leads to the approximate expression of $e(t)$. Recalling (2.3.108), (2.3.109), (2.3.110), it is derived that

$$\frac{\partial \,\text{cov}(Y_z, Y_t)}{\partial z}|_{z=t-} = a(t)\,\text{var}(Y_t) + \sigma^2(t). \tag{2.3.114}$$

And from (2.3.108), (2.3.111) and (2.3.104), we know

$$L'_y(t) = L'(t) - \mathbb{E}'(X_t) = L'(t) - (a(t)\mathbb{E}(X_t) + b(t)). \tag{2.3.115}$$

From (2.3.104),the following equation holds

$$e(t) \approx -L'_y(t) + L_y(t)(\frac{\sigma^2(t)}{\text{var}(Y_t)} + a(t)). \tag{2.3.116}$$

Summarize all the above, we achieve at Equation (2.3.113). □

It is interesting to find the following conclusion.

Proposition 2.3.22. *The Durbin's approximation (2.3.113) is the same with the tangent approximation (2.3.58) for the OU process.*

Proof. Calculate (2.3.113) explicitly with the expressions in (2.3.108), and compare it with the result in (2.3.58), the conclusion comes naturally. □

Remarks :

1. This proposition provides the connection among different approximations of the first passage density. In Durbin's approximation, it ignores the probability that the process have crossed the boundary before the current time. In tangent approximation, it approximates the first passage density based on the local approximation of the boundary by a quasi-linear boundary. From the final result, it is interesting to see that these two ideas are essentially the same.
2. This connection is observed for Brownian motion by Durbin in [26].

Durbin also provides another series' expression of the first passage density $g(t)$ for Brownian motion based on the following integral equation in [27].

Lemma 2.3.23. *For Brownian motion B_t with the first passage boundary $L(t)$, $g(t|0,0)$ can be solved from the following equation :*

$$g(t|0,0) = (\frac{L(t)}{t} - L'(t))p(L(t),t|0,0) - \int_0^t g(u|0,0)\left[\frac{L(t)-L(u)}{t-u} - L'(t)\right] p(L(t),t|L(u),u)du. \tag{2.3.117}$$

Proof. see [27]. □

Remark : In this integral equation, the first item is the tangent approximation / Durbin's approximation. And it is also possible to extend this result to the OU process, which leads to the discussion of a non-singular Volterra integral equation in the next section.

2.4 Nonsingular Volterra Integral Equation of Second Kind

In the last section, we have investigated the first passage problem based on the initial-boundary value problem for the Fokker-Planck equation. This section contributes to the consideration of first passage problems based on integral equations in Section 2.2. And the discussion on such an issue is divided into several steps in this section :

1. The kernel in Fortet's equation in Section 2.2 is singular. See Section 2.4.1.
2. To avoid the singularity, we consider using the sum of Fortet's equation and the master equation in Section 2.2 under some conditions. See Section 2.4.2.
3. The resulted non-singular Volterra integral equation of second kind can be used for numerical calculation. See Section 2.4.3.
4. The Volterra integral equation leads to a series' expression of first passage density under some conditions. The corresponding truncated approximation is presented for the first passage distribution. See Section 2.4.4.

2.4.1 Singularity of Transition Kernel

Following results in [10], we will derive a non-singular Volterra integral equation to numerically solve the first passage density based on (2.2.1).

The first thing we should notice is that $\lim_{s\to t} p(L(t),t|L(s),s) = +\infty$. Actually by L'Hôpital's rule and using notations given in (1.3.3), we have

$$\begin{aligned}&\lim_{t\to s}\frac{(L(t)e^{\alpha(t,s)}+\beta(t,s))-L(s)}{4\gamma(t,s)}\\ &=\lim_{t\to s}\frac{(L'(t)-L(t)a(t)-b(t))e^{\alpha(t,s)}}{4c(t)e^{2\alpha(t,s)}} = \frac{L'(s)-L(s)a(s)-b(s)}{4c(s)}.\end{aligned} \tag{2.4.1}$$

Therefore, we have

$$\lim_{t\to s}\frac{((L(t)e^{\alpha(t,s)}+\beta(t,s))-L(s))^2}{4\gamma(t,s)} = 0, \tag{2.4.2}$$

which contributes to

$$\begin{aligned}\lim_{s\to t} p(L(t),t|L(s),s) &= \lim_{t\to s}\frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\exp(-\frac{((L(t)e^{\alpha(t,s)}+\beta(t,s))-L(s))^2}{4\gamma(t,s)}) \\ &= \lim_{t\to s}\frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}} = +\infty.\end{aligned} \tag{2.4.3}$$

Moreover, the following equation holds :

$$p_x(x,t|L(s),s) = \frac{-2(xe^{\alpha(t,s)}+\beta(t,s)-L(s))e^{2\alpha(t,s)}}{8\sqrt{\pi\gamma^3(t,s)}}\exp(-\frac{((xe^{\alpha(t,s)}+\beta(t,s))-L(s))^2}{4\gamma(t,s)}). \tag{2.4.4}$$

Noticing the equation (2.4.1), we have

$$\begin{aligned}\lim_{s\to t} p_x(L(t),t|L(s),s) &= \lim_{t\to s}\frac{-2(xe^{\alpha(t,s)}+\beta(t,s)-L(s))e^{2\alpha(t,s)}}{8\sqrt{\pi\gamma^3(t,s)}} \\ &= \frac{L'(t)-L(t)a(t)-b(t)}{4c(t)}\lim_{t\to s}\frac{1}{\sqrt{\pi\gamma(t,s)}} = \infty.\end{aligned} \tag{2.4.5}$$

So in the following subsection, how to avoid the singularity for $\lim_{s\to t} p(L(t),t|L(s),s)$ is analyzed to establish a non-singular kernel.

2.4.2 Avoiding the Singularity

Because the kernel $p(L(t),t|L(s),s)$ is singular as stated in last subsection, it could be a difficult point for numerical issues. Therefore in this subsection it is discussed how to derive a non-singular integral equation. The idea comes from the one in [10], where a Volterra integral equation (VIE) of second kind is derived for the FPT densities based on Lemma 2.2.1. It is specially designed such that in some special cases the singularity of transition kernel can be avoided, whose numerical solution therefore can be derived naturally.

The technical discussion is divided into several steps as follows :

1. The master equation (2.2.7) leads to a Volterra integral equation of second kind, see (2.4.6).
2. By introducing a coefficient function $k(t)$, we consider the new integral equation by the sum of (2.4.6) and Fortet's equation with a new kernel $K(x,t|y,s)$ defined in (2.4.9).
3. The kernel function $K(x,t|y,s)$ defined in (2.4.9) can be expressed explicitly in (2.4.12) and (2.4.13).
4. Requiring $\lim_{s\to t} K(L(t),t|L(s),s) = 0$ leads a non-singular kernel $K(x,t|y,s)$ where $k(t)$ is chosen by (2.4.15).

And first we introduce a new integral equation for $g(t|x_0,t_0)$.

Lemma 2.4.1. *Suppose $L(t)\in C[t_0,+\infty)$, $F(x,t|y,s)=\int_{-\infty}^{x} p(z,t|y,s)dz$, x_0 is constrained in $(-\infty,L(t_0))$, then the following equation holds :*

$$g(t|x_0,t_0) = -2\frac{\partial F(L(t),t|x_0,t_0)}{\partial t} + 2\int_{t_0}^{t} g(s|x_0,t_0)\frac{\partial F(L(t),t|L(s),s)}{\partial t}ds. \tag{2.4.6}$$

Proof. Differentiate (2.2.7) with respect to t for $x = L(t)$, then we have :

$$-\frac{\partial F(L(t),t|x_0,t_0)}{\partial t} = g(t|x_0,t_0)(1-\lim_{s\to t} F(L(t),t|L(s),s)) - \int_{t_0}^{t} g(s|x_0,t_0)\frac{\partial F(L(t),t|L(s),s)}{\partial t}ds. \tag{2.4.7}$$

Moreover, with notations in (1.3.3), we know

$$\lim_{s\to t} F(L(t),t|L(s),s) = \lim_{s\to t}\Phi\left(\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)}{\sqrt{2\gamma(t,s)}}\right). \tag{2.4.8}$$

Obviously, $\lim_{s\to t}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)}{\sqrt{2\gamma(t,s)}} = \lim_{s\to t}\sqrt{\gamma(t,s)/2}\frac{L'(t)-a(t)L(t)-b(t)}{c(t)} = 0$. Therefore $\lim_{s\to t} F(L(t),t|L(s),s) = \frac{1}{2}$. Combining with (2.4.7), (2.4.6) holds. □

From all the above, $g(t|x_0,t_0)$ satisfies a Volterra integral equation of second kind as follows :

Proposition 2.4.2. *Suppose $L(t) \in C[t_0,+\infty), k(t) \in C[t_0,+\infty), x_0$ is constrained in $(-\infty, L(t_0))$, setting for all $y \in R$ and $t_0 \le s < t$,*

$$K(L(t),t|y,s) = \frac{\partial F(L(t),t|y,s)}{\partial t} + k(t)p(L(t),t|y,s), \tag{2.4.9}$$

then $g(t|x_0,t_0)$ satisfies :

$$g(t|x_0,t_0) = -2K(L(t),t|x_0,t_0) + 2\int_{t_0}^{t} g(s|x_0,t_0)K(L(t),t|L(s),s)ds. \tag{2.4.10}$$

Proof. From the definition of $K(L(t),t|y,s)$ and Equations (2.4.6),(2.2.1), we know that

$$\begin{aligned}
&-2K(L(t),t|x_0,t_0) + 2\int_{t_0}^{t} g(s|x_0,t_0)K(L(t),t|L(s),s)ds \\
&= -2\frac{\partial F(L(t),t|x_0,t_0)}{\partial t} + 2\int_{t_0}^{t} g(s|x_0,t_0)\frac{\partial F(L(t),t|L(s),s)}{\partial t}ds \\
&- k(t)\left[p(L(t),t|x_0,t_0) - \int_{t_0}^{t} g(s|x_0,t_0)p(L(t),t|L(s),s)ds\right] \\
&= g(t|x_0,t_0).
\end{aligned} \tag{2.4.11}$$

Therefore, (2.4.10) holds. □

The above Volterra integral equation seems numerically efficient with a good iterative structure. However, noticing that $\lim_{s\to t} p(L(t),t|L(s),s) = +\infty$, so the kernel $K(L(t),t|L(s),s)$ is generally singular when $s \to t$. To remove this singularity, we follow the method in [10], and firstly we introduce the following lemma :

Lemma 2.4.3. *Under the condition of theorem 2.4.2, for $t_0 < s < t$, we have :*

$$K(L(t),t|y,s) = p(L(t),t|y,s)H(t,s,y), \tag{2.4.12}$$

where

$$H(t,s,y) = -a(t)L(t) - b(t) - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-y}{2(\gamma(t,s))} + k(t). \tag{2.4.13}$$

Proof. From (2.1.3), we have :

$$\begin{aligned}\frac{\partial F(L(t),t|y,s)}{\partial t} &= \frac{\partial \int_{-\infty}^{z} p(x,t|y,s)dx}{\partial t}|_{z=L(t)} = \int_{-\infty}^{L(t)} \frac{\partial p(x,t|y,s)}{\partial t}dx \\ &= \int_{-\infty}^{L(t)} \left[-\frac{\partial(a(t)x+b(t))p(x,t|y,s)}{\partial x} + c(t)\frac{\partial p(x,t|y,s)}{\partial x^2}\right]dx \\ &= [-(a(t)x+b(t))p(x,t|y,s) + c(t)p_x(x,t|y,s)]_{-\infty}^{L(t)} \\ &= -(a(t)L(t)+b(t))p(L(t),t|y,s) + c(t)p_x(L(t),t|y,s) \\ &= \left[-(a(t)L(t)+b(t)) - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-y}{2\gamma(t,s)}\right]p(L(t),t|y,s).\end{aligned} \tag{2.4.14}$$

From above equations, the lemma obviously holds. □

From the above lemma, we have the following theorem.

Proposition 2.4.4. *If $K(L(t),t|L(s),s)$ is given as in theorem 2.4.2, and $L(t) \in C^2[t_0,+\infty)$, then $\lim_{s\to t} K(L(t),t|L(s),s) = 0$ iff :*

$$k(t) = \frac{a(t)L(t)+b(t)+L'(t)}{2} \tag{2.4.15}$$

Moreover under the above condition,

$$K(L(t),t|y,s) = \left[\frac{L'(t)-a(t)L(t)-b(t)}{2} - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-y}{2(\gamma(t,s))}\right]p(L(t),t|y,s). \tag{2.4.16}$$

Proof. On the one hand, if $K(L(t),t|L(s),s)$ is given by (2.4.16), then by L'Hôpital's rule and (2.4.1), we know that :

$$\lim_{s\to t} c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)}{2(\gamma(t,s))} = \frac{L'(t)-a(t)L(t)-b(t)}{2}. \tag{2.4.17}$$

Moreover, from Lemma (2.4.3), we can deduce that :

$$\begin{aligned}&\lim_{s\to t} K(L(t),t|L(s),s) \\ &= \lim_{s\to t}\left[\frac{L'(t)-a(t)L(t)-b(t)}{2} - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)}{2(\gamma(t,s))}\right]p(L(t),t|L(s),s) \\ &= \frac{1}{\sqrt{4\pi}}\exp(-\frac{L'(t)-a(t)L(t)-b(t)}{2c(t)})\lim_{s\to t}\left[\frac{L'(t)-a(t)L(t)-b(t)}{2\sqrt{r(t,s)}} - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)}{2(\gamma^{\frac{3}{2}}(t,s))}\right]\end{aligned} \tag{2.4.18}$$

From (2.4.18), $\lim_{s\to t} K(L(t),t|L(s),s) = 0$ holds.

On the other hand, it is supposed that $\lim_{s\to t} K(L(t),t|L(s),s) = 0$ holds. Noticing that $\lim_{s\to t} p(L(t),t|L(s),s) = +\infty$, therefore to promise $\lim_{s\to t} K(L(t),t|L(s),s) = 0$, $\lim_{s\to t} H(t,s,L(s))$ must be 0. Moreover from (2.4.17), $k(t) = \frac{a(t)L(t)+b(t)+L'(t)}{2}$ holds. □

We conclude at the end by two interesting propositions.

Proposition 2.4.5. *The first item of the non-singular Volterra integral equation, i.e.* $-2K(L(t),t|y,s)$ *as defined in* (2.4.16)*, is the same with the tangent approximation* (2.3.58) *for the OU process.*

Proof. Calculate (2.3.58) explicitly, it is natural to find this conclusion. □

Remark : The result here can be seen as a parallel consideration with the one in [27].

Proposition 2.4.6. *The boundary is quasi-linear iff* $K(L(t),t|L(s),s) = 0$*, where* K *is defined in* (2.4.16)*.*

Proof. On one hand, if the boundary is quasi-linear as in (2.3.21), then $K(L(t),t|L(s),s) = 0$ comes directly.

On the other hand, if $K(L(t),t|L(s),s) = 0$, we have

$$(L'(t) - a(t)L(t) - b(t))e^{\alpha(t,s)}\gamma(t,s) - c(t)e^{2\alpha(t,s)}(L(t)e^{\alpha(t,s)} + \beta(t,s) - L(s)) = 0. \quad (2.4.19)$$

That is to say,

$$\frac{\partial}{\partial t}\left(\frac{L(t)e^{\alpha(t,s)} + \beta(t,s) - L(s)}{\gamma(t,s)}\right) = 0. \quad (2.4.20)$$

Therefore there exists a constant $C \in \mathbb{R}$ (which may depend on the initial time s), such that

$$\frac{L(t)e^{\alpha(t,s)} + \beta(t,s) - L(s)}{\gamma(t,s)} = C. \quad (2.4.21)$$

This expression coincides with the definition of quasi-linear in (2.3.21). □

2.4.3 Numerical Solutions to Volterra Integral Equations

From all the above, to numerically solve (2.4.10), we here choose the non-singular kernel $K(L(t),t|y,s)$ as in theorem 2.4.4, then corresponding numerical schemes are analyzed in [10]. And in this thesis we don't concern much on numerical analysis rigorously, and the scheme to numerical integrals throughout this thesis is adopted as compound trapezoid rule.

Based on (2.4.10), the following numerical scheme is proposed X_t defined in (2.1.1) with a constant start $x_0 < L(t_0)$. Suppose the time grid for $[t_0,T]$ is given by $\{t_i\}_{i=0}^n$, it is aimed to estimate $g_i = g(t_i), i = 1,2,...,n$. Here $K(x,t|y,s)$ is given by (2.4.16).

$$\begin{cases} g_0 = 0, g_1 = -2K(L(t_1),t_1|x_0,t_0) \\ g_k = -2K(L(t_k),t_k|x_0,t_0) \\ \quad + \sum_{j=0}^{k-1}(t_{j+1}-t_j)(g_jK(L(t_k),t_k|L(t_j),t_j) + g_{j+1}K(L(t_k),t_k|L(t_{j+1}),t_{j+1})), k = 1,2,3... \end{cases} \quad (2.4.22)$$

Corresponding pseudocode is given in Algorithm 1.

Although in previous analysis, we don't distinguish between FPT and randomized FPT, as they don't show difference for analysis. However when numerical issues are discussed, randomized FPT should be treated slightly differently due to unexplicit expression of transition pdf.

Algorithm 1 First passage density estimation : non-singular VIE with constant start

Require: Time grid, $\{t_i\}_{i=0}^n$ for $[t_0, T]$; Transition variable, G ; Failure level $L(t)$; Parameters of the process, $a(t), b(t), \sigma(t), x_0$; Related functions, $\alpha(t,s), \beta(t,s), \gamma(t,s)$; Transition pdf, $p(x,t|y,s)$; Non-singular kernel, $K(x,t|y,s)$;

Ensure: First passage density at time t_i, $\{g_i\}_{i=0}^n$;

$g_0 = 0,\ g_1 = -2K(L(t_1), t_1|x_0, t_0)$;
for $k = 2, k < n, k++$ **do**
  $G = -2K(L(t_k), t_k|x_0, t_0)$;
  for $j = 1, j < k, j++$ **do**
    $G = G + 2g_j \times (t_j - t_{j-1}) \times K(L(t_k), t_k|L(t_j), t_j)$;
  end for
  $g_k = G$;
end for
return $\{g_i\}_{i=0}^n$;

Based on (2.1.5) and (2.1.7), we truncate the pdf of x_0 on $[l_b, L(t_0)]$. l_b can be set to enough small to achieve required accuracy. For instance, for $x_0 \sim \mathbb{N}(\mu, \sigma^2)$, it would be enough to set $l_b = 6\sigma$. Then given a time grid $\{s_i\}_{i=0}^q$ for $[l_b, L(t_0)]$, we can approximate $K(x,t|x_0,t_0)$ in (2.4.16) as

$$K_q(x,t|x_0,t_0) := \frac{1}{2}\sum_{i=0}^{q-1}(s_{i+1}-s_i)(K(x,t|s_{i+1},t_0)f_0(s_{i+1}) + K(x,t|s_i,t_0)f_0(s_i)). \quad (2.4.23)$$

Remarks : More rigorous consideration should be discussed based on other tools, e.g. optimal quantization [50].

Therefore when the time grid for $[t_0, T]$ is given by $\{t_i\}_{i=0}^n$, $g_i = g(t_i), i = 1, 2, ..., n$ is estimated still based on (2.4.16) with (2.4.23).

$$\begin{cases} g_0 = 0, g_1 = -2K_q(L(t_1), t_1|x_0, t_0) \\ g_k = -2K_q(L(t_k), t_k|x_0, t_0) \\ \quad + \sum_{j=0}^{k-1}(t_{j+1}-t_j)(g_j K(L(t_k), t_k|L(t_j), t_j) + g_{j+1}K(L(t_k), t_k|L(t_{j+1}), t_{j+1})), k = 1, 2, 3... \end{cases} \quad (2.4.24)$$

Corresponding pseudocode is given in Algorithm 2.

2.4.4 Truncated Approximation from Series' Expression

In this subsection, several series' solutions will be given based on previous integral equations. And the discussion here is devoted to :

1. The master equation (2.2.7) leads to a Volterra integral equation, whose distribution function can be solved by a series' solution as in (2.4.29).

Moreover, following classical results for Volterra integral equation, a series solution is given, and the proof can be divided into several steps :

1. It is shown that the first passage distribution $G(t|x_0, t_0)$ satisfies a self-mapping related to an integral operator.

Algorithm 2 First passage density estimation : non-singular VIE with random start

Require: Time grid, $\{s_i\}_{i=0}^{q}$ for $[l_b, L(t_0)]$,$\{t_i\}_{i=0}^{n}$ for $[t_0, T]$; Transition variables, E, G ; Parameters of the process, $a(t), b(t), \sigma(t), f_0(t)$; Related functions, $\alpha(t,s), \beta(t,s), \gamma(t,s)$; Transition pdf, $p(x,t|y,s)$; Non-singular kernel, $K(x,t|y,s)$; Failure level $L(t)$;

Ensure: First passage density at time t_i, $\{g_i\}_{i=0}^{n}$;

$g_0 = 0, E = 0$;
for $i = 0, i < q, i++$ **do**
  $E = E + (s_{i+1} - s_i)(K(L_{t_1}, t_1|s_{i+1}, t_0)f_0(s_{i+1}) + K(L(t_1), t_1|s_i, t_0)f_0(s_i))$;
end for
$g_1 = -E$;
for $k = 2, k < n, k++$ **do**
  $E = 0$;
  for $i = 0, i < q, i++$ **do**
    $E = E + (s_{i+1} - s_i)(K(L(t_k), t_k|s_{i+1}, t_0)f_0(s_{i+1}) + K(L(t_k), t_k|s_i, t_0)f_0(s_i))$;
  end for
  $G = -E$;
  for $j = 1, j < k, j++$ **do**
    $G = G + 2g_j \times (t_j - t_{j-1}) \times K(L(t_k), t_k|L(t_j), t_j)$;
  end for
  $g_k = G$;
end for
return $\{g_i\}_{i=0}^{n}$;

2. Under the condition given by (2.4.27) and (2.4.28), the mapping is proved to be a contraction.
3. By the fixed-point theorem, the solution for $G(t|x_0, t_0)$ can be expressed by a series' form which is given in (2.4.29). This could be a way to give an approximate solution based on the truncated form, which is discussed also in [27].
4. To keep the simplicity of calculation and also high accuracy, we combine the one-term truncation with the previously derived tangent approximation in Proposition 2.3.9.

Derivation

Denote the first passage distribution function as $G(t|y,s) = P(\tau_{y,s} \leq t)$, it leads to the following Voterra integral equation of second kind.

Lemma 2.4.7. *Suppose $L(t) \in C[t_0, +\infty)$, $F(x,t|y,s) = \int_{-\infty}^{x} p(z,t|y,s)dz$, $f_0(x)$ is defined in $(-\infty, L(t_0))$, then the following equation holds :*

$$G(t|x_0,t_0) = 2(1 - F(L(t),t|x_0,t_0)) - 2\int_{t_0}^{t} G(s|x_0,t_0)\frac{dF(L(t),t|L(s),s)}{ds}ds. \tag{2.4.25}$$

Proof. By partial integral, for (2.2.7) with $x = L(t)$, it can be expressed by :

$$1-F(L(t),t|x_0,t_0) = G(t|x_0,t_0)(1-\lim_{s\to t} F(L(t),t|L(s),s)) + \int_{t_0}^{t} G(s|x_0,t_0)\frac{dF(L(t),t|L(s),s)}{ds}ds. \tag{2.4.26}$$

Moreover, as $\lim_{s\to t} F(L(t),t|L(s),s) = \frac{1}{2}$, (2.4.25) comes naturally from (2.4.26). □

Proposition 2.4.8. *For the OU process X_t with first passage boundary $L(t) \in C^1[t_0, +\infty)$, $L(t_0) > x_0$ satisfying for any $s \in [t_0, t]$, a constant $\kappa \in \mathbb{R}$,*

$$L(t)e^{\alpha(t,t_0)} - L(t_0) + \beta(t, t_0) < \kappa\sqrt{2\gamma(t, t_0)}. \tag{2.4.27}$$

and also

$$\frac{d\bar{F}(L(t), t|L(s), s)}{ds} \geq 0, \tag{2.4.28}$$

where $\bar{F}(x, t|y, s) = 1 - F(x, t|y, s)$.

Let $G(t|x_0, t_0)$ denote the distribution function of the FPT $\tau := \inf_{t>0}\{X_t \geq L(t)|x_0, t_0\}$. Then we have

$$G(t|x_0, t_0) = h(t) + \sum_{n=1}^{\infty}(\int_{t_0}^{t} K_n(t, s)h(s)ds) \tag{2.4.29}$$

where the series converges uniformly over all $t \geq t_0$ and

$$\begin{aligned} h(t) &= 2\bar{F}(L(t), t|x_0, t_0), \\ K_1(t, s) &= 2\frac{d\bar{F}(L(t), t|L(s), s)}{ds}, \\ K_{n+1}(t, s) &= \int_s^t K_1(t, r)K_n(r, s)dr, \end{aligned} \tag{2.4.30}$$

for $0 \leq s < t$ and $n \geq 1$.

Proof. The proof is given by using fixed point theorem [57].

First, from Equation (2.4.25), by notations in (2.4.30), it comes to

$$G(t|x_0, t_0) = h(t) + \int_{t_0}^{t} G(s|x_0, t_0)K_1(t, s)ds \tag{2.4.31}$$

Define a mapping T on $B(\mathbb{R}_+)$ by setting

$$(T(G))(t) = h(t) + \int_{t_0}^{t} G(s)K_1(t, s)ds \tag{2.4.32}$$

for $G \in B(\mathbb{R}_+)$. Here $B(*)$ is Banach space. Then Equation (2.4.31) turns to

$$T(G(t|x_0, t_0)) = G(t|x_0, t_0), \tag{2.4.33}$$

Therefore it remains to solve (2.4.33) in $B(\mathbb{R}_+)$. To verify the fixed-point theorem, it needs to show T is a contraction from $B(\mathbb{R}_+)$ to $B(\mathbb{R}_+)$ under the sup norm $\|h\|_\infty := \sup_{t\geq t_0} |h(t)|$.

Second, we have

$$\begin{aligned} \|T(G_1) - T(G_2)\|_\infty &= \sup_{t\geq t_0} |T(G_1 - G_2)(t)| \\ &= \sup_{t\geq t_0} |\int_{t_0}^{t} (G_1(s) - G_2(s))K_1(t, s)ds| \\ &\leq \sup_{t\geq t_0}(\int_{t_0}^{t} |K_1(t, s)|ds) \, \|G_1 - G_2\|_\infty \end{aligned} \tag{2.4.34}$$

So from (2.4.27),

$$\begin{aligned}\sup_{t\geq t_0}\Big(\int_{t_0}^{t}|K_1(t,s)|ds\Big) &\leq 2\int_{t_0}^{t}\frac{d\bar{F}(L(t),t|L(s),s)}{ds}ds\\ &=\sup_{t\geq t_0}2\Big(\frac{1}{2}-\bar{F}(L(t),t|L(t_0),t_0)\Big)\\ &=\sup_{t\geq t_0}2\Big(\Phi\Big(\frac{L(t)e^{\alpha(t,t_0)}+\beta(t,t_0)-L(t_0)}{\sqrt{2\gamma(t,t_0)}}\Big)-\frac{1}{2}\Big)\leq 2\Big(\Phi(\kappa)-\frac{1}{2}\Big)<1\end{aligned} \tag{2.4.35}$$

So summarize all the above, T is a contraction from the Banach space $B(\mathbb{R}_+)$ to itself, and by the fixed-point theorem there exists a unique solution $g(t)$ in $B(\mathbb{R}_+)$ satisfying (2.4.32) which is the desired first passage density $g(t|x_0,t_0)$.

Third, the series solution in (2.4.29) can be derived directly from (2.4.32) and the well known formula for the resolvent of the integral operator $K=T-h$ with the kernel K_1 :

$$(I-K)^{-1}=\sum_{n=0}^{\infty}K^n. \tag{2.4.36}$$

This completes the proof.

□

Truncated Approximation

This series' solution leads directly to the approximation of first passage distribution. For example, truncated at the first item, it returns

$$G(t|x_0,t_0)\approx 2\bar{F}(L(t),t|x_0,t_0)+4\int_{t_0}^{t}\bar{F}(L(s),s|x_0,t_0)d\bar{F}(L(t),t|L(s),s). \tag{2.4.37}$$

This one-term truncated approximation can hardly return an accurate result, and it would be very difficult to calculate the truncation with many items. As this involves multi-integral, the calculation based on discretized schemes could introduce new errors which increase rapidly as the truncation is with more items. Therefore to promise the accuracy, the calculation increase rapidly to compensate the errors.

Noticing the operator $T(G)$ in (2.4.32) is a contraction, therefore if the starting point is near the real solution, even the one-term truncation approximation can also return an accurate result. Therefore our natural idea comes to use a previously derived tangent approximation in Proposition 2.3.9 as the starting point :

$$H(t):=\Phi\Big(\frac{-\tilde{L}(t_0)-\tilde{C}(t,t_0)\gamma(t,t_0)+x_0}{\sqrt{2\gamma(t,t_0)}}\Big)+e^{\tilde{C}(t,t_0)(x_0-L(t_0))}\Phi\Big(\frac{\tilde{C}(t,t_0)\gamma(t,t_0)-\tilde{L}(t_0)+x_0}{\sqrt{2\gamma(t,t_0)}}\Big). \tag{2.4.38}$$

Then the first passage distribution is approximated by

$$G(t|x_0,t_0)\approx 2\bar{F}(L(t),t|x_0,t_0)+2\int_{t_0}^{t}H(s)d\bar{F}(L(t),t|L(s),s). \tag{2.4.39}$$

To calculate the value of $G(t|x_0,t_0)$, the iterative steps can be applied like the procedures in Algorithm 1 for the non-singular VIE. For a mesh $\{t_i\}_{i=0}^{N}$,

TABLE 2.1 – Boundary requirements of different methods

Method	piecewise continuity	continuity	differentiablity
Simple MC	not required	not required	not required
Piecewise MC	required	not required	not required
Non-singular VIE	required	required	required
Tangent approximations	required	required	required
Linear programming	required	required	not required
Truncated approximation	required	required	required

1. $G(t_0|x_0, t_0) = 0$;
2. Using the right-point rectangle scheme,

$$G(t_i|x_0, t_0) \approx 2\bar{F}(L(t_i), t_i|x_0, t_0) + 2\sum_{j=1}^{i} H(t_j)(\bar{F}(L(t_i), t_i|L(t_j), t_j) - \bar{F}(L(t_i), t_i|L(t_{j-1}), t_{j-1})). \tag{2.4.40}$$

Remark : The approximation (2.4.39) is not easy to be calculated, as an integral is involved which needs discretization in the calculation. But we should notice even the calculation is tedious, it is still an explicit expression with only the information on the boundary $L(t)$. Therefore it could be expected as an approximation when the boundary $L(t)$ is pending, such as in the case of maintenance optimization which will be discussed in Chapter 4. Such an approximation is expected to compensate the accuracy of purely tangent approximation when the tangent approximation doesn't perform well.

2.5 Simulation Tests

In previous sections, we have investigated the first passage problems from different angles with corresponding solving methods, and in general they can be divided into 3 categories : Monte-Carlo methods, numerical algorithms and analytical approximation. These methods can be distinguished in a theoretical view based on the classification of considering boundaries :

1. Differentiable boundary. All methods proposed in this chapter are utilizable for those differentiable boundaries.
2. Continuous boundary. When the global differentiability of the boundary cannot be promised, those methods involving the derivative cannot be used directly.
3. Piecewise continuous boundary. When the boundary is only piecewise continuous, only Monte-Carlo simulation methods are utilizable, or other methods need tedious transformations.

These facts are obvious from those hypotheses used for the derivation of different methods, which are summarized in Table 2.1 [2].

2. Here simple MC and piecewise quasi-linear MC are presented in Section 2.3.6. Non-singular VIE is given in Section 2.4.3. Tangent approximations are given in Section 2.3.3, including Durbin's approximation in Section 2.3.7. Linear programming is given in Section 2.3.5. Truncated approximation is given in Section 2.4.4.

However, even theoretically distinguishing proposed methods could be interesting, the following simulation tests are still essential to consider the calculation correctness, accuracy and also efficiency.

Therefore the simulation tests in this section will be organized based on this consideration, and the discussion is divided specifically into the following points :

1. The accuracy and efficiency of different methods are compared in the case of a drifted Brownian motion.
2. The accuracy and efficiency of different methods are compared in the case of a time-changed Brownian motion.
3. Different methods are tested for a general OU process, where results are compared based on the one calculated from the Volterra integral equation method.
4. Some cases are presented where the proposed methods have bad performance.

The simulation tests are all done in a laptop with a computing core i7-3520M, 12G memory, where the software environment is the R software with the operating system of Windows 7.

2.5.1 Drifted Brownian Motion

In this subsection, proposed methods are tested for the drifted Brownian motion, and it concerns the following points :

1. The influence of step sizes for the non-singular VIE and the truncated approximation is discussed from the view of accuracy and efficiency.
2. The influence of number of produced trajectories and time step-sizes for Monte-Carlo simulation methods is discussed from the view of accuracy and efficiency.
3. The influence of chosen point-measures and also fitting times for the linear programming approach is discussed from the view of accuracy and efficiency.

Experimental Setup

The first passage problems can hardly be derived analytically except for some special cases, so to verify previously stated methods, we first consider to compare different methods with the analytical solutions of the first passage problem for the drifted Brownian motion $X_t = \mu t + \sigma dB_t$, $\mu, \sigma \in \mathbb{R}$. The drifted Brownian motion is also a well-known model to describe degradation processes with fluctuations, and it can be found in many applications in PHM [66] and maintenance [44]. It is also a special case of the OU process $dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t$, with corresponding coefficients to be $a(t) = 0, b(t) = \mu, \sigma(t) = \sigma$.

Suppose $X_t = t + B_t$ is considered with the failure level as $L = 10$. In such a case, the first passage density satisfies the inverse Gaussian distribution in Corollary 2.3.6, which is given explicitly by :

$$g(t) = \frac{L \exp(-(t-L)^2/(2t))}{\sqrt{2\pi t^3}}. \tag{2.5.1}$$

Corresponding first passage distribution is given by

$$G(t) = \Phi(\frac{-L+t}{\sqrt{t}}) + e^{2L}\Phi(\frac{-t-L}{\sqrt{t}}) \tag{2.5.2}$$

The pdf and cdf functions are shown in Figure 2.4(a) and Figure 2.4(b) respectively.

Moreover, as for the drifted Brownian motion, it is noticed that in this setup, the tangent approximations / Durbin's approximation degenerate to the exact result, such that these two trivial issues will be ignored in the following simulation tests. So we will mainly discuss numerical solutions to integral equations, Monte-Carlo simulation methods and the linear programming approach.

Integral Equation Methods

Supposing the step-size $\Delta t > 0$ is chosen and assuming $t_0 = 0$, denote $t_n = n\Delta t$ for $n = 0, 1, 2, ...$, then we can fulfill the numerical algorithm based on the non-singular Volterra integral equation in Section 2.4.3 and also the truncated approximation in Section 2.4.4. Here we want to investigate the influence of time step-size Δt for these two methods, on the estimation accuracy and the calculation time.

Moreover as only the first passage density $\{g_i\}_{i=0}^N$ is given based on the VIE method, it is of interest to produce further corresponding cdf $\{G_i\}_{i=0}^N$ from the iterative steps based on the right-point formula :

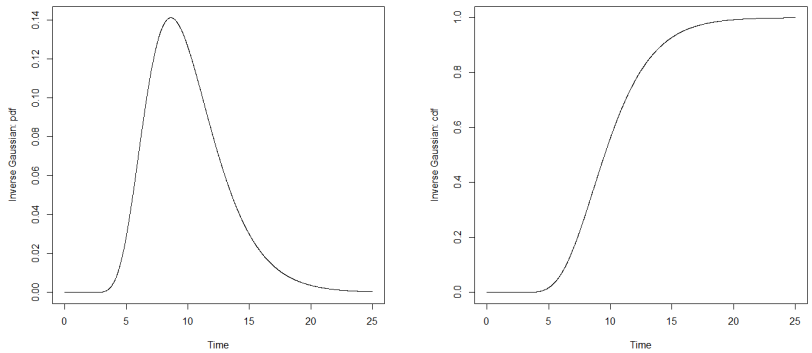
1. $G_0 = 0$,
2. $G_{i+1} = G_i + \Delta t g_{i+1}, i = 0, 1, ..., N-1$.

The calculation time of the methods with three different step-size $0.05, 0.02, 0.01$, which are given in Table 2.2. Corresponding errors of estimated first passage distribution are shown in 2.5.

Summarize the results, we can conclude

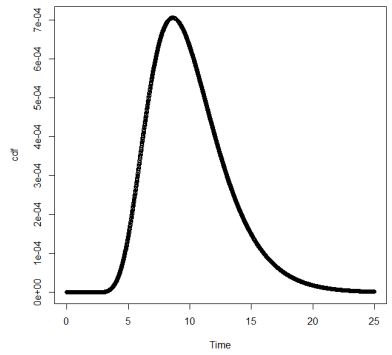
1. The non-singular VIE and truncated approximation have the same level of calculation time and also accuracy with the same step-size.
2. When the step-size is smaller, the estimate is more accurate.

3. The time is given based on the calculation on all the time mesh points.

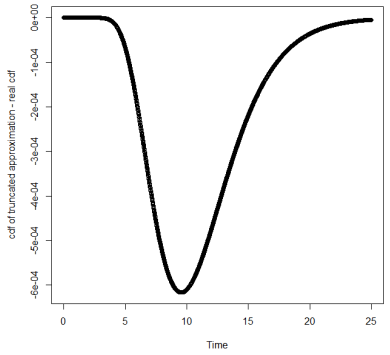


(a) Inverse Gaussian : pdf (b) Inverse Gaussian : cdf

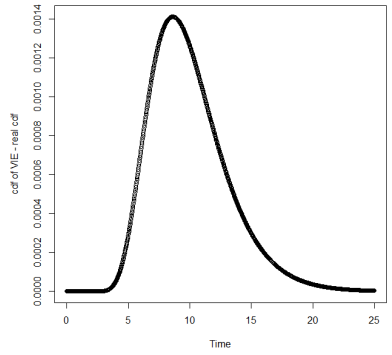
FIGURE 2.4 – The pdf and cdf of the first passage time of $X_t = t + B_t$ with $L = 10$.



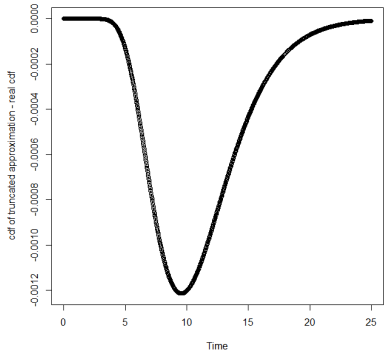
(a) Errors of the VIE method on cdf, step-size $\Delta t = 0.01$, maximal deviation $\approx 7e-04$.



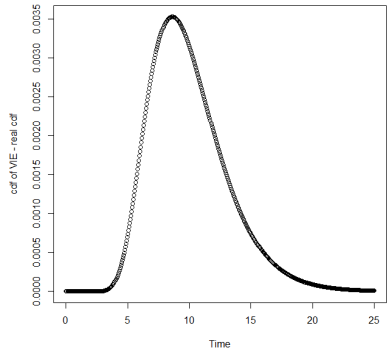
(b) Errors of the truncated approximation on cdf, step-size $\Delta t = 0.02$, maximal deviation $\approx -6e-04$.



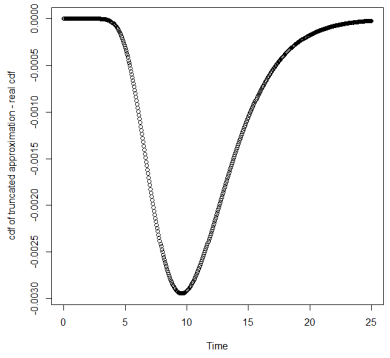
(c) Errors of the VIE method on cdf, step-size $\Delta t = 0.02$, maximal deviation ≈ 0.00014.



(d) Errors of the truncated approximation on cdf, step-size $\Delta t = 0.02$, maximal deviation ≈ 0.00012.



(e) Errors of the VIE method on cdf, step-size $\Delta t = 0.05$, maximal deviation ≈ 0.00035.



(f) Errors of the truncated approximation on cdf, step-size $\Delta t = 0.05$, maximal deviation ≈ 0.00030.

FIGURE 2.5 – Errors of different methods compared with the real first passage density for the drifted Motion $X_t = t + B_t$.

TABLE 2.2 – Calculation time of different methods with different time step-sizes

Method	$\Delta t = 0.05$	$\Delta t = 0.02$	$\Delta t = 0.01$
Non-singular VIE	12.97374 secs	36.13607 secs	6.167119 mins
Truncated approximation	31.25279 secs	2.337634 mins	7.941354 mins
Exact solution [3]	0.7410431 secs	0.980056 secs	1.020058 secs

TABLE 2.3 – Calculation time of Monte-Carlo methods in different situations

Method	$\Delta t = 0.05$	$\Delta t = 0.02$	$\Delta t = 0.01$
Simple MC (200 trajectories)	4.971285 secs	3.578205 secs	7.245414 secs
Piecewise MC (200 trajectories)	9.364536 secs	15.34588 secs	26.2165 secs
Simple MC (2000 trajectories)	14.55883 secs	22.48129 secs	38.01617 secs
Piecewise MC (2000 trajectories)	32.24885 secs	1.184968 mins	1.927094 mins
Simple MC (20000 trajectories)	1.34541 mins	4.31843 mins	7.133408 mins
Piecewise MC (20000 trajectories)	3.90794 mins	8.581241 mins	16.3463 mins

Monte-Carlo Simulation Methods

Supposing the step-size $\Delta t > 0$ is chosen and assuming $t_0 = 0$, denote $t_n = n\Delta t$ for $n = 0, 1, 2, ...$, $X_n = X_{t_n}$ and $X_0 = x_0$. Constrained in the consideration of the drifted Brownian motion $t + B_t$, the Euler-Maruyama scheme for Monte-Carlo simulation is updated to

$$X_{n+1} = X_n + \Delta t + \Delta B_t, \tag{2.5.3}$$

where $\Delta B_t \sim \mathcal{N}(0, \Delta t)$.

Here the calculation time of the methods with three different step-size $0.05, 0.02, 0.01$ and also with three different numbers of trajectories $200, 2000, 20000$, which are given in Table 2.3 [4]. Corresponding errors are shown respectively in Figure 2.6, 2.7 and 2.8.

Summarize the results, it is concluded that

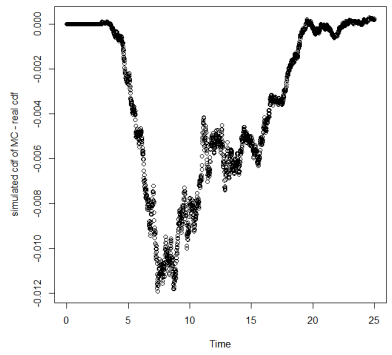
1. The accuracy is improved by producing more trajectories with the same step-size for Monte-Carlo simulation methods.

2. When the step-size is smaller, the estimate is more accurate.

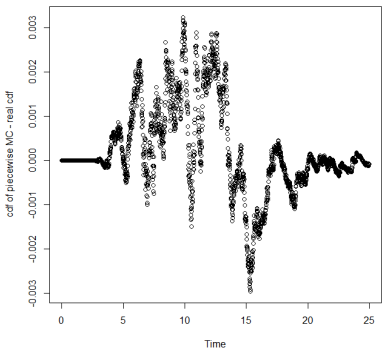
3. The piecewise quasi-linear MC method is more accurate than the simple MC method, with the same number of trajectories and the same step-size. But it also costs more time to do extra calculation, however in general the piecewise quasi-linear MC method costs the same level of time with the simple MC method.

4. When enough trajectories are produced (20000 trajetories), the Monte-Carlo simulation methods is still less accurate than the integral equation methods with the same step-size.

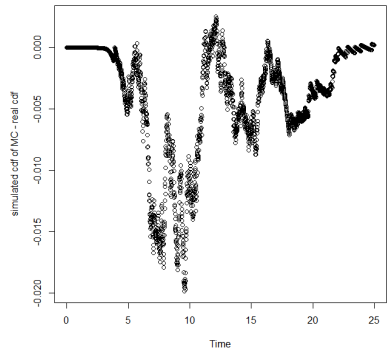
4. Here the calculation time and errors are given based on trajectories produced in one test, and it may vary due to different simulation tests. This influence would be slight when enough trajectories are produced



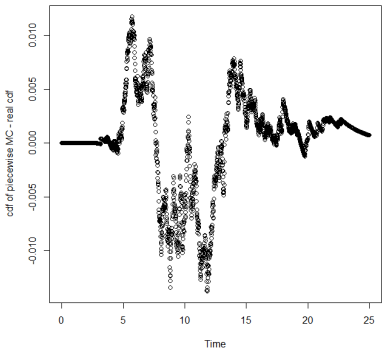
(a) Errors of simple MC, 20000 trajectories, , maximal deviation ≈ 0.012



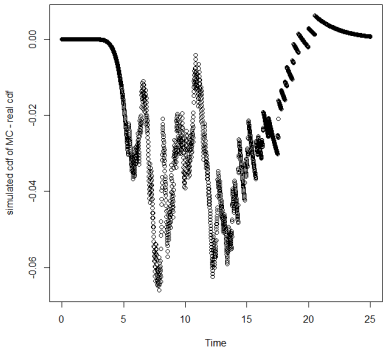
(b) Errors of piecewise MC, 20000 trajectories, maximal deviation ≈ 0.003



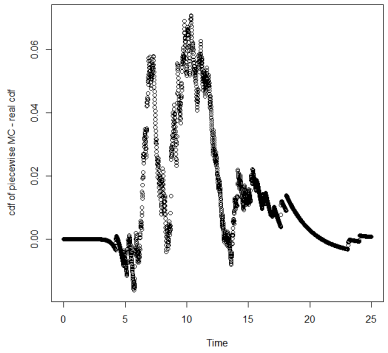
(c) Errors of simple MC, 2000 trajectories, maximal deviation ≈ 0.02



(d) Errors of piecewise MC, 2000 trajectories, maximal deviation ≈ 0.01

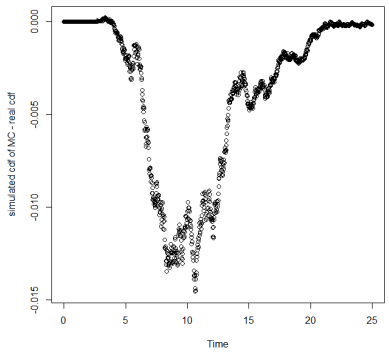


(e) Errors of simple MC, 200 trajectories, maximal deviation ≈ 0.06

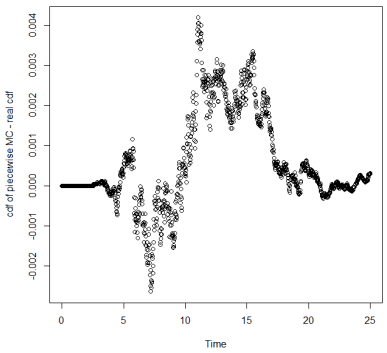


(f) Errors of piecewise MC, 200 trajectories, maximal deviation ≈ 0.06

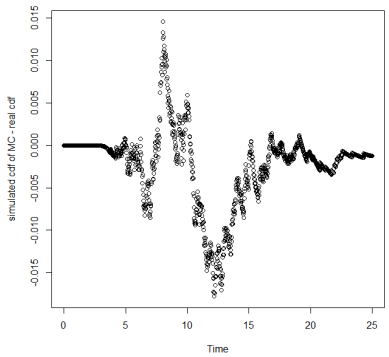
FIGURE 2.6 – Errors of Monte-Carlo simulation methods on cdf for the drifted Motion $X_t = t + B_t$, step-size $\Delta t = 0.01$



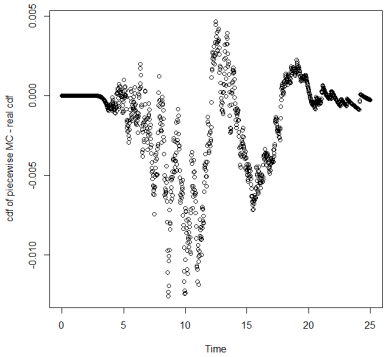
(a) Errors of simple MC, 20000 trajectories, maximal deviation ≈ 0.015



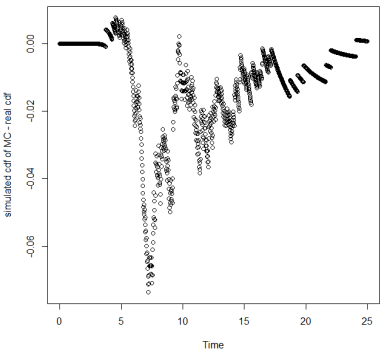
(b) Errors of piecewise MC, 20000 trajectories, maximal deviation ≈ 0.004



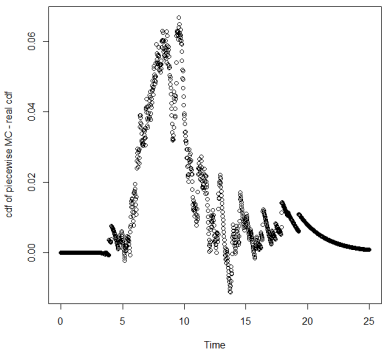
(c) Errors of simple MC, 2000 trajectories, maximal deviation ≈ 0.015



(d) Errors of piecewise MC, 2000 trajectories, maximal deviation ≈ 0.005

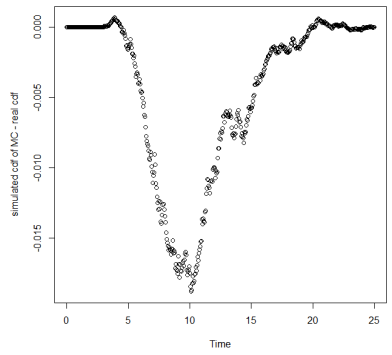


(e) Errors of simple MC, 200 trajectories, maximal deviation ≈ 0.07

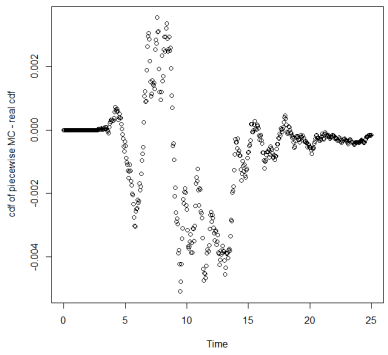


(f) Errors of piecewise MC, 200 trajectories, maximal deviation ≈ 0.06

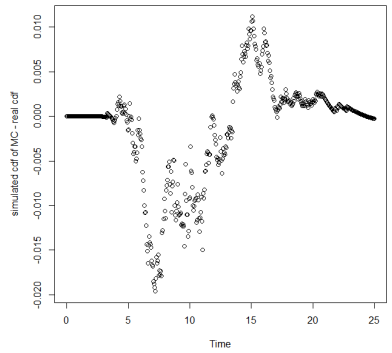
FIGURE 2.7 – Errors of Monte-Carlo simulation methods on cdf for the drifted Motion $X_t = t + B_t$, step-size $\Delta t = 0.02$



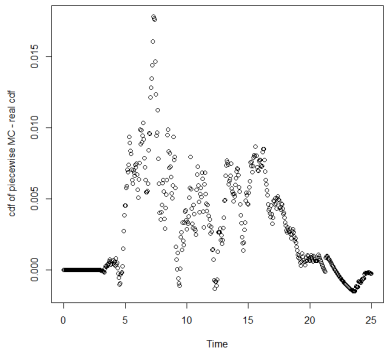
(a) Errors of simple MC, 20000 trajectories, maximal deviation ≈ 0.02



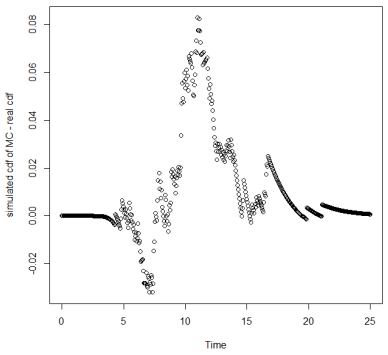
(b) Errors of piecewise MC, 20000 trajectories, maximal deviation ≈ 0.005



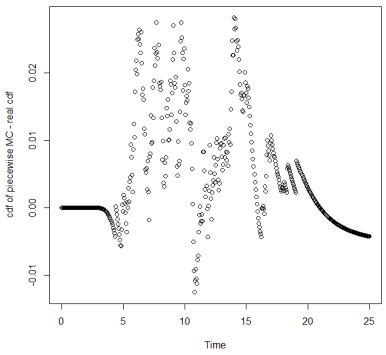
(c) Errors of simple MC, 2000 trajectories, maximal deviation ≈ 0.02



(d) Errors of piecewise MC, 2000 trajectories, maximal deviation ≈ 0.015



(e) Errors of simple MC, 200 trajectories, maximal deviation ≈ 0.08



(f) Errors of piecewise MC, 200 trajectories, maximal deviation ≈ 0.02

FIGURE 2.8 – Errors of Monte-Carlo simulation methods on cdf for the drifted Motion $X_t = t + B_t$, step-size $\Delta t = 0.05$

Linear Programming Approach

Recall the method of linear programming proposed in Section 2.3.5. An approximation for the first passage density is proposed based on

$$p(L(t),t|y,s)-\sum_{i=1}^{N}\xi_i p(L(t),t|L(s)+r_i,s)\approx 0, \tag{2.5.4}$$

for some selected points $\{r_i\}_{i=1}^n$ to estimate $\{\xi_i\}_{i=1}^n$ in (2.3.67). In such a case, the first passage density is given in (2.3.63), and corresponding first passage distribution is given in (2.3.64). Later the estimation of $\{\xi_i\}_{i=1}^n$ is to promise (2.5.4) at several times $\{t_i\}$ such that the error can be minimized on these times.

From previous statements, numerical algorithms involve the choice of time step-sizes, Monte-Carlo simulation methods involve not only the time step-size, but also those produced trajectories. However things are totally different when it comes to the linear programming approach. Actually, from the derivation of the method, it is found that this method is influenced more by the choice of point measures and the fitting times.

Therefore the time mesh is chosen as $t_i = i\Delta t, i \in \mathbb{N}$ in the interval $[0, 25]$, with $\Delta t = 0.05$, the following issues are concerned here

1. The different performance when the linear programming approach is solved for those times in a smaller interval.
2. The different performance when the point measures vary.

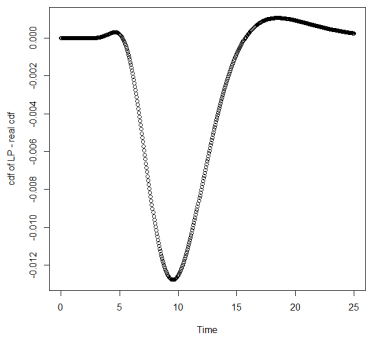
The calculation times of the linear programming approach in different situations are given in Table 2.4. The objective values of the linear programming approach in different situations are given in Table 2.5. Corresponding errors of estimated first passage distribution are shown in Figure 2.9.

Summarize the results for linear programming, it is concluded that

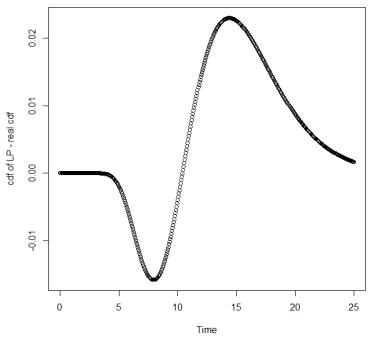
1. The accuracy is improved if the points r_i are properly selected. Especially in this case, the exact result is in the parametric form as adopted in linear programming. Therefore when a point is chosen as the boundary, the fitting result is just the exact result. And in general cases, it is suggested to choose the points around the mean first passage time.
2. The fitting times influence the estimate a lot, and it is suggested to consider to fit the linear programming around the mean first passage time.

Based on all the above results, among the three methods

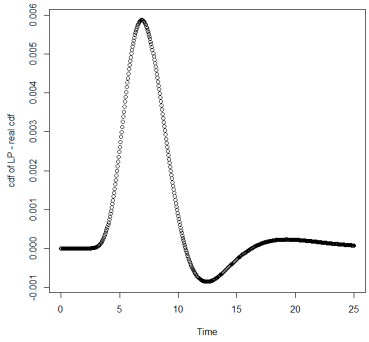
1. The integral equation methods show a good approximation accuracy.
2. The piecewise Monte-Carlo method is superior than the simple Monte-Carlo method. With few trajectories, the piecewise Monte-Carlo method shows also a good accuracy. However with the same time step-size, it can hardly reach the same accuracy of the integral equation methods.
3. The linear programming approach depends a lot on the preliminary guess on the fitting times and chosen points. But this approximation is analytical. Moreover, if the real first passage density is of the parametric form, it is possible to derive an exact result from the linear programming approach.



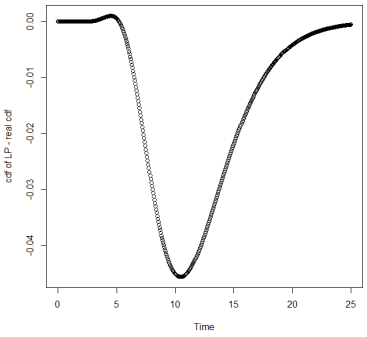
(a) Errors of the linear programming approach on cdf, fitting times between 5 and 15, chosen points 0.5+3i,i=0,...,8.



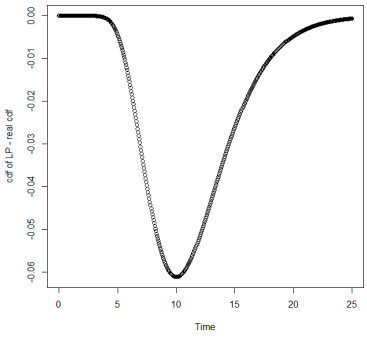
(b) Errors of the truncated approximation on cdf, fitting times between 0 and 10, chosen points 0.5+3i,i=0,...,8.



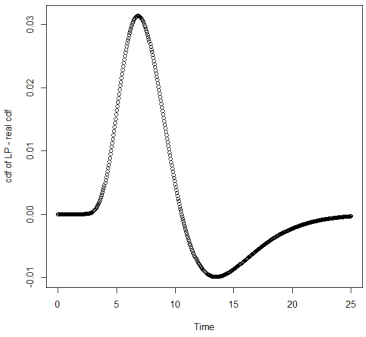
(c) Errors of the linear programming approach on cdf, fitting times between 10 and 25, chosen points 0.5+3i,i=0,...,8.



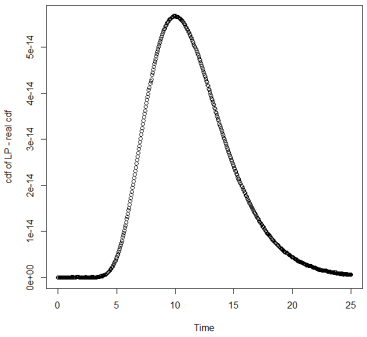
(d) Errors of the linear programming approach on cdf, fitting times between 5 and 15, chosen points 0.5+i,i=0,...,8.



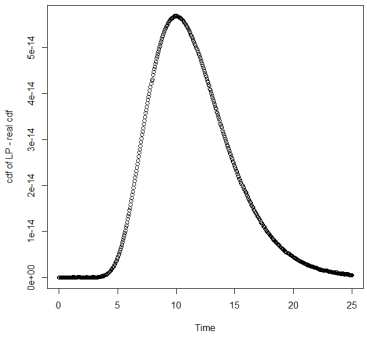
(e) Errors of the truncated approximation on cdf, fitting times between 0 and 10, chosen points 0.5+i,i=0,...,8.



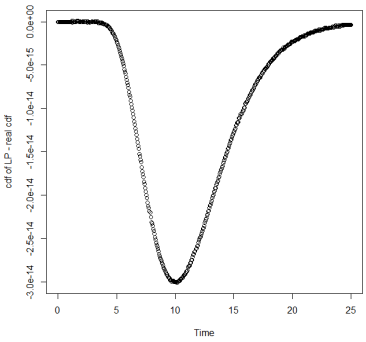
(f) Errors of the linear programming approach on cdf, fitting times between 10 and 25, chosen points 0.5+i,i=0,...,8.



(g) Errors of the linear programming approach on cdf, fitting times between 5 and 15, chosen points 10+i,i=0,...,8.



(h) Errors of the truncated approximation on cdf, fitting times between 0 and 10, chosen points 10+i,i=0,...,8.



(i) Errors of the linear programming approach on cdf, fitting times between 10 and 25, chosen points 10+i,i=0,...,8.

FIGURE 2.9 – Errors of the first passage distribution of linear programming approach for the drifted Motion $X_t = t + B_t$.

TABLE 2.4 – Calculation time of different situations for the linear programming approach

	fitting times in $[5, 15]$	fitting times in $(0, 10]$	fitting times in $[10, 25]$
$r_i = 0.5 + 3i, i = 0, ..., 8$	4.395251 secs	2.908166 secs	4.329247 secs
$r_i = 0.5 + i, i = 0, ..., 8$	4.105235 secs	4.490256 secs	4.495257 secs
$r_i = 10 + i, i = 0, ..., 8$	4.071233 secs	4.478257 secs	4.532259 secs

TABLE 2.5 – Objective values of different situations for the linear programming approach

	fitting times in $[5, 15]$	fitting times in $(0, 10]$	fitting times in $[10, 25]$
$r_i = 0.5 + 3i, i = 0, ..., 8$	3.317761	4.252861	0.6275354
$r_i = 0.5 + i, i = 0, ..., 8$	12.35105	8.728594	2.773304
$r_i = 10 + i, i = 0, ..., 8$	0	0	0

2.5.2 Time-changed Brownian Motion

As introduced in (2.3.39), a time-changed Brownian motion is introduced to model non-linear tendency of the degradation process. Especially its first passage density is explicit, and therefore it could be interesting as it can simplify corresponding analysis in reliability analysis. And here we will consider one time-changed Brownian motion X_t as follows :

$$dX_t = 0.5tdt + \sqrt{t}dB_t, X_0 = 0, t \geq 0. \tag{2.5.5}$$

Its transition pdf $p(x, t|y, s)$ when $t > s$ is given by

$$p(x, t|y, s) = \frac{1}{\sqrt{\pi(t^2 - s^2)}} \exp(-\frac{(x - y - (t^2 - s^2)/4)^2}{(t^2 - s^2)}), \tag{2.5.6}$$

such that the transition distribution is given by

$$F(x, t|y, s) = \Phi(\frac{x - y - (t^2 - s^2)/4}{\sqrt{(t^2 - s^2)/2}}). \tag{2.5.7}$$

And given the failure level $L = 10$, corresponding first passage density is given by

$$g(t|0, 0) = \frac{2L}{t} p(L, t|0, 0). \tag{2.5.8}$$

Correspondingly, the first passage distribution is given by

$$G(t|0, 0) = \Phi(\frac{-L + t^2/4}{\sqrt{t^2/2}}) + e^{-(y - L(s))}\Phi(\frac{-t^2/4 - L}{\sqrt{t^2/2}}), \tag{2.5.9}$$

where $\Phi(*)$ is the normal distribution function. The pdf and cdf of the first passage time are shown in Figure 2.10(a) and Figure 2.10(b) respectively.

Moreover, it is noticed that in this setup, the tangent approximations / Durbin's approximation degenerate to the exact result, such that these two trivial issues will be ignored in the following simulation tests. And we will mainly discuss numerical solutions to integral equations, Monte-Carlo simulation methods and the linear programming approach.

TABLE 2.6 – Calculation time of different methods with different time step-sizes

Method	$\Delta t = 0.05$	$\Delta t = 0.02$	$\Delta t = 0.01$
Non-singular VIE	11.08863 secs	1.019258 mins	3.333924 mins
Truncated approximation	21.21021 secs	3.023623 mins	10.12706 mins
Exact solution [5]	0.9240532 secs	1.003058 secs	1.183067 secs

Integral Equation Methods

Supposing the step-size $\Delta t > 0$ is chosen and assuming $t_0 = 0$, denote $t_n = n\Delta t$ for $n = 0, 1, 2, ...$, then we can fulfill the numerical algorithm based on the non-singular Volterra integral equation in Section 2.4.3 and also the truncated approximation in Section 2.4.4. Here we want to investigate the influence of time step-size Δt for these two methods, on the estimation accuracy and the calculation time.

Moreover as only the first passage density $\{g_i\}_{i=0}^N$ is given based on the VIE method, it is of interest to produce further corresponding cdf $\{G_i\}_{i=0}^N$ from the iterative steps based on the right-point formula :

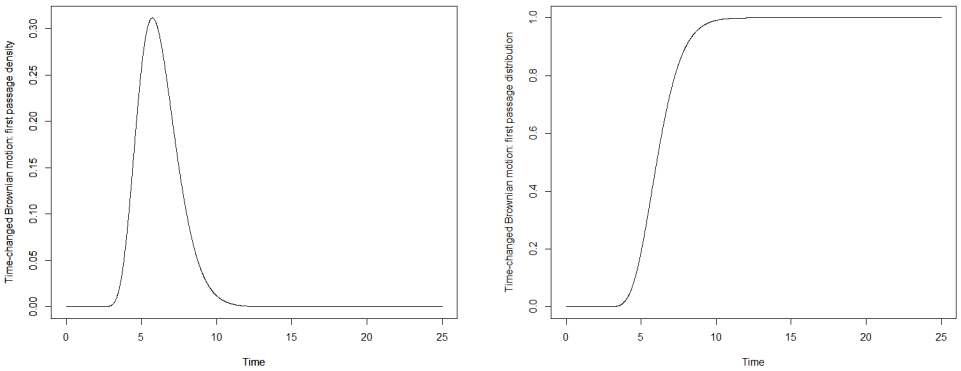
1. $G_0 = 0$,
2. $G_{i+1} = G_i + \Delta t g_{i+1}, i = 0, 1, ..., N-1$.

The calculation time of the methods with three different step-size $0.05, 0.02, 0.01$, which are given in Table 2.6. Corresponding errors of estimated first passage distribution are shown in Figure 2.11.

Summarize the results, we can conclude

1. The non-singular VIE and truncated approximation have the same level of calculation time and also accuracy with the same step-size.
2. When the step-size is smaller, the estimate is more accurate.

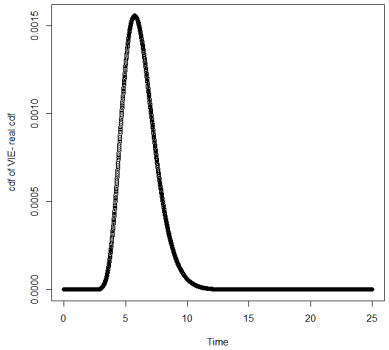
5. The time is given based on the calculation on all the time mesh points.



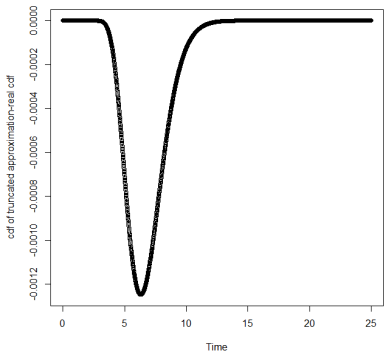
(a) The first passage density of time-changed Brownian motion

(b) The first passage density of time-changed Brownian motion

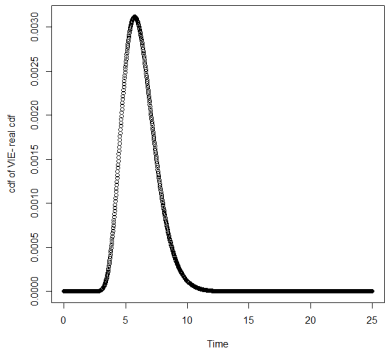
FIGURE 2.10 – The pdf and cdf of the first passage time of $dX_t = 0.5tdt + \sqrt{t}B_t$ with $L = 10$.



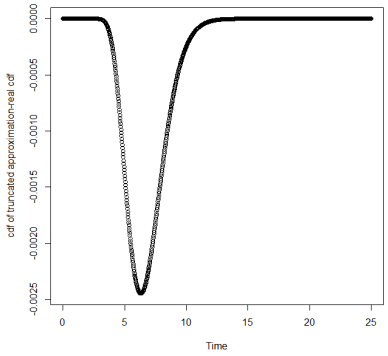
(a) Errors of the VIE method on cdf, step-size $\Delta t = 0.01$, maximal deviation ≈ 0.0015.



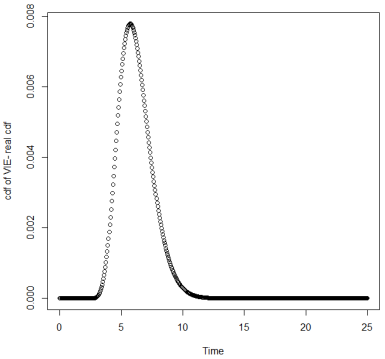
(b) Errors of the truncated approximation on cdf, step-size $\Delta t = 0.01$, maximal deviation ≈ -0.0012.



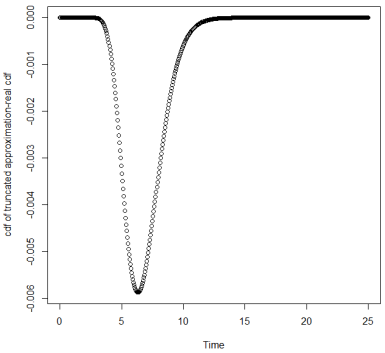
(c) Errors of the VIE method on cdf, step-size $\Delta t = 0.02$, maximal deviation ≈ 0.003.



(d) Errors of the truncated approximation on cdf, step-size $\Delta t = 0.02$, maximal deviation ≈ -0.0025.



(e) Errors of the VIE method on cdf, step-size $\Delta t = 0.05$, maximal deviation ≈ 0.008.



(f) Errors of the truncated approximation on cdf, step-size $\Delta t = 0.05$, maximal deviation ≈ -0.008.

FIGURE 2.11 – Errors of different methods compared with the real first passage density for the time-changed Brownian motion $dX_t = 0.5tdt + \sqrt{t}dB_t$

TABLE 2.7 – Calculation time of Monte-Carlo methods in different situations

Method	$\Delta t = 0.05$	$\Delta t = 0.02$	$\Delta t = 0.01$
Simple MC (200 trajectories)	6.569376 secs	6.297361 secs	9.576548 secs
Piecewise MC (200 trajectories)	8.770501 secs	11.30465 secs	18.15304 secs
Simple MC (2000 trajectories)	4.816276 secs	28.84765 secs	31.10478 secs
Piecewise MC (2000 trajectories)	31.75282 secs	47.44571 secs	1.501519 mins
Simple MC (20000 trajectories)	1.584791 mins	2.68857 mins	5.702576 mins
Piecewise MC (20000 trajectories)	2.172774 mins	8.341027 mins	14.55055 mins

Monte-Carlo Simulation Methods

Supposing the step-size $\Delta t > 0$ is chosen and assuming $t_0 = 0$, denote $t_n = n\Delta t$ for $n = 0, 1, 2, ...$, $X_n = X_{t_n}$ and $X_0 = x_0$. Constrained in the consideration of the time-changed Brownian motion $dX_t = 0.5tdt + \sqrt{t}dB_t$, the Euler-Maruyama scheme for Monte-Carlo simulation is updated to

$$X_{n+1} = X_n + 0.5t_n\Delta t + \sqrt{t_n}\Delta B_t, \tag{2.5.10}$$

where $\Delta B_t \sim \mathcal{N}(0, \Delta t)$.

Here the calculation time of the methods with three different step-size $0.05, 0.02, 0.01$ and also with three different numbers of trajectories $200, 2000, 20000$, which are given in Table 2.7 [6]. Corresponding errors are shown respectively in Figures 2.12, 2.13 and 2.14.

Summarize the results, it is concluded that

1. The accuracy is improved by producing more trajectories with the same step-size for Monte-Carlo simulation methods.
2. When the step-size is smaller, the estimate is more accurate.
3. The piecewise quasi-linear MC method is more accurate than the simple MC method, with the same number of trajectories and the same step-size. But it also costs more time to do extra calculation, however in general the piecewise quasi-linear MC method costs the same level of time with the simple MC method.
4. When enough trajectories are produced (20000 trajetories), the Monte-Carlo simulation methods is still less accurate than the integral equation methods with the same step-size.

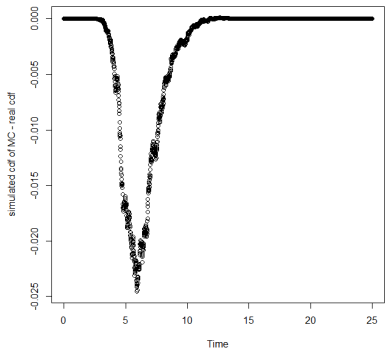
Linear Programming Approach

Recall the method of linear programming proposed in Section 2.3.5. An approximation for the first passage density is proposed based on

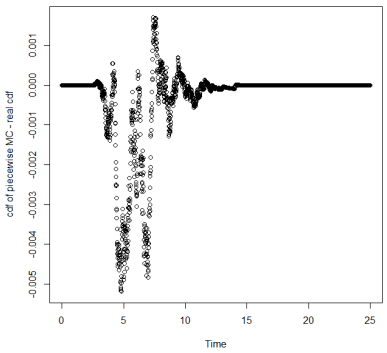
$$p(L(t), t|y, s) - \sum_{i=1}^{N} \xi_i p(L(t), t|L(s) + r_i, s) \approx 0, \tag{2.5.11}$$

for some selected points $\{r_i\}_{i=1}^n$ to estimate $\{\xi_i\}_{i=1}^n$ in (2.3.67). In such a case, the first passage density is given in (2.3.63), and corresponding first passage distribution is given in (2.3.64).

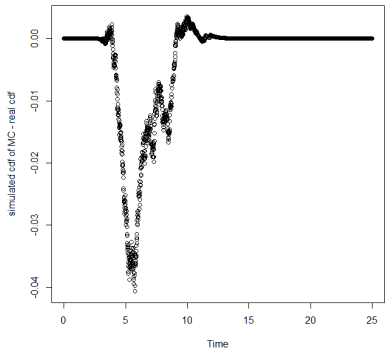
6. Here the calculation time and errors are given based on trajectories produced in one test, and it may vary due to different simulation tests. This influence would be slight when enough trajectories are produced



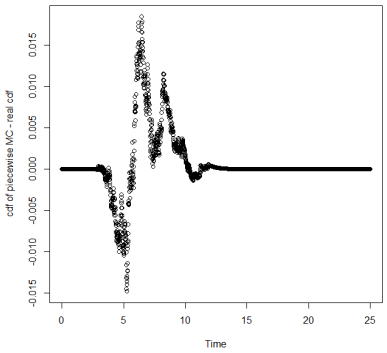
(a) Errors of simple MC, 20000 trajectories, maximal deviation ≈ -0.025



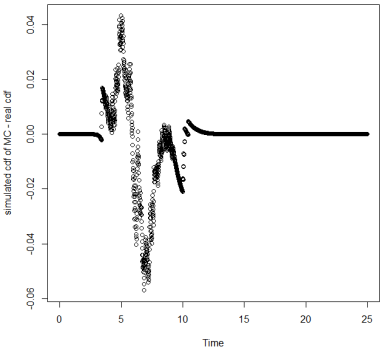
(b) Errors of piecewise MC, 20000 trajectories, maximal deviation ≈ 0.005



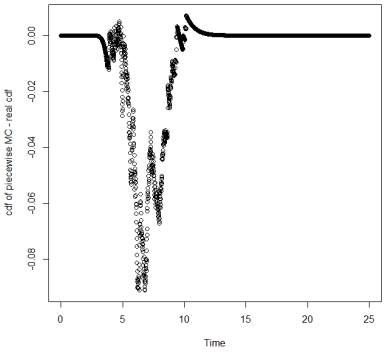
(c) Errors of simple MC, 2000 trajectories, maximal deviation ≈ 0.04



(d) Errors of piecewise MC, 2000 trajectories, maximal deviation ≈ 0.015

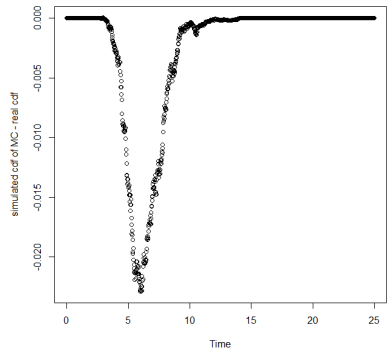


(e) Errors of simple MC, 200 trajectories, maximal deviation ≈ -0.06

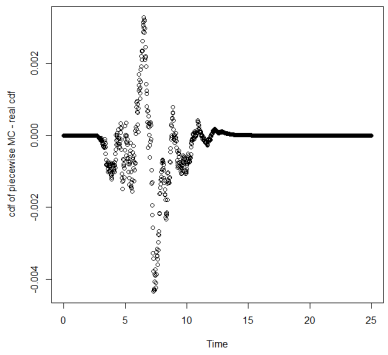


(f) Errors of piecewise MC, 200 trajectories, maximal deviation ≈ -0.08

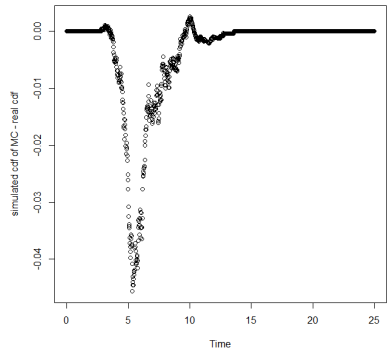
FIGURE 2.12 – Errors of Monte-Carlo simulation methods on cdf for the time-changed Brownian motion $dX_t = 0.5tdt + \sqrt{t}dB_t$, step-size $\Delta t = 0.01$



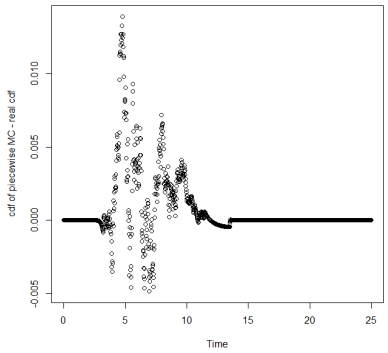
(a) Errors of simple MC, 20000 trajectories, maximal deviation ≈ -0.02.



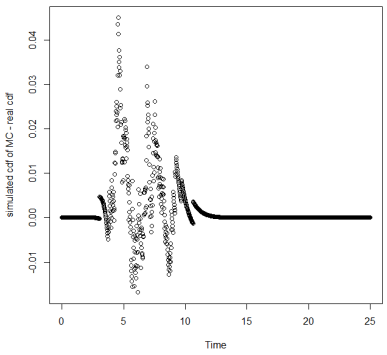
(b) Errors of piecewise MC, 20000 trajectories, maximal deviation ≈ -0.004.



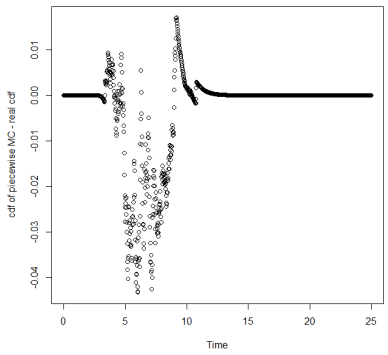
(c) Errors of simple MC, 2000 trajectories, maximal deviation ≈ -0.04.



(d) Errors of piecewise MC, 2000 trajectories, maximal deviation ≈ 0.01.

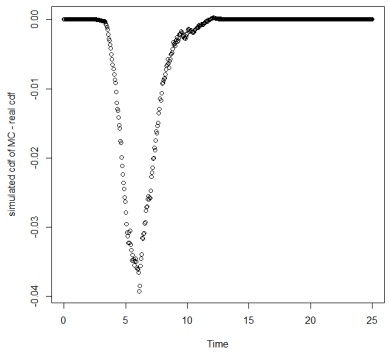


(e) Errors of simple MC, 200 trajectories, maximal deviation ≈ 0.04.

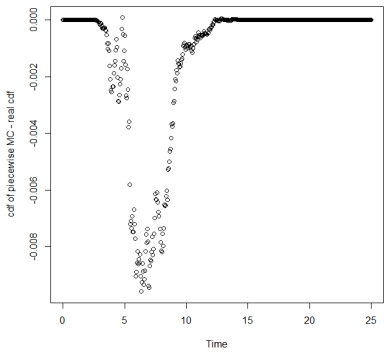


(f) Errors of piecewise MC, 200 trajectories, maximal deviation ≈ -0.04.

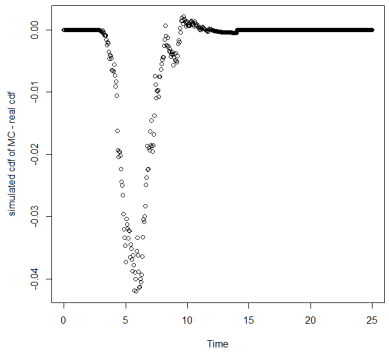
FIGURE 2.13 – Errors of Monte-Carlo simulation methods on cdf for the time-changed Brownian motion $dX_t = 0.5tdt + \sqrt{t}dB_t$, step-size $\Delta t = 0.02$



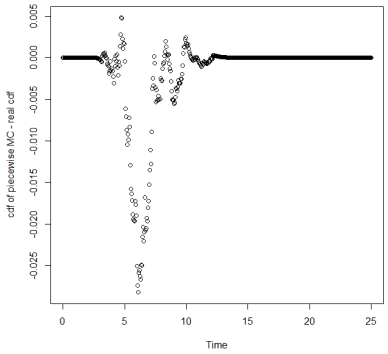
(a) Errors of simple MC, 20000 trajectories, maximal deviation ≈ -0.04.



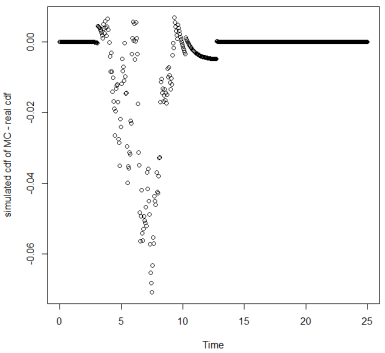
(b) Errors of piecewise MC, 20000 trajectories, maximal deviation ≈ -0.008.



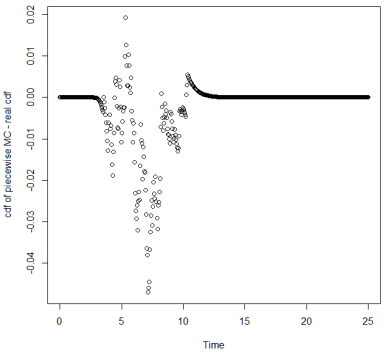
(c) Errors of simple MC, 2000 trajectories, maximal deviation ≈ -0.04.



(d) Errors of piecewise MC, 2000 trajectories, maximal deviation ≈ -0.025.



(e) Errors of simple MC, 200 trajectories, maximal deviation ≈ -0.06.



(f) Errors of piecewise MC, 200 trajectories, maximal deviation ≈ -0.04.

FIGURE 2.14 – Errors of Monte-Carlo simulation methods on cdf for the time-changed Brownian motion $dX_t = 0.5tdt + \sqrt{t}dB_t$, step-size $\Delta t = 0.05$

TABLE 2.8 – Calculation time of different situations for the linear programming approach

	fitting times in $[5, 15]$	fitting times in $(0, 10]$	fitting times in $[10, 25]$
$r_i = 0.5 + 3i, i = 0, ..., 8$	3.372193 secs	2.949169 secs	2.948169 secs
$r_i = 0.5 + i, i = 0, ..., 8$	4.645266 secs	4.527259 secs	3.852221 secs
$r_i = 10 + i, i = 0, ..., 8$	4.356249 secs	3.893223 secs	3.67221 secs

Later the estimation of $\{\xi_i\}_{i=1}^n$ is to promise (2.5.11) at several times $\{t_i\}$ such that the error can be minimized on these times.

From previous statements, numerical algorithms involve the choice of time step-sizes, Monte-Carlo simulation methods involve not only the time step-size, but also those produced trajectories. However things are totally different when it comes to the linear programming approach. Actually, from the derivation of the method, it is found that this method is influenced more by the choice of point measures and the fitting times.

Therefore the time mesh is chosen as $t_i = i\Delta t, i \in \mathbb{N}$ in the interval $[0, 25]$, with $\Delta t = 0.05$, the following issues are concerned here

1. The different performance when the linear programming approach is solved for those times in a smaller interval.

2. The different performance when the point measures vary.

The calculation times of the linear programming approach in different situations are given in Table 2.8. The objective values of the linear programming approach in different situations are given in Table 2.9. Corresponding errors of estimated first passage distribution are shown in Figure 2.15.

Summarize the results for linear programming, it is concluded that

1. The accuracy is improved if the points r_i are properly selected. Especially in this case, the exact result is in the parametric form as adopted in linear programming. Therefore when a point is chosen as the boundary, the fitting result is just the exact result. And in general cases, it is suggested to choose the points around the mean first passage time.

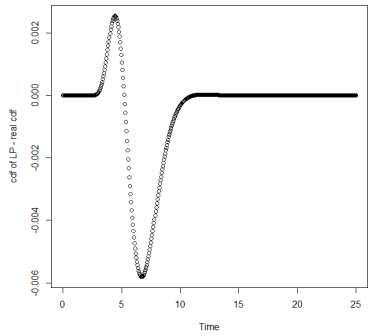
2. The fitting times influence the estimate a lot, and it is suggested to consider to fit the linear programming around the mean first passage time.

Based on all the above results, among the three methods

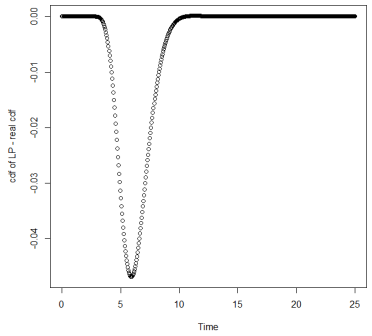
1. The integral equation methods show a good approximation accuracy.

2. The piecewise Monte-Carlo method is superior than the simple Monte-Carlo method. With few trajectories, the piecewise Monte-Carlo method shows also a good accuracy. However with the same time step-size, it can hardly reach the same accuracy of the integral equation methods.

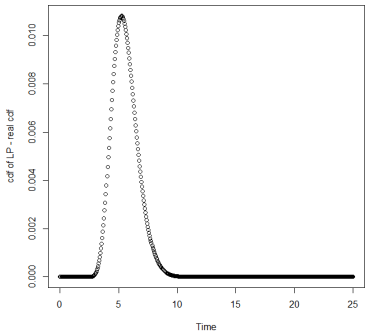
3. The linear programming approach depends a lot on the preliminary guess on the fitting times and chosen points. But this approximation is analytical. Moreover, if the real first passage density is of the parametric form, it is possible to derive an exact result from the linear programming approach.



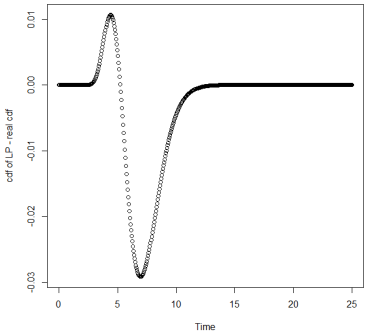
(a) Errors of the linear programming approach on cdf, fitting times between 5 and 15, chosen points 0.5+3i,i=0,...,8.



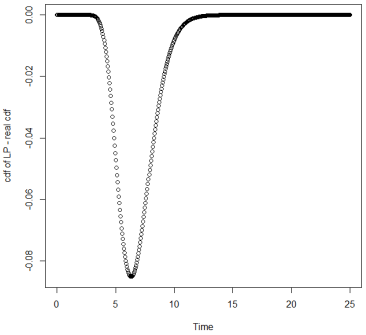
(b) Errors of the truncated approximation on cdf, fitting times between 0 and 10, chosen points 0.5+3i,i=0,...,8.



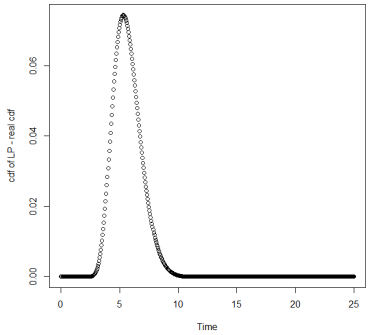
(c) Errors of the linear programming approach on cdf, fitting times between 10 and 25, chosen points 0.5+3i,i=0,...,8.



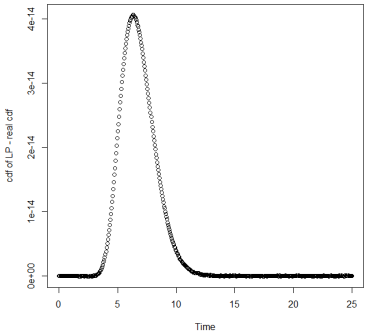
(d) Errors of the linear programming approach on cdf, fitting times between 5 and 15, chosen points 0.5+i,i=0,...,8.



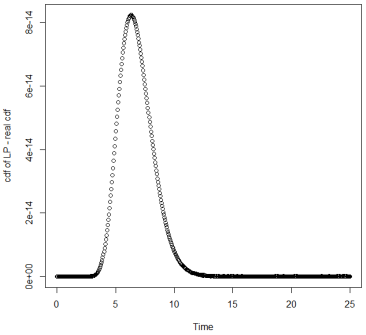
(e) Errors of the truncated approximation on cdf, fitting times between 0 and 10, chosen points 0.5+i,i=0,...,8.



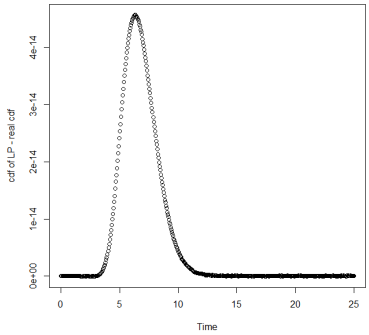
(f) Errors of the linear programming approach on cdf, fitting times between 10 and 25, chosen points 0.5+i,i=0,...,8.



(g) Errors of the linear programming approach on cdf, fitting times between 5 and 15, chosen points 10+i,i=0,...,8.



(h) Errors of the truncated approximation on cdf, fitting times between 0 and 10, chosen points 10+i,i=0,...,8.



(i) Errors of the linear programming approach on cdf, fitting times between 10 and 25, chosen points 10+i,i=0,...,8.

FIGURE 2.15 – Errors of the first passage distribution of linear programming approach for the tim-changed Brownian motion $dX_t = 0.5tdt + \sqrt{t}dB_t$.

TABLE 2.9 – Objective values of different situations for the linear programming approach

	fitting times in $[5, 15]$	fitting times in $(0, 10]$	fitting times in $[10, 25]$
$r_i = 0.5 + 3i, i = 0, ..., 8$	0.3249917	2.841034	7.309698e-05
$r_i = 0.5 + i, i = 0, ..., 8$	1.857895	6.08458	0.003083371
$r_i = 10 + i, i = 0, ..., 8$	0	1.81146e-08	1.967621e-07

2.5.3 A General OU process

Experimental Setup

Hereafter we will consider a case described in (1.2) which is estimated from the degradation data of a passive component in power plants. In Chapter 1, we have discussed the modeling work, and the following process X_t is given :

$$dX_t = (-r(X_t - m(t)) + m'(t))dt + \sigma dB_t, t \geq 0, X_0 = x_0, \tag{2.5.12}$$

where $r = 0.1806708, \sigma = 2.4640884, m(t) = 2.4402845((t+1)^{0.8892020} - 1) + x_0, x_0 = 2.8074561$. Some trajectories of the process (2.5.12) are produced based on Monte-Carlo simulation, which are shown in Figure 1.4(b).

Other issues related to the first passage problem are chosen as follows :

1. The failure level is chosen as a constant 10.
2. The considered time interval is [0,25].
3. The step-size Δt to discretize the time interval is 0.01.

Nonsingular Volterra Integral Equation

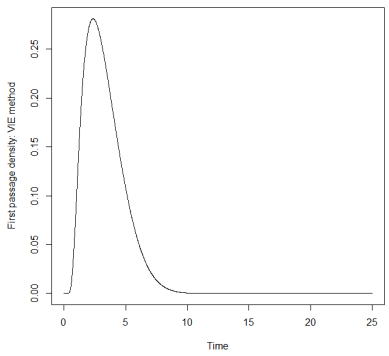
The nonsingular Volterra integral equation (VIE) (2.4.10) is adopted to calculate the first passage density, and the results are shown in Figure 2.16(a). Correspondingly the pdf is compared with the fitted pdf from the simple Monte-Carlo simulation (50000 times simulation), see Figure 2.16(b).

The corresponding calculated cdf is also compared with the simulated cdf from the simple Monte-Carlo method, see Figure 2.16(c) and 2.16(d).

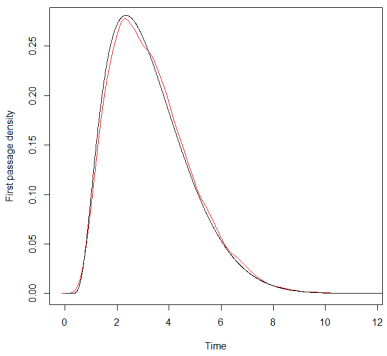
Piecewise Quasi-Linear Monte-Carlo Simulation

Proposed in Section 2.3.6, it is suggested to consider another MC simulation method from approximating the original boundary $L(t)$ to a piecewise quasi-linear boundary. As stated before, when the time step-size is very small, this method provides an approximation of first passage distribution to the real one.

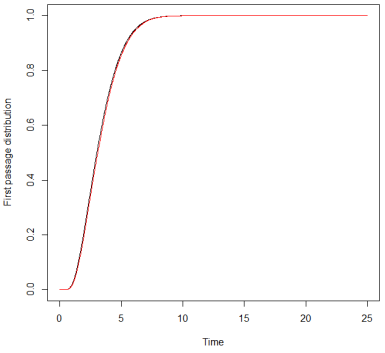
We have produced 20000 trajectories to calculate the piecewise quasi-linear Monte-Carlo simulation. The corresponding calculated cdf is also compared with the results from the VIE method, see Figure 2.16(c) and 2.16(d).



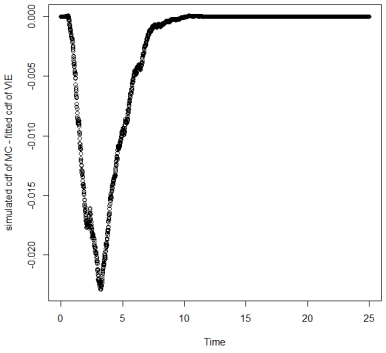
(a) First passage density : VIE method



(b) First passage density : VIE-black, MC-red

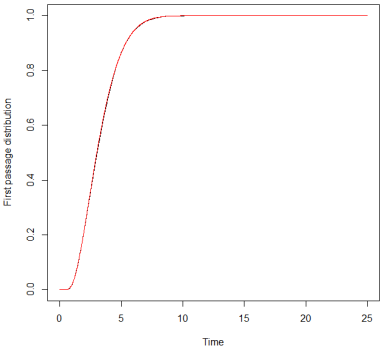


(c) First passage distribution : VIE-black, MC-red

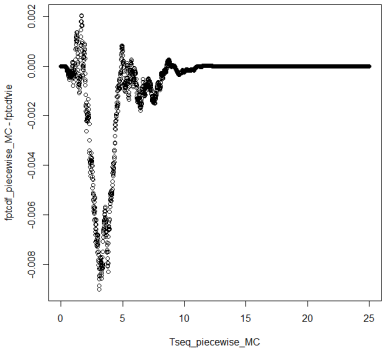


(d) Difference between VIE and simple MC on first passage distribution, maximal deviation ≈ -0.02.

FIGURE 2.16 – The VIE method, compared with the simple Monte-Carlo method



(a) First passage distribution : piecewise MC-black, VIE-red



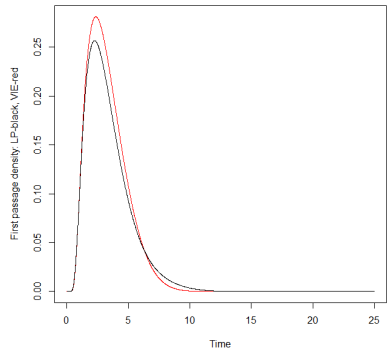
(b) Difference between piecewise quasi-linear MC simulation and VIE method, maximal deviation ≈ -0.008.

FIGURE 2.17 – The piecewise quasi-linear Monte-Carlo method, compared with the results from VIE method

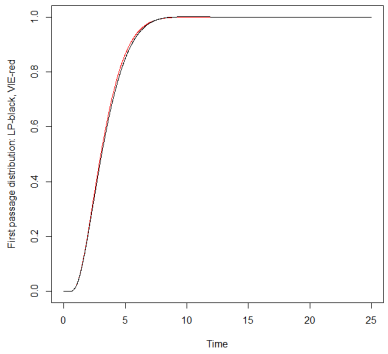
Linear Programming Approach

Continuing previous consideration in this simulation test, we concern two issues : the fitting times in consideration, and the chosen points $\{r_i\}_{i=1}^n$. The weight P is chosen with all elements to be 1. Here suppose the time mesh $\{t_i\}_{i=1}^m$ in the time interval $[0, 25]$ with a step-size 0.01 is considered. And later we choose those times in a smaller interval, such that the estimation based on this smaller interval is compared with the estimation based on [0,25].

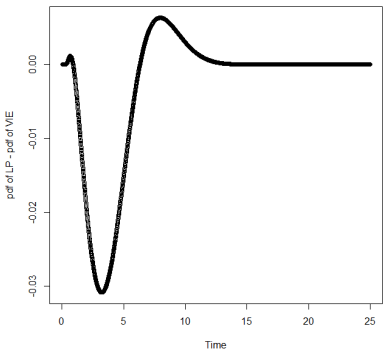
A first illustration on the result of linear programming approach is presented for the case with $r_i = 0.5 + i, i = 0, 1, .., 9$, and the time interval considered for the linear programming is $[1, 7]$. The objective value is achieved at 4.996158, corresponding to the coefficient series $(\xi_i) = (0.00000, 0.00000, 0.00000, 0.00000, 0.00000, 0.00000, 1.755968, 6.324348, 0.00000, 0.00000)$. The pdf and cdf by linear programming are shown in Figure 2.18(a) and 2.18(b), compared with the ones from the VIE method. The difference on pdf and cdf between the linear programming the VIE method are shown in Figure 2.18(c) and 2.18(d).



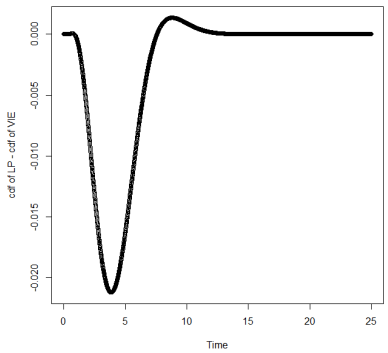
(a) First passage density : LP-black,VIE-red



(b) First passage distribution : LP-black, VIE-red



(c) Difference between LP and VIE on First passage density, maximal deviation ≈ -0.03.

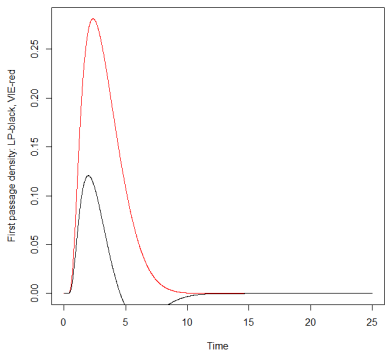


(d) Difference between LP and VIE on first passage distribution, maximal deviation ≈ -0.02.

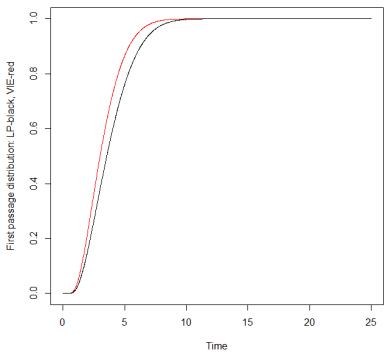
FIGURE 2.18 – The linear programming approach, compared with the VIE method : considered time $[1, 7]$, $r_i = 0.5 + i, i = 0, 1, .., 9$.

Now it is supposed that the time interval is considered as $(0, 20]$. Then the calculation gives the objective value is achieved at 41.29026, corresponding to the coefficient series $(\xi_i) =$

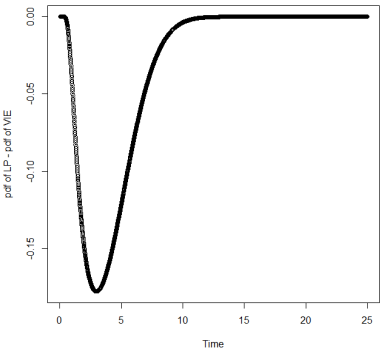
$(0.00000, 0.000000, 0.00000, 0.00000, 0.00000, 0.00000, 0.1038629, 1.7883317, 0.00000, 0.00000)$. The pdf and cdf by linear programming are shown in Figure 2.19(a) and 2.19(b), compared with the ones from the VIE method. The difference on pdf and cdf between the linear programming the VIE method are shown in Figure 2.19(c) and 2.19(d). From the pictures, we can see the considered time interval is extremely important, as the returned results show a great error compared to the consideration in $[1, 7]$.



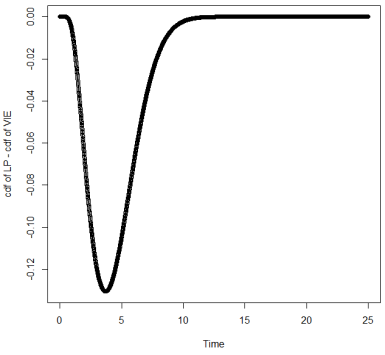
(a) First passage density : LP-black,VIE-red



(b) First passage distribution : LP-black, VIE-red



(c) Difference between LP and VIE on First passage density, maximal deviation ≈ -0.15.



(d) Difference between LP and VIE on first passage distribution, maximal deviation ≈ -0.12.

FIGURE 2.19 – The linear programming approach, compared with the VIE method : considered time $(0, 20]$, $r_i = 0.5 + i, i = 0, 1, .., 9$.

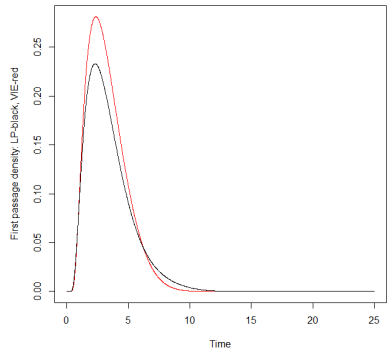
Now keep the considered time interval $[1, 7]$, but with changed values on $(r_i) = (0.5, 3.5, 6.5, 9.5)$. Then the calculation returns the objective value 8.166429, corresponding to the coefficient series $(\xi_i) = (0.000000, 0.000000, 3.714276, 5.796250)$. The pdf and cdf by linear programming are shown in Figure 2.20(a) and 2.20(b), compared with the ones from the VIE method. The difference on pdf and cdf between the linear programming and the VIE method are shown in Figure 2.20(c) and 2.20(d). We can see compared with the case with 10-point approximation, the accuracy of 4-point approximation is only slightly decreasing (based on the hypothesis of the accuracy of the VIE method).

So we have several conclusions here :

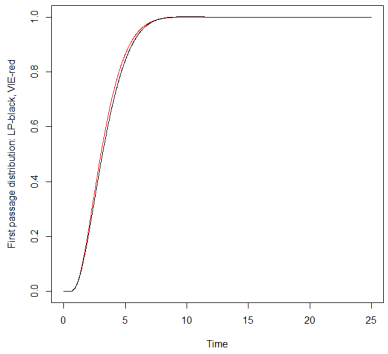
- The accuracy depends a lot on the chosen time series, where too many near-zero values

will influence the result a lot. So even we don't know the first passage density before the calculation, a first step is essential to estimate and analyze such that the chosen time series is around the mean first passage time.

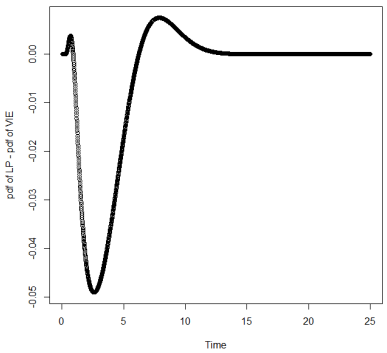
- The selected $\{r_i\}_{i=1}^n$ should be chosen also around the process' mean first passage time such that they could get more weight. It could be seen that even we include more consideration for different times, but only the ones around the mean first passage time return obvious weight values.
- Therefore a preliminary analysis on mean first passage time could be helpful to give an accurate estimate of first passage time based on linear programming. What should be emphasized is that this accurate approximation is not numerical. It is analytical, such that it could be used for later analysis.



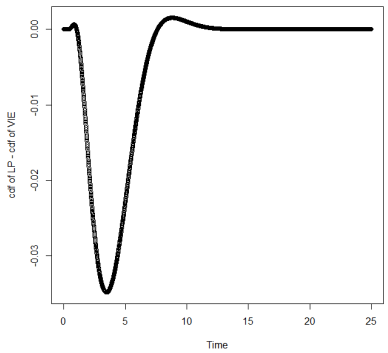
(a) First passage density : LP-black,VIE-red



(b) First passage distribution : LP-black, VIE-red



(c) Difference between LP and VIE on First passage density, maximal deviation ≈ -0.05.



(d) Difference between LP and VIE on first passage distribution, maximal deviation ≈ -0.03.

FIGURE 2.20 – The linear programming approach, compared with the VIE method : considered time $[1, 7]$, $(r_i) = (0.5, 3.5, 6.5, 9.5)$.

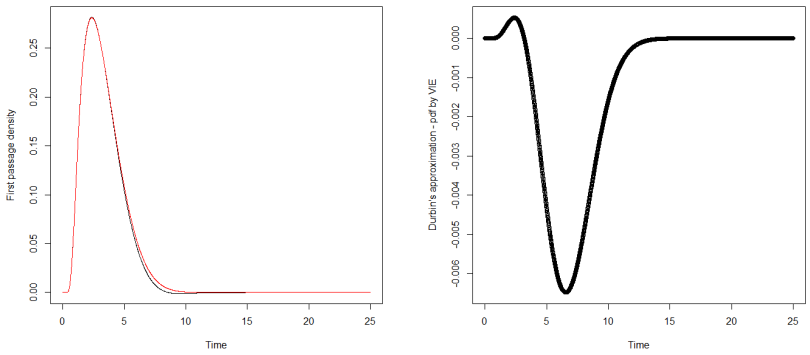
Durbin's Approximation

The Durbin's approximation is compared with the VIE method in Figure 2.21(a) and 2.21(b). The maximal difference with the results of VIE method is about 0.006, therefore this approximation is accurate in this case.

The Tangent Approximation

As stated in Section 2.3.3, when time is not far away from the current observation, the corresponding first passage density and distribution can be approximated by a tangent method given in (2.3.58). In a previous analysis, it is found that this tangent approximation is the same with the one of Durbin's approximation. And therefore here we just test the approximate first passage distribution. The approximation is compared with the results derived from VIE methods.

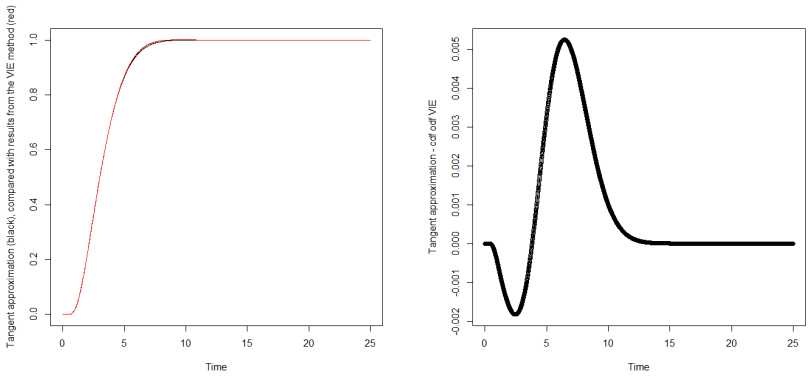
The produced cdfs and the difference with the VIE method are presented in Figure 2.22(a) and 2.22(b). The maximal difference with the results of VIE method is about 0.005, therefore this approximation is accurate in this case.



(a) First passage density : Durbin-black, VIE-red

(b) Difference on first passage density between VIE and Durbin's approximation, maximal deviation ≈ -0.006.

FIGURE 2.21 – Durbin's approximation, compared with the VIE method



(a) First passage distribution : Tangent approximation-black, VIE-red

(b) Difference on first passage distribution between VIE and the tangent approximation, maximal deviation $\approx$ 0.005.

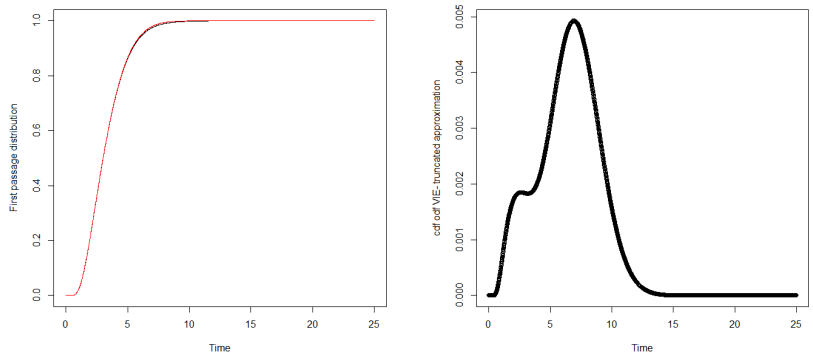
FIGURE 2.22 – Tangent approximation, compared with the VIE method

Truncated Approximation

It is presented in Section 2.4.4 that the first passage distribution can be estimated from a truncated approximation if the initial point $H(s)$ is chosen as the tangent approximation,

$$G(t|x_0,t_0) \approx 2\bar{F}(L(t),t|x_0,t_0) + 2\int_{t_0}^{t} H(s)d\bar{F}(L(t),t|L(s),s). \tag{2.5.13}$$

The first passage distribution produced by 2.5.13 is given in Figure 2.23(a). In Figure 2.23(b), the difference of (2.5.13) and the results from VIE method is presented.



(a) First passage distribution : Truncation approximation-black, VIE-red

(b) Difference on first passage distribution between VIE and the truncated approximation, maximal deviation ≈ 0.005.

FIGURE 2.23 – Truncated approximation starting with tangent approximation, compared with the VIE method

Quasi-Linear Boundary

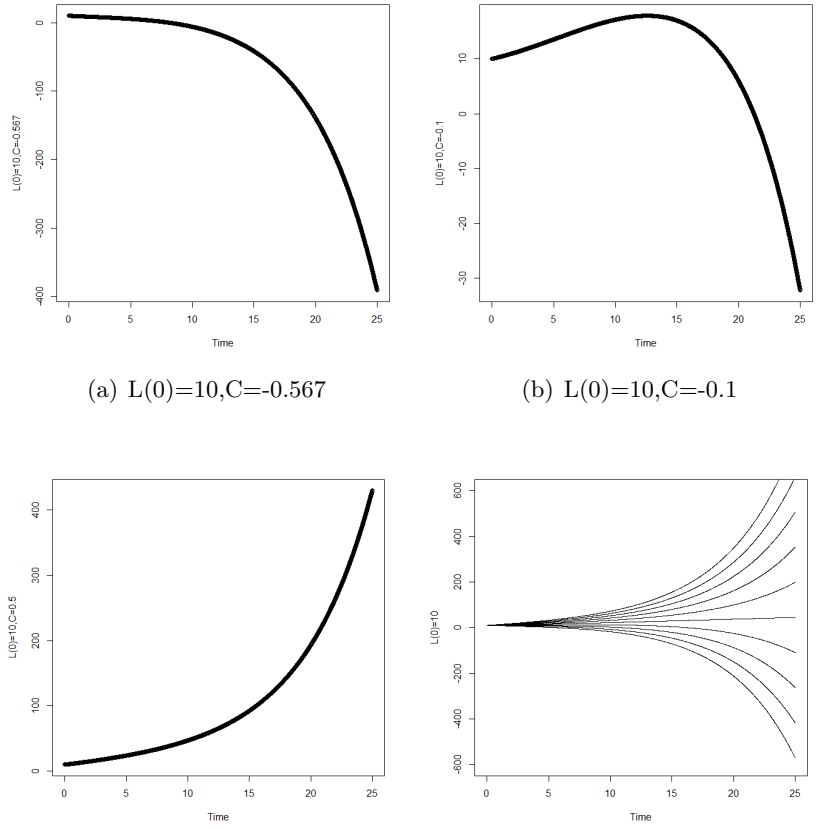
As the first passage problems can be solved explicitly under the quasi-linear boundaries, such that it aims to present the appearance of these boundaries when involved parameters are changing. It is also of interest to see what shape of boundaries can be applied to real cases. The quasi-linear boundary is a two-parameter boundary with the adjustable initial boundary and one shape variable. It is expected to have more flexility than the constant boundary. And as stated in Corollary 2.3.5, $C < 0$ is a more practical choice such that the system will be failed at the end.

With the form $L(t) = e^{-\alpha(t,0)}(L(0) - \beta(t,0) + C\gamma(t,0)$, and $L(0) = 10$, by changing the parameter C, following figures are presented, see Figure 2.24(a), 2.24(b), 2.24(c) and 2.24(d).

A great variability can be seen in the results. This is a very important point in real applications where failure records are presented. When fitting these failure records, the flexility of the failure level increases the fitting goodness of the first passage time. In such a way, the gap between the failure records and first passage times is expected to be small.

At a first glance, a rapid increasing rate of some boundaries seems unrealistic. However these boundaries can still fit well those "significant" failure records. That is to say, the boundary is chosen such that the corresponding first passage time has a significant pdf for

the most part of failure records. We can ignore the influence of the failure level for those unusual events.



(a) L(0)=10,C=-0.567 (b) L(0)=10,C=-0.1

(c) L(0)=10,C=0.5 (d) L(0)=10,C=-1+0.2*i, i=1,2,...,10

FIGURE 2.24 – Difference appearance of quasi-linear boundaries

2.5.4 Prognostics of Different Models

In this subsection, the discussion in Chapter 1 is continued on the topic of prognostics of first passage failures. In Chapter 1, we have presented three different models to fit the degradation data-set and have compared their fitting goodness. In this subsection, these three models will be considered respectively for the prediction of system failures based on first passage times.

The first passage time is considered for constant failure levels 15, 35 and 50 respectively. Later supposing a new component is put into operation and according to the three models given in Tables 1.1, 1.2 and 1.3, the pdf and cdf for the first passage time are calculated based on the integral equation (2.4.10) and (2.4.22).

The corresponding first passage density and distribution can be observed directly in Figures 2.25 and 2.26. Also corresponding MTTF is estimated and results are listed in Table 2.10. In these pictures, the solid line stands for M_{LD}, the dashed one stands for M_{OU}, the mixed one stands for M_{OUR}.

Some conclusions are derived here

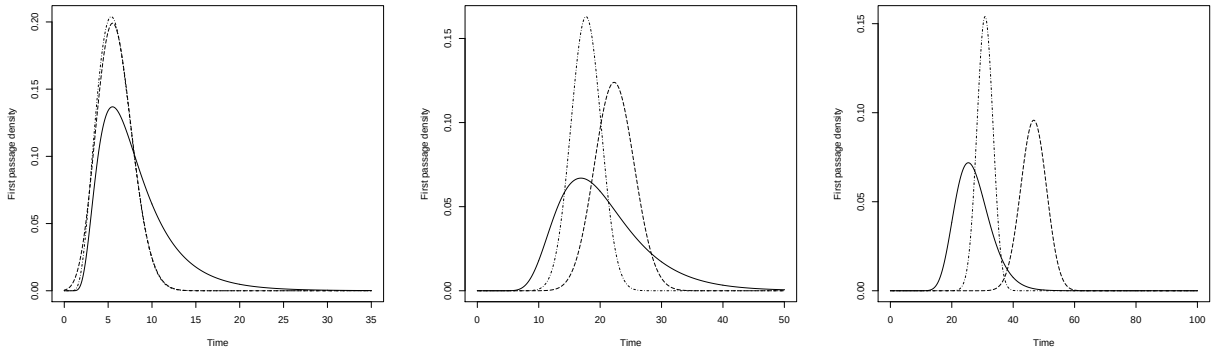
1. The comparison between M_{OU} and M_{LD} : when the model is with more fluctuations, the pdf of the first passage time is more spread-out ; when the model is with less fluctuations, the pdf of the first passage time is more peaky.
2. The comparison between M_{OU} and M_{OUR} : The initial uncertainties show little influence for a low boundary ($L = 15$). But for a high boundary ($L = 35$ and $L = 55$), the pdf of M_{OU} is more peaky than M_{OUR} because less uncertainties are introduced.

Therefore if the model's uncertainties is over-estimated, then more probability is distributed to the less possible events, especially those failed states. This could lead to a more conservative RUL estimate.

This conservative estimate may lead to an earlier estimate of the failure time by first passage failures. This phenomenon is observed, and is tried to be explained by last passage failures in [6]. However in this paper, we tried to reduce the gap between the real failure and first passage failure by adjusting the model's statistical properties properly.

TABLE 2.10 – MTTF prediction based on first passage failure

boundary	$M_{\mathbf{LD}}$	$M_{\mathbf{OU}}$	$M_{\mathbf{OUR}}$
15	6.321825	5.825432	5.800996
35	16.7787	17.69869	22.35642
55	27.19735	30.79393	46.59552



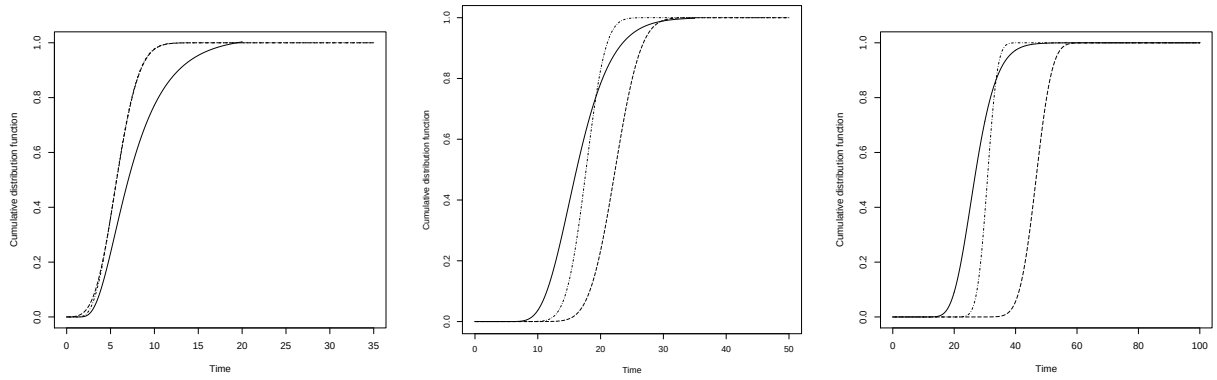
(a) Boundary=15 (b) Boundary=35 (c) Boundary=55

FIGURE 2.25 – first passage density function : $- \cdot - M_{\mathrm{OU}}$, $- - M_{\mathrm{OUR}}$, $-M_{\mathrm{LD}}$.

2.5.5 The Global Accuracy of Durbin's Approximation

Although we didn't analyze the global errors for Durbin's approximation analytically, some simulation tests will be done to see its performance illustratively. As suggested by [25], the global accuracy is influenced by the probability of crossing boundary before the current time. Therefore we will consider the global accuracy based on the folowing experiment set-up :

1. The boundary is set to be a constant $L = 10$.
2. The process is chosen as in (2.5.12) with a polynomial function $m(t) = 2.4402845((t+1)^{0.8892020} - 1) + 2.8074561$ with variable r and σ. Here the variance $v(t) = \frac{\sigma^2}{2r}(1 - \exp(-2rt))$, and $\mathrm{cov}(X_t, X_s) = \frac{\sigma^2}{2r}\exp(-r(t+s))(\exp(2rs) - 1)$.



(a) Boundary=15 (b) Boundary=35 (c) Boundary=55

FIGURE 2.26 – cumulative passage probability : $-\cdot-M_{\mathrm{OU}}$, $--M_{\mathrm{OUR}}$, $-M_{\mathrm{LD}}$.

Having chosen the processes with the same mean function, we compare the results of Durbin's approximation with the results from the VIE method. And corresponding results for different σ and r are given as in Figure 2.27.

The approximation is achieved by ignoring the probability of crossing boundary before the current time. Therefore introducing more fluctuations into the model would increase this probability. This would lead to a larger error of Durbin's approximation.

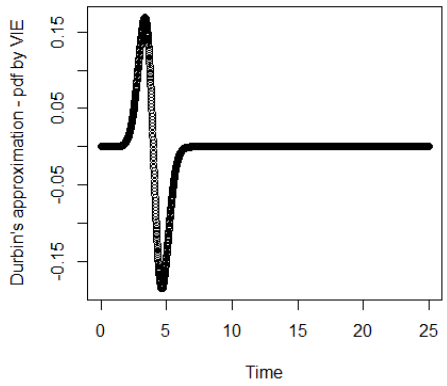
Furthermore, it is noticed that $\sigma^2/2r$ is the asymptotic variance, and r is the asymptotic correlation coefficient. As stated in Chapter 1, the correlation evaluates fluctuations in one trajectory of the process, while the variance evaluates the total fluctuations in the process. Therefore from the results, some conclusions can be derived :

1. The global accuracy of Durbin's approximation can be promised.
2. When the correlation is stronger (r is smaller), the Durbin's approximation is more accurate.
3. When the variance is smaller ($\frac{\sigma^2}{2r}$ is smaller), the Durbin's approximation is more accurate.
4. The influence of variance is smaller than the influence of correlation. In Figure 2.27(f), with a large variance and a strong correlation, the result is still accurate.

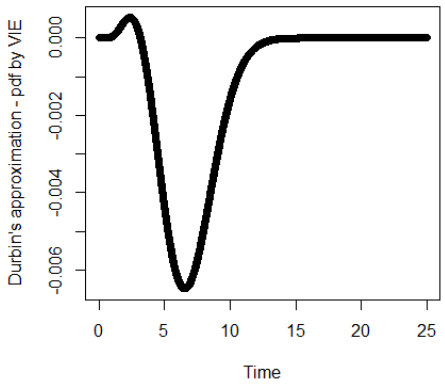
2.5.6 Summary

In this section, proposed methods are fulfilled in simulation tests, the advantages and disadvantages of proposed methods are commented here. As different methods are proposed with different concerns, the time-efficiency and accuracy are hard to be discussed with the same standard. So these comments are more from a general consideration.

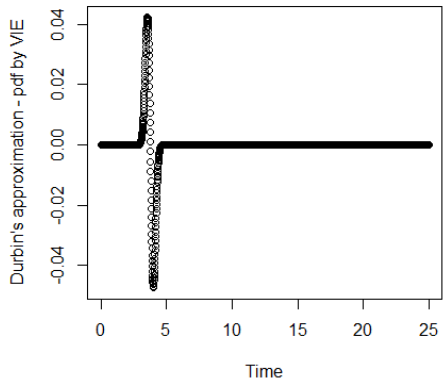
1. The non-singular Volterra integral equation method refers to an ordinary numerical technique to solve an integral equation, so its accuracy is increasing as the chosen step-size decreases. The results are numerical, and the accuracy can be reached by sacrificing the time-efficiency due to different requirements in applications. However, as it is an iterative numerical solving procedure based on the time mesh, it cannot



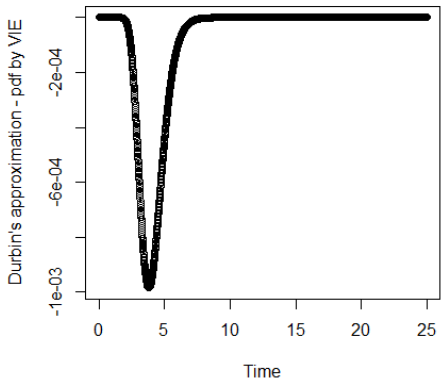
(a) $\sigma = 2.4640884, r = 1.806708, \sigma^2/2r = 1.68033$, maximal deviation$\approx 0.15$.



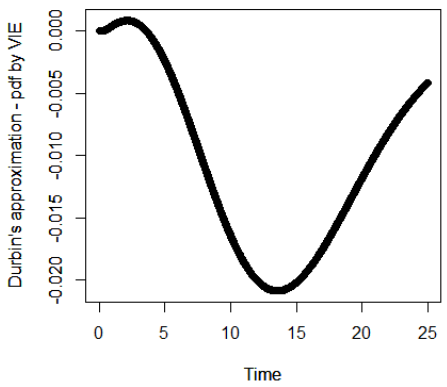
(b) $\sigma = 2.4640884, r = 0.1806708, \sigma^2/2r = 16.8033$, maximal deviation$\approx 0.006$.



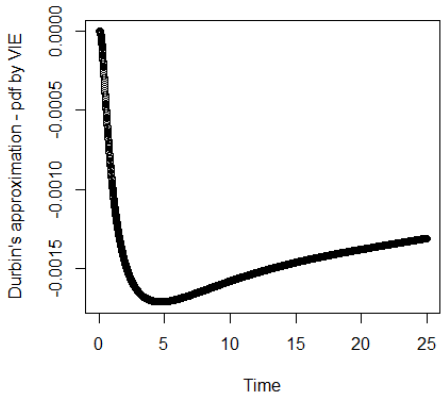
(c) $\sigma = 0.7792132, r = 1.806708, \sigma^2/2r = 0.168033$, maximal deviation$\approx 0.04$.



(d) $\sigma = 0.7792132, r = 0.01806708, \sigma^2/2r = 16.8033$, maximal deviation$\approx -0.001$.



(e) $\sigma = 7.792132, r = 0.1806708, \sigma^2/2r = 168.033$, maximal deviation$\approx 0.02$.



(f) $\sigma = 7.792132, r = 0.01806708, \sigma^2/2r = 1680.33$, maximal deviation$\approx 0.0.0015$.

FIGURE 2.27 – Difference between Durbin's approximation and the VIE method, for processes with the same mean $m(t) = 2.4402845((t+1)^{0.8892020} - 1) + 2.8074561$.

quickly react to return the value at a large time. From the simulation tests, this method balances well the accuracy and efficiency.

2. The piecewise Monte-Carlo simulation method produces many times of computer simulations to calculate the corresponding average. This method is influenced by the time step-size and the number of produced trajectories. And when producing enough trajectories, the accuracy is still less than integral equation methods under the fixed step size. The piecewise quasi-linear MC method costs a little more time than the simple MC, but it reaches also a more accurate estimate. It could be a substitute of Volterra integral equation method when the boundary is without differentiability.

3. The tangent approximations / Durbin's approximation are analytical approximations which can return the approximate value of first passage density immediately. And the result is not bad in general, and it is increasingly accurate as the probability that the process has crossed the boundary before the considered time diminishes [25]. It is already used in the field of PHM [66].The error of this approximation depends on the similarity of boundary to the quasi-linear boundary. When the boundary is quasi-linear, the approximation turns to be the exact solution.

4. The linear programming approach provides an analytical approximation which can return the approximate value of first passage density immediately. The accuracy of this method depends a lot on the chosen points and fitting times, which are both suggested to be around its mean first passage time. This needs preliminary judgement on the mean first passage time. This is due to the similarity of the real distribution to a linear combination of two chosen distribution functions. When the initial guess is good enough, it is expected to be accurate.

5. The truncated series' solution is explicit in an integral form, which needs to be approximated when the it is calculated. So the calculation afford is between totally analytical expression and totally numerical algorithms. And combining the tangent approximation with the truncated approximation is a trial to make up the cases when tangent approximations don't perform well. The accuracy and efficiency seem at the same level with non-singular Volterra integral equation method.

2.6 Conclusions and Perspectives

In this chapter, providing that the stochastic degradation process is expressed by the OU process X_t (2.1.1) $dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t$, the system failure is described by first passage failure to a given failure level $L(t)$. That is to say, the failure time is defined by conditional FPT $\tau_{y,s} := \inf_{t\geq s}\{X_t \geq L(t)|y, s\}$. It is of great interest to investigate $\tau_{y,s}$ in the field of reliability analysis, PHM and condition-based maintenance. And in this chapter different methods are derived to estimate RUL density $g(t|y, s) := \frac{\partial P(\tau_{y,s}\leq t)}{\partial t}$ and also corresponding cdf $P(\tau_{y,s} \leq t)$.

A key characteristic to distinguish RUL estimation in reliability engineering with other first passage problems is the requirement of condition-based prognosis, that is to say, updating prognosis of system failures based on the current observation. This task induces the consideration of $\tau_{y,s}$, i.e. the first passage time conditioned on y at time s throughout this thesis. Also the heterogeneity and non-stationarity of the time-dependent OU process induces plenty of difficulty to follow existing analysis. And this chapter have reproduced several classical ideas for this time-dependent OU process to provide a foundation for later analysis.

Different methods have been derived to calculate the first passage density or first passage distribution, and they are categorized into three classes :

1. Analytical approximations. In this chapter, we have investigated the first passage problem for the OU process from different angles, and some analysis turns out to be good approximations meanwhile. In these methods, the tangent approximation derived from the method of images is of most interest. The Durbin's approximation is proved to be the same with the tangent approximation. And a truncated approximation is derived based on the series solution of integral equation. The linear programming approach is derived directly from the method of images, and it is accurate when the chosen points are well-posed.
2. Monte-Carlo method. Based on the analysis of the OU process, a novel piecewise quasi-linear Monte-Carlo simulation method is derived, which shows superior than the simple Monte-Carlo simulation method.
3. Numerical algorithms. Based on Fortet's equation, a non-singular Volterra integral equation is derived to calculate the first passage density of the OU process.

The first passage problem induces an interdisciplinary consideration among stochastic analysis, PDE, integral equation, numerical analysis etc.. And the analysis in this chapter starts with $w(x,t|y,s)$, however a possibility to follow other analysis exists. Relaxing the constraint of crossing boundaries by using the explicit expression from the method of images leads to many insightful results [43].

The view of $u(x,t|y,s) := P(\tau_{y,s} > t, X_t < x|y,s)$ leads to a free boundary problem in (2.3.9), which is essential for a rigorous analysis of inverse first passage problem [13, 12]. This will be addressed in the next chapter. Also it is a starting point for optimal stopping problems [58].

Derived approximate expressions of first passage density are preferred in engineering problems than the numerical results, as in applied problems the computability could be more emphasized than the accuracy. With the explicit RUL estimate, the heuristic framework of maintenance optimization can be possibly reproduced, which will be explained in Chapter 4.

Chapitre 3

Failure Level Estimation via Inverse First Passage Problems

Several results in this chapter appear in [19, 21, 22].

When the stochastic degradation process is established from inspection records, the failure level itself describes system failures when first passage failures are considered. In previous literature the failure level is generally treated as physical barriers or experts' opinions, based on which failure prognosis can hardly fit existing failure records. Therefore in this chapter to make up such a gap, it is considered from a pure data-analysis based on inverse first passage problem to do the failure level estimation.

Suppose the lifetime distribution is given or estimated from failure records, the inverse first passage problem aims to reproduce the failure level under which the first passage time of the given stochastic process can have the same distribution with the given lifetime distribution. Therefore if we consider the system failure as the first passage failure, the failure records can provide a data-analysis method to estimate the failure level, such that the gap between failure records and prognosis of system failures are made up.

The whole chapter is organized as follows. In Section 3.1, IFPT problem will be introduced. In Section 3.2, an result on limit at zero for Brownian motion is introduced, and this result will be extended to the time-dependent OU process at a given time. In Section 3.3, the estimation of failure level based on Fortet's equation will be presented. In Section 3.4, the estimation of failure level based on the Master equation will be presented. In Section 3.5, simulation tests for presented results will be done. Conclusions and prospectives will be given in Section 3.6.

3.1 Introduction

3.1.1 Background

In real cases, it is natural to use observed information of the considered system to describe and predict system failures. And in this chapter, the observed information is considered only from two sources : failure information on the recorded failures and inspection information on the system's state. Emphasizing on different sources of system information, the failure can be described also from two different angles : one is from statistical analysis of failure records leading to lifetime distribution [40], the other is from system dynamic models leading to first passage failure [1, 41, 89]. These two views are rather independent in previous literature, and this chapter is therefore triggered to merge together the failure information and inspection information. The technical task is fulfilled by estimating the failure level for first passage failures based on inverse first passage problems (IFPT) [12, 21, 97].

Based on inspection information of detectable degradation indicators such as temperature, vibration signal, crack length etc., stochastic degradation models can be established e.g.

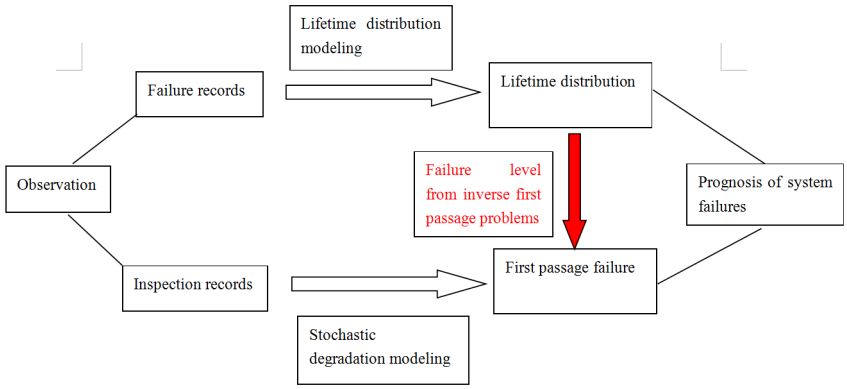


FIGURE 3.1 – The introduction of inverse first passage problems into PHM.

Gamma process [84], drifted Brownian motion [44] and OU process [20]. Furthermore to connect the degradation process with system failures, it is natural to introduce a failure level such that the system failure is defined as the first passage failure. That is to say, the failure time is described by the time when the degradation indicator reaches the failure level for the first time.

It is quite common to determine the failure level based on physical barriers or experts' judgments. However corresponding first passage failures can hardly fit existing failure records due to models' fitting goodness and uncertainties. Moreover in some cases, degradation indicators are derived from data-analysis techniques, e.g. principle component analysis from multiple indicators such that no physical meaning is presented [78]. Therefore in the latter situation, physical barriers and experts' experience can contribute little to estimating the failure level.

The above inconsistency between first passage failures and existing failure records induces our consideration based on IFPT problems in this chapter. Suppose the lifetime distribution is given or estimated from failure records, the inverse first passage problem aims to reproduce the failure level under which the first passage time of the given stochastic process can have the same lifetime distribution [1]. Solving such a problem maximizes simultaneously the utility of failure information and degradation modeling, and it is valuable in model-based prognosis and system health management (PHM). An illustration of the role of inverse first passage problems can be seen in Figure 3.1.

In the previous literature, some other efforts exist to consider simultaneously the failure records and inspection records. However most of them are from the parametric view and don't reflect the real system dynamic. The proportional hazard regression (also known as Cox model) [14] tries to jointly fit the inspection data and survival records from the view of hazard rate. This consideration emphasizes on the failure information, and the structure between the inspection records and the survival records can hardly be explained in a direct way.

For some degradation processes with explicit first passage density, e.g. drifted Brownian motion and Gamma process [84], the failure level can be estimated by fitting failure records. This parametric consideration adopts a strong pre-judgment for the failure level, and it cannot

1. The definition of inverse first passage problem here could be confusing with the inverse problem for random first passage problem [38, 83], that is to say, to find the initial distribution function to reproduce the randomized first passage density for a given boundary.

work for general cases without explicit first passage density.

The explicit first passage density of Brownian motion induces also the consideration of marker processes [81, 90, 41], where the system is described by the observed marker process and a hidden degradation process. Here the hidden process is expressed by a drifted Brownian motion [90]. In such a case, the failure level and the hidden degradation process are jointly determined from failure records and inspection records. It is remarked that this consideration should not be confused with the filtering problem (e.g. [68]). The hidden process and the estimated failure level don't reflect the real system dynamic.

Specifically in this chapter, the stochastic degradation model is supposed as a time-dependent Ornstein-Uhlenbeck (OU) process [2, 20, 92]. This process shows good properties in degradation modeling, especially from its mean-reversion and controllable flexibility on mean, variance and covariance [20]. Moreover, although the general IFPT problem is still open, the IFPT problem for diffusion processes (including the time-dependent OU process) is proved to be well-posed from recent results by Chen etc. [12, 13]. From [12], the solution to the IFPT problem for the time-dependent OU process exists with uniqueness (under the sense of viscosity solution). These results provide the solid mathematical foundation to discuss the IFPT problem in this chapter.

3.1.2 Problem Statement

Supposing the degradation process is described by a stochastic process $X_t, t \geq 0$ defined on a complete probability space $(\Omega, \mathfrak{F}, \mathbf{P})$, the following time-dependent Ornstein-Uhlenbeck process X_t will be considered throughout this chapter [2, 20, 92] :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0, \tag{3.1.1}$$

where $a(t), b(t), \sigma(t) \in C^1(R^+)$, B_t is a standard Brownian motion and $X(0) = x_0$ is a constant.

Remarks : The model in this chapter is discussed without initial uncertainties.

It is derived from Equation (3.1.1) for the initial observation $(x_0, 0)$ that

$$X_t = X_t^{x_0,0} = e^{-\alpha(t,0)}\left[x_0 - \beta(t,0) + \int_0^t \sigma(u)e^{\alpha(u,0)}dB_u\right], \quad t \geq 0, \tag{3.1.2}$$

where α, β, γ are defined in (1.3.3).

The transition pdf $p(x, t|y, s), t > s$ of X_t can be solved directly when y is deterministic :

$$p(x,t|y,s) = \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}} \exp(-\frac{(xe^{\alpha(t,s)} + \beta(t,s) - y)^2}{4\gamma(t,s)}), \tag{3.1.3}$$

where α, β, γ are defined in (1.3.3).

Furthermore the conditional first passage time $\tau_{y,s}$ of X_t based on the observation (y, s) to a given failure level $L(t) \in C^1[s, +\infty)$ is defined as :

$$\tau_{y,s} := \inf_{t \geq s}\{t|X_t \geq L(t)\}. \tag{3.1.4}$$

The failure level influences a lot for PHM based on such a model-based first passage failure, and it is hard to promise its consistence to observed failure records when the failure

level is estimated from physical barriers or experts' judgments. Therefore the question comes naturally : whether the failure level can be estimated in a way such that the first passage failure can fit well the observed failure records ? This is true when the degradation process is considered as the OU process, where corresponding failure level is given by solving the IFPT problem.

In this chapter, the pdf of $\tau_{x_0,0}$ is denoted by $p_{\tau_{x_0,0}}(t)$ depending on $L(t)$, and the lifetime density function $g(t)$ is supposed to be given or estimated from failure records. Then the IFPT problem is to find the failure level $L(t)$ such that

$$p_{\tau_{x_0,0}}(t) = g(t). \tag{3.1.5}$$

Or equivalently, given $G(t)$ as the lifetime distribution function, the IFPT problem is to find $L(t)$ such that

$$P(\tau_{x_0,0} \leq t) = G(t). \tag{3.1.6}$$

From [12, 13], the existence and uniqueness of the solution to this IFPT problem for a general diffusion process are promised. However the analytical solutions are impossible except for few cases, therefore this chapter concentrates on how to numerically solve the IFPT problem, and several methods will be derived in the following.

3.2 Initial Boundary Estimation

As the information in the IFPT problem are only the lifetime distribution and the stochastic degradation process, therefore to connect such information with the failure level, we first consider the asymptotic analysis of the first passage distribution. Interestingly based on such an analysis, the failure level near the initial time can be approximately estimated. In the following we call the failure level recovered near the initial time as initial boundary. In later analysis, we will extend this initial boundary estimation to the boundary at later times by an iterative procedure.

In this section, it concerns two points as follows :

1. A preliminary result to estimate the initial boundary in [13] for the Brownian motion is introduced.
2. The result in [13] is extended to the OU process by the technique of time-change.

3.2.1 Preliminary Results on Brownian Motion

Several preliminary results have been already derived for Brownian motion. For example, the initial estimate based on pdf is proposed for Brownian motion in [56]. However we are more interested in the following result. As proved in [13], there is an estimate of $\lim_{t\to 0} L(t)$ for Brownian motion which is given by the following Lemma.

Lemma 3.2.1. *[13] For a standard Brownian motion B_t, $L(t)$ is the corresponding upper boundary for the first passage time $\tau := \inf_{t\geq 0}\{B_t \geq L(t)\}$, $G(t) := P(\tau \leq t)$. Assume that*

$$\lim_{t\to+0} \sup \frac{G(t)}{tG'(t)} < +\infty, \tag{3.2.1}$$

then

$$\lim_{t\to+0} \frac{L(t)}{\sqrt{-2t\log G(t)}} = 1. \tag{3.2.2}$$

Proof. See [13], and the sign is changed from the symmetry of Brownian motion to fit the discussion of upper boundaries in this thesis. □

3.2.2 First Passage Time via Time-change

To apply results of Brownian motion to OU process, in this subsection a connection will stated for FPTs between OU process and Brownian Motion. First, it is noticed that OU process can be treated as a Gauss-Markov process under a time-change which has been given in Section 1.3.2. Denote $\rho(t) := 2\gamma(t,0), m(t) := e^{-\alpha(t,0)}(y - \beta(t,0)), n(t) := e^{-\alpha(t,0)}$, and suppose the initial value is y at time 0. It comes that

$$X_t = e^{-\alpha(t,0)}\left[y - \beta(t,0) + \int_0^t \sigma(u)e^{\alpha(u,0)}dB_u\right] = m(t) + n(t)B_{\rho(t)}, \tag{3.2.3}$$

with $\rho(0) = 0$.

This time-change establishes the connection between OU process and Brownian motion, and the connection between their first passage times can be derived as in the following lemma. This can be also found in [87].

Lemma 3.2.2. *Suppose* $\tau := \inf_{t\geq 0}\{X_t \geq L(t)\}$ *and* $\tau^* := \inf_{s\geq 0}\{B_s \geq (L(\rho^{-1}(s)) - m(\rho^{-1}(s)))/n(\rho^{-1}(s))\}$. *And denote respectively the pdfs for* τ *and* τ^* *by* $p_\tau(t) := \frac{dP(\tau\leq t)}{dt}$ *and* $p_{\tau^*}(t) := \frac{dP(\tau^*\leq t)}{dt}$. *We have*

$$p_\tau(t) = p_{\tau^*}(\rho(t))\rho'(t) \tag{3.2.4}$$

Proof. Noticing $\rho(t)$ is strictly increasing if $\sigma(t) \neq 0$, the inverse ρ^{-1} exists.

$$\begin{aligned} P(\tau > t) &= P(X_s < L(s), s \leq t) = P(m(s) + n(s)B_{\rho(s)} < L(s), s \leq t) \\ &= P(B_s < (L(\rho^{-1}(s)) - m(\rho^{-1}(s)))/n(\rho^{-1}(s)), s \leq \rho(t)) \\ &= P(\tau^* > \rho(t)). \end{aligned} \tag{3.2.5}$$

Differentiate Equation (3.2.5), it comes naturally to (3.2.4). □

Remark : even if the lemma presented here is helpful to calculate the first passage time of the OU process based on the time-change, it is not convenient to calculate the first passage problem in general cases except when the inverse of time-change ρ^{-1} is explicit. And that is also the motivation why we reproduce the tedious analysis on the first passage time of the OU process rather than calculating based on the time-change from the results for Brownian motion.

Based on Lemma 3.2.2, the following proposition comes :

Proposition 3.2.3. *Suppose under the passage boundary* $L(t)$, *the corresponding first passage density* $p_\tau(t)$ *for* X_t *is given. And under* $L_1(t)$, *the corresponding first passage density* $p_{\tau^*}(t)$ *is given for Brownian motion. If* $p_\tau(t)$ *and* $p_{\tau^*}(t)$ *satisfy the equation* (3.2.4), *then we have*

$$L(t) = L_1(\rho(t))n(t) + m(t) \tag{3.2.6}$$

Proof. The existence and uniqueness of the solution to the IFPT problem for diffusion process have been derived in [13]. And here we just state that if (3.2.4) holds, then $L_1(\rho(t))n(t)+m(t)$ would be one solution satisfying the IFPT problem for the time-dependent OU process (3.1.1).

Actually, denote $L_2(t) = L_1(\rho(t))n(t) + m(t)$, then we have

$$\tau^* = \inf_{s>0}\{B_s \geq L_1(s)\} = \inf_{s>0}\{B_s \geq (L_2(\rho^{-1}(s)) - m(\rho^{-1}(s)))/n(\rho^{-1}(s))\}. \tag{3.2.7}$$

From Lemma 3.2.2, we know that for $\tau_1 := \inf_{t>0}\{X_t \geq L_2(t)\}$,

$$p_{\tau_1}(t) = p_{\tau^*}(\rho(t))\rho'(t).$$

From the given condition (3.2.4), we have $p_\tau = p_{\tau^*}(\rho(t))\rho'(t) = p_{\tau_1}$, and due to the uniqueness of the solution to the IFPT problem, it comes to $L = L_2$. □

Following Lemmas 3.2.1, Lemma 3.2.2 and Proposition 3.2.3, the following proposition is given to estimate the initial boundary for the FPT $\tau_{x_0,0}$ of $X_t^{x_0,0}$ to the upper boundary $L(t)$.

Proposition 3.2.4. *Following notations in (3.2.3), denote* $\tau := \inf_{t>0}\{m(t) + n(t)B_{\rho(t)} \geq L(t)\}$, *suppose* $G(t) := P(\tau \leq t)$ *is already known. Assume that*

$$\lim_{t\to+0}\sup\frac{\rho'(t)G(t)}{\rho(t)G'(t)} < +\infty. \tag{3.2.8}$$

Then

$$\lim_{t\to+0}\frac{L(t) - m(t)}{n(t)\sqrt{-2\rho(t)\log G(t)}} = 1. \tag{3.2.9}$$

Proof. From Lemma 3.2.2, for Brownian motion B_t induced by (3.2.3), $\tau^* := \inf_{s\geq 0}\{B_s \geq (L(\rho^{-1}(s)) - m(\rho^{-1}(s)))/n(\rho^{-1}(s))\}$, the following equation holds :

$$G_{\tau^*}(\rho(t)) := P(\tau^* \leq \rho(t)) = P(\tau \leq t) = G(t). \tag{3.2.10}$$

From Lemma 3.2.1 and the hypothesis (3.2.8), noticing $t \to \rho(t)$ is a bijection from $\mathbb{R}^+$ to $\mathbb{R}^+$, then

$$\lim_{\rho(t)\to+0}\sup\frac{G_{\tau^*}(\rho(t))}{\rho(t)G'_{\tau^*}(x)|_{x=\rho(t)}} = \lim_{t\to+0}\sup\frac{\rho'(t)G(t)}{\rho(t)G'(t)} < +\infty. \tag{3.2.11}$$

Therefore the condition (3.2.8) in Lemma 3.2.1 is satisfied for Brownian motion $B_{\rho(t)}$. Then by Lemma 3.2.1, its boundary $L_1(\rho(t))$ can be estimated :

$$\lim_{t\to+0}\frac{L_1(\rho(t))}{\sqrt{-2\rho(t)\log G_{\tau^*}(\rho(t))}} = 1. \tag{3.2.12}$$

By Proposition 3.2.3, the result can be translated to $X_t = m(t)+n(t)B_{\rho(t)}$ directly, which leads to Equation (3.2.9). □

Corollary 3.2.5. *Consider* $G(t|y,s) := P(\tau_{y,s} \leq t)$ *and* $g(t|y,s) = \frac{\partial G(t|y,s)}{\partial t}$ *are known for the process* $X_t = e^{-\alpha(t,s)}(y - \beta(t,s) + \int_s^t \sigma^2(z)e^{\alpha(z,s)}dB_z)$, *assume that*

$$\lim_{t\to+s}\sup\frac{\gamma'(t,s)G(t|y,s)}{\gamma(t,s)g(t|y,s)} < +\infty, \tag{3.2.13}$$

then

$$\lim_{t\to+s}\frac{L(t)e^{\alpha(t,s)} + \beta(t,s) - y}{\sqrt{-4\gamma(t,s)\log G(t|y,s)}} = 1. \tag{3.2.14}$$

Proof. From Proposition 3.2.4, set the initial value as (y, s), denote $\rho(t) := 2\gamma(t, s), m(t) := e^{-\alpha(t,s)}(y - \beta(t, s)), n(t) = e^{-\alpha(t,s)}$, the result comes naturally. □

Remark : This limit gives the estimation of $L(t)$ when t is near 0, such that

$$L(t) \approx e^{-\alpha(t,0)}(x_0 - \beta(t, 0) + \sqrt{-4\gamma(t, 0) \log G(t)}). \tag{3.2.15}$$

However, due to the calculation limit of the computer to calculate $\lim_{t\to 0} \log G(t) = \infty$, the estimation may not exist when t is too small. Therefore it should be careful to test this estimate several times to balance the accuracy and the computability.

3.3 Integral Equation Method

3.3.1 Fortet's Equation

In this subsection, the following conditions are considered :

$$\sigma(t) > 0, a(t) \leq 0, \forall t \in [0, +\infty). \tag{3.3.1}$$

On one hand, the first passage density p_τ to the failure time τ satisfies Fortet's equation corresponding to the failure level $L(t)$ [57] :

$$p(x, t|x_0, 0) = \int_0^t p_\tau(s)p(x, t|L(s), s)ds, \quad \forall x \geq L(t), \tag{3.3.2}$$

where $p(x, t|y, s)$ is the transition pdf defined in (3.1.3).

On the other hand, the IFPT problem is proved to be a well-posed problem such that if the lifetime pdf $g(t)$ is given, the solution $L(t)$ should also satisfy Equation (3.3.2), i.e.

$$p(x, t|x_0, 0) = \int_0^t g(s)p(x, t|L(s), s)ds, \quad \forall x \geq L(t), \tag{3.3.3}$$

therefore the remaining task is to solve such an equation to find the failure level $L(t)$.

An iterative procedure is adopted in this section to solve Equation (3.3.3). Supposing $L(s)$ is known for all $s \in [0, t)$, let denote in this section

$$\Gamma(x, t) := p(x, t|x_0, 0), \quad \Psi(x, t) := \int_0^t g(s)p(x, t|L(s), s)ds \tag{3.3.4}$$

From (3.3.3), the current boundary $L(t)$ will just be the minimal zero solution of the following equation :

$$Z(x, t) := \Gamma(x, t) - \Psi(x, t) = 0. \tag{3.3.5}$$

That is to say,

$$L(t) = \inf_{x\in R}\{x|Z(x, t) = 0\}. \tag{3.3.6}$$

Here the original IFPT problem has been converted to the problem of solving Equation (3.3.6). For such a topic, various classical numerical schemes can be possibly used such as secant method, Newton's method, bisection method etc. However to promise the efficiency and correctness of the calculation, more analysis is necessary. Especially in this iterative solving procedure, the following questions are essential :

- How is $\Psi(x, t)$ expressed in a computable form ?
- How is L(t) verified as the minimal solution of Equation (3.3.5) ?

In the last section, the initial boundary has been estimated explicitly, and the above two questions will be discussed in this section.

3.3.2 Discretization of $\Psi(x,t)$

The integral $\Psi(x,t)$ in (3.3.4) can hardly be expressed explicitly, which induces the need of approximate expression. However as $\lim_{s\to t} p(L(t),t|L(s),s) = +\infty$, $\Psi(x,t)$ is with a singular point at t. So for later numerical calculation, the integral $\Psi(x,t)$ should be considered carefully.

From (1.3.3), it is noticed first that :

$$\frac{\partial\gamma(t,s)}{\partial s} = \int_s^t c(u)e^{\alpha(u,s)} \times 2a(s)du - c(s) = 2a(s)\gamma(t,s) - c(s). \tag{3.3.7}$$

Noticing as in Chapter 1, $c(t) = \frac{\sigma^2(t)}{2} > 0$, therefore $\gamma(t,s) = \int_s^t c(u)e^{\alpha(u,s)}du$ is a decreasing function for the variable s. That is to say, the following inequality holds :

$$\frac{\partial\gamma(t,s)}{\partial s} = 2a(s)\gamma(t,s) - c(s) \leq \eta < 0, \tag{3.3.8}$$

for some $\eta \in \mathbb{R}$ when s is in any closed set in $\mathbb{R}^+$. This is derived from the constraint (3.3.1) for $a(s)$ and $\sigma(s)$.

Furthermore Equation (3.3.7) contributes to :

$$d\sqrt{\gamma(t,s)} = \frac{1}{2\sqrt{\gamma(t,s)}}\frac{\partial\gamma(t,s)}{\partial s}ds. \tag{3.3.9}$$

Based on (3.3.8) and (3.3.9), now another expression for $\Psi(x,t)$ is given :

$$\begin{aligned}\Psi(x,t) &= \int_0^t g(s)p(x,t|L(s),s)ds \\ &= \int_0^t g(s)\frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\exp(-\frac{((xe^{\alpha(t,s)}+\beta(t,s))-L(s))^2}{4\gamma(t,s)})ds \\ &= \int_0^t g(s)\frac{e^{\alpha(t,s)}}{\sqrt{\pi}(2a(s)\gamma(t,s)-c(s))}\exp(-\frac{((xe^{\alpha(t,s)}+\beta(t,s))-L(s))^2}{4\gamma(t,s)})d\sqrt{\gamma(t,s)}\end{aligned} \tag{3.3.10}$$

Remark : the singularity in Equation (3.3.10) is avoided as for any $s \in [0,t]$,

$$|\frac{g(s)e^{\alpha(t,s)}}{\sqrt{\pi}(2a(s)\gamma(t,s)-c(s))}\exp(-\frac{((xe^{\alpha(t,s)}+\beta(t,s))-L(s))^2}{4\gamma(t,s)})| \leq \frac{g(s)e^{\alpha(t,s)}}{\sqrt{\pi}|2a(s)\gamma(t,s)-c(s)|}, \tag{3.3.11}$$

which is bounded from Equation (3.3.8).

Following Equation (3.3.10), denoting

$$\Omega(x,t|y,s) := \frac{e^{\alpha(t,s)}}{\sqrt{\pi}(2a(s)\gamma(t,s)-c(s))}\exp(-\frac{((xe^{\alpha(t,s)}+\beta(t,s))-y)^2}{4\gamma(t,s)}), \tag{3.3.12}$$

an approximate expression by compound trapezoid scheme is adopted for $\Psi(x,t)$:

$$\Psi_n(x,t) \approx \frac{1}{2}\sum_{i=0}^{n-1}(\sqrt{\gamma(t,t_{i+1})} - \sqrt{\gamma(t,t_i)})\,[g(t_{i+1})\Omega(x,t|L(t_{i+1}),t_{i+1}) + g(t_i)\Omega(x,t|L(t_i),t_i)] \tag{3.3.13}$$

where $n \in N, t_0 = 0, t_n = t, \{t_i\}_{i=0}^n$ are mesh points for $[0, t]$. In such a form, it is noticed some items vanish, especially $\Omega(L(t), t|L(t), t) = \frac{-1}{\sqrt{\pi}c(t)}$ from Equation (2.4.1) and $g(0) = 0$.

We should notice that the original equation $\Gamma(x, t) - \Psi(x, t) = 0$ is now approximated by a new equation $\Gamma(x, t) - \Psi_n(x, t) = 0$, a natural question is whether the solution to the approximate equation can also approximate the real solution to the original equation. This question is beyond our concentration here, but it is reasonable in general. A partial answer can refer to approximation optimization in Appendix B.

3.3.3 The Minimal solution to $Z(x, t) = 0$

In this subsection, we will address how to verify that the solution is the minimal solution of Equation (3.3.5). Due to the properties of the function $Z(x, t)$, it is found that this verification can be promised by a simple step in the numerical algorithm. Actually, we have

- $Z(x, t) = 0, \forall x \geq L(t)$.
- $\frac{\partial Z(x,t)}{\partial x}|_{x=L(t)-} = -g(t)/c(t) < 0$, when $g(t) \neq 0$.

The first point is just Fortet's equation (2.2.1). And the second point is derived from Equation (2.2.4) and (2.3.14).

Therefore $Z(x, t) > 0$ holds for $x \in (L(t) - \epsilon, L(t))$ with a small ϵ. When searching $L(t)$ by the secant method, if we start with a S near the real solution $L(t)$ such that $Z(S, t) > 0$, the iteration converges naturally to $L(t)$ due to the monotonicity of $Z(x, t)$ w.r.t x without the risk of overestimation.

3.3.4 Numerical Scheme

Based on previous results, a numerical scheme is proposed as follows. Considering the IFPT problem on $[0, T]$, which is divided by $\{t_i\}_{i=0}^N$, where N is finite, the approximate value L_i will be calculated for the failure level $L(t_i)$ based on the secant method. And in such an algorithm $Z(x, t_i) \approx \Gamma(x, t) - \Psi_n(x, t)$ is applied by (3.3.13).

1. Initial failure level : $L_0 = x_0$, and by (3.2.9) for t_1 small enough,
$$L_1 \approx e^{-\alpha(t_1,0)}(x_0 - \beta(t_1, 0) + \sqrt{-4\gamma(t_1, 0)\log(g(t_1)t_1)}.$$
2. Initial guess : for $i \geq 2$, suppose $L_1, ..., L_{i-1}$ are calculated, then a good initial estimate of L_i is essential for later calculation. However there is a risk of over-estimating of L_i as $Z(x, t_i) = 0, \forall x \geq L(t_i)$. So the initial guess S_i^0 for L_i is given by L_{i-1} from the continuity of $L(t)$, and it is modified by the following procedure :
   - $S_i^0 \leftarrow L_{i-1}$;
   - while $(Z(S_i^0, t_i) \leq \epsilon)$, $\quad S_i^0 \leftarrow S_i^0 - c \times (t_i - t_{i-1})$,

   where c is positive to control the reduction of initial guess, and ϵ is an enough small digit.
3. Secant Method : following the estimate S_i^0 in Step 2, we calculate iteratively S_i^j for L_i :
$$\begin{cases} S_i^1 = S_i^0 - c \times (t_i - t_{i-1}), \\ S_i^j = S_i^{j-1} - Z(S_i^{j-1}, t_i)\frac{S_i^{j-1} - S_i^{j-2}}{Z(S_i^{j-1}, t_i) - Z(S_i^{j-2}, j-2)}, \quad j \geq 2. \end{cases} \tag{3.3.14}$$
   The iteratively solving process is stopped when S_i^j is enough close to the previous iterative value :
$$|S_i^j - S_i^{j-1}| \leq \epsilon. \tag{3.3.15}$$

Then the value of S_i^j will be assigned to L_i.

4. Iterative procedure : after L_i is calculated, let $i \leftarrow i+1$, and turn back to Step 2.

3.4 Another Integral Equation

3.4.1 The Master Equation

The previously stated numerical algorithm adopts a variable-change trick to discretize the singular integral $\Psi(x,t)$. And this can also be avoided in other ways. In this section the master equation (2.2.7) will be adopted to calculate the failure level. Such an idea is reproduced based on [57, 97].

From Lemma 2.2.2 for X_t in (3.1.1) with an initial start $(x_0, 0)$, (2.2.7) holds,

$$1 - F(x,t|x_0,0) = \int_0^t (1 - F(x,t|L(s),s))g(s)ds, \quad \forall x \geq L(t). \tag{3.4.1}$$

Here note $\Phi(*)$ as the cdf of normal distribution,

$$F(x,t|y,s) = \Phi(\frac{xe^{\alpha(t,s)} + \beta(t,s) - y}{2\sqrt{\gamma(t,s)}}). \tag{3.4.2}$$

As $\Phi(*)$ is an easily calculated function, Equation (3.4.1) is expressed in a computable form based on (3.4.2) :

$$1 - \Phi(\frac{xe^{\alpha(t,0)} + \beta(t,0) - x_0}{2\sqrt{\gamma(t,0)}}) = \int_0^t (1 - \Phi(\frac{xe^{\alpha(t,z)} + \beta(t,z) - L(z)}{2\sqrt{\gamma(t,z)}}))g(z)dz, \quad \forall x \geq L(t). \tag{3.4.3}$$

Following iterative procedures in Section 3.3, to solve the equation (3.4.3). Supposing $L(s)$ is known for all $s \in [0,t)$, let denote in this section

$$\tilde{\Omega}(x,t) := 1 - \Phi(\frac{xe^{\alpha(t,0)} + \beta(t,0) - x_0}{2\sqrt{\gamma(t,0)}}), \quad \tilde{\Psi}(x,t) := \int_0^t (1 - \Phi(\frac{xe^{\alpha(t,z)} + \beta(t,z) - L(z)}{2\sqrt{\gamma(t,z)}}))g(z)dz, \tag{3.4.4}$$

From (3.4.3), the current boundary $L(t)$ will just be the minimal zero solution of the following equation :

$$\tilde{Z}(x,t) := \tilde{\Omega}(x,t) - \tilde{\Psi}(x,t) = 0. \tag{3.4.5}$$

That is to say,

$$L(t) = \inf_{x \in R}\{\tilde{Z}(x,t) = 0\}. \tag{3.4.6}$$

Here the original IFPT problem has been converted to the problem of solving Equation (3.4.6).

Not similar with the singular integral in (3.3.4), Equation (3.4.4) is without problems to calculate. However the integral $\Psi(x,t)$ in (3.4.4) can also hardly be expressed explicitly, which induces the need of approximate expression. By compound trapezoid scheme, denoting

$$\tilde{\Omega}(x,t|y,s) := 1 - \Phi(\frac{xe^{\alpha(t,s)} + \beta(t,s) - y}{2\sqrt{\gamma(t,s)}}), \tag{3.4.7}$$

an approximate expression is proposed for $\tilde{\Psi}(x,t)$ based on the compound trapezoid rule :

$$\tilde{\Psi}_n(x,t) \approx \frac{1}{2}\sum_{i=0}^{n-1}(t_{i+1}-t_i)\left[g(t_{i+1})\tilde{\Omega}(x,t|L(t_{i+1}),t_{i+1}) + g(t_i)\tilde{\Omega}(x,t|L(t_i),t_i)\right], \tag{3.4.8}$$

where $n \in N, t_0 = 0, t_n = t, \{t_i\}_{i=0}^n$ are mesh points for $[0,t]$. In such a form, some items vanish, especially $\tilde{\Omega}(L(t),t|L(t),t) = 1/2$ from Lemma 2.4.1, $g(0) = 0$.

We should notice the original equation $\tilde{\Gamma}(x,t) - \tilde{\Psi}(x,t) = 0$ is now approximated by a new equation $\tilde{\Gamma}(x,t) - \tilde{\Psi}_n(x,t) = 0$, a natural question is whether the solution to the approximate equation can also approximate the real solution to the original equation. This question is not our concentration here, but it is reasonable in general. A partial answer can refer to approximation optimization in Appendix B.

3.4.2 Numerical Scheme

Based on previous results, a numerical scheme is proposed as follows. The main procedures are the same with the one in Section 3.3, however with different interior technical calculation. Considering the IFPT problem on $[0,T]$, which is divided by $\{t_i\}_{i=0}^N$, where N is finite, the approximate value L_i will be calculated for the failure level $L(t_i)$ based on secant method. And in such an algorithm $\tilde{Z}(x,t_i) \approx \Gamma(x,t) - \tilde{\Psi}_n(x,t)$ is applied by (3.4.8).

1. Initial failure level : $L_0 = x_0$, and by (3.2.9) for t_1 small enough,
$$L_1 \approx e^{-\alpha(t_1,0)}(x_0 - \beta(t_1,0) + \sqrt{-4\gamma(t_1,0)\log(g(t_1)t_1)}.$$
2. Initial guess : for $i \geq 2$, suppose $L_1, ..., L_{i-1}$ are calculated, then a good initial estimate of L_i is essential for later calculation. However there is a risk of over-estimating of L_i as $\tilde{Z}(x,t_i) = 0, \forall x \geq L(t_i)$. So the initial guess S_i^0 for L_i is given by L_{i-1} from the continuity of $L(t)$, then it is modified by the following procedure :
   - $S_i^0 \leftarrow L_{i-1}$;
   - while$(\tilde{Z}(S_i^0,t_i) \leq \epsilon)$, $\quad S_i^0 \leftarrow S_i^0 - c \times (t_i - t_{i-1})$,

   where c is positive to control the reduction of initial guess, and ϵ is an enough small digit.
3. Secant Method : following the estimate S_i^0 in Step 2, we calculate iteratively S_i^j for L_i :
$$\begin{cases} S_i^1 = S_i^0 - c \times (t_i - t_{i-1}), \\ S_i^j = S_i^{j-1} - \tilde{Z}(S_i^{j-1},t_i)\frac{S_i^{j-1} - S_i^{j-2}}{\tilde{Z}(S_i^{j-1},t_i) - \tilde{Z}(S_i^{j-2},j-2)}, \quad j \geq 2. \end{cases} \tag{3.4.9}$$
   The iteratively solving process is stopped when S_i^j is enough close to the previous iterative value :
$$|S_i^j - S_i^{j-1}| \leq \epsilon. \tag{3.4.10}$$
   Then the value of S_i^j will be assigned to L_i.
4. Iterative procedure : after L_i is calculated, let $i \leftarrow i+1$, and turn back to Step 2.

3.5 Simulation Tests

3.5.1 Experimental Setup

Hereafter we will consider a model described in Table 1.2 which is estimated from the degradation data of a passive component in power plants. In Chapter 1, we have discussed

the modeling work, and the following process X_t is given :

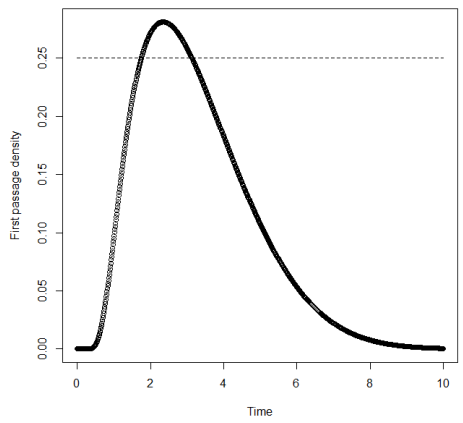
$$dX_t = (-r(X_t - m(t)) + m'(t))dt + \sigma dB_t, t \geq 0, X_0 = x_0, \tag{3.5.1}$$

where $r = 0.1806708, \sigma = 2.4640884, m(t) = 2.4402845((t+1)^{0.8892020} - 1) + x_0, x_0 = 2.8074561$. Some trajectories of the process are produced based on Monte-Carlo simulation, which are shown in Figure 1.4(b).

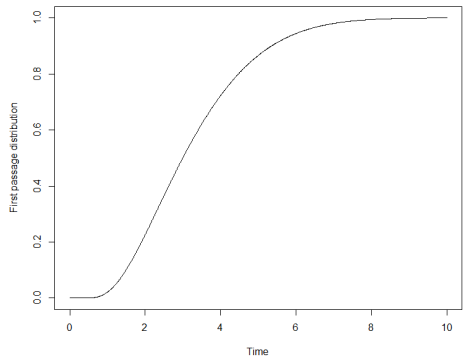
Other issues related to the first passage problem are chosen as follows :

1. The considered time interval is $[0, 10]$.
2. The step-size Δt to discretize the time interval is 0.01.
3. The lifetime density $g(t)$ and distribution $G(t)$ are produced based on the first passage time to the boundary $L(t) \equiv 10$.

The pdf and cdf of the first passage time have been calculated based on the Volterra integral equation method (2.4.10). And their appearance is shown in Figure 3.2(a) and 3.2(b).



(a) First passage density to $L(t) = 10$



(b) First passage distribution to $L(t) = 10$

FIGURE 3.2 – The produced lifetime pdf and cdf for calculation.

In the following simulation tests, it is aimed to reproduce the original failure level $L = 10$ by solving the IFPT problem.

Remark : The time-dependent boundary is not discussed, as it can be transformed to the process itself based on the variable-change in (2.3.15).

3.5.2 The Estimate of Initial Boundary

The first issue to test is the accuracy of the initial boundary given in (3.2.15), i.e.

$$L(t) \approx e^{-\alpha(t,0)}(x_0 - \beta(t,0) + \sqrt{-4\gamma(t,0)\log G(t)}). \tag{3.5.2}$$

And here we want to see its performance when different times are proposed. The estimated failure level from the initial estimation can be seen in Figure 3.3(a) and a special consideration near zero can be seen in Figure 3.3(b). We should notice when time is near 0, the estimate is with a very high accuracy. Actually the estimate returns $L(0.01) \approx 9.979214$. Later calculation will be based on this initial boundary value.

3.5.3 Integral Equation Methods

The Fortet's Equation Method

The Fortet's equation method is fulfilled here, and the reproduced failure level is given in Figure 3.4(a). Corresponding errors are shown in Figure 3.4(b).

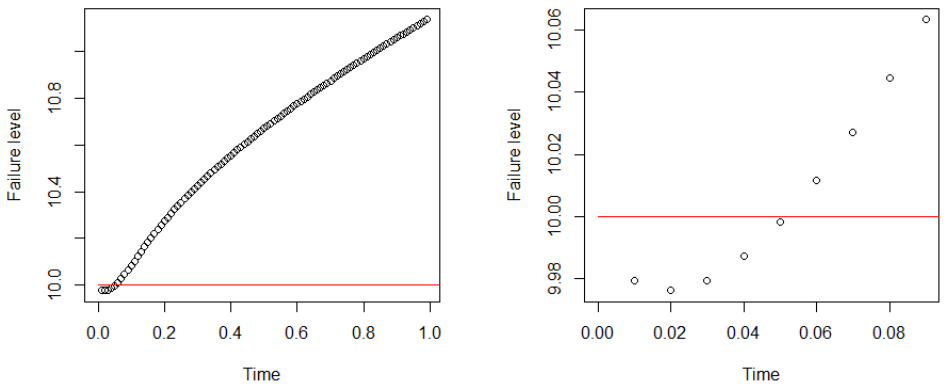
From the result, we can derive

1. The error of the estimate increases rapidly as the time increases. This is due to the discretization error of the numerical integral, which is based on the difference of the changed time $\sqrt{\gamma(t,s)}$. And we should notice with the equal time step-size, this discretization will induce rapidly increasing errors as $\sqrt{\gamma(t,s)}$ is exponentially increasing. This would be improved by consider non-equal time step-size, and design the step-size according to the time-change.
2. The estimate has a sudden drop near the initial time. This is because of the natural limit of the numerically solving procedure. As it is impossible for the computer to recognize $Z(x,t) = 0$, instead it is frequent to judge $Z(x,t) < \epsilon$ for a small ϵ. This treatment will tolerate small errors. Therefore at the beginning, the initial guess for the solving procedure of the boundary is so good such that $Z(x,t) < \epsilon$ for the initial guess and the solving is stopped immediately and returns the result. But when the initial guess doesn't satisfy $Z(x,t) < \epsilon$ for some time, the solving procedure starts to search a result which is different with the initial guess. That is why a sudden drop appears in the estimate.

The Master Equation Method

The master equation method is fulfilled here, and the reproduced failure level is given in Figure 3.5(a). Corresponding errors are shown in Figure 3.5(b).

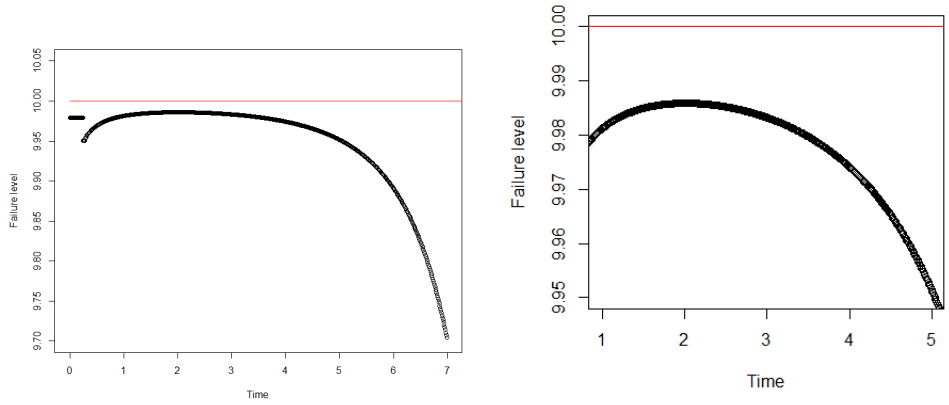
1. The estimate is very accurate and performs much better than the Fortet's equation method. This verifies also the guess for the error of the Fortet's equation, as here



(a) Estimated initial boundary (black) compared with the original boundary (red)

(b) The performance near zero, $L(0.01) \approx 9.98$.

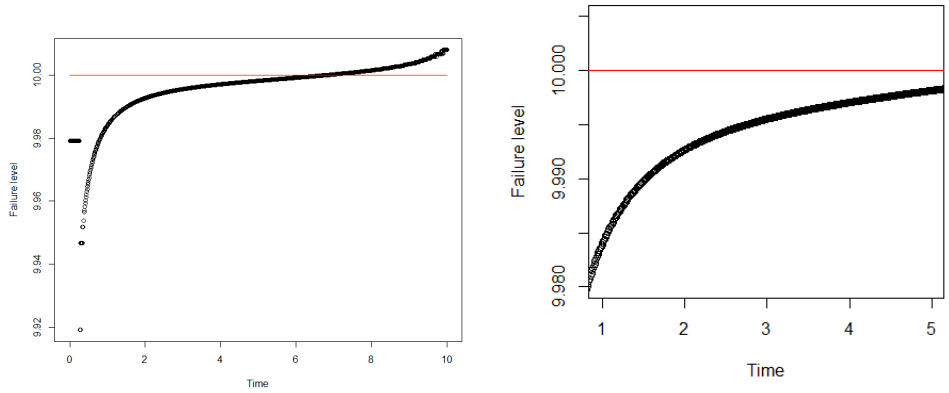
FIGURE 3.3 – The estimation on the initial boundary



(a) Estimated boundary (black) compared with the original boundary (red)

(b) The estimate near the peak of the lifetime distribution.

FIGURE 3.4 – The estimation on the boundary based on Fortet's equation.



(a) Estimated boundary (black) compared with the original boundary (red)

(b) The estimate near the peak of the lifetime distribution.

FIGURE 3.5 – The estimation on the boundary based on the Master equation.

the discretization error of the numerical integral is much less than the ones appeared in Fortet's equation method. Here when discretizing, the time is not changed and it increases linearly.

2. There is also a sudden drop for the estimates near the initial time. The reason has been explained before in the Fortet's equation method.

3.6 Conclusions and Perspectives

In this chapter, the inverse first passage problem is introduced for the time-dependent OU process. Such a problem aims to reproduce the failure level such that corresponding first passage failure can fit given failure records. This makes up the gap between the prognosis based on first passage failures and existing failure records. Integral equation methods based on Fortet's equation and master equation have been proposed in this chapter to solve this problem respectively. Corresponding numerical tests are done to verify proposed algorithms.

The effort to make up the gap between failure records and inspection records are made from various considerations. Generally speaking, 2 ways exist to fulfill this task when boundary-crossing failures are considered. One is to modify the description of system failures, such as introducing last passage failure [59], passage duration [9] and killed FPT [29]. Such efforts consider the failure level is fixed but the failure description itself can be modified.

The other is to treat the failure level as pending which can be determined based on given failure records. IFPT problem is therefore introduced to reproduce the failure level. We should notice the OU process is a simple process, such that it would be much more difficult when it comes to other diffusion processes. To extend the inverse first passage problem into other processes, such as jump processes or jump-diffusion processes could also be interesting.

Chapitre 4

Maintenance Optimization for Continuously Monitored Systems

In this chapter, the condition-based maintenance based on the OU process will be discussed for continuously monitored systems. This is the direct application of Chapter 1 on the modeling work of degradation processes, Chapter 2 and Chapter 3 on the description and prognosis of system failures. These issues are used for condition-based maintenance to prevent the system failures.

Based on monitored system conditions and the prognosis of system failures, the condition-based maintenance introduces preventive maintenance such that the balance can be achieved between operation costs and disastrous results caused by system failures. In this chapter, corresponding maintenance optimization problems are discussed based on the time-dependent OU process and the hypothesis of continuously monitored system. Due to the unexplicit expression for prognosis of system failures, classical heuristic optimization procedures cannot be fulfilled. Therefore the approximate first passage density is introduced to fulfill the maintenance optimization.

The whole chapter is organized as follows. In Section 4.1, basic facts for condition-based maintenance, description on the continuously monitored system and regenerative processes will be introduced. In Section 4.2, the maintenance optimization problems will be proposed for the time-dependent OU process based on the cost criterion and the availability criterion. In Section 4.3, the idea of approximation-optimization is fulfilled to solve maintenance optimization problems for three cases : the drifted Brownian motion, the time-changed Brownian motion and the general OU process. In Section 4.4, several simulation tests will be done to solve these maintenance optimization problems. In Section 4.5, conclusions and perspectives will be given.

4.1 Introduction

4.1.1 Background

The condition-based maintenance is a current active field in the field of reliability engineering. And it is generally treated as maintenance when need arises. Compared to the planned maintenance, personally condition-based maintenance shows its power on three points :

1. Because of the technical developments, sensors or detection techniques become more convenient and also cheaper. So more and more inspection data appears in modern industrial systems, how to combine this immense observed information to improve the system performance, especially on preventing a system failure could be utilizable.
2. As the modern industrial system becomes more and more complex, the uncertainties in the system state are also increasing. These uncertainties induce a high variation in

the system states, such that the system state can hardly be predicted, or the prediction is with a high uncertainty also. Therefore the planned maintenance, which is totally based on the prediction of system states in its lifetime, can hardly process such a situation. And actually the planned maintenance gives an intuitive and simple solution to such a problem : proposing more strict strategies to prevent the risk of ignoring the uncertainty contained in the prediction.

3. It should be noticed that even the long-term system state can hardly be predicted with a high accuracy, the prediction in a short period could be trustable based on the current observation. Independent of the prediction techniques, the prediction can basically be around the current observation in a short period, if the system state is supposed to have some continuity. This also gives the reasonability to consider a more accurate control based on the current observation.

Moreover in the framework of condition-based maintenance, 2 kinds of maintenance actions are of interest : preventive maintenance and corrective maintenance. And correspondingly 2 kinds of system failures are considered : real failure and virtual failure [1]. When a real failure occurs, the corrective maintenance is performed. When a virtual failure occurs, the preventive maintenance is performed.

The key idea of maintenance optimization for continuously monitored systems is then to arrange the preventive maintenance appropriately such that proposed objective functions such as operation cost, system availability etc. can be optimized. Moreover in this thesis, the real failure and virtual failure are all considered as first passage failures discussed in Chapter 2 and Chapter 3. Therefore in this chapter, the first passage failure to the corresponding virtual failure level determines the preventive maintenance. Such hypotheses simplify the original problem of arranging the preventive maintenance to the optimization of the virtual failure level based on some objective functions.

Therefore, the strong connections among maintenance optimization, stochastic degradation modeling and prognosis of system failures are revealed : the stochastic degradation modeling provides a way to consider first passage failures, the first passage failures simplify the maintenance optimization problems, and the maintenance optimization problems return a virtual failure level below the real failure level which helps to control the system performance.

An important hypothesis on the maintenance action adopted in this chapter is "as good as new". That is to say, after each maintenance operation the system state returns to the new state. This induces the natural consideration of regenerative theory and simplifies the maintenance optimization problems.

However, it is not an easy work to move directly from the previous analysis on first passage times to the maintenance optimization. Due to the demand on a computable form of the objective function with the virtual failure level in the maintenance optimization problems, it is inconvenient to apply those numerical results which require a pre-set failure level to the optimization problem.

So in this chapter, the work is presented based on the approximation-optimization framework. Previously derived approximation of the first passage density will be introduced into the objective function and later the induced approximate expression will be optimized. We will first consider two cases with explicit first passage density : drifted Brownian motion and time-changed Brownian motion. For the general OU process, the approximate expression from Section 2.3.7 will be considered.

1. The virtual failure is generally treated as an alarm before the real failure.

4.1.2 System Description

In this chapter, we suppose the system state is inspected continuously. Following the hypotheses in [63, 7, 44], we consider maintenance optimization problems based on the minimum cost and the maximum availability respectively.

During the lifetime of the system, supposing the degradation process is described by a one-dimensional stochastic process $\{X_t, t \geq 0\}$ defined on a complete probability space $(\Omega, \mathfrak{F}, \mathbf{P})$, the following time-dependent Ornstein-Uhlenbeck process X_t will be considered throughout this chapter [20, 92, 2] :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0, \tag{4.1.1}$$

where $a(t), b(t), \sigma(t) \in C^1(R^+)$, $\sigma(t)$ is positive, B_t is a standard Brownian motion and $X_0 = x_0$ is a deterministic initial value.

Therefore X_t can be expressed explicitly by (1.3.3).

$$X_t = e^{-\alpha(t,0)}\left[x_0 - \beta(t,0) + \int_0^t \sigma(s)e^{\alpha(s,0)}dB_s\right]. \tag{4.1.2}$$

Remark : Although in maintenance problems, it is a general hypothesis to propose an initial value 0 which is translated as the new state. But this treatment is not appropriate to consider the OU process. This is due to

1. The degradation process is supposed to be independent with the real physical state, and derived from data-analysis, such that 0 doesn't have an explicit meaning in general.
2. Technically, the considered OU process is not with independent increments, such that the initial value matters the later prediction a lot.

To consider optimization problems, it is specifically supposed that :

1. The system state is monitored continuously.
2. The real failure occurs instantly when the monitored system state reaches a pre-set failure level L_r, where the failure time is denoted as $\tau_r := \inf_{t>0}\{x_t \geq L_r\}$. Also when a real failure occurs, a maintenance is decided and the system state is immediately unavailable.
3. The virtual failure occurs instantly when the monitored system state reaches a failure level $L_v \in [x_0, L_r]$, where the failure time is denoted as $\tau_v := \inf_{t>0}\{x_t \geq L_v\}$. Also when a virtual failure occurs, a maintenance is decided. But the system still works until the maintenance starts or until the real failure occurs.
4. Maintenance will be performed after a deterministic delay time κ when it is decided, and the maintenance operation has a random duration λ. When the maintenance is finished, the system state is as good as new.
5. Between τ_v and $\tau_v + \kappa$, the system deteriorates and possibly reaches the real failure level during this period. If a real failure occurs ($\tau_r \leq \tau_v + \kappa$), the system is off from failure time until the end of maintenance operation $\tau_v + \kappa + \lambda$. If a real failure doesn't occur ($\tau_r > \tau_r + \kappa$), the system is unavailable from the time $\tau_v + \kappa$ until the end of maintenance operation $\tau_v + \kappa + \lambda$.

In such a framework, maintenance optimization problem is to find a value for the virtual failure level L_v such that the objective function can be minimized. Corresponding to different emphases in application, the objective function can vary in a large range, such as cost-based optimization [63, 34], availability-based optimization [7, 44] etc..

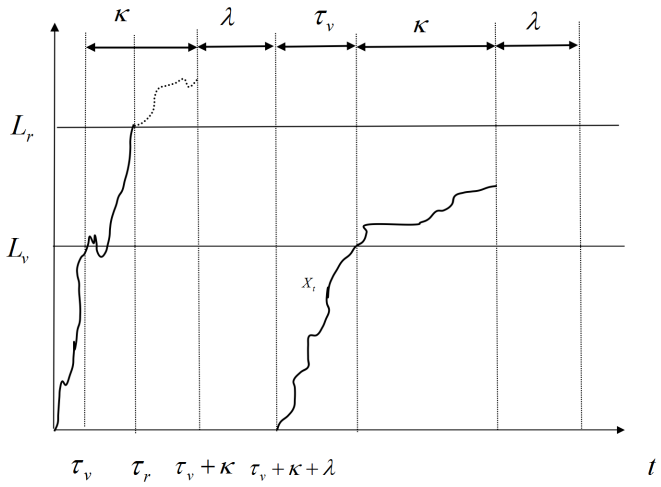


FIGURE 4.1 – Description of the maintenance policy.

TABLE 4.1 – Problem setup for the maintenance optimization

Degradation process	X_t
Maintenance delay	κ
Maintenance duration	λ
Virtual failure level	L_v
Real failure level	L_r
Virtural failure time	τ_v
Real failure time	τ_r

Summarize all the above, the problem setup is given in Table 4.1.2. And also it is illustrated in Figure 4.1.

4.1.3 Preliminary Knowledge on Regenerative Process

When the hypothesis of "as good as new" is considered for maintenance operations, the whole degradation process can be naturally considered as a regenerative process. This is because system states can return to the new state after each maintenance, which are supposed to be independent of the previous states. Another reason why maintenance optimization problems are deeply involved with the regenerative theory is that the objective function can hardly proposed based on finite domains due to random effects, where time-average properties could be utilizable from the renewal-reward theorem [69]. So in this subsection, it is aimed to introduce the following issues :

1. The definition of regenerative processes.
2. The first passage times of the OU process can construct a regenerative process.
3. The renewal-reward theorem for the regenerative process.

Definition

Definition 4.1.1. [69] A stochastic process $\{X_t, t \geq 0\}$ is called regenerative if (by enlarging the probability space if necessary) there is a random variable $R_1 > 0$ such that

1. $\{X_{t+R_1} : t \geq 0\}$ is independent of $\{\{X_t, t < R_1\}, R_1\}$, and
2. $\{X_{t+R_1} : t \geq 0\}$ is stochastically equivalent to $\{X_t : t \geq 0\}$.

Here R_1 is generally called a regeneration epoch, and say that X_t regenerates or starts over at this point.

An alternative definition of regenerative processes is also given in [69, 72], which is given as follows.

Definition 4.1.2. [69] A stochastic process $\{X_t, t \geq 0\}$ is called regenerative if (by enlarging the probability space if necessary) there is a random variable $R_1 > 0$ such that

1. $\{X_{t+R_1} : t \geq 0\}$ is independent of R_1, and
2. $\{X_{t+R_1} : t \geq 0\}$ is stochastically equivalent to $\{X_t : t \geq 0\}$.

The Regenerative Process with the Component as the OU process

Now consider $\{X_t, t \geq 0\}$ to be the OU process given in (4.1.1) with intial value x_0. For a failure level $L > x_0$, the following consideration comes naturally.

Let $\tau_1(L) := \inf_{t>0}\{X_t = L\}$ denote the first passage time of X_t to the failure level L. And define $Z_1 := \inf_{t>\tau_1(L)}\{X_t = L\}$. Recursively define for $k \geq 1$:

$$\tau_{k+1}(L) = \inf_{t>Z_k}\{X_t = L\},\ Z_{k+1} := \inf_{t>\tau_{k+1}(L)}\{X_t = L\}. \tag{4.1.3}$$

Following the statement in [69] for Brownian motion, by the continuity of the OU process and corresponding strong Markov property. $\{Z_k : k \geq 1\}$ is a well-defined embedded renewal process for X_t that is null recurrent. Note that X_t is delayed regenerative with respect to the embedded renewal process $\{\tau_k(L) : k \geq 1\}$.

Moreover since the OU process is with continuous trajectories, the only possible regeneration epochs R_1 are hitting times to a fixed state : continuity together with the definition 4.1.1 implies that $X(R_1)$ is independent of itself which implies that there exists a L such that $P(Y_{R_1} = L) = 1$.

Renewal-Reward Theorem

For the regenerative process, a renewal-reward theorem can be of interest to translate the average rate function in long-term period to the one in one regenerative cycle. This is divided into two steps as follows :

1. The renewal-reward theorem is introduced.
2. The limiting behaviour of regenerative processes is introduced.

The renewal-reward theorem is given from the statement in [85]. Denote the system rewards rate as $f(x)$ whenever it is in state x, then the total reward up to time t is given by $\int_0^t f(X_s)ds$, where $\{f(X_t), t \geq 0\}$ is a regenerative process. The renewal-reward theorem states that the average reward rate in long-term period can be given by the limiting behavior of regenerative processes.

Lemma 4.1.3.

$$\lim_{t\to+\infty}\int_0^t f(X(s))ds/t = \int_{-\infty}^{+\infty} f(y)dF(y) = \mathbb{E}(f(X_\infty)), \tag{4.1.4}$$

where $F(y)$ is the limiting distribution of the regenerative process X_t.

And the limiting distribution of the regenerative process is given by the following lemma.

Lemma 4.1.4.

$$F(y) = \frac{1}{\mathbb{E}(R_1)}\mathbb{E}(\int_0^{R_1} I_{X_s\in(-\infty,y]}ds), \tag{4.1.5}$$

where R_1 is the regeneration epoch, and I_A is the indicator function of A.

Moreover the result on the indicator function can be extended to a more general function, and it turns back to the original problem.

Lemma 4.1.5. *If the state space E is a complete separable metric space, for a measurable function $f : E \to \mathbb{R}$, we have that*

$$\lim_{t\to+\infty}\mathbb{E}(f(X_t)) = \frac{1}{\mathbb{E}(R_1)}\mathbb{E}(\int_0^{R_1} f(X_s)ds). \tag{4.1.6}$$

Therefore this classical result has turned the long-term average reward into the consideration in one regeneration cycle. It induces also directly a computable form of the objective function in maintenance optimization in the next section.

4.2 Evaluation Criterions

4.2.1 Cost-Based Maintenance Optimization

To consider the cost-based maintenance optimization, it is supposed further that :

1. A maintenance is performed with a cost C_m.
2. The inactivity of the system has a cost per unit of time C_u.

Under above hypotheses, the process X_t describing the system state is a regenerative process with regenerative times which are times of maintenance. At a time of maintenance the process returns to x_0, and the random evolution of the system after the maintenance does not depend on the past.

Following the consideration in [63], where the long run expected cost per unit of time is defined as the mean cost on a cycle divided by the mean duration of a cycle :

$$\begin{aligned} EC_\infty &= \frac{C_m + C_u\mathbb{E}((\tau_v+\kappa+\lambda-\tau_r)I_{\tau_r\le\tau_v+\kappa} + \lambda I_{\tau_r>\tau_v+\kappa})}{\mathbb{E}(\tau_v+\kappa+\lambda)} \\ &= \frac{C_m + C_u(\mathbb{E}\lambda+\kappa-\mathbb{E}(\inf(\kappa,\tau_r-\tau_v)))}{\mathbb{E}(\tau_v)+\kappa+\mathbb{E}\lambda}. \end{aligned} \tag{4.2.1}$$

Then the cost-based problem is to solve L_v from the following problem :

$$\text{argmin}_{L_v} EC_\infty, \quad s.t. \quad x_0 \le L_v \le L_r. \tag{4.2.2}$$

The remaining work is to give a computable expression of EC_∞ with the virtual failure level L_v. Following the analysis in [63], we first calculate $\mathbb{E}(\inf(\sigma, \tau_r - \tau_v))$, related to the joint pdf $p_{\tau_v,\tau_r}(x, y)$ of (τ_v, τ_r).

As it is constrained that $L_v \le L_r$, then obviously $\tau_r \ge \tau_v$. So the joint distribution of (τ_v, τ_r) is concentrated on $\{x \le y\}$. Different to Gamma process considered in [63], the time-dependent OU process considered here is continuous, and therefore the probability of $\tau_r = \tau_v$

is zero. We only need to consider the survival function of (τ_v, τ_r) on $x < y$ based on the strong Markov property and also the continuity of the OU process X_t :

$$\begin{aligned} P(\tau_v > x, \tau_r > y) &= P(\tau_r > y) - P(\tau_v \leq x, \tau_r > y) = P(\tau_r > y) - \int_0^x P(\tau_r > y | \tau_v = z) g_{\tau_v}(z|x_0, t_0) dz \\ &= \int_y^{+\infty} g_{\tau_r}(s|x_0, t_0) ds - \int_0^x \int_y^{+\infty} g_{\tau_r}(\eta|\tau_v = z) g_{\tau_v}(z|x_0, t_0) d\eta dz \\ &= \int_y^{+\infty} g_{\tau_r}(s|x_0, t_0) ds - \int_0^x \int_y^{+\infty} g_{\tau_r}(\eta|L_v(z), z) g_{\tau_v}(z|x_0, t_0) d\eta dz, \end{aligned} \tag{4.2.3}$$

where $g_{\tau_r}(\eta|y, s)$ expresses the conditional first passage density to L_r for X_t based on (y, s), $g_{\tau_v}(x|y, s)$ for the conditional first passage density to L_v for X_t based on (y, s).

From (4.2.3), its density function is equal to :

$$p_{\tau_v, \tau_r}(x, y) = \frac{\partial^2}{\partial x \partial y}(P(\tau_v > x, \tau_r > y)) = g_{\tau_r}(y|L_v(x), x) g_{\tau_v}(x|x_0, 0). \tag{4.2.4}$$

Furthermore, we have for $t \geq 0$, the survival function $\bar{H}(t)$ of $\tau_r - \tau_v$ is given :

$$\bar{H}(t) = \int \int_{y-x>t} p_{\tau_v, \tau_r}(x, y) dx dy = \int_t^{+\infty} \int_0^{y-t} g_{\tau_r}(y|L_v(x), x) g_{\tau_v}(x|x_0, 0) dx dy. \tag{4.2.5}$$

Finally we get

$$\mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) = \int_0^{\kappa} \bar{H}(s) ds. \tag{4.2.6}$$

Based on all the above we actually have expressed $EC_{+\infty}$ in a closed form by the conditional first passage density for τ_v and τ_r, and more specific computation on such issues with direct connection for L_v will be investigated in the next section for the optimization in (4.2.2).

4.2.2 Availability-Based Maintenance Optimization

It is also an emphasis in the field of reliability engineering to maximize the system availability. And therefore the following asymptotic unavailability criterion can also be of interest [7].

Supposing X_t is the regenerative process based on the OU process with the regenerative times defined in (4.1.3), the asymptotic unavailability rate is given by the mean off time on a cycle divided by the mean duration of a cycle. This is due to the renewal-reward theorem.

$$\begin{aligned} U_\infty &= \frac{\mathbb{E}(\lambda I_{\tau_r > \tau_v + r} + (\tau_v + \kappa - \tau_r + \lambda) I_{\tau_r \leq \tau_v + \kappa})}{\mathbb{E}(\tau_v + \kappa + \lambda)} \\ &= \frac{\mathbb{E}\lambda + \mathbb{E}((\tau_v + \kappa - \tau_r) I_{\tau_r \leq \tau_v + \kappa})}{\mathbb{E}(\tau_v + \kappa + \lambda)} \\ &= \frac{\mathbb{E}\lambda + \kappa - \mathbb{E}(\inf(\kappa, \tau_r - \tau_v))}{\mathbb{E}(\tau_v) + \kappa + \mathbb{E}\lambda}. \end{aligned} \tag{4.2.7}$$

Then the availability-based problem is to solve L_v from the following problem :

$$\text{argmin}_{L_v} U_\infty, \quad s.t. \quad x_0 \leq L_v \leq L_r. \tag{4.2.8}$$

It is noted that essentially there is no technical difference between the cost-based optimization and availability-based optimization, as the core issues in (4.2.1) and (4.2.7) remain the same.

4.2.3 Linking the Maintenance Duration with the System State

It is a natural consideration to suppose that the time of maintenance operation depends on system states such that in [7], it is suggested to consider that

$$\mathbb{E}\lambda = \lambda_1 + \lambda_2 \mathbb{E}(X_{\tau_v+\kappa}), \tag{4.2.9}$$

where λ_1, λ_2 are deterministic values. The initial idea for such a proposition is to consider that the maintenance duration is longer when the system degradation is more severe.

However instead of (4.2.9), in this chapter another discounted version for the duration of maintenance operation would be considered as follows,

$$\mathbb{E}\lambda = \lambda_1 + \lambda_2 \mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)}). \tag{4.2.10}$$

This is due to the calculation difficulty for (4.2.9) by the nonlinearity of the process. And also (4.2.10) didn't introduce any issue related to the process, but just a weight $e^{\alpha(t,0)}$ to X_t. Therefore it still can reproduce the original idea of linking the system state with the maintenance duration. However the relationship turns from a linear link to a nonlinear link.

To proceed the estimation on $\mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)})$, the following lemma on the martingale properties of X_t is essential.

Lemma 4.2.1. *Recalling the process X_t in (4.1.2), the process $X_t e^{\alpha(t,0)} + \beta(t,0) - x_0$ is a mean-zero Gaussian martingale.*

Proof. Note that under the time-change given in Proposition 1.3.1, $X_t e^{\alpha(t,0)} + \beta(t,0) - x_0$ is a Gauss-Markov process. And this result comes naturally from Exercise 1.13, Page 86 in [61]. □

Based on Lemma 4.2.1, the following proposition is given for the expression of $\mathbb{E}(X_{\tau_v+\kappa})$.

Proposition 4.2.2. *The discounted mean degradation level at the entrance time $\tau_v + \kappa$ is given by*

$$\mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)}) = -\mathbb{E}(\beta(\tau_v+\kappa,0)) + x_0. \tag{4.2.11}$$

Proof. By Lemma 4.2.1, the process $X_t e^{\alpha(t,0)} + \beta(t,0) - x_0$ is a Gaussian martingale. Therefore for the natural filtration $\mathcal{F}_t$ of the process X_t, we have

$$\mathbb{E}(X_t e^{\alpha(t,0)} + \beta(t,0) - x_0 | \mathcal{F}_s) = X_s e^{\alpha(s,0)} + \beta(s,0) - x_0. \tag{4.2.12}$$

Therefore by the optional stopping theorem with the stopping time $\tau_v + \kappa$, we know that

$$\mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)} + \beta(\tau_v+\kappa,0) - x_0) = 0. \tag{4.2.13}$$

This leads to

$$\mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)}) = -\mathbb{E}(\beta(\tau_v+\kappa,0)) + x_0. \tag{4.2.14}$$

□

To calculate $\mathbb{E}(\beta(\tau_v + \kappa, 0))$, the following proposition could be utilizable.

Proposition 4.2.3. *For any function $f \in C^1[s, +\infty)$, and denote the conditional first passage time of a continuous process X_t to the boundary $L(t)$ by $\tau_{y,s} := \inf_{t>s}\{X_t^{y,s} \geq L(t)\}$,*

$$\mathbb{E}(f(\tau_{y,s})) = f(s) + \int_s^{+\infty} P(\tau_{y,s} > t) df(t). \tag{4.2.15}$$

Especially,

$$\mathbb{E}(\tau_{y,s}) = s + \int_s^{+\infty} P(\tau_{y,s} > t) dt. \tag{4.2.16}$$

Proof. Noticing $P(\tau_{y,s} > s) = 1,\ P(\tau_{y,s} > +\infty) = 0$,

$$\begin{aligned}\mathbb{E}(f(\tau_{y,s})) &= \int_s^{+\infty} f(t) g(t|y,s) dt = -\int_s^{+\infty} f(t) dP(\tau_{y,s} > t) \\ &= -f(t)P(\tau_{y,s} > t)|_{t=s}^{+\infty} + \int_s^{+\infty} P(\tau_{y,s} > t) df(t) \\ &= f(s) + \int_s^{+\infty} P(\tau_{y,s} > t) df(t).\end{aligned} \tag{4.2.17}$$

□

4.3 Approximation-Optimization

In Section 4.2, it is presented that the core issues in the maintenance optimization problems are to calculate the mean first passage time for the OU process X_t to the virtual failure level L_v and $\mathbb{E}(\inf(\kappa, \tau_r - \tau_v))$. However it is difficult to have an explicit expression of the first passage density for the OU process like Gamma process [7], such that in this section we will emphasize on the possibility to derive results from approximate expressions.

Here we suppose L_r and L_v are all constant boundaries in this section. And as described before the real failure level L_r is given, and L_v is pending to be optimized from the maintenance optimization problems. And the key idea in this section is to propose some computable expressions whose accuracy doesn't depend on the pending L_v, such that the optimization procedure for L_v will not be influenced by the approximation. Such an approximation-optimization procedure is widely in engineering problems where the objective function cannot be expressed explicitly, more consideration can be referred to Appendix B.

And the discussion in this section concerns the following points :

1. It is presented that the drifted Brownian motion can be considered in an explicit way for the presented maintenance optimization problems.
2. It is presented that when the time-changed Brownian motion is applied, the maintenance optimization problem can be solved based on an approximation-optimization way. We should notice the first passage density is still explicit in this case, what we need do is to do approximation for the integral involved in the objective functions.
3. For a general OU process, the situation is similar. And we try to reproduce the approximation-optimization procedure for the time-changed Brownian motion. We first approximate the first passage density by a parametric form related to the virtual failure level, and solve the optimization problems based on corresponding objective functions. This leads to the estimate of the failure level directly. And here the derived results in Section 2.3.7 are applied in the calculation.

4.3.1 Drifted Brownian motion

Basic Properties

In this subsection, a direct analysis based on Brownian motion will be presented. And our consideration is limited to the drifted Brownian motion given by $X_t = \mu t + \sigma B_t$ with the initial value 0 at 0, $\mu, \sigma > 0$, and corresponding virtual and real failure level L_v and L_r, whose transition pdf is given by

$$p(x,t|y,s) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \exp(-\frac{(x-y-\mu(t-s))^2}{2\sigma^2(t-s)}). \tag{4.3.1}$$

Its first passage density to a constant boundary $L > 0$ based on (y,s) satisfies inverse Gaussian distribution :

$$g(t|y,s) = \frac{L-y}{\sqrt{2\pi\sigma^2(t-s)^3}} \exp(-\frac{(L-\mu(t-s)-y)^2}{2\sigma^2(t-s)}). \tag{4.3.2}$$

Corresponding first passage distribution $G(t|y,s)$ is given from (2.3.36) for $C = -\frac{2\mu}{\sigma^2}$

$$\begin{aligned} G(t|y,s) &= 1 - F(L,t|y,s) + e^{\frac{2\mu}{\sigma^2}(y-L)} F(L,t|2L-y,s) \\ &= \Phi(\frac{\mu(t-s)-L+y}{\sigma\sqrt{t-s}}) + e^{\frac{2\mu}{\sigma^2}(y-L)} \Phi(\frac{-\mu(t-s)-L+y}{\sigma\sqrt{t-s}}), \end{aligned} \tag{4.3.3}$$

where $\Phi(*)$ is the normal distribution function.

The mean first passage time is given by

$$\mathbb{E}(\tau_{y,s}) = \int_s^{+\infty} t g(t|y,s) dt = \int_0^{+\infty} (z+s) g(z|y,0) dz = s + \frac{L-y}{\mu}. \tag{4.3.4}$$

Calculation on $\mathbb{E}(\inf(\kappa, \tau_r - \tau_v))$

It is noticed that the Brownian motion is with continuous trajectories such that it can simplify the calculation on $\mathbb{E}(\inf(\kappa, \tau_r - \tau_v))$. This simplification is due to the continuity and strong Markov propertie of Brownian motion.

It should be noticed by the homogeneous property of X_t that $\tau_r - \tau_v$ should have the same probability law with the first passage time of X_t to the boundary $L_r - L_v$, which is denoted by τ_{r-v}.

Therefore we have

$$\mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) = \mathbb{E}(\inf(\kappa, \tau_{r-v})) = \int_0^{\kappa} (1 - G_{\tau_{r-v}}(z|0,0)) dz. \tag{4.3.5}$$

It is then derived from (2.3.36) that

$$1 - G_{\tau_{r-v}}(t|0,0) = F(L_r - L_v, t|0,0) - e^{\frac{2\mu}{\sigma^2}(L_v - L_r)} F(L_r - L_v, t|2(L_r - L_v), 0). \tag{4.3.6}$$

Therefore, we have

$$\begin{aligned} \mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) &= \mathbb{E}(\inf(\kappa, \tau_{r-v})) \\ &= \int_0^{\kappa} F(L_r - L_v, z|0,0) dz - e^{\frac{2\mu}{\sigma^2}(L_v - L_r)} \int_0^{\kappa} F(L_r - L_v, z|2(L_r - L_v), 0) dz. \end{aligned} \tag{4.3.7}$$

And also if we propose to connect the maintenance duration with the degradation level, then we have $\mathbb{E}\lambda = \lambda_1 + \lambda_2\mathbb{E}(\tau_v + \kappa)$ with

$$\mathbb{E}(X_{\tau_v+\kappa}) = \mu(\mathbb{E}(\tau_v) + \kappa). \tag{4.3.8}$$

Therefore it is achieved at a computable form for the asymptotic unavailability criterion

$$\begin{aligned} U_\infty &= \frac{\mathbb{E}\lambda + \kappa - \mathbb{E}(\inf(\kappa, \tau_r - \tau_v))}{\mathbb{E}(\tau_v) + \kappa + \mathbb{E}\lambda} \\ &= \frac{\lambda_1 + \lambda_2\mathbb{E}(\tau_v + \kappa) + \kappa - \int_0^\kappa (1 - G_{\tau_{r-v}}(z|0,0))dz}{\mathbb{E}(\tau_v) + \kappa + \lambda_1 + \lambda_2\mathbb{E}(\tau_v + \kappa)}. \end{aligned} \tag{4.3.9}$$

Correspondingly the asymptotic cost rate is given by

$$EC_\infty = \frac{C_m + C_u(\lambda_1 + \lambda_2\mathbb{E}(\tau_v + \kappa) + \kappa - \int_0^\kappa (1 - G_{\tau_{r-v}}(z|0,0))dz)}{\mathbb{E}(\tau_v) + \kappa + \lambda_1 + \lambda_2\mathbb{E}(\tau_v + \kappa)}. \tag{4.3.10}$$

Only one integral is involved which needs further approximation, and the compound trapezoid rule is adopted to approximate this issue. Then the optimization error only depends on the error from numerical integrals, and this can be compensated generally by considering more accurate numerical algorithms with smaller step-sizes.

4.3.2 Time-changed Brownian motion

The time-changed Brownian motion will be discussed in this subsection on its application in maintenance optimization problems. And also the core issue is also to calculate $\mathbb{E}(\tau_v)$ and $\mathbb{E}(\inf(\kappa, \tau_r - \tau_v))$.

Recall in (2.3.39), the time-changed Brownian motion is discussed with the following form.

$$dX_t = \frac{C}{2}\sigma^2(t)dt + \sigma(t)dB_t, X_0 = 0, t \geq 0 \tag{4.3.11}$$

where $C > 0$, and $\sigma(t) > 0$ is a continuous function.

The notations in (1.3.3) for (2.3.39) are updated to

$$\alpha(t,s) = 0, \gamma(t,s) = \frac{1}{2}\int_s^t \sigma^2(z)dz, \beta(t,s) = -C \times \gamma(t,s), \tag{4.3.12}$$

Remark : Although under a time-change $\Gamma(t) = 2\gamma(t,0)$, X_t can be treated in the form of $X_t = \frac{C}{2}\Gamma(t) + B_{\Gamma(t)}$, such that much simpler analysis can be performed based on the properties of Brownian motion. But this treatment can only lead directly to $\mathbb{E}(\Gamma(\tau_v))$ and also the probability law of $\Gamma(\tau_r) - \Gamma(\tau_v)$, which are still far from the desired issues $\mathbb{E}(\tau_v)$ and $\mathbb{E}(\inf(\kappa, \tau_r - \tau_v))$. Therefore the following analysis is not trivial to derive some approximate expressions for calculation.

Basic Properties

For the time-changed Brownian motion, its transition pdf is given by

$$p(x,t|y,s) = \frac{1}{\sqrt{4\pi\gamma(t,s)}}\exp(-\frac{(x - y - C\gamma(t,s))^2}{4\gamma(t,s)}), \tag{4.3.13}$$

such that the transition distribution is given by

$$F(x,t|y,s) = \Phi(\frac{x-y-C\gamma(t,s)}{\sqrt{2\gamma(t,s)}}). \tag{4.3.14}$$

Furthermore, for a constant failure level L, its first passage density is given by (2.3.44)

$$g(t|y,s) = \frac{\sigma^2(t)(L-y)}{2\gamma(t,s)} p(L,t|y,s). \tag{4.3.15}$$

Correspondingly, the first passage distribution is given by (2.3.45)

$$G(t|y,s) = \Phi(\frac{-L+C\gamma(t,s)+y}{\sqrt{2\gamma(t,s)}}) + e^{-C(y-L)}\Phi(\frac{-C\gamma(t,s)-L+y}{\sqrt{2\gamma(t,s)}}), \tag{4.3.16}$$

where $\Phi(*)$ is the normal distribution function.

These analytical properties lead to the calculation in the following.

Mean First Passage Time

From Proposition (4.2.3), for the time-changed Brownian motion, it comes to

$$\mathbb{E}(\tau_v) = \int_0^{+\infty} (1 - G_{\tau_v}(z|0,0))dz. \tag{4.3.17}$$

And also for the mean duration of maintenance action,

$$\begin{aligned}\mathbb{E}(X_{\tau_v+\kappa}) &= -\mathbb{E}(\beta(\tau_v+\kappa,0)) = -\int_0^{+\infty}(1-G_{\tau_v}(z|0,0))d\beta(z+k,0) \\ &= \frac{C}{2}\int_0^{+\infty}(1-G_{\tau_v}(z|0,0))\sigma^2(z+k,0)dz\end{aligned} \tag{4.3.18}$$

To allow a computable form of these two issues, we should consider a truncated form for the integral on infinite interval. But the difficulty exists for the value of the virtual failure level L_v which is unknown before the optimization. And to judge whether the truncated integral is enough good is not obvious.

Here it is suggested to consider truncating the integral based on the analysis of τ_r as $\tau_r \geq \tau_v$. It is obvious therefore the value of $(1-G_{\tau_r}(z|y,s))$ is always larger than $(1-G_{\tau_v}(z|y,s))$. If we truncate the interval based on the judgement of $(1-G_{\tau_r}(z|y,s))$, it should be enough accurate for the smaller function $(1-G_{\tau_r}(z|y,s))$ even it is with a pending parameter L_v. And the truncated integral for $(1-G_{\tau_r}(z|y,s))$ should be at least as good as the truncated integral for $(1-G_{\tau_r}(z|y,s))$.

So summarize all the above, the computation follows the following procedures :

1. Set a tolerance error $\epsilon > 0$ and a control value $h > 0$, and choose a large value T such that $1 - G_{\tau_r}(T|y,s) < \epsilon$ and $\int_T^{T+h}(1-G_{\tau_r}(z|y,s))dz < \epsilon$. The same treatment is applied to $\int_0^{+\infty}(1-G_{\tau_v}(z|y,0))\sigma^2(z+\kappa)dz$, but it needs the consideration on $(1-G_{\tau_r}(z|y,s))\sigma^2(z+\kappa)$.

2. When the end-time T is chosen, such that the approximations are given :

$$\int_s^{\infty}(1-G_{\tau_r}(z|y,s))dz \approx \int_s^{T}(1-G_{\tau_r}(z|y,s))dz, \tag{4.3.19}$$

and

$$\int_s^{\infty}(1-G_{\tau_r}(z|y,s))\sigma^2(z+\kappa)dz \approx \int_s^{T}(1-G_{\tau_r}(z|y,s))\sigma^2(z+\kappa)dz. \tag{4.3.20}$$

3. The integral in a finite interval is further approximated by numerical algorithms, such as compound trapezoid rule.

Following the above procedures, we can achieve an enough accurate approximate expression for the mean regenerative time and mean maintenance duration, which is an explicit function on the pending L_v. The approximation accuracy could be slightly influenced by the knowledge on L_v, and only the error from the numerical integral scheme contributes.

Calculation on $\mathbb{E}(\inf\kappa, \tau_r - \tau_v)$

From (4.2.6), we know

$$\mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) = \int_0^{\kappa}\bar{H}(s)ds \tag{4.3.21}$$

where

$$\bar{H}(t) = \int_t^{+\infty}\int_0^{y-t} g_{\tau_r}(y|L_v,x)g_{\tau_v}(x|0,0)dxdy. \tag{4.3.22}$$

Although as stated before, $g_{\tau_r}(y|L_v(x),x)$ and $g_{\tau_v}(x|0,0)$ can be given explicitly, but the integral on the infinite interval is still a problem for the optimization. Following the previous stated analysis based on the fact that $\tau_r \geq \tau_v$, we will also propose to truncate the infinite interval based on τ_r.

Actually, from the Fortet's equation (2.2.1),

$$p(L_r,t|0,0) = \int_0^t p(L_r,t|L_v,z)g_{\tau_v}(z|0,0)dz. \tag{4.3.23}$$

Therefore from (4.3.15), we have

$$\int_0^{y-t} g_{\tau_r}(y|L_v,x)g_{\tau_v}(x|0,0)dx = \int_0^{y-t}\frac{\sigma^2(y)(L_r-L_v)}{2\gamma(y,x)}p(L_r,t|L_v,x)g_{\tau_v}(x|0,0)dx. \tag{4.3.24}$$

Noticing $x \in [0, y-t]$ such that $\gamma(y,x) \in [\gamma(y,y-t),\gamma(y,0)]$, and $L_r - L_v \in [0, L_r - x_0]$ therefore from (2.2.1), we should have

$$\begin{aligned}\int_0^{y-t} g_{\tau_r}(y|L_v,x)g_{\tau_v}(x|0,0)dx &\leq \frac{\sigma^2(y)(L_r-x_0)}{2\gamma(y,y-t)}\int_0^{y-t}p(L_r,t|L_v,x)g_{\tau_v}(x|0,0)dx \\ &= \frac{\sigma^2(y)(L_r-x_0)}{2\gamma(y,y-t)}p(L_r,y-t|0,0).\end{aligned} \tag{4.3.25}$$

Now we can propose the truncated approximation for $\mathbb{E}(\inf(\kappa, \tau_r - \tau_v))$ as follows :

1. Set a tolerance error $\epsilon > 0$ and a control value $h > 0$, and choose a large value $T > t$ such that $\frac{\sigma^2(T)(L_r - x_0)}{2\gamma(T,T-t)} p(L_r, T-t|0,0) < \epsilon$ and $\int_T^{T+h} \frac{\sigma^2(z)(L_r-x_0)}{2\gamma(z,z-t)} p(L_r, z-t|0,0)dz < \epsilon$.
2. When the end-time T is chosen, the approximation is given :

$$\bar{H}(t) \approx \int_t^T \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|0,0) dx dy. \tag{4.3.26}$$

3. The integral in a finite interval for $\int_0^\kappa \bar{H}(s)ds$ is further approximated by numerical algorithms, such as compound trapezoid rule.

Following the above procedures, we can achieve an enough accurate approximate expression for $\int_0^\kappa \bar{H}(s)ds$, which is an explicit function on the pending L_v. The approximation accuracy could be slightly influenced by the knowledge on L_v, and only the error from the numerical integral scheme contributes. This could be compensated by more accurate approximation schemes on the integrals.

4.3.3 The General OU Process

Durbin's Approximation

In Section 2.3.7, an approximate expression for first passage density function has been proposed. Using such an approximation to calculate the optimization problem, it comes to the same heuristic procedures for maintenance optimization as in the one for Brownian motion [44] and the one for Gamma process [7].

Recalling the process X_t given by (4.1.1),

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, X_0 = x_0, t \geq 0. \tag{4.3.27}$$

An approximate expression for the first passage density to the boundary $L(t)$ is given by (2.3.113) :

$$g(t|y,s) \approx \left[a(t)\mathbb{E}(X_t) + b(t) - L'(t) + (L(t) - \mathbb{E}(X_t))(\frac{\sigma^2(t)}{\text{var}(X_t)} + a(t)) \right] \times p(L(t), t|y, s), \tag{4.3.28}$$

where

$$\begin{aligned} \mathbb{E}(X_t) &= e^{-\alpha(t,s)} \Big(y - \beta(t,s) \Big), \\ \text{var}\,(X_t) &= e^{-2\alpha(t,s)} \int_s^t \sigma^2(u) e^{2\alpha(u,s)} du. \end{aligned} \tag{4.3.29}$$

Especially for the constant failure level L, it is noted that

$$g(t|y,s) \approx \left[a(t)\mathbb{E}(X_t) + b(t) + (L - \mathbb{E}(X_t))(\frac{\sigma^2(t)}{\text{var}(X_t)} + a(t)) \right] p(L, t|y, s). \tag{4.3.30}$$

The expression in (4.3.30) leads directly to the expression of first passage densities $g_{\tau_r}(t|y,s)$ and $g_{\tau_v}(t|y,s)$ by considering L_r and L_v respectively.

Mean First Passage Time

Now we have an explicit expression for the approximate first passage density (4.3.30), which leads directly to the mean first passage time and the proposed mean maintenance duration. From Proposition (4.2.3), for the OU process, it comes to

$$\mathbb{E}(\tau_v) = \int_0^{+\infty} z g_{\tau_v}(z|x_0, 0))dz. \tag{4.3.31}$$

And also the mean duration of maintenance action,

$$\mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)}) = -\mathbb{E}(\beta(\tau_v+\kappa, 0)) + x_0 = x_0 - \int_0^{+\infty} g_{\tau_v}(z|x_0, 0))\beta(z+k, 0)dz. \tag{4.3.32}$$

To allow a computable form of these two issues, we should consider a truncated form for the integral on infinite interval. But the difficulty exists for the value of the virtual failure level L_v which is unknown before the optimization. And to judge whether the truncated integral is enough good is not obvious.

Following the previous stated analysis for the time-changed Brownian motion (4.3.25), using the fact that $\tau_r \geq \tau_v$, we can propose a controllable approximation to truncate the infinite interval based on τ_r.

Calculation on $\mathbb{E}(\inf \kappa, \tau_r - \tau_v)$

From (4.2.6), we know

$$\mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) = \int_0^{\kappa} \bar{H}(s)ds \tag{4.3.33}$$

where

$$\bar{H}(t) = \int_t^{+\infty} \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|x_0, 0)dxdy. \tag{4.3.34}$$

It follows two steps to give an approximation of $\bar{H}(t)$:

1. Truncate the infinite interval with an end-time $T > 0$, such that

$$\bar{H}(t) \approx \int_t^{T} \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|x_0, 0)dxdy. \tag{4.3.35}$$

2. Approximate $\int_t^T \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|x_0, 0)dxdy$ by numerical algorithms, such as compound trapezoid rule.

Such an approximation has two kinds of errors, which are stated as follows :

1. Errors in control. As stated before, $g_{\tau_r}(y|L_v(x), x)$ and $g_{\tau_v}(x|0, 0)$ can be approximated by (4.3.30) explicitly, what we need to do is to truncate the infinity interval and propose approximations for the integral in a finite interval. Following the previous stated analysis for the time-changed Brownian motion (4.3.25), using the fact that $\tau_r \geq \tau_v$, we can propose a controllable approximation to truncate the infinite interval based on τ_r.

TABLE 4.2 – Experimental setup for the maintenance optimization based on the drifted Brownian motion

Process	κ	λ_1	λ_2	C_m	C_u	L_r
$X_t = 2t + 10B_t$	2	2	0.2	100	40	20

TABLE 4.3 – The maintenance optimization results for the drifted Brownian motion

criterion	optimal value	optimal L_v
asymptotic cost rate EC_∞	31.55786	17.18251
asymptotic unavailability U_∞	0.6832056	2.088506

2. Errors out of control. We should notice the accuracy of the approximate first passage density (4.3.30) can be promised for itself, but when it comes to an approximation for the multiple integral $\bar{H}(t)$ involving many approximate first passage density values, it is hard to have an estimate for the overall error for such an approximation. Therefore even we here do this approximation, we cannot propose an accurate control right now. And it induces more consideration on the original first passage problem in the future.

Following the above procedures, we can achieve an enough accurate approximate expression for $\int_0^\kappa \bar{H}(s)ds$, which is an explicit function on the pending L_v. The approximation accuracy can be promised without the knowledge on L_v.

4.4 Simulation Tests

4.4.1 Drifted Brownian motion

The first example is chosen as the drifted Brownian motion, and here we suppose the considered process as the drifted Brownian motion $X_t = 2t + 10B_t$. And other parameters are chosen as in Table 4.4.1.

The optimization is performed based on the asymptotic unavailability (4.2.7) and also the asymptotic cost rate (4.2.1). And the results are presented in Table 4.4.1. Correspondingly, the asymptotic unavailability and asymptotic cost rate are presented in Figure 4.2(a) and Figure 4.2(b) respectively.

4.4.2 Time-Changed Brownian Motion

Continuing the setup in Section 2.5, here we will consider a time-changed Brownian motion X_t as follows :

$$dX_t = 0.5tdt + \sqrt{t}dB_t, X_0 = 0, t \geq 0. \tag{4.4.1}$$

And other parameters for the simulation test are chosen as in Table 4.4.2.

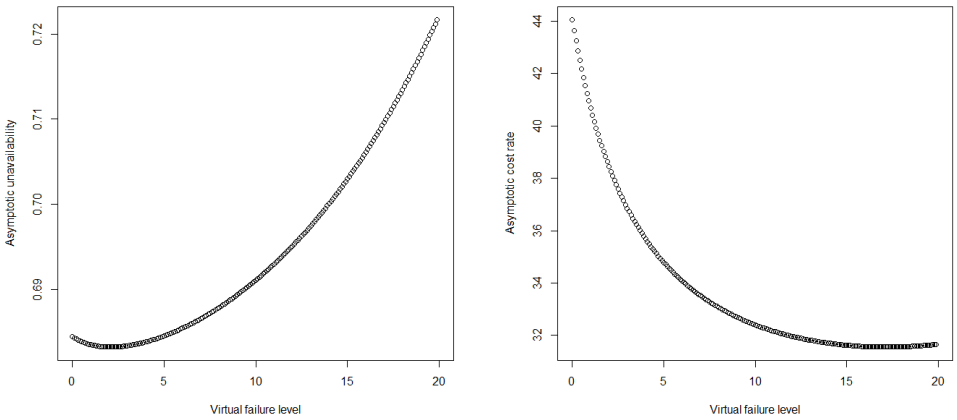
TABLE 4.4 – Experimental setup for the maintenance optimization based on the time-changed Brownian motion

Process	κ	λ_1	λ_2	C_m	C_u	L_r
$dX_t = 0.5tdt + \sqrt{t}dB_t$	2	2	0.2	100	40	10 .

TABLE 4.5 – The maintenance optimization results for the time-changed Brownian motion

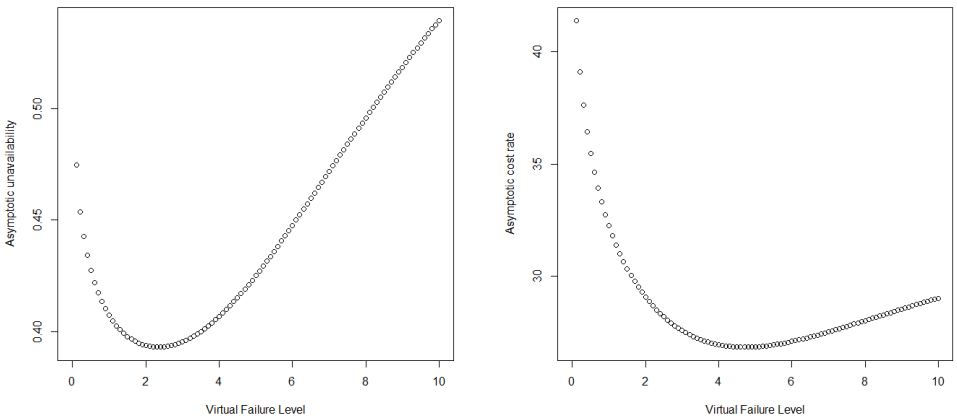
criterion	optimal value	optimal L_v
asymptotic cost rate EC_∞	26.84566	4.710752
asymptotic unavailability U_∞	0.3929153	2.3366

The optimization is performed based on the asymptotic unavailability (4.2.7) and also the asymptotic cost rate (4.2.1). And the results are presented in Table 4.4.2. Correspondingly, the asymptotic unavailability and asymptotic cost rate are presented in Figure 4.3(a) and Figure 4.3(b) respectively.



(a) Asymptotic unavailability rate (b) Asymptotic cost rate

FIGURE 4.2 – The performance of systems under different criterions for $X_t = 2t + 10B_t$



(a) Asymptotic unavailability rate (b) Asymptotic cost rate

FIGURE 4.3 – The performance of systems under different criterions for $dX_t = 0.5tdt + \sqrt{t}dB_t$

TABLE 4.6 – Experimental setup for the maintenance optimization based on the OU process $dX_t = (-r(X_t - m(t)) + m'(t))dt + \sigma dB_t$

Process	κ	λ_1	λ_2	C_m	C_u	L_r	x_0
$dX_t = (-r(X_t - m(t)) + m'(t))dt + \sigma dB_t$	2	2	0.2	10	5	10	2.8074561.

TABLE 4.7 – The maintenance optimization results for the OU process

criterion	optimal value	optimal L_v
asymptotic cost rate EC_∞	4.587412	8.007235
asymptotic unavailability U_∞	0.6886225	4.050314

4.4.3 Ornstein-Uhlenbeck Process

Continuing the discussion for M_{OU} in Section 1.5.3 with corresponding parameters fitted from a real degradation data-set given in Table 1.4, we here will consider a OU process X_t as follows :

$$dX_t = (-r(X_t - m(t)) + m'(t))dt + \sigma dB_t, t \geq 0, X_0 = x_0, \tag{4.4.2}$$

where $r = 0.1806708, \sigma = 2.4640884, m(t) = 2.4402845((t+1)^{0.8892020} - 1) + x_0, x_0 = 2.8074561$. Some trajectories of the process (2.5.12) are produced based on Monte-Carlo simulation, which are shown in Figure 1.4(b). And other parameters for the simulation test are chosen as in Table 4.4.3.

The optimization is performed based on the asymptotic unavailability (4.2.7) and also the asymptotic cost rate (4.2.1). And the results are presented in Table 4.4.3. Correspondingly, the asymptotic unavailability and asymptotic cost rate are presented in Figure 4.4(a) and Figure 4.4(b) respectively.

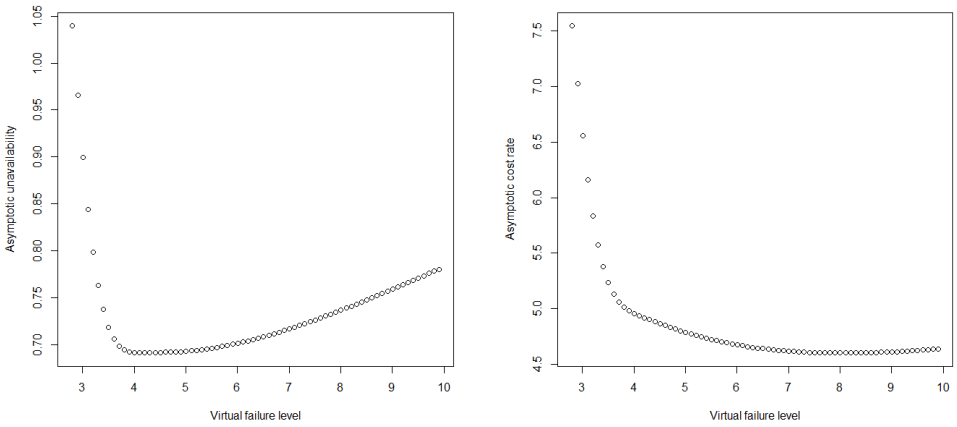
As stated before, due to the approximate expression of first passage density in the calculation, the overall error of the calculation can hardly be controlled. And this can be easily observed in the calculation results themselves. From Figure 4.4(a), it is noticed when time is small, a value around 1.05 is returned while the real one should be smaller than 1, this overestimated value will give the confusion to use such a result directly. This is due to the negative value produced by the Durbin's approximation, although this value tends to zero when the time is large.

So we redo the experiment by simply setting those negative values to zero, and new figures are produced, see Figure 4.5(a) and 4.5(b). And under this treatment, the optimal value for the cost criterion and unavailability criterion are given in Table 4.4.3.

Even with errors, the trend of the object functions can be described in a computable form, such that corresponding optimal value for the virtual failure level could be treated as an independent reference compared with experts' judgements.

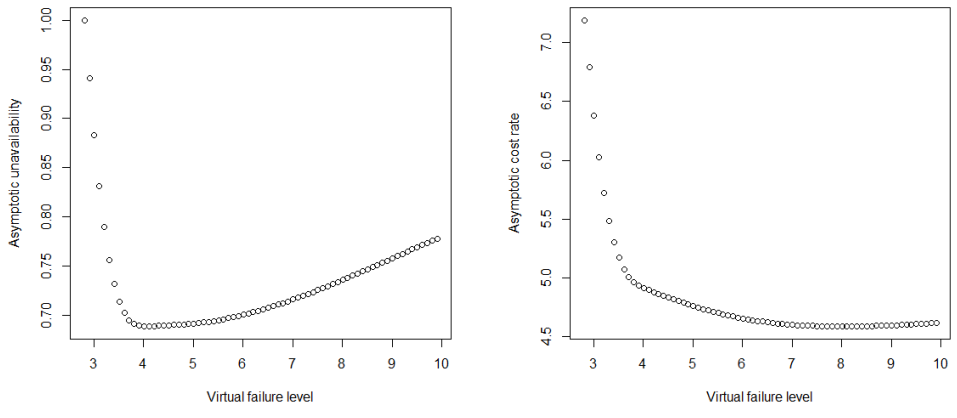
4.5 Conclusions and Perspectives

In this chapter, we presented the maintenance optimization problem for a continuously monitored system, where it is aimed to arrange properly the preventive maintenance before corrective maintenance. The previous discussion on first passage problems are extended to this application. Under the description of first passage failures, the maintenance optimization



(a) Asymptotic unavailability rate (b) Asymptotic cost rate

FIGURE 4.4 – The performance of systems under different criterions for $dX_t = (a(X_t - m(t)) + m'(t))dt + \sigma dB_t$, rough calculation.



(a) Asymptotic Unavailability (b) Asymptotic Cost rate

FIGURE 4.5 – The performance of systems under different criterions for $dX_t = (a(X_t - m(t)) + m'(t))dt + \sigma dB_t$, modified calculation.

problem is simplified to find an optimal virtual failure level as a decision rule.

The consideration is based on an approximation-optimization way, and three different processes are discussed : the drifted Brownian motion, the time-changed Brownian motion and the genenral Ornstein-Uhlenbeck process. The former two cases are with explicit first passage density, but the latter is with unexplicit expression such that the Durbin's approximation is introduced to fulfill the approximation-optimization. The maintenance problem is also considered from two different criterions : the asymptotic cost rate and the asymptotic unavailability rate. Moreover the link between the maintenance operation duration and the degradation level is considered in the problem.

Using analytical approximations of first passage density in maintenance optimization problems allows more possibilities rather than the continuously monitored system. A natural extension is to consider optimizing the inspection interval at the same time with the virtual failure level when the system is not continuously monitored [34].

Conclusions and Perspectives

Conclusions

In this thesis, we have discussed four issues, aiming to describe, predict and prevent system failures :

1. stochastic degradation modeling based on a time-dependent Ornstein-Uhlenbeck (OU) process ;
2. prognosis of system failures via first passage problems ;
3. failure level estimation via inverse first passage problems ;
4. maintenance optimization for a continuously monitored system.

The whole work adopts a model-based view based on first passage failures, and the illustration of the framework is presented in Figure 4.6.

In the part about *stochastic degradation modeling based on a time-dependent OU process*, we have introduced the following time-dependent OU process in the form of stochastic differential equation :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0. \tag{4.5.1}$$

It is shown that the time-dependent OU process is a good choice for degradation modeling from different aspects :

1. The OU process is with great flexility to model statistical properties of degradation records, especially on the mean, variance and covariance.
2. The mean-reverting property of the OU process localizes the fluctuations introduced in the degradation process. This is a good point to connect the stochastic modeling with justified experts' opinions and physical laws.

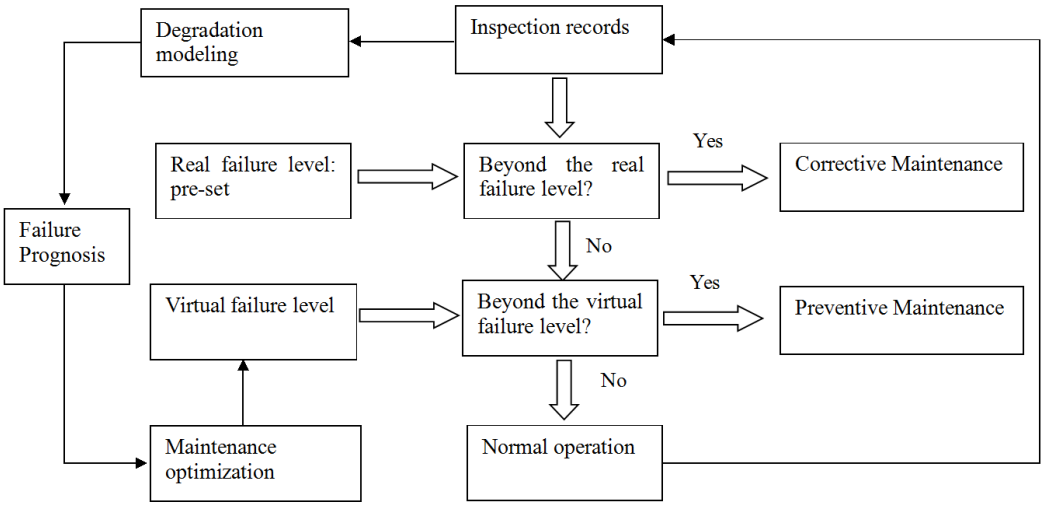


FIGURE 4.6 – System failures : description as first passage failures, prediction and prevention.

3. The OU process is with an explicit transition pdf, such that the parameter estimation from degradation records is not hard to be fulfilled.
4. A case study for a degradation data-set from passive components in power plants is performed to compare the OU process with a nonlinear-drift, linear-diffusion process.

In the part about *prognosis of system failures via first passage problems*, it concerns the residual useful lifetime estimation from conditional first passage problems. And we have developed different techniques to estimate the corresponding first passage density or the first passage distribution from two different views : partial differential equations and integral equations. These techniques can be generally categorized into three classes : analytical approximations, numerical algorithms and Monte-Carlo methods. Specifically, the following issues have been investigated.

1. By solving an initial-boundary value problem for Fokker-Planck equation, it is found that the first passage problems can be solved explicitly under a quasi-linear boundary defined in Proposition 2.3.4.
2. The quasi-linear boundary leads to an explicit tangent approximation for the first passage distribution in Proposition 2.3.9, which is proved to have a global accuracy under some conditions. This is proved to be the same with Durbin's approximation in Proposition 2.3.21.
3. The general consideration of the method of images leads to the parametric approximation of first passage density, which is solved by a linear programming approach (2.3.67). A special case of quasi-Daniels boundary is considered in Proposition 2.3.12.
4. A piecewise quasi-linear Monte-Carlo method is derived in (2.3.6).
5. A non-singular Volterra integral equation leads to a numerical algorithm (2.4.22).
6. Simulation tests are done to compare these methods in three cases : drifted Brownian motion, time-changed Brownian motion and a general OU process.

In the part about *failure level estimation via inverse first passage problems*, the gap between first passage failures and existing failure records is considered from a pure data-analysis view. To make up the gap, the inverse first passage is introduced, which is solved later based on integral equation methods. The role of the inverse first passage problem is illustrated in Figure 4.7. Specifically, the following issues have been discussed :

1. The initial boundary estimation in (3.2.15) for the OU process is derived from a preliminary result for Brownian motion.
2. The inverse first passage problem is solved from the Fortet's equation and the master equation respectively via an iterative procedure.
3. A simulation test is done to reproduce the failure level from simulated first passage density and distribution in a case study.

In the part about *maintenance optimization for a continuously monitored system*, it aims to propose a decision rule to perform preventive maintenance before corrective maintenance such that the objective function can be optimized. Continuing the discussion based on first passage failures, the technical target is simplified to find an optimal virtual failure level. Specifically, the following issues have been fulfilled :

1. The OU process has been introduced into condition-based maintenance for a continuously monitored system.
2. The maintenance optimization problem for the OU process has been solved in an approximation-optimization way, under the criterion of the asymptotic cost rate and asymptotic unavailability rate.

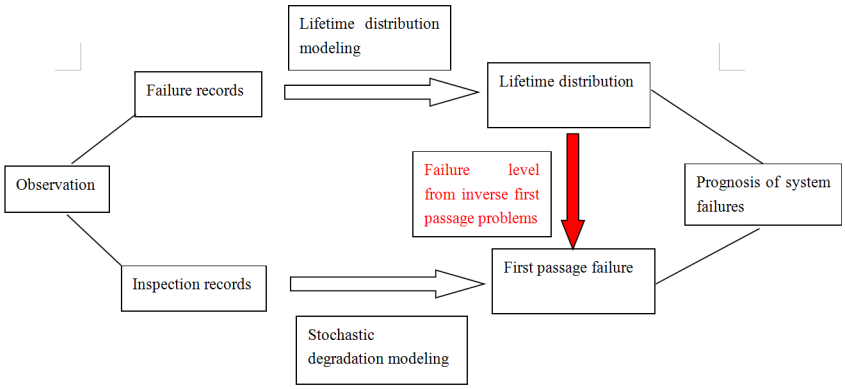


FIGURE 4.7 – The introduction of inverse first passage problems into PHM.

3. This maintenance optimization problem is discussed respectively for the drifted Brownian motion, time-changed Brownian motion and the OU process in simulation tests.

Perspectives

The perspectives of this thesis can be considered from also four aspects : degradation modeling, failure prognosis, failure level estimation and maintenance optimization.

The mean-reversion in stochastic degradation modeling implies the possibility to consider 2-stage models furthermore as in [75] and also Section 1.3.3. That is to say, the real degradation process is treated as a should-be processes given by physical laws. To make up the gap between this should-be process and observed data, the co-effect of mean-reversion and exciting-noise leads to a modified model.

Also the mean-reversion could be a starting point to introduce Levy-driven OU processes into stochastic degradation modeling. This is triggered by considering the risk of sudden change of system states or jumps in mathematical descriptions. However this is from a modeling view, it would be difficult to consider prognosis of system failures and involve more mathematical tools.

To do failure prognosis, the first passage problem is introduced for the OU process. This induces an interdisciplinary consideration among stochastic analysis, PDE, integral equation, numerical analysis etc.. The analysis in this thesis starts with $w(x,t|y,s) := \frac{\partial P(\tau_{y,s}>t, X_t<x|X_s=y)}{\partial x}$, but a possibility to follow other analysis exists. Relaxing the constraint of crossing boundaries by using the explicit expression from the method of images leads to many insightful results [43].

The view of $u(x,t|y,s) := P(\tau_{y,s} > t, X_t < x|X_s = y)$ leads to a free boundary problem in (2.3.9), which is essential for a rigorous analysis of inverse first passage problems [13, 12]. Also it leads to a starting point of optimal stopping problems [58].

The first passage failure to the pre-set failure level can hardly fit existing failure records. The effort to make up the gap between failure records and first passage failures are made from various considerations. Generally speaking, 2 ways exist to fulfill this task when boundary-crossing failures are considered. One is to modify the description of system failures, such as introducing last passage failure [59], passage duration [9] and killed FPT [29]. Such efforts consider the failure level is fixed but the failure description itself can be modified.

The other is the inverse first passage problem we have discussed in this thesis. We should notice the OU process is a simple process, such that it would be much more difficult when it comes to other diffusion processes. To extend the inverse first passage problem to other processes, e.g. jump processes or jump-diffusion processes could also be interesting.

Using analytical approximations of first passage density in maintenance optimization problems allows more possibilities rather than the continuously monitored system. A natural extension is to consider optimizing the inspection interval at the same time with the virtual failure level when the system is not continuously monitored [34].

Annexe A

Fokker-Planck Equation

A.1 Derivation

We here will derive the Fokker Planck Equation (2.1.3) for X_t satisfying the equation (2.1.1), the derivation of Fokker Planck equation for general SDEs may follow the same procedure. The following statement is based on the one by Einstein from Kamers-Moyal expansion [28, 37], other statements can refer to [62, 33].

First, it is noticed that X_t is a Markov process as it is a solution to Itô's SDE.

Second, as X_t is a Markov process, we consider three times t_1, t_2, t_3 where $t_0 \leq t_1 \leq t_2 \leq t_3, \Delta > 0$, and the transition pdfs satisfies the Chapman-Kolmogorov equation

$$p(x, t_3|y, t_1) = \int_{-\infty}^{+\infty} p(x, t_3|z, t_2)p(z, t_2|y, t_1)dz \tag{A.1.1}$$

Third, $\forall h(z) \in C^{\infty}(R)$ with compact support (i.e. it vanishes at ∞), for $\Delta > 0$, we have

$$\int_{-\infty}^{+\infty} h(x)\frac{\partial p(x, t|y, s)}{\partial t}dx = \lim_{\Delta\to 0}\frac{1}{\Delta}\int_{-\infty}^{+\infty} h(x)(p(x, t+\Delta|y, s) - p(x, t|y, s))dx. \tag{A.1.2}$$

Applying (A.1.1) into (A.1.2), the right side of (A.1.2) can be written by

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x)p(x, t+\Delta|u, t)p(u, t|y, s)dudx - \int_{-\infty}^{+\infty} h(x)p(x, t|y, s)dx\right]. \tag{A.1.3}$$

Noticing $\int_{-\infty}^{+\infty} p(x, t+\Delta|u, t)dx = 1$ and changing the notation x to u in the second term in (A.1.3), then we get

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\left[\int_{-\infty}^{+\infty} p(u, t|y, s)\int_{-\infty}^{+\infty}(h(x) - h(u))p(x, t+\Delta|u, t)dudx\right]. \tag{A.1.4}$$

Furthermore, based on Taylor's expansion for $h(x)$ at u, we have

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\left[\int_{-\infty}^{+\infty} p(u, t|y, s)\int_{-\infty}^{+\infty}\sum_{n=1}^{+\infty} h^n(u)\frac{(x-u)^n}{n!}p(x, t+\Delta|u, t)dudx\right]. \tag{A.1.5}$$

Noting

$$D^n(u, t) = \frac{1}{n!}\lim_{\Delta\to 0}\frac{1}{\Delta}\int_{-\infty}^{+\infty}(x-u)^n p(x, t+\Delta|u, t)dx, \tag{A.1.6}$$

it follows that

$$\int_{-\infty}^{+\infty} h(x)\frac{\partial p(x, t|y, s)}{\partial t}dx = \int_{-\infty}^{+\infty} p(u, t|y, s)\sum_{n=1}^{+\infty} D^n(u, t)h^n(u)du. \tag{A.1.7}$$

Integrating each term in the right side by parts n times, and using the assumptions on h, $\forall n \in N$, we know

$$\int_{-\infty}^{+\infty} p(u,t|y,s)D^n(u,t)h^n(u)du = \int_{-\infty}^{+\infty} h(u)(-\frac{\partial}{\partial u})^n \left[D^n(u,t)p(u,t|y,s)\right] du. \tag{A.1.8}$$

Substitute (A.1.8) into (A.1.7), we have

$$\int_{-\infty}^{+\infty} h(u)\{\frac{\partial p(u,t|y,s)}{\partial t} - \sum_{n=1}^{+\infty}(-\frac{\partial}{\partial u})^n \left[D^n(u,t)p(u,t|y,s)\right]\}du = 0 \tag{A.1.9}$$

As h is arbitrarily chosen, therefore $\forall u$, we should have

$$\frac{\partial p(u,t|y,s)}{\partial t} = \sum_{n=1}^{+\infty}(-\frac{\partial}{\partial u})^n \left[D^n(u,t)p(u,t|y,s)\right]. \tag{A.1.10}$$

Remark : The equation (A.1.10) is called commonly as Kramers-Moyal expansion.

Fourth, the remaining work is for calculating $D^n(u,t)$ in the case of OU process. Actually given X_t as in (2.1.1), we can write (A.1.6) in another form

$$\begin{aligned} D^n(u,t) &= \frac{1}{n!}\lim_{\Delta\to 0}\frac{1}{\Delta}\int_{-\infty}^{+\infty}(x-u)^n p(x,t+\Delta|u,t)dx \\ &= \frac{1}{n!}\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}((X_{t+\Delta}-X_t)^n)|_{X_t=u} \\ &= \frac{1}{n!}\frac{\mathbb{E}((dX_t)^n)}{dt}|_{X_t=u}, \end{aligned} \tag{A.1.11}$$

where $\mathbb{E}((X_{t+\Delta}-X_t)^n)|_{X_t=u}$ is the conditional expectation of $(X_{t+\Delta}-X_t)^n$ based on $X_t = u$.

Recalling the equation (2.1.1), we have

$$dX_t = (a_t X_t + b_t)dt + \sigma_t dB_t. \tag{A.1.12}$$

Therefore from $\mathbb{E}(dB_t \times dB_t) = dt, \mathbb{E}(dB_t \times dt) = 0, dt \times dt = 0, \mathbb{E}(dB_t^n \times d_t^m) = 0, n+m \geq 3$, we have

$$(dX_t)^n = \sum_{i=0}^{n} C_n^i (a_t X_t + b_t)^i \sigma_t^{n-i}(dB_t^i \times dt^{n-i}). \tag{A.1.13}$$

Equation (A.1.13) leads to

$$D^1(u,t) = a(t)u + b(t), D^2(u,t) = \sigma^2(t), D^n(u,t) = 0, n \geq 3. \tag{A.1.14}$$

Substitute (A.1.14) into (A.1.10), we then arrive at the Fokker Planck equation (2.1.3)

$$\frac{\partial p(u,t|y,s)}{\partial t} = (-\frac{\partial}{\partial u})\left[(a(t)u + b(t))p(u,t|y,s)\right] + \frac{\sigma^2(t)}{2}\frac{\partial^2}{\partial u}p(u,t|y,s). \tag{A.1.15}$$

Fifth, noting $p(u,t|y,s)$ is discussed based on the initial condition $X_s = y$, therefore when the distribution of y is given by f_y, the initial condition for (A.1.10) is $p(u,s|y,s) = f_y(u)$.

A.2 Time-Space Transform Method for the Initial-Value Problem

As stated in (2.1.3), the transition pdf of an OU process satisfies a linear Fokker-Planck equation. Here a time-space transform method to solve (2.1.3) is explained as follows. And it is adopted in many literatures, eg. [62]. And we recall the linear Fokker-Planck equation in (2.1.3) :

$$\frac{\partial p(x,t|y,s)}{\partial t} = \{-\frac{\partial(a(t)x+b(t))}{\partial x} + c(t)\frac{\partial^2}{\partial^2 x}\}p(x,t|y,s), \tag{A.2.1}$$

where $c(t) > 0$ and $p(\pm\infty, t|y,s) = 0, p(x,s|y,s) = f_y(x), t \geq s$. In the following we denote $p(x,t) := p(x,t|y,s)$.

Then we can do a variable transform as follows :

$$\begin{cases} \psi(x,t) = x\exp(\alpha(t,s)) + \beta(t,s); \\ \phi(t) = 2\gamma(t,s), \end{cases} \tag{A.2.2}$$

where $\alpha(t,s), \beta(t,s), \gamma(t,s)$ are given by (1.3.3).

Moreover, $c(t) > 0$ promises that $\phi(t)$ is monotone and therefore $(\tilde{t}, \tilde{x}) = (\phi(t), \psi(x,t)) : \mathbb{R}\times\mathbb{R} \to \mathbb{R}\times\mathbb{R}$ is a bijective. Then we define a new function $\eta(\tilde{x},\tilde{t}) = p((\tilde{x} - \beta(\phi^{-1}(\tilde{t}),s))e^{-\alpha(\phi^{-1}(\tilde{t}),s)}, \phi^{-1}(\tilde{t}))$, satisfying $\eta(\psi(x,t),\phi(t)) = p(x,t)$. For this new function, we should notice its natural boundary conditions $\eta(\pm\infty,\tilde{t}) = p(\pm\infty,\tilde{t}) = 0, \tilde{t} \geq \phi(s)$, and the initial condition $\eta(\tilde{x},\phi(s)) = p(\tilde{x},s) = f_y(\tilde{x}), \tilde{x} \in R$.

With this expression, for $\eta(\psi(x,t),\phi(t))$ the following equations hold from (1.3.3) and (A.2.2) :

$$\begin{cases} \frac{\partial p(x,t)}{\partial t} = \frac{\partial\eta}{\partial\tilde{x}}\frac{\partial\psi(x,t)}{\partial t} + \frac{\partial\eta}{\partial\tilde{t}}\frac{\partial\phi(t)}{\partial t} = \{(-xa(t)e^{\alpha(t,s)} - b(t)e^{\alpha(t,s)})\frac{\partial\eta}{\partial\tilde{x}} + 2c(t)e^{2\alpha(t,s)}\frac{\partial\eta}{\partial\tilde{t}}\}|_{\tilde{t}=\phi(t),\tilde{x}=\psi(x,t)}; \\ \frac{\partial p(x,t)}{\partial x} = \frac{\partial\eta}{\partial\tilde{x}}\frac{\partial\psi(x,t)}{\partial x} + \frac{\partial\eta}{\partial\tilde{t}}\frac{\partial\phi(t)}{\partial x} = e^{\alpha(t,s)}\frac{\partial\eta}{\partial\tilde{x}}|_{\tilde{t}=\phi(t),\tilde{x}=\psi(x,t)}; \\ \frac{\partial^2 p(x,t)}{\partial x^2} = e^{2\alpha(t,s)}\frac{\partial^2\eta}{\partial\tilde{x}^2}|_{\tilde{t}=\phi(t),\tilde{x}=\psi(x,t)}. \end{cases} \tag{A.2.3}$$

Substitute (A.2.3) into (A.2.1), under the condition $\tilde{t} = \phi(t), \tilde{x} = \psi(x,t)$, it comes to :

$$\frac{\partial\eta}{\partial\tilde{t}} = \frac{1}{2}\frac{\partial^2\eta}{\partial\tilde{x}^2} \tag{A.2.4}$$

where the boundary condition is $\eta(\pm\infty,\tilde{t}) = 0$, the initial condition is $\eta(\tilde{x},s) = f_y(\tilde{x})$.

For the equation (A.2.4), it is just the familiar heat equation, whose solution can be given easily by :

$$\eta(\tilde{x},\tilde{t}) = \frac{1}{\sqrt{2\pi\tilde{t}}}\int_{-\infty}^{+\infty}\exp(-\frac{(\tilde{x}-u)^2}{2\tilde{t}})f_y(u)du \tag{A.2.5}$$

Noticing the transform $\tilde{t} = \phi(t), \tilde{x} = \psi(x,t)$, inverse the above transforms, and the solution $p(x,t)$ to (2.1.3) is given :

$$\begin{aligned} p(x,t|y,s) = \eta(\psi(x,t),\phi(t)) &= \frac{e^{\alpha(t,s)}}{\sqrt{2\pi\phi(t)}}\int_{-\infty}^{+\infty}\exp(-\frac{(\psi(x,t)-u)^2}{2\phi(t)})f_y(u)du \\ &= \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\int_{-\infty}^{+\infty}\exp(-\frac{(xe^{\alpha(t,s)}+\beta(t,s)-u)^2}{4\gamma(t,s)})f_y(u)du. \end{aligned} \tag{A.2.6}$$

A.3 The Method of Images for Initial-Boundary Value Problems

Initial-Boundary Value Problems

Now we will turn to the initial-boundary value problem for Equation (A.2.1). The initial-boundary value problem for the constrained equation is presented to search a solution $w(x,t|y,s)$ satisfying (A.2.1) with the boundary condition at a one-sided, upper absorbing boundary $L(t)$ given by

$$w(L(t),t|y,s)=0, w(-\infty,t|y,s)=0, \tag{A.3.1}$$

with also the initial condition

$$w(x,s|y,s)=f_y(x),\ \ for\ x\leq L(s). \tag{A.3.2}$$

Method of Images

The method of images provides an illustrative way to solve initial-boundary value problems $w(x,t|y,s)$ when corresponding initial-value problem $p(x,t|y,s)$ is explicit. The key idea is to extend the initial condition for $x\leq L(s)$ into the whole real line, such that a new unbounded problem is presented.

Suppose the function $p(x,t|r,s)$ satisfies the initial-value problem for the Fokker-Planck equation (A.2.1) with a deterministic initial condition r, and it can be solved explicitly by (A.2.6). Then for $w(x,t|y,s)$, introducing a pending function $U(x)$ to the undefined internal $x>L(s)$, we can verify that the following equation satisfies the initial-value problem (A.2.1).

$$w(x,t|y,s)=\int_{-\infty}^{L(s)}p(x,t|r,s)f_y(r)dr-\int_{L(s)}^{+\infty}p(x,t|r,s)U(r)dr, \tag{A.3.3}$$

where $w(x,s|y,s)=f_y(x), x\leq L(s)$.

To esimate the pending function $U(x)$, it is referred to the boundary condition $w(L(t),t|y,s)=0$ such that

$$p(L(t),t|y,s)-\int_{L(s)}^{+\infty}p(L(t),t|r,s)U(r)dr=0,\forall t\geq s. \tag{A.3.4}$$

This is a Fredholm integral equation of first kind, and in general there is no explicit solution.

Annexe B

Approximation Optimization

In previous discussions on inverse first passage problems and also maintenance optimization problems, plenty of approximate expressions have been adopted to substitute the original objective functions in the optimization. Such a treatment is utilizable, and it appears frequently in similar situations where exact expressions for objective functions are not available. To see whether this approximation-optimization is appropriate, the discussion will be introduced in this section. A framework on sequential approximation optimization (SAO) can also be referred to as in [39].

Instead of a specific consideration, we will present a fundamental result on the convergence of the supreme of a function sequence. Suppose the objective function is given by $h(\theta)$ with θ in a compact space Θ. We want to find a θ^* such that $h(\theta^*) = \sup_{\theta\in\Theta} h(\theta)$. Then when approximating the original $h(\theta)$ by a series $h_q(\theta)$, for each q, we can do the optimization to get an optimal parameter θ_q. It is natural to guess that under some conditions, $\theta_q \to \theta^*$. Actually this is verified by the following lemma.

Lemma B.0.1. *Given a real function $h(\theta) \in C(\Theta)$ and a real function sequence $\{h_q(\theta)\}_{q=1}^{+\infty} \in C(\Theta)$, where Θ is a compact sub-space of R^n, we suppose that $\forall\theta \in \Theta$, uniformly $\lim_{q\to+\infty} h_q(\theta) = h(\theta)$. Then if $\{\theta_q\}_{q=1}^{+\infty}$ satisfy $h_q(\theta_q) = \sup_{\theta\in\Theta} h_q(\theta)$, a sub-sequence $\{\theta_{q_k}\}_{k=1}^{+\infty}$ exists such that $\lim_{k\to+\infty} \theta_{q_k} = \theta^* \in \Theta$, moreover the following equation holds :*

$$h(\theta^*) = \sup_{\theta\in\Theta} h(\theta). \tag{B.0.1}$$

Proof. To prove (B.0.1), we suppose there is a $\theta^1 \in \Theta$ such that $h(\theta^1) = \sup_{\theta\in\Theta} h(\theta)$. Then as uniformly $h_q \to h$, $\forall\epsilon > 0, \exists Q \in N, \forall q > Q$, we have :

$$|h_q(\theta_q) - h(\theta_q)| < \epsilon,\ |h_q(\theta^1) - h(\theta^1)| < \epsilon. \tag{B.0.2}$$

Moreover, from (B.0.2), we have

$$h(\theta^1) - \epsilon < h_q(\theta^1) \le h_q(\theta_q) < h(\theta_q) + \epsilon \le h(\theta^1) + \epsilon. \tag{B.0.3}$$

That is to say,

$$|h_q(\theta_q) - h(\theta^1)| \le \epsilon. \tag{B.0.4}$$

Therefore, $h_q(\theta_q) \to h(\theta^1)$ when $q \to +\infty$. Noticing that Θ is compact, there exists a subsequence $\{\theta_{q_k}\}_{k=1}^{+\infty}$ such that $\theta_{q_k} \to \theta^*$. Furthermore, the following inequality holds :

$$|h(\theta^1) - h(\theta^*)| \le |h(\theta^1) - h_{q_k}(\theta_{q_k})| + |h_{q_k}(\theta_{q_k}) - h(\theta_{q_k})| + |h(\theta^*) - h(\theta_{q_k})| \tag{B.0.5}$$

Let $k \to +\infty$ in (B.0.5), then from (B.0.4), (B.0.2) and the continuity of $h(\theta)$, $h(\theta^*) = h(\theta^1) = \sup_{\theta\in\Theta} h(\theta)$, i.e. (B.0.1) holds. □

If the uniqueness of θ^1 is given, then the uniqueness of θ^* is promised by (B.0.5) vice versa. Actually if we find a convergent sub-sequence θ_{q_k}, then the lemma says that the limit will just be one of the optimal parameters for the original function. Especially this is true when $h(\theta)$ is convex (or concave), naturally we have the following proposition.

Proposition B.0.2. *Under the conditions in Lemma B.0.1, if h is convex (or concave), then there exists a point θ^*, $\theta_k \to \theta^*$. Moreover $h(\theta^*) = \sup_{\theta\in\Theta} h(\theta)$.*

Proof. From Lemma B.0.1, a sub-sequence $\{\theta_{q_k}\}_{k=1}^{+\infty}$ for h_q exists such that $\lim_{k\to+\infty} \theta_{q_k} = \theta^*$, then $h(\theta^*) = \sup_{\theta\in\Theta} h(\theta)$. If there exists another sub-sequence $\{\theta^*_{q_k}\}_{k=1}^{+\infty}$ exists such that $\lim_{k\to+\infty} \theta^*_{q_k} = \theta^{**}$, $h(\theta^{**}) = \sup_{\theta\in\Theta} h(\theta)$ holds. As h is a convex function, only one global optimization result is returned such that $\theta^{**} = \theta^*$. □

Remark : in another word, when the approximation for the original function is enough accurate, the result of approximation-optimization can be treated an approximate solution for the original optimization problem.

In this thesis, the approximation occurs almost all in the cases of discretizing the integral. Therefore to validate the uniform convergence of the approximate integral schemes such that the condition in Lemma B.0.1 is satisfied, the following lemma is introduced for the compound trapezoid scheme (which can be found easily in any textbook on numerical analysis) :

Lemma B.0.3. *For $f \in C^2[a,b]$, $q \in N$,and $\Delta := \frac{b-a}{q}$, then the following equation holds :*

$$\int_a^b f(x)dx - \Delta\left[\frac{1}{2}f(a) + \sum_{i=1}^{q-1} f(a+i\Delta) + f(b)\right] = -\frac{(b-a)^3}{12q^2}f''(\zeta), \tag{B.0.6}$$

for some $\zeta \in (a,b)$.

Furthermore, if f is with parameters θ in Θ, define $h(\theta) = \int_a^b f(x|\theta)dx$, and $h_q(\theta) := \Delta\left[\frac{1}{2}f(a|\theta) + \sum_{i=1}^{q-1} f(a+i\Delta|\theta) + f(b|\theta)\right]$, then

$$h_q(\theta) \to h(\theta),\ uniformly\ when\ q \to \infty. \tag{B.0.7}$$

Proof. The proof is omitted. □

Annexe C

Résume de Thèse en Français

C.1 Introdution

C.1.1 Contexte

L'objet central de ces travaux de thèse est l'utilisation d'informations de surveillance pour décrire, prédire et prévenir la défaillance de systèmes. Ces travaux s'inscrivent ainsi dans une démarche de *Prognostic and Health Management* (PHM) [70] visant à augmenter la disponibilité et la sécurité d'un système en se basant sur une maintenance conditionnelle [16, 34] du système, fonction de son état courant.

Ces travaux de thèse s'organisent autour de trois problématiques qui sont :

- la description de la défaillance,
- la prédiction de l'instant de défaillance,
- la prévention de la défaillance.

L'approche adoptée repose sur l'hypothèse que le processus de dégradation du système peut être modélisé [1]. Sur la base de ce modèle et des observations réalisées sur le système, l'état de santé courant du système peut être évalué et l'évolution de cet état peut être prédit. Si par ailleurs, la défaillance du système est associée au temps d'atteinte d'un état de santé critique, la prédiction de l'instant de défaillance du système peut également être réalisée. Ces informations peuvent enfin être exploitées dans le cadre d'une démarche d'optimisation de la maintenance.

C.1.2 Organisation du document

Cette thèse est organisé en quatre parties :

1. la modélisation stochastique de la dégradation,
2. le pronostic de l'instant de défaillance du système,
3. l'estimation du niveau dégradation associé à la défaillance,
4. l'optimisation de la maintenance.

Ces différentes parties sont liées entre elles et tentent de décrire, prévoir et prévenir la défaillance du système.

Dans la première partie, un modèle stochastique de la dégradation s'appuyant sur un processus d'Ornstein-Uhlenbeck (OU) dépendant du temps et sur l'exploitation conjointe des donnés d'inspection est proposé. Les qualités de ce modèle sont démontrées au travers de ses propriétés statistiques qui permettent d'ajuster de manière indépendante la moyenne, la variance et la corrélation. Une propriété de "convergence" vers la moyenne est ensuite exploitée pour interpréter la corrélation temporelle des fluctuations autour d'une tendance globale de dégradation. Puis, s'appuyant sur une technique de maximisation de la vraisemblance, une

méthode d'estimation des paramètres de ce modèle est proposée. Enfin, un cas d'application portant sur l'étude de la dégradation d'un composant passif de central électrique est traité.

La deuxième partie de la thèse est consacrée au pronostic de l'instant de défaillance du système en s'appuyant sur un processus OU dépendant du temps. Cet instant de défaillance est défini comme le premier temps d'atteinte d'un état de dégradation, critique i.e. d'un état de santé inacceptable.. L'estimation de cet instant de défaillance est abordé selon deux approches : *i)* équations aux dérivés partielles, *ii)* équations intégrales. Ces approches conduisent à différentes techniques d'estimation qui peuvent être classées selon le schéma suivant : les techniques d'approximation analytique, les techniques d'approximation numérique, les techniques de simulation de Monté-Carlo.

Des essais numériques destinés au calcul de la densité de l'instant de défaillance et permettant la confrontation de ces différentes techniques, concluent cette seconde partie de la thèse.

L'estimation du niveau dégradation associé à la défaillance, que l'on appelle dans la suite par commodité niveau de défaillance, est l'objet de la troisième partie du document. Classiquement, ce niveau de défaillance est déterminé sur la base de caractéristiques physiques ou d'avis d'experts. Cependant ce niveau de défaillance "théorique" n'est pas toujours cohérent avec les données associées à des défaillances réelles. L'accent est donc mis sur la réduction de cet écart. Pour ce faire, la loi de la durée de vie est supposée connue ou, tout au moins, estimée sur la base de données de défaillances. Le niveau de défaillance peut alors être déterminé de telle sorte que le processus stochastique de dégradation considéré conduise à une distribution du premier temps d'atteinte du niveau de défaillance qui corresponde à la densité estimée.

La quatrième partie est dédiée à l'optimisation de la maintenance lorsque le processus de dégradation considéré est un processus OU dépendant du temps. S'appuyant sur les données de surveillance continue du système et sur le pronostic de l'instant de défaillance, un compromis peut être trouvé entre les coûts de maintenance préventive et ceux associés à une défaillance du système. Dans ce contexte, la formulation non explicite du pronostic de l'instant de défaillance ne permet pas d'exploiter les techniques d'optimisation classiques. Une approximation de la densité de l'instant de défaillance est donc proposée pour mener à bien cette étape d'optimisation.

C.1.3 Contributions principales

Les contributions principales de ces travaux sont les suivantes :

- le processus OU dépendant du temps est proposé pour la modélisation stochastique de la dégradation dans le cadre d'une démarche de *Prognostic and Health Management*.
- le prognostic de l'instant de défaillance est considéré d'un point de vue pratique. En d'autres termes on s'interroge sur la possibilité d'utiliser et d'exploiter des techniques de calcul existantes et on cherche à tester leur mise en oeuvre dans des configurations rélles ou réalistes. Dans ce contexte, des approximations analytiques, des techniques de calcul numériques, et une méthode de Monté-Carlo quasi-linéaire par morceaux sont proposés.
- la détermination du seuil de dégradation critique menant à la défaillance est vu et

traité comme un problème inverse.
- l'optimisation de la maintenance d'un système surveillé en continu et dont la dégradation est modélisée par un processus OU fonction du temps est présentée.

C.2 Processus d'Ornstein-Uhlenbeck dépendant du temps pour la modélisation de la dégradation

Un processus d'Ornstein-Uhlenbeck (OU) dépendant du temps est proposé pour modéliser la dégradation. Ce processus présente l'avantage d'avoir une moyenne, une variance et une corrélation qui peuvent être fixés de manière indépendante. Une propriété de "convergence" vers la moyenne est également étudié pour permettre d'interpréter des fluctuations temporaires autour d'une tendance générale. L'estimation des paramètres par maximum de vraisemblance est présentée. Plusieurs essais numériques sur données réelles sont présentés également.

Dans cette thèse, nous nous concentrons sur les systèmes soumis à des processus de dégradation progressifs tels que la fissuration par fatigue, la corrosion, l'érosion, *etc.* D'une manière générale, les données issues d'inspections de ce type de dégradation se caractérisent par des évolutions (taille, profondeur, *etc.*) différentes pour un même système et un même environnement. Ces différences sont généralement associées, d'une part, aux incertitudes inhérentes aux équipements de mesures et, d'autre part, aux mécanismes internes de détérioration. Les modèles physiques et purement déterministes de ces mécanismes de dégradation ne peuvent pas expliquer et représenter ces variations.

Lorsque la modélisation du mécanisme de dégradation par une approche physique s'avère insuffisante, elle peut être abordée comme un problème de régression ou de filtrage pour un modèle choisi. Dans ce cas, la flexibilité du modèle est cruciale pour approcher correctement les données de dégradation disponibles. Dans ce contexte, nous proposons de modéliser le système de détérioration par un processus stochastique en temps continu que nous appelons processus de dégradation. Une hypothèse réaliste est que les fluctuations temporaires de la dégradation peuvent être tolérées tandis que la tendance générale de dégradation doit être conservée.

Pour les processus couramment utilisés tels que le processus Gamma [84] ou le mouvement brownien [66], la tendance de la moyenne (ou la "dérive") peut être choisie relativement librement. En revanche, la variance dépend fortement du processus envisagé. L'accent a donc été mis sur la mise en cohérence des propriétés statistiques du processus de dégradation considéré et des données de dégradation disponibles ainsi que la possible prise en compte d'avis d'experts. L'élaboration d'un modèle permettant de choisir librement à la fois la tendance de la moyenne et la variance est la principale contribution de cette section et justifie l'utilisation d'un processus d'Ornstein-Uhlenbeck dépendant du temps.

La dégradation est souvent vue comme un phénomène purement monotone pouvant être modélisé par des processus stochastiques tels que le processus de Lévy ou les processus de sauts incluant notamment les processus Gamma [84], et inverse Gaussien [88]. Ceci vient notamment de l'hypothèse que la dégradation résulte de l'accumulation de petits chocs sans réparation ou auto-réparation [75]. La propriété de monotonie conduit à une formulation explicite de l'instant de défaillance (il s'agit d'un temps d'atteinte) pour un

niveau de défaillance constant. Cette formulation facilite amplement l'estimation de l'instant de défaillance, de la durée de vie résiduelle (*Remaining Useful Life*, RUL) et du temps moyen avant la défaillance (*Mean Time To Failure*, MTTF), *etc.* Toutefois, un processus monotone ne peut modéliser correctement les fluctuations observées dans les données de dégradation. Pour décrire ces fluctuations, une première approche consiste à les imputer à des phénomènes stochastiques externes tels que les erreurs de mesure. Cela conduit à introduire dans les modèles de dégradation précédents (physique et stochastique) un bruit extérieur. L'estimation du processus de dégradation devient alors un problème de régression pour les modèles déterministes et un problème de filtrage pour les modèles stochastiques. Par exemple, Le Son et al. introduisent un bruit additif gaussien indépendant sur un modèle de processus Gamma [78].

Une seconde approche pour décrire ces fluctuations est de les imputer à des phénomènes internes au mécanise de dégradation. On peut citer par exemple dans le cas de la propagation de fissure de fatigue, la non-homogénéité des matériaux, les variations de charges (incertitudes) ou les effets thermiques [52, 73, 75, 76]. Cette approche conduit à introduire des phénomènes aléatoires sur des systèmes dynamiques : l'ajout de bruit blanc est communément acceptée et permet d'introduire des fluctuations gaussiennes corrélées que l'on retrouve notamment dans les données de dégradation. On parle généralement alors de modèle de diffusion. Ces modèles de diffusion sont établis selon des lois physiques spécifiques telles qu'ils peuvent difficilement être appliqués à un processus de dégradation général où les mécanismes de dégradation ne sont pas, ou pas encore bien connus. De plus, l'estimation du premier temps d'atteinte d'un niveau de défaillance est délicate dans le cas général des processus de diffusion. Aussi, très souvent, les modèles de diffusion effectivement considérés sont réduits au mouvement Brownien [54, 45] où le premier temps d'atteinte d'un niveau de défaillance peut être explicitement obtenu par la distribution gaussienne inverse. La linéarité du mouvement brownien limitant son application, Si et al. [66], Tseng et Peng [82] ont considéré un processus avec une dérive non linéaire et une diffusion linéaire. L'estimation de la RUL repose alors sur une expression approchée de la densité de premier passage d'un processus de Gauss [25]. Un modèle plus général, le processus de Gauss-Markov [23] a été récemment repris par Peng et Tseng [55]. L'estimation de la RUL est obtenue alors par la résolution numérique d'une équation intégrale. Pour l'ensemble de ces modèles, on peut regretter cependant que l'introduction des fluctuations autour d'une tendance moyenne amène à considérer un processus dont la variance n'est pas vraiment "contrôlable" : on ne dispose pas de suffisament de degrès de liberté pour ajuster à la fois la moyenne et la variance de manière "optimale" aux données.

La suite de ce chapitre est donc consacrée à la définition d'un modèle dynamique présentant un bon compromis entre capacité d'ajustement et efficacité d'exploitation pour l'estimation de la RUL. Le processus d'Ornstein-Uhlenbeck (OU) dépendant du temps s'avère être un bon candidat et ses propriétés sont tout d'abord présentées pour justifier ce choix. Les capacités de contrôle de critères statistiques tels que la moyenne, la variance et la corrélation seront discutées, de même que l'estimation du premier temps d'atteinte d'un niveau de défaillance donné (estimation de la RUL). La propriété de "convergence" vers la moyenne sera ensuite présentée. Elle pourra être interprétée comme un mécanisme d'auto réparation permettant de compenser les incertitudes inhérentes au processus OU dépendant du temps. Elle permettra également de mettre en avant le fait que même si des fluctuations existent au sein des données de dégradation, ces fluctuations sont nécessairement temporaires et le comportement moyen tend forcément vers une dégradation monotone.

Pour conclure ce premier châpitre, on précise ci-dessous les propriétés du système considéré dans la suite de ces travaux :

- la dégradation du système est supposée graduelle, sans choc important, de telle sorte que le processus de dégradation puisse être supposé continu.
- les relevés de dégradation sont entachés de bruit et sont associés à un processus de dégradation sous-jacent ayant une tendance monotone à long terme.
- les bruits des données de dégradation sont Gaussien et proviennent de mécanismes internes. Ils sont susceptibles de s'accumuler et d'influencer les relevés de données de dégradation suivant.
- Il existe des mécanismes "d'autoréparation" permettant au système de ne pas s'éloigner du processus de dégradation sous-jacent.

C.3 Le processus d'Ornstein-Uhlenbeck

Propriétés principales

Le processus OU est un des rares processus pouvant être traité de manière explicite par des lois de probabilité issues d'équations différentielles stochastiques. L'application d'un tel processus dans le domaine de la fiabilité est intéressant pour la constitution d'un bon compromis entre capacité de modélisation (d'ajustement) et calculabilité. Le processus OU est largement utilisé dans le domaine des marchés financiers pour décrire notamment la dynamique de systèmes se stabilisant sur un point d'équilibre. Il ne peut cependant pas être directement appliqué à la modélisation de la dégradation étant donné que sa moyenne est constante ce qui va à l'encontre de la tendance générale d'un phénomène de dégradation. Deux modifications sont donc proposées pour compenser cet inconvénient :

1. l'ajout d'une dérive, comme proposé dans [47] ;
2. la prise en compte de coefficients dépendants du temps, comme proposé dans [2].

De ce qui précède, il résulte qu'il faudra traiter l'équation stochastique différentielle d'Itô [92] :

$$dY_t = (a(t)Y_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0. \tag{C.3.1}$$

où a, b et σ sont des fonctions assez lisses et B_s est le mouvement brownien standard. La valeur initiale Y_0 est une variable aléatoire indépendante de B_t (y compris dans le cas déterministe où la distribution tend vers une mesure de Dirac).

Les notations suivantes seront, par ailleurs, considérées tout au long de ce document :

$$\begin{aligned} \alpha(t,s) &= -\int_s^t a(u)du, \\ \beta(t,s) &= -\int_s^t b(u)e^{\alpha(u,s)}du, \\ \gamma(t,s) &= \int_s^t c(u)e^{2\alpha(u,s)}du, \end{aligned} \tag{C.3.2}$$

où $c(t) = \frac{\sigma^2(t)}{2} > 0$.

L'équation (C.3.1) est trop générale pour être directement connectée à la modélisation de dégradation, mais plusieurs propriétés intéressantes peuvent être soulignées :

- une forme explicite de Y_t peut être déduite de l'équation (C.3.1). Les notations sont reprises de l'équation (C.3.2) :

$$Y_t = e^{-\alpha(t,0)}\left[Y_0 - \beta(t,0) + \int_0^t \sigma(s)e^{\alpha(s,0)}dB_s\right]. \tag{C.3.3}$$

- la moyenne du processus est donnée par :

$$\mathbb{E}(Y_t) = e^{-\alpha(t,0)}\Big(\mathbb{E}(Y_0) - \beta(t,0)\Big). \tag{C.3.4}$$

- la covariance du processus est donnée par :

$$\operatorname{cov}(Y_t, Y_s) = e^{-(\alpha(t,0)+\alpha(s,0))}\left\{\operatorname{var}(Y_0) + \int_0^{t\wedge s} \sigma^2(u)e^{2\alpha(u,0)}du\right\}, \tag{C.3.5}$$

- la variance correspondante s'écrit donc :

$$\operatorname{var}(Y_t) = e^{-2\alpha(t,0)}\left\{\operatorname{var}(Y_0) + \int_0^{t} \sigma^2(u)e^{2\alpha(u,0)}du\right\}. \tag{C.3.6}$$

- le processus de "coefficient de corrélation" est le suivant :

$$\rho_{Y_t,Y_s} = \frac{\operatorname{cov}(Y_t,Y_s)}{\sqrt{\operatorname{var}(Y_t)\operatorname{var}(Y_s)}} = \frac{\sqrt{\operatorname{var}(Y_0) + \int_0^s \sigma^2(u)e^{2\alpha(u,0)}du}}{\sqrt{\operatorname{var}(Y_0) + \int_0^t \sigma^2(u)e^{2\alpha(u,0)}du}}, t \geq s \tag{C.3.7}$$

Enfin, il est important de préciser que le processus OU est markovien à trajectoires continues.

Un processus OU fonction du temps

A partir de maintenant, un processus OU particulier est considéré. Il est noté X_t au lieu de Y_t, et est décrit par l'équation différentielle stochastique d'Itô suivante :

$$dX_t = (a(t)X_t + m'(t) - a(t)m(t))dt + \sigma(t)dB_t, \quad t \geq 0, \tag{C.3.8}$$

où X_0 est une variable aléatoire $\mathbb{E}(X_0) = m(0), \operatorname{var}(X_0) = v(0)$.

Le processus X_t est non-stationnaire, inhomogène en temps et son expression explicite peut être formellement dérivée de l'équation (C.3.3) :

$$X_t = e^{-\alpha(t,0)}\Big(X_0 - \beta(t,0) + \int_0^t \sigma(u)e^{\alpha(u,0)}dB_u\Big), \tag{C.3.9}$$

avec les notations suivantes mises à jour :

$$\alpha(t,s) = -\int_s^t a(u)du \text{ et } \beta(t,s) = m(s) - m(t)e^{\alpha(t,s)}.$$

Il peut être déduit naturellement de l'équation (C.3.9) que $\mathbb{E}(X_t) = m(t)$.

C.3.1 Propriété de convergence vers la moyenne

On peut montrer que le processus X_t est exactement déterminé par la connaissance des fonctions $m(t)$, $v(t)$ et $\sigma(t)$ qui décrivent respectivement sa moyenne, sa variance et la corrélation à long terme. Sur le plan statistique, ce processus semble pertinent pour la modélisation d'observations non monotones. En outre, le processus OU fonction du temps présente une propriété de retour à la moyenne qui fait qu'à long terme la grandeur X_t converge vers sa valeur moyenne. En effet, supposons que le processus de dégradation a été parfaitement observé à l'instant s avec $X_s = y$. De l'équation (C.3.9), l'espérance $E(X_t^{y,s})$, $t \geq s$ peut être écrite comme :

$$\mathbb{E}(X_t^{y,s}) = m(t) + (y - m(s)) \exp(\int_s^t a(u)du). \tag{C.3.10}$$

Lorsque $a(t)$ est négatif et satisfait l'équation $\int_s^\infty a(u)du = -\infty$, l'influence de l'observation réalisée à l'instant s sur l'estimation de la valeur moyenne du processus tend à s'annuler au fur et à mesure que l'on s'éloigne de s.

C.3.2 Estimation des paramètres du modèle

Dans cette section, la problème de l'inférence statistique du processus OU est traité. La première étape consiste à choisir certaines fonctions paramétriques pour $m(t)$, $v(t)$ et $\sigma(t)$. Dans la suite, le vecteur de paramètres utiles sera noté θ et les fonctions correspondantes respectivement $m(t|\theta)$, $v(t|\theta)$ et $\sigma(t|\theta)$. Ces fonctions sont dépendantes du cas d'étude et peuvent être choisies au sein d'une liste prédéterminée, ce choix s'appuyant sur les statistiques descriptives des données disponibles ou sur des avis d'experts.

Soit $X_{t;\theta}$ le processus stochastique décrit par l'équation (C.3.8). La densité de probabilité $p(x,t|\theta) := p(x,t|x_s,s;\theta)$, $t > s \geq 0$ peut être déduite de l'équation stochastique intégrale d'Itô [80].

Il peut être montré que l'expression analytique de $p(x,t|y,s;\theta)$ est donnée par :

$$p(x,t|y,s;\theta) = \frac{e^{\alpha(t,s|\theta)}}{\sqrt{4\pi\gamma(t,s|\theta)}} \exp(-\frac{(xe^{\alpha(t,s|\theta)} + \beta(t,s|\theta) - x_s)^2}{4\gamma(t,s|\theta)}), \tag{C.3.11}$$

où $\alpha(t,s|\theta), \beta(t,s|\theta), \gamma(t,s|\theta)$ et $c(u|\theta) := \frac{\sigma(u|\theta)^2}{2}$ sont donnés :

$$\begin{aligned} \alpha(t,s|\theta) &= -\int_s^t a(u|\theta)du, \\ \beta(t,s|\theta) &= m(s|\theta) - m(t|\theta)e^{\alpha(t,s|\theta)}, \\ \gamma(t,s|\theta) &= \int_s^t c(u|\theta)e^{2\alpha(u,s|\theta)}du. \end{aligned} \tag{C.3.12}$$

Dans la suite, on suppose que les données de dégradation proviennent de n composants indépendants. On note m_i le nombre de données recueillis sur l'i-ième composant avec $(1 \leq i \leq n)$, et (x_{ij}, t_{ij}) l'enregistrement à l'instant j du i-ième composant avec $j \in \{1, \ldots, m_i\}$. On note dans le cas général que X_0 suit une distribution $F_{0,\theta}$ et que la distribution de la transition correspondante, basée sur la condition initiale $p(x,t|X_0,0;\theta)$ peut être écrite comme :

$$p(x,t|X_0,0;\theta) = \int_{-\infty}^{+\infty} p(x,t|u,0)dF_{0,\theta}(u) \tag{C.3.13}$$

Si X_0 est considéré comme déterministe, sa loi de probabilité devient une mesure de Dirac avec $t_{i0} = 0$ et $x_{i0} = x_0$ pour tous $i \in \{1, \ldots, n\}$. Comme le processus stochastique considéré est un processus de Markov, la densité de probabilité conditionnelle d'un enregistrement connaissant les enregistrements précédents de la même trajectoire peut être simplifiée :

$$p(x_{ij}, t_{ij}|x_{i(j-1)}, t_{i(j-1)}; \theta) = p(x_{ij}, t_{ij}|\{(x_{iz}, t_{iz})\}_{z=0}^{j-1}; \theta).$$

La fonction log-vraisemblance pour la i-ième trajectoirevaut alors :

$$\log L_i(\theta) = \sum_{j=1}^{m_i-1} \log\Big(p\big(x_{i(j+1)}, t_{i(j+1)}|x_{ij}, t_{ij}; \theta\big)\Big) + \log(p(x_{i1}, t_{i1}|X_0, 0; \theta). \tag{C.3.14}$$

Si l'on considère l'indépendance des différentes trajectoires (des différents composants), la fonction log-vraisemblance pour l'ensemble des données disponibles devient :

$$\log L(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{m_i-1} \log\Big(p\big(x_{i(j+1)}, t_{i(j+1)}|x_{ij}, t_{ij}; \theta\big)\Big) + \sum_{i=1}^{n} \log(p(x_{i1}, t_{i1}|X_0, 0; \theta)). \tag{C.3.15}$$

L'estimation du maximum de vraisemblance $\theta^* = \text{argmax}_\theta \log L(\theta)$ des paramètres du modèle peut être obtenue par maximisation de la fonction du log-vraisemblance (C.3.15). Une solution analytique est généralement délicate à obtenir et des solutions numériques telles que la méthode Nelder-Mead mises en oeuvre.

C.4 Pronostic de défaillance du système et temps d'atteinte

Le pronostic de l'instant de défaillance du système est traité dans ce chapitre. Ce pronostic s'appuie sur le processus OU dépendant du temps venant d'être présenté et qui est supposé ici être le processus stochastique de dégradation du système considéré. Ce chapitre comporte plusieurs résultats publiés [20, 18, 17] et met l'accent sur l'estimation des lois de probabilité de l'instant de premier passage (*First Passage Time*, FPT) d'un niveau de dégradation donnée (précédemment nommé niveau de défaillance). On parelra également de temps d'atteinte. Différentes méthodes sont proposées, celles-ci sont notamment connectée à l'équation de Fortet [32, 57].

C.4.1 Introduction

Contexte

L'objectif de ce chapitre est de proposer des méthodes permettant l'estimation de la durée de vie résiduelle d'un système (*Remainig Useful Life*, RUL). Cette RUL est souvent utilisée comme un indicateur de décision clé dans les démarches de type PHM. Elle est communément définie comme : la durée de vie résiduelle d'un système à l'instant t étant donné toutes les informations disponibles jusqu'au temps t.

Le problème central est donc de déterminer quand la défaillance du système va se produire. Pour rappel, nous nous focalisons dans ce document sur les phénomènes de dégradation

cumulatifs (corrosion, érosion, propagation de fissure, *etc.*) qui "s'intègrent" jusqu'à la rupture mécanique du système. Dans ce contexte, la notion de temps d'atteinte est primordiale, elle correspond à la première atteinte, par accumulation de dégradation, d'un niveau de dégradation donné associé généralement à la défaillance du système (on parle également de niveau de défaillance [1]).

En fiabilité, lorsque des processus stochastiques monotones croissants sont considérés, il est aisé d'obtenir la loi du temps d'atteinte en raison de la propriété de monotonie. C'est le cas pour le processus Gamma [84], le processus de Poisson ou le processus Gaussien inverse [93].

Dans le cas général de l'utilisation des modèles de diffusion, l'obtention de la loi du temps d'atteinte est plus délicate : la propriété de monotonie n'étant plus acquise. On pourra dans ce domaine notamment se référer au travail bibliographique réalisé dans [24, 83]. Les résultats les plus intéressants sont obtenus lorsqu'un mouvement Brownien et une limite linéaire sont considérés. Dans ce cas, la loi du temps d'atteinte peut être analytiquement déterminée et prend la forme d'une distribution Gaussienne inverse. Le développement d'algorithmes efficaces fait également l'objet de nombreux travaux. Ceux-ci s'appuient généralement *i)* soit sur l'équation intégrale [57], *ii)* soit sur l'équation de Fokker-Planck [43, 53].

Dans la suite de ce chapitre, ces deux dernières approches sont examinées et adaptées au processus OU. Une connexion entre ces deux approches est proposée au travers de l'équation de Fortet [24, 83]. Ces deux approches conduisent à différentes méthodes pour l'estimation de la loi du temps d'atteinte au sein desquelles trois catégories peuvent être identifiées :

1. les approximations analytiques,
2. les algorithmes numériques,
3. la simulation Monté-Carlo.

Notons, pour achever la présentation du contexte de ces travaux, que ce chapitre ne porte pas seulement sur le problème de l'estimation de la loi du temps d'atteinte mais également sur celui de l'estimation de la loi conditionnelle du temps d'atteinte. En effet, l'orientation PHM de nos travaux nous conduit à mettre jour le pronostic de durée de vie résiduelle du système sur la base de observations nouvellement réalisées. Comme nous le verrons dans la suite, cette "extension" n'est pas directe lorsque le processus de dégradation considéré est non-homogène et non stationnaire.

Temps d'atteinte ou instant de premier passage

Nous considérons le processus X_t tel que défini équation pas l'(C.3.1). L'instant de premier passage $\tau_{y,s}$ lorsque X_t atteint le niveau de défaillance prédéterminé $L(t)$, basée sur l'observation (y, s), c'est-à-dire sachant $X_s = y$ est donnée par la relation suivante :

$$\tau_{y,s} := \inf_{t \geq s} \Big\{ t | X_t \geq L(t) \Big\}, \tag{C.4.1}$$

Dans la suite, nous supposons : $P(\tau_{y,s} < \infty) = 1$ et, par ailleurs, que lorsque $y \geq L(s)$ alors $\tau_{y,s} = s$.

1. On trouve dans la littérature de nombreuses expressions similaires : seuil d'alarme, limite de sécurité, barrière de sécurité, passage level, *etc. Niveau de défaillance* est le terme qui sera utilisé dans le contexte "applicatif" ; *Limite de passage* est le terme qui sera utiliés dans le contexte mathématique.

La durée de vie résiduelle basée sur l'observation actuelle (y, s) est naturellement définie par la relation $\text{RUL}_s := \tau_{y,s} - s$. En conséquence, le temps moyen avant la défaillance (MTTF_s) conditionnellement à l'observation actuelle (y, s) est défini par $E(\tau_{y,s}) - s$.

En outre, en l'absence d'observation parfaite, par exemple en raison de petites fissures indétectables [17], quand la valeur initial x_0 est traitée comme une variable aléatoire avec la fonction de distribution F_0, l'instant de premier passage τ_{x_0,t_0} peut être considéré comme un temps d'atteinte aléatoire (*Randomized First Passage Time*, RFPT) comme décrit dans [38, 83, 48]. Sa loi est calculée pour $t \geq t_0$ par :

(C.4.1) :

$$\mathbb{P}(\tau_{x_0,t_0} \leq t) = \int_{-\infty}^{L(t_0)} \mathbb{P}(\tau_{u,t_0} \leq t) dF_0(u) + \int_{L(t_0)}^{+\infty} dF_0(u). \tag{C.4.2}$$

La densité de probabilité et la fonction de répartition de $\tau_{y,s}$ sont respectivement notées :

$$g(t|y,s) := \frac{\partial P(\tau_{y,s} \leq t | X_s = y)}{\partial t}, \quad G(t|y,s) = P(\tau_{y,s} \leq t | X_s = y). \tag{C.4.3}$$

On note également :

$$u(x,t|y,s) := P(X_t < x, \tau_{y,s} > t|y,s), \quad w(x,t|X_s = y) = \frac{\partial u(x,t|y,s)}{\partial x}. \tag{C.4.4}$$

C.4.2 Méthodes de résolution

Limite quasi-linéaire

Proposition C.4.1. *Pour le processus X_t donné par l'équation (C.3.1) avec la condition initiale (y, s) où y est une variable aléatoire définie sur $(-\infty, L(s)]$, la loi du temps d'atteinte $g(t|y, s)$ satisfait :*

$$g(t|y,s) = \frac{c(t)e^{2\alpha(t,s)}}{\sqrt{4\pi\gamma^3(t,s)}} \int_{-\infty}^{L(s)} (L(s) - r) \exp\left(-\frac{(\beta(t,s) + L(t)e^{\alpha(t,s)} - r)^2}{4\gamma(t,s)}\right) f_y(r)dr, \tag{C.4.5}$$

si la frontière $L(t)$ satisfait :

$$L(t) = e^{-\alpha(t,s)}(L(s) - \beta(t,s) + C\gamma(t,s)), \tag{C.4.6}$$

où $C \in \mathbb{R}$.

En outre, la fonction de répartition $G(t|y, s) := P(\tau_{y,s} \leq t)$ satisfait :

$$G(t|y,s) = \int_{-\infty}^{L(s)} \left[\Phi\left(\frac{-L(s) - C\gamma(t,s) + r}{\sqrt{2\gamma(t,s)}}\right) + e^{C(r-L(s))}\Phi\left(\frac{C\gamma(t,s) - L(s) + r}{\sqrt{2\gamma(t,s)}}\right)\right] f_y(r)dr \tag{C.4.7}$$

où $\Phi(*)$ est la distribution de la loi normale.

Remarque : Dans la suite nous ferons référence à la limite (C.4.6) comme à la limite quasi-linéaire.

Approximation de la tangente

Le principe de la méthode d'approximation de la tangente est d'approximer la limite réelle par une limite linéaire. Cette méthode est utilisée pour les problèmes de temps d'atteinte de mouvements Brownien et est notamment étudiée dans [43]. Nous appliquons ici cette démarche au processus OU en nous appuyant sur la limite quasi linéaire identifiée par l'équation (C.4.6).

Ainsi, nous souhaitons approximer la limite $L(t)$ par une équation de la forme de celle donnée (C.4.6) : cela est possible *via* l'approximation de la tangente.

Ainsi, quand $L(t) \in C^1[s,t]$, on peut noter lorsque s est proche de t que :

$$\tilde{C}(t,s) := \frac{L'(t) - a(t)L(t) - b(t)}{c(t)e^{\alpha(t,s)}} \approx \frac{L(t)e^{\alpha(t,s)} - L(s) + \beta(t,s)}{\gamma(t,s)}, \tag{C.4.8}$$

tel que pour $z \in [s,t]$, nous pouvons rapprocher la limite originale $L(z)$ par :

$$\tilde{L}(z) = L(t)e^{\alpha(t,z)} + \beta(t,z) - \tilde{C}(t,s)\gamma(t,z). \tag{C.4.9}$$

La densité et la fonction de répartition du temps d'atteinte peuvent être données par la proposition suivante.

Proposition C.4.2. *Si pour un T constant, la limite $L(t) \in C^1[s,T]$, pour $\tilde{C}, \tilde{L}$ définies respectivement par les équations (C.4.8) et (C.4.9),*

$$\tilde{G}(t|y,s) := F(L(t),t|y,s) - e^{\tilde{C}(t,s)(y-\tilde{L}(s))}F(L(t),t|2\tilde{L}(s)-y,s) \tag{C.4.10}$$

Nous avons alors :

$$G(t|y,s) = \tilde{G}(t|y,s)(1+o(\gamma(t,s))),\ t \to s. \tag{C.4.11}$$

Proposition C.4.3. *Sous certaines conditions, pour $\tilde{C}, \tilde{L}$ définis respectivement par les équations (C.4.8) et (C.4.9) :*

$$\tilde{g}(t|y,s) = \frac{c(t)e^{\alpha(t,s)}}{\gamma(t,s)}(\tilde{L}(s)-y)p(L(t),t|y,s). \tag{C.4.12}$$

alors :

$$g(t|y,s) = \tilde{g}(t|y,s)(1+o(1)), t \to s. \tag{C.4.13}$$

Approximation paramétrique

$w(x,t|y,s)$ peut être généralement explicité par la méthode des images pour un mesure $U(r)dr$ donnée :

$$w(x,t|y,s) = p(x,t|y,s) - \int_{L(s)}^{+\infty} p(x,t|r,s)U(r)dr. \tag{C.4.14}$$

D'après l'idée de [95], supposons que des points de maillage soient donnés $\{r_i\}_{i=1}^N$ sur l'intervalle $(0,+\infty)$ puis, qu'une approximation de $U(r)$ soit introduite pour chaque point de sorte que :

$$U(r)dr = \sum_{i=1}^{N} \xi_i \delta_{L(s)+r_i}(dr), \tag{C.4.15}$$

où $\delta_y(*)$ est la mesure de Dirac centrée sur y, et où les $\{\xi_i\}_{i=1}^N$ sont non négatifs et à déterminer.

Nous obtenons l'approximation suivante pour $w(x,t|y,s)$:

$$w(x,t|y,s) \approx p(x,t|y,s) - \sum_{i=1}^{N} \xi_i p(x,t|L(s)+r_i,s). \tag{C.4.16}$$

Cette approximation permet d'obtenir l'expression analytique suivante de la densité du FPT :

$$g(t|y,s) \approx \frac{c(t)e^{2\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}}\{\int_{-\infty}^{L(s)} \frac{\beta(t,s)+L(t)e^{\alpha(t,s)}-r}{2\gamma(t,s)} \exp(-\frac{(\beta(t,s)+L(t)e^{\alpha(t,s)}-r)^2}{4\gamma(t,s)}) f_y(r)dr$$
$$-\sum_{i=1}^{N} \xi_i \frac{\beta(t,s)+L(t)e^{\alpha(t,s)}-L(s)-r_i}{2\gamma(t,s)} \exp(-\frac{(L(t)e^{\alpha(t,s)}+\beta(t,s)-L(s)-r_i)^2}{4\gamma(t,s)})\} \tag{C.4.17}$$

La fonction de répartition $G(t|y,s)$ qui en découle, est donnée par :

$$1-G(t|y,s) \approx F(L(t),t|y,s) - \sum_{i=1}^{N} \xi_i F(L(t),t|L(s)+r_i,s)). \tag{C.4.18}$$

Quand un maillage $\{t_j\}_{j=1}^M$ est considéré pour $(0,T]$, $M \geq N$, la tâche restante est de déterminer les ξ_i afin que le terme d'erreur pour l'approximation de ces points de maillage puisse être minimisé.

Comme indiqué précédemment, le terme d'erreur $\epsilon(t)$ peut être approximé par la relation suivante :

$$\epsilon(t) = h(t) - \sum_{i=1}^{N} \xi_i K(r_i,t) \tag{C.4.19}$$

Lorsque la valeur initiale y à l'instant s est déterministe, nous avons alors :

$$h(t) = p(L(t),t|y,s), K(r,t) = p(L(t),t|L(s)+r,s). \tag{C.4.20}$$

Comme indiqué dans [95], un problème de programmation linéaire peut être posé pour minimiser le terme d'erreur $\epsilon(t)$. Soit le vecteur P de dimension M de poids positifs $W = (\xi_i)_{M\times 1}$, le vecteur de dimension N noté $E = (\epsilon(t_i))_{N\times 1}$, le vecteur de dimension N noté $H = (h(t_i))_{N\times 1}$, la matrice de dimensions $M \times N$ notée $\tilde{K} = (K(r_i,t_j))_{M\times N}$.

L'estimation des ξ_i est alors donnée par le problème de programmation linéaire suivant :

$$\begin{aligned} &\text{Minimize} \quad Z = P^T E, \\ &\text{subject to :} \quad E + W^T \tilde{K} = H, \quad E \geq 0, \quad W \geq 0. \end{aligned} \tag{C.4.21}$$

Un cas particulier est de considérer une approximation à 2 paramètres dans l'équation (C.4.15) conduisant à la limite de quasi-Daniels [15]. Ceci, fait l'objet de la proposition suivante, où la limite $L(t)$ sachant $L(s)$ est introduite plutôt qu'une fonction explicite.

Proposition C.4.4. *Pour le processus X_t donné équation (C.3.1) avec pour valeur initiale (y, s), lorsque la limite $L(t)$ est donnée par la relation suivante :*

$$L(t) = e^{-\alpha(t,s)}\{L(s) - \beta(t,s) + \frac{\gamma(t,s)}{y - L(s)} \log(\frac{\xi_1 + \sqrt{\xi_1^2 + 4\xi_2 \exp(-\frac{2(y-L(s))^2}{\gamma(t,s)})}}{2})\} \quad \text{(C.4.22)}$$

où $\xi_1, \xi_2 \in \mathbb{R}$ et $L(s) > y$ sont des paramètres ajustables.

La densité de premier passage correspondante est donnée par :

$$\begin{aligned} g(t|y,s) = \frac{c(t)e^{\alpha(t,s)}}{2\gamma(t,s)}\{&(\beta(t,s) + L(t)e^{\alpha(t,s)} - y)p(L(t), t|y, s) \\ &- \sum_{i=1}^{2} \xi_i(\beta(t,s) + L(t)e^{\alpha(t,s)} - L(s) - r_i)p(L(t), t|L(s) + r_i, s)\}, \end{aligned} \quad \text{(C.4.23)}$$

où $r_1 = L(s) - y, r_2 = 3(L(s) - y)$.

Remarque : nous appelons dans la suite la limite donnée ici comme la limite de quasi-Daniels.

Simulation de Monte-Carlo quasi-linéaire par morceaux

Pour le mouvement Brownien, une limite linéaire par morceaux peut être proposée de sorte que la densité du temps d'atteinte peut être exprimée analytiquement [86, 97]. Une telle idée est reprise ici pour le processus OU.

On considère un maillage $\{t_i\}_{i=1}^N$ avec $t_0 = s, t_N = T$ et on définit $\Lambda_N = \max_{i=0,...,N-1}\{t_{i+1} - t_i\}$. Soit $L(t_i)$ la limite effective au point t_i, on peut définir l'approximation par morceau $\tilde{L}(t)$ sur l'intervalle $(t_i, t_{i+1}]$ sous la forme : $\tilde{L}(t) = e^{-\alpha(t,t_i)}(L(t_i) - \beta(t, t_i) + \eta_i\gamma(t, t_i)), t \in (t_i, t_{i+1}]$ avec $\eta_i = (L(t_{i+1})e^{\alpha(t_{i+1},t_i)} + \beta(t_{i+1}, t_i) - L(t_i))/\gamma(t_{i+1}, t_i)$.

FIGURE C.1 – Approximating the original boundary (black) by the piecewise quasi-linear boundary (red).

Lorsque Λ_N est assez petit, la limite par morceaux est, par construction, proche de la limite effective. La distribution de premier passage calculée sur la base de cette limite approximée se rapproche donc de la distribution effective de premier passage.

Ainsi, en adoptant les algorithmes numériques appropriés pour X_t (cf. équation (C.5.1)) tel que, pour ce qui suit, le schéma d'Euler-Maruyama [36], la simulation de Monte-Carlo peut être adoptée pour produire l'estimation de $G(t|y, s)$.

Supposons ainsi qu'une trajectoire soit décrite par $\{z_i\}_{i=0}^N$ où z_i est la valeur de $X(t_i)$ avec :

$$z_0 = x_0, \quad z_{i+1} = (a(t_i)z_i + b(t_i))(t_{i+1} - t_i) + \sigma(t_i)\Delta B_i, \quad i = 0, ..., N-1, \quad \text{(C.4.24)}$$

avec $\Delta B_i \sim N(0, t_{i+1} - t_i)$.

Définissons par ailleurs la fonction $\Upsilon_k(t_0, t_1, ..., t_k, z_0, z_1, ..., z_k)$, $k \geq 1$ par

$$\Upsilon_k(t_0, t_1, ..., t_k, z_0, z_1, ..., z_k) = \Pi_{i=0}^{k-1} I(z_{i+1} < L(t_{i+1}))(1 - \exp(-\frac{(z_i - L(t_i))(z_{i+1} - L(t_{i+1}))e^{\alpha(t_{i+1},t_i)}}{\gamma(t_{i+1}, t_i)})), \quad \text{(C.4.25)}$$

où $I(*)$ est la fonction indicatrice.

Ainsi pour un maillage $\{t_i\}_{i=1}^N$ donné et à condition que suffisamment de trajectoires (soit ici J trajectoires) soient simulées, une expression approchée de $G(t|y,s)$ est donnée par :

1. $G(t_0|y,s) = 0$, avec $z_0 = y, t_0 = s$.
2. Pour $k \geq 1$,

$$G(t_k|y,s) \approx 1 - \frac{1}{J}\sum_{j=1}^{J} \Upsilon_k(s, t_1, ..., t_k, y, z_1^j, ..., , z_k^j). \tag{C.4.26}$$

où z_i^j représente la valeur de X_{t_i} pour la j-ième trajectoire.

Equation intégrale non-singulière Volterra

Sur la base des résultats présentés dans [10] nous pouvons dériver une équation intégrale de Volterra non singulière pour déterminer numériquement la densité FPT.

En fait $g(t|x_0,t_0)$ satisfait une équation intégrale de Volterra de second type comme suit :

Proposition C.4.5. *Soit $L(t) \in C[t_0,+\infty), k(t) \in C[t_0,+\infty), f_0(x)$ défini sur $(-\infty, L(t_0))$, pour tous $y \in R$ et $t_0 \leq s < t$,*

$$K(L(t),t|y,s) = \frac{\partial F(L(t),t|y,s)}{\partial t} + k(t)p(L(t),t|y,s), \tag{C.4.27}$$

si $x_0 < L(t_0), g(t|x_0,t_0)$ satisfait :

$$g(t|x_0,t_0) = -2K(L(t),t|x_0,t_0) + 2\int_{t_0}^{t} g(s|x_0,t_0)K(L(t),t|L(s),s)ds. \tag{C.4.28}$$

on peut montrer que :

$$K(L(t),t|y,s) = p(L(t),t|y,s)H(t,s,y), \tag{C.4.29}$$

où

$$H(t,s,y) = -a(t)L(t) - b(t) - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)} + \beta(t,s) - y}{2(\gamma(t,s))} + k(t). \tag{C.4.30}$$

Proposition C.4.6. *Si $K(L(t),t|L(s),s)$ est défini comme précédemment, si $L(t) \in C^1[t_0,+\infty)$ et si $\lim_{s\to t} K(L(t),t|L(s),s) = 0$ alors :*

$$k(t) = \frac{a(t)L(t) + b(t) + L'(t)}{2} \tag{C.4.31}$$

De plus sous les conditions précédentes, nous avons :

$$K(L(t),t|y,s) = \left[\frac{L'(t) - a(t)L(t) - b(t)}{2} - c(t)e^{\alpha(t,s)}\frac{L(t)e^{\alpha(t,s)} + \beta(t,s) - y}{2(\gamma(t,s))}\right] p(L(t),t|y,s). \tag{C.4.32}$$

C.5 Estimation du niveau de défaillance via un problème inverse

Plusieurs résultats de ce chapitre sont en cours de publication [21] ou on été publiés [19, 22].

Dans la littérature fiabiliste relative aux phénomènes de dégradation, le niveau de défaillance est généralement déterminé d'après des modèles physiques ou des avis d'experts, en fixant un niveau critique inacceptable sur l'indicateur et le modèle de dégradation retenus. Cette démarche est justifiée pour les systèmes sur lesquels aucune panne n'est réellement observée. Nous nous intéressons ici au lien qui peut être fait entre ce niveau critique et l'observation effective de dates de défaillance lorsqu'il s'avère qu'elles sont observables et disponibles. En d'autres termes, il s'agit de poser le problème inverse suivant : sachant que l'on s'est donné un modèle de dégradation sur la base d'observations de dégradations et que l'on a part ailleurs observé des dates de pannes, quel est le niveau critique de dégradation qu'il faut fixer pour obtenir une distribution de panne qui correspond effectivement aux dates de pannes collectées ?

Problème Inverse et temps d'atteinte

En supposant que le phénomène de dégradation est décrit par un processus stochastique $X_t, t \geq 0$ défini sur un espace de probabilité complet $(\Omega, \mathfrak{F}, \mathbf{P})$, le processus d'Ornstein-Uhlenbeck dépendant du temps noté X_t sera considéré dans ce chapitre [20, 92, 2] :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0, \tag{C.5.1}$$

où $a(t), b(t), \sigma(t) \in C^1(R^+)$, $\sigma(t)$ est positif, B_t est un mouvement brownien standard et $X(0) = x_0$ est une constante.

Remarques : Dans ce chapitre le modèle est étudié sans incertitudes initiales.

On peut déduire de l'équation (C.5.1) pour l'observation initiale $(x_0, 0)$ que :

$$X_t = X_t^{x_0,0} = e^{-\alpha(t,0)}\left[x_0 - \beta(t,0) + \int_0^t \sigma(u)e^{\alpha(u,0)}dB_u\right], \quad t \geq 0, \tag{C.5.2}$$

où α, β, γ sont définies dans (C.3.2).

La densité de probabilité $p(x, t|y, s), t > s$ de X_t peut être explicitée par résolution directe lorsque y est déterministe :

$$p(x,t|y,s) = \frac{e^{\alpha(t,s)}}{\sqrt{4\pi\gamma(t,s)}} \exp\left(-\frac{(xe^{\alpha(t,s)} + \beta(t,s) - y)^2}{4\gamma(t,s)}\right). \tag{C.5.3}$$

En outre, l'instant de premier passage $\tau_{y,s}$ de X_t connaissant le niveau de dégradation à l'instant s (noté $X_s = (y, s)$) et connaissant le niveau de défaillance à t (noté $L(t) \in C^1[s, +\infty)$) est défini comme :

$$\tau_{y,s} := \inf_{t \geq s}\{t|X_t \geq L(t)\}. \tag{C.5.4}$$

La date de défaillance étant définie comme l'instant de premier passage du seuil $L(t)$ pour le processus X_t, le problème inverse que nous posons pour déterminer $L(t)$ est le suivant (on

notera dans la suite IFPT pour First Inverse Passage Time). La densité de probabilité de $\tau_{x_0,0}$ notée $p_{\tau_{x_0,0}}(t)$ est définie en fonction de $L(t)$ d'une part, et la loi de la durée de vie $g(t)$ est estimée par ailleurs à partir des observations de pannes. Alors la problème de IFPT est de trouver le niveau de défaillance $L(t)$ tel que :

$$p_{\tau_{x_0,0}}(t) = g(t). \tag{C.5.5}$$

Ou de manière équivalente, étant donnée $G(t)$ la fonction de répartition de la durée de vie, la problème IFPT est de trouver $L(t)$ telle que

$$P(\tau_{x_0,0} \leq t) = G(t). \tag{C.5.6}$$

Dans [13, 12], l'existence et l'unicité de la solution à ce problème IFPT pour un processus de diffusion général sont étudiés. Une analyse connexe pour le mouvement brownien peut être trouvée dans [43]. Toutefois, il n'existe pas de solution analytique, sauf pour quelques cas particuliers. Nous nous concentrons donc dans ce chapitre sur la résolution numériquement du problème IFPT, et plusieurs méthodes sont proposées.

C.5.1 Estimation de la limite initiale

Il est intéressant de noter que l'on dispose d'une approximation explicite pour la valeur de $L(t)$ au voisinage de zéro (on parlera alors de Limite Initiale), en tant que solution du problème inverse. L'idée est d'utiliser cette approximation comme valeur intiale d'un processus itératif qui permet de construire $L(t)$.

L'estimation de $L(t)$ au voisinnage de 0 est donnée par :

$$L(t) \approx e^{-\alpha(t,0)}(x_0 - \beta(t,0) + \sqrt{-4\gamma(t,0)\log G(t)}). \tag{C.5.7}$$

Toutefois, en raison des limites de calcul et du fait que $G(0) = 0$, lorsque t est trop faible, l'estimation n'est pas définie. L'implémentation de cette méthode nécessite donc quelques précautions.

C.5.2 Equation Intégrale

La densité de l'instant de premier passage p_τ satisfait l'équation de Fortel correspondant au niveau de défaillance $l(t)$ [57] :

$$p(x,t|x_0,0) = \int_0^t p_\tau(s)p(x,t|l(s),s)ds, \quad \forall x \geq l(t), \tag{C.5.8}$$

où $p(x,t|y,s)$ est la densité de probabilité définie dans (C.5.3). Comme indiqué précédemment, la problème de IFPT est un problème bien posé de sorte que si la loi de durée de vie $g(t)$ est donnée, la solution $L(t)$ vérifie l'équation (C.5.8) :

$$p(x,t|x_0,0) = \int_0^t g(s)p(x,t|L(s),s)ds, \quad \forall x \geq L(t). \tag{C.5.9}$$

Il reste donc à résoudre une telle équation pour trouver le niveau de défaillance $L(t)$.

Une procédure itérative est adoptée pour résoudre l'équation (C.5.9). Supposons $L(s)$ connu pour tous $s \in [0,t)$. Notons dans cette section

$$\Gamma(x,t) := p(x,t|x_0,0), \quad \Psi(x,t) := \int_0^t g(s)p(x,t|L(s),s)ds, \tag{C.5.10}$$

qui sont connus au temps t. A partir de (C.5.9), la limite actuelle $L(t)$ sera alors la solution de l'équation suivante :

$$Z(x,t) := \Gamma(x,t) - \Psi(x,t) = 0. \tag{C.5.11}$$

C'est à dire

$$L(t) = \inf_{x \in R}\{x | Z(x,t) = 0\}. \tag{C.5.12}$$

Ici, la problème IFPT original a été converti en la résolution de l'équation (C.5.12).

C.5.3 Equation Intégrale - Cont'd

Pour X_t défini par (C.5.1) avec des conditions initiales $(x_0, 0)$, on a :

$$1 - F(x,t|x_0,0) = \int_0^t (1 - F(x,t|L(s),s))g(s)ds, \quad \forall x \geq L(t). \tag{C.5.13}$$

Si on note $\Phi(*)$ la distribution de la loi normale,

$$F(x,t|y,s) = \Phi(\frac{xe^{\alpha(t,s)} + \beta(t,s) - y}{2\sqrt{\gamma(t,s)}}). \tag{C.5.14}$$

L'équation (C.5.13) peut alors s'exprimer sous une forme calculable basée sur (C.5.14) :

$$1 - \Phi(\frac{xe^{\alpha(t,0)} + \beta(t,0) - x_0}{2\sqrt{\gamma(t,0)}}) = \int_0^t (1 - \Phi(\frac{xe^{\alpha(t,z)} + \beta(t,z) - L(z)}{2\sqrt{\gamma(t,z)}}))g(z)dz, \quad \forall x \geq L(t). \tag{C.5.15}$$

Après une procédure itérative, on peut résoudre l'équation (C.5.15). Supposons que $L(s)$ est connue pour tous $s \in [0,t)$, et notons :

$$\tilde{\Omega}(x,t) := 1 - \Phi(\frac{xe^{\alpha(t,0)} + \beta(t,0) - x_0}{2\sqrt{\gamma(t,0)}}), \quad \tilde{\Psi}(x,t) := \int_0^t (1 - \Phi(\frac{xe^{\alpha(t,z)} + \beta(t,z) - L(z)}{2\sqrt{\gamma(t,z)}}))g(z)dz, \tag{C.5.16}$$

qui sont connus au temps t.A partir de l'équation (C.5.15), la limite actuelle $L(t)$ sera la solution de l'équation suivante :

$$\tilde{Z}(x,t) := \tilde{\Omega}(x,t) - \tilde{\Psi}(x,t) = 0. \tag{C.5.17}$$

C'est à dire,

$$L(t) = \inf_{x \in R}\{\tilde{Z}(x,t) = 0\}. \tag{C.5.18}$$

Ici, la problème IFPT original consiste à trouver une solution à l'équation (C.5.18).

C.6 Optimisation de la maintenance des systèmes surveillés en continu

Dans ce chapitre, une politique de maintenance conditionnelle basée sur le processus OU sera étudiée pour un système surveillé en continu. C'est l'application directe du travail de modélisation précédent concernant la dégradation, la description et le pronostic des défaillances. Ce châpitre a pour objectif de montrer comment les résultats précédents relatifs à la modélisation de la dégradation et au calcul de la durée de vie résiduelle peuvent être utilisés pour la modélisation et l'optimisation de politiques de maintenance. On considère que le système est surveillé en continu (X_t est donc connu à chaque instant) et on s'intéresse à la mise en oeuvre d'une politique de maintenance conditionnelle.

C.6.1 Introduction

Background

La maintenance conditionnelle a pour vocation d'utiliser la connaissance que l'on a de l'état du système pour entreprendre des actions de maintenance préventive. Son utilisation est actuellement en plein essort pour les raisons principales suivantes :

1. de plus en plus de capteurs ou plus généralement de techniques de surveillance sont disponibles, bon marché et implementable. Il est donc raisonnable de penser que de plus en plus d'informations seront disponibles sur l'état de santé du système.
2. les incertitudes liées à l'état du système sont réelles et l'on ne peut se satisfaire de modèles simples ne les prenant pas en compte. Ces incertitudes peuvent induire une forte variance dans la prédiction des états futurs et une maintenance calendaire ou basée sur l'âge, qui sont optimisées sur la prévision moyenne des états du système peuvent difficilement être efficaces dans ce contexte.
3. il faut remarquer enfin que l'état du système à long terme peut difficilement être prédit avec une grande précision, indépendamment des techniques de prévision. Cela renforce l'intérêt d'un contrôle plus fréquent et idéalement, d'un contrôle quasi continu.

Rappelons qu'en général, deux types d'actions de maintenance sont étudiées : la maintenance préventive et la maintenance corrective. En conséquence deux types de temps d'atteinte sont considérés ici : l'un qui corrrespond à la véritable défaillance du système et l'autre qui correspond à une défaillance virtuelle [2]. Quand un véritable échec se produit, une maintenance corrective est effectuée. Quand une panne virtuelle se produit, une maintenance préventive est effectuée. L'idée clé de l'optimisation de la maintenance des systèmes surveillés en continu est alors d'organiser une maintenance préventive appropriée, et d'optimiser les fonctions objectifs proposées telles que le coût de fonctionnement, la disponibilité du système, etc... en fonction du niveau de défaillance virtuelle.

Une hypothèse importante au sujet de l'action de maintenance adoptée dans ce chapitre est que le système est remis à neuf après chaque opération. Ceci permet d'utiliser les outils relatifs à la théorie du renouvellement et simplifie les problèmes d'optimisation de la maintenance. Nous avons d'abord considéré deux cas avec une densité explicite de l'instant de premier passage : mouvement brownien avec dérive linéaire et le mouvement brownien avec dérive non linéaire. Pour le processus OU général, une expression approchée sera considérée.

Description du système

Dans ce chapitre, nous supposons que l'état du système est surveillé en permanence. Suivant les hypothèses émises dans [63, 7, 44], nous considérons les problèmes d'optimisation de maintenance en fonction du coût minimum et de la disponibilité maximale respectivement.

On suppose par ailleurs que :

1. L'état du système est surveillé en permanence.
2. Une véritable défaillance se produit instantanément lorsque l'état du système atteint un niveau de défaillance pré-établi constant L_r, et l'instant de défaillance est noté $\tau_r := \inf_{t>0}\{t|x_t \geq L_r\}$. Quand une véritable défaillance se produit, une action de maintenance est décidé et l'état du système est indisponible.

2. La défaillance virtuelle est généralement traitée comme une alarme avant la véritable défaillance

3. Une défaillance virtuelle se produit instantanément lorsque l'état du système atteint un niveau de défaillance constant $L_v \in [x_0, L_r]$, où l'instant de défaillance est noté $\tau_v := \inf_{t>0}\{t|x_t \geq L_v\}$. Le système fonctionne encore jusqu'à ce que la maintenance commence ou jusqu'à ce que la véritable défaillance se produise.
4. La maintenance préventive ou corrective sera effectuée après un délais déterministe κ , et l'opération de maintenance a une durée aléatoire λ. Lorsque la maintenance est terminée, l'état du système est comme neuf.
5. Entre τ_v et $\tau_v + \kappa$, le système se détériore et peut atteindre le niveau de véritable défaillance au cours de cette période. Si une véritable défaillance se produit ($\tau_r \leq \tau_v+\kappa$), le système est hors service jusqu'à la fin de l'opération de maintenance $\tau_v+\kappa+\lambda$. Si une véritable défaillance ne se produit pas ($\tau_r > \tau_r + \kappa$), le système n'est pas disponible à partir du moment $\tau_v + \kappa$ jusqu'à la fin de l'opération de maintenance $\tau_v + \kappa + \lambda$.

Dans un tel cadre, la problème d'optimisation de la maintenance consiste à trouver une valeur pour le niveau de défaillance virtuel L_v tel que la fonction objectif peut être minimisée ou maximisée (fonction objectif de type coûts [63, 34], ou de tyoe disponibilité [7, 44] etc ..)

C.6.2 Critères d'évaluation

Optimisation de la maintenance basée sur les coûts

Pour envisager l'optimisation de la maintenance basée sur les coûts, il est supposé en outre que :

1. une maintenance est effectuée avec un coût C_m.
2. l'inactivité du système a un coût par unité de temps C_u.

Sous les hypothèses ci-dessus, le processus X_t décrivant l'état du système est un processus de régénération avec des instants de régénération qui sont les périodes de maintenance. Au moment de la maintenance le processus revient à x_0, et l'évolution aléatoire du système après la maintenance ne dépend pas du passé. D'après l'étude proposée dans [63], on a le coût à long terme prévu par unité de temps qui est égal au coût moyen sur un cycle divisé par la durée moyenne d'un cycle :

$$EC_\infty = \frac{C_m + C_u(\mathbb{E}\lambda + \kappa - \mathbb{E}(\inf(\kappa, \tau_r - \tau_v)))}{\mathbb{E}(\tau_v) + \kappa + \mathbb{E}\lambda}. \tag{C.6.1}$$

Le problème est alors de trouver le seuil L_v optimal tel que :

$$\text{argmin}_{L_v} EC_\infty, \quad s.t. \quad x_0 \leq L_v \leq L_r. \tag{C.6.2}$$

Il faut pour cela trouver une expression calculable de EC_∞ en fonction du niveau de défaillance virtuel L_v. Cela passe par le calcul de $E(\inf(\sigma, \tau_r - \tau_v))$, lié à la densité jointe $p_{\tau_v,\tau_r}(x, y)$ du couple (τ_v, τ_r). En outre, nous savons que pour $t \geq 0$, la fonction de survie $\bar{H}(t)$ de $\tau_r - \tau_v$ est donnée par :

$$\bar{H}(t) = \int\int_{y-x>t} p_{\tau_v,\tau_r}(x, y)dxdy = \int_t^{+\infty}\int_0^{y-t} g_{\tau_r}(y|L_v(x), x)g_{\tau_v}(x|0, 0)dxdy, \tag{C.6.3}$$

où $g_{\tau_r}(\eta|y, s)$ est la densité conditionnelle de premier passage pour L_r sachant (y, s), et $g_{\tau_v}(x|y, s)$ est la densité conditionnelle de premièr passage pour L_v sachant (y, s).

Finalement :

$$\mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) = \int_0^{\kappa} \bar{H}(s)ds \tag{C.6.4}$$

Sur la base de tout ce qui précède, nous avons effectivement exprimé explicitement $EC_{+\infty}$ en fonction de la densité de premier passage conditionnelle pour τ_v et τ_r. Le calcul plus précis sur ces questions sera examiné dans la section suivante pour l'optimisation dans (C.6.2).

Optimisation de la maintenance basée sur la disponibilité

Dans le domaine de l'ingénierie et de la fiabilité, il peut être pertinent également d'optimiser la disponibilité du système [7]. On utilise alors un critère asymptotique d'indisponibilité. Supposons que X_t est le processus de régénération de type OU avec des instants de régénération définis par les instants de maintenance. Le taux d'indisponibilité asymptotique est donné par le temps moyen d'indisponibilité sur un cycle, divisé par la durée moyenne d'un cycle.

$$U_{\infty} = \frac{\mathbb{E}\lambda + \kappa - \mathbb{E}(\inf(\kappa, \tau_r - \tau_v))}{\mathbb{E}(\tau_v) + \kappa + \mathbb{E}\lambda}. \tag{C.6.5}$$

Alors la problème d'optimisation des équipements consiste à trouver le seuil L_v tel que :

$$\text{argmin}_{L_v} U_{\infty}, \quad s.t. \quad x_0 \leq L_v \leq L_r. \tag{C.6.6}$$

Il est à noter que, intrinsèquement, il n'y a pas de différence technique entre l'optimisation basée sur les coûts et l'optimisation basée sur la disponibilité, parce que les questions fondamentales dans (C.6.1) et (C.6.5) restent les mêmes.

Durée de la maintenance et état du système

C'est une idée naturelle de supposer que la durée de la maintenance dépend de l'état du système de sorte que dans [7], il est suggéré de considérer que

$$\mathbb{E}\lambda = \lambda_1 + \lambda_2 \mathbb{E}(X_{\tau_v+\kappa}), \tag{C.6.7}$$

où λ_1, λ_2 sont des valeurs déterministes. L'idée initiale de cette proposition est de considérer que la durée de la maintenance est plus longue lorsque la dégradation du système est plus importante. Cependant, au lieu de (C.6.7), dans ce chapitre une autre version actualisée de la durée de la maintenance est choisie car la non linéarité du processus de dégradation rend la première version difficilement calculable.

$$\mathbb{E}\lambda = \lambda_1 + \lambda_2 \mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)}). \tag{C.6.8}$$

L'expression de $\mathbb{E}(X_{\tau_v+\kappa})$ est alors donnée par l'équation suivante :

Proposition C.6.1.

$$\mathbb{E}(X_{\tau_v+\kappa} e^{\alpha(\tau_v+\kappa,0)}) = -\mathbb{E}(\beta(\tau_v + \kappa, 0)) + x_0. \tag{C.6.9}$$

C.6.3 Approximation-optimization

Dans le contexte précédent, une expression approchée de la loi du temps d'atteinte a été proposée. Sur la base de cette approximation, la problème d'optimisation de la maintenance consiste à mettre en oeuvre des heuristiques similaires à celles mises en ouvre pour le mouvement brownien dans [44] et pour le processus Gamma dans [7].

Rappelons que le processus X_t donnée par(C.5.1) a la frome suivante :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, X_0 = x_0, t \geq 0. \tag{C.6.10}$$

Une expression approchée de la densité de premier passage à la limite $L(t)$ est donnée par (2.3.113) :

$$g(t|y,s) \approx \left[a(t)\mathbb{E}(X_t) + b(t) - L'(t) + (L(t) - \mathbb{E}(X_t))(\frac{\sigma^2(t)}{\text{var}(X_t)} + a(t))\right] \times p(L(t), t|y, s), \tag{C.6.11}$$

où

$$\begin{aligned} \mathbb{E}(X_t) &= e^{-\alpha(t,s)}\bigg(y - \beta(t,s)\bigg), \\ \text{var}\,(X_t) &= e^{-2\alpha(t,s)}\int_s^t \sigma^2(u)e^{2\alpha(u,s)}du. \end{aligned} \tag{C.6.12}$$

Pour un niveau de dégradation constant L, on a :

$$g(t|y,s) \approx \left[a(t)\mathbb{E}(X_t) + b(t) + (L - \mathbb{E}(X_t))(\frac{\sigma^2(t)}{\text{var}(X_t)} + a(t))\right] p(L, t|y, s). \tag{C.6.13}$$

L'expression (C.6.13) mène directement à l'expression de la densité de premier passage $g_{\tau_r}(t|y,s)$ et $g_{\tau_v}(t|y,s)$ en considérant L_r et L_r respectivement.

Temps moyen de premier passage

Nous avons maintenant une expression explicite pour la loi du temps d'atteinte (C.6.13), qui mène directement au temps moyen de premier passage et à la durée moyenne de la maintenance. Pour le processus OU, il s'agit de

$$\mathbb{E}(\tau_v) = \int_0^{+\infty} z g_{\tau_v}(z|x_0, 0))dz. \tag{C.6.14}$$

et :

$$\mathbb{E}(X_{\tau_v+\kappa}e^{\alpha(\tau_v+\kappa,0)}) = -\mathbb{E}(\beta(\tau_v + \kappa, 0)) + x_0 = x_0 - \int_0^{+\infty} g_{\tau_v}(z|x_0, 0))\beta(z + k, 0)dz. \tag{C.6.15}$$

Pour une forme calculable de ces deux quantités, nous devrions envisager une forme tronquée de l'intégrale. Mais la difficulté existe pour la valeur du niveau virtuel de défaillance L_v qui n'est pas connu avant l'optimisation. Juger si l'intégrale tronquée sera assez bonne n'est pas évident. Mais en utilisant le fait que $\tau_r \geq \tau_v$, nous pouvons proposer une approximation contrôlable pour tronquer l'intervalle infini basé sur τ_r.

Calcul sur $\mathbb{E}(\inf \kappa, \tau_r - \tau_v)$

A partir de (C.6.4), nous savons que :

$$\mathbb{E}(\inf(\kappa, \tau_r - \tau_v)) = \int_0^{\kappa} \bar{H}(s)ds \tag{C.6.16}$$

où

$$\bar{H}(t) = \int_t^{+\infty} \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|x_0, 0)dxdy. \tag{C.6.17}$$

Il s'ensuit deux mesures pour donner une approximation de $\bar{H}(t)$:

1. Tronquer le domaine d'intégration avec un temps de fin $T > 0$, de sorte que

$$\bar{H}(t) \approx \int_t^{T} \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|x_0, 0)dxdy. \tag{C.6.18}$$

2. calculer numériquement $\int_t^T \int_0^{y-t} g_{\tau_r}(y|L_v, x) g_{\tau_v}(x|x_0, 0)dxdy$.

En suivant la procédure ci-dessus, nous pouvons parvenir à une expression approchée suffisamment précise pour $\int_0^{\kappa} \bar{H}(s)ds$, qui est une fonction explicite du temps d'atteinte de L_v. La précision de l'approximation peut être assurée quelle que soit L_v.

C.7 Conclusions et perspectives

Conclusions

Dans cette thèse, nous avons discuté de quatre grands thèmes, visant à décrire, prédire et prévenir les défaillances du système :

1. la modélisation de la dégradation stochastique basée sur un processus d'Ornstein-Uhlenbeck (OU) dépendant du temps ;
2. le pronostic des défaillances d'un système via un problème de temps d'atteinte ;
3. l'estimation du niveau de défaillance via un problème inverse de temps d'atteinte ;
4. l'optimisation de la maintenance d'un système surveilé en continu.

Dans la partie de *Modélisation de dégradation basée sur un processus OU en fonction du temps*, nous avons introduit le processus OU dépendant du temps avec l'équation différentielle stochastique suivante :

$$dX_t = (a(t)X_t + b(t))dt + \sigma(t)dB_t, \quad t \geq 0. \tag{C.7.1}$$

Nous avons montré que ce processus est un bon choix pour la modélisation de la dégradation pour les raisons suivantes :

1. le processus OU permet d'avoir suffisament de degrès de liberté pour adapter de manière indépendante la moyenne, la variance et la covariance aux données.
2. Le processus OU converge vers sa valeur moyenne, ce qui permet de localiser dans le temps l'effet des fluctuations introduites pas le bruit blanc. Cela peut s'avérer positif et en parfait accord avec les avis d'experts et les lois physiques associées au phénomème de dégradation.

3. Le processus OU a une densité de probabilité explicite, ce qui rend aisé l'estimation des paramètres à partir des données de dégradation..
4. Une étude de cas réel est proposée concernant des composants passifs dans les centrales électriques.

Dans la partie de *Pronostic de défaillances et temps d'atteinte*, on présente l'estimation de la durée de vie conditionnelle d'un système à partir du modèle de dégradation proposé dans le châpitre précédent. Nous avons développé différentes techniques à partir de deux points de vue différents : équations aux dérivées partielles et équations intégrales. Ces techniques peuvent être généralement classées en trois catégories : approximation analytique, algorithmes numériques et les méthodes de Monte-Carlo.

Dans la partie de *Estimation du niveau de défaillance via un problème inverse*, on s'interroge sur l'adéquation entre le niveau de défaillance choisi pour le modèle de dégradation et l'observation de dates de défaillances. Plus précisément, les questions suivantes ont été traitées :

1. estimation de la limite initiale (C.5.7) pour le processus OU est résultat préliminaire pour le mouvement brownien.
2. problème inverse de premier passage résolu à partir de l'équation de Fortet et l'équation Master respectivement par l'intermédiaire d'une procédure itérative.
3. tests en simulation pour reproduire le niveau de défaillance et la distribution du temps d'atteinte..

Dans la partie de *Optimisation de la maintenance pour une système surveillé en continu*, on propose une règle de décision pour effectuer la maintenance préventive avant la maintenance corrective de telle sorte que la fonction objectif peut être optimisée. L'objectif est de trouver trouver un niveau optimal de défaillance virtuelle. Plus précisément, les questions suivantes ont été étudiées :

1. le problème d'optimisation de la maintenance pour le processus OU a été résolu à l'aide d'une approximation, en minimisant un critère de taux asymptotique de coût et un critère de taux asymptotique d'indisponibilité.
2. le problème d'optimisation de la maintenance est étudié respectivement pour le mouvement brownien avec dérive, pour le mouvement brownien avec dérive non linéaire et pour le processus OU.

Perspectives

Les perspectives de cette thèse peuvent être considérées égaelement selon 4 grands axes : modélisation de la dégradation, pronostic de défaillance, estimation de niveau de défaillance et optimisation de la maintenance.

La convergence vers la moyenne du processus OU offre la possibilité d'étudier les phénomènes en respectant l'idée suivante : il existe un processus de dégradation réel traité comme un processus sous-jacent et qui peut être déterminé par des lois physiques. Cette propriété de "retour à la moyenne" pourrait être un point de départ pour introduire des processus OU de type Levy dans la modélisation de la dégradation stochastique.

Pour la pronostic de défaillance, la problème de temps d'atteinte est introduit pour le processus OU. Cela induit un exercice multidisciplinaire entre analyse stochastique, PDE, équation intégrale, analyse numérique, etc .. L'analyse dans cette thèse commence par $w(x,t|y,s) := \frac{\partial P(\tau_y,s>t,X_t<x|X_s=y)}{\partial x}$, cependant la possibilité de suivre une autre analyse existe.

Et la méthode des images conduit à de nombreux résultats également pertinents qu'il pourrait être intéressant d'explorer davantage. [43].

Concernnant le niveau de défaillance et l'étude d'un problème inverse, il faut remarquer que le processus OU est un processus simple, de sorte qu'il est sans doute difficile d'étendre les résultats obtenus à d'autres processus de diffusion. Toutefois, il semble pertinent de s'intéresser à des processus tels que les processus de saut ou les processus saut-diffusion.

Enfin, concernant la maintenance, les approximations analytiques de la densité du temps d'atteinte dans les problèmes d'optimisation de la maintenance doivent permettre d'envisager des politiaues plus élaborées. Une extension naturelle est de considérer que l système n'est plus surveillé en continu et de s'intéresser à l'optimisation de l'intervalle entre deux inspections. [34].

Bibliographie

[1] Aalen, O. O., and Gjessing, H. K. Understanding the shape of the hazard rate : a process point of view (with comments and a rejoinder by the authors). *Statistical Science 16*, 1 (2001), 1–22.

[2] Aalen, O. O., and Gjessing, H. K. Survival models based on the ornstein-uhlenbeck process. *Lifetime Data Analysis 10*, 4 (2004), 407–423.

[3] Akaike, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on 19*, 6 (1974), 716–723.

[4] Anonymous. Internal technical report. Tech. rep., Électricité de France, 2011.

[5] Avellaneda, M., and Jingyi, Z. Distance to default. *Risk 14*, 2 (2001), pp. 125–129.

[6] Barker, C. T., and Newby, M. J. Optimal non-periodic inspection for a multivariate degradation model. *Reliability Engineering & System Safety 94* (2009), 33–43.

[7] Bérenguer, C., Grall, A., and Dieulle, L. Maintenance policy for a continuously monitored deteriorating system. *Probability in the Engineering and Information Sciences 4* (2003), 235–250.

[8] Blain, C. *Modélisation des dégradations de composants passifs par processus stochastiques et prise en compte des incertitudes*. PhD thesis, Université de Technologie de Troyes, 2008.

[9] Bo, L., Wang, Y., and Yang, X. First passage times of (reflected) ornstein-uhlenbeck processes over random jump boundaries. *Journal of Applied Probability 48*, 3 (09 2011), 723–732.

[10] Buonocore, A., Nobile, A. G., and Ricciardi, L. M. A New Integral Equation for the Evaluation of First-Passage-Time Probability Densities. *Advances in applied probability 19*, 4 (1987), 784–800.

[11] C.F.Lo, and C.H.Hui. Computing the first passage time density of a time-dependent ornstein-uhlenbeck process to a moving boundary. *Applied Mathematics Letters 19* (2006), 1399–1405.

[12] Chen, X., Cheng, L., Chadam, J., and Saunders, D. Existence and uniqueness of solutions to the inverse boundary crossing problem for diffusions. *Annals of Applied Probability 21*, 5 (2011), 1663–1693.

[13] Cheng, L., Chen, X., Chadam, J., and Saunders, D. Analysis of an Inverse First Passage Problem from Risk Management. *SIAM J. Math. Anal. 38*, 3 (2006), 845–873.

[14] Cox, D. R. Regression models and life-tables. *Journal of the Royal Statistical Society : Series B (Statistical Methodology) 34*, 2 (1972), 187–220.

[15] Daniels, H. E. Approximating the first crossing-time density for a curved boundary. *Bernoulli 2*, 2 (06 1996), 133–143.

[16] Dekker, R. Applications of maintenance optimization models : a review and analysis. *Reliability Engineering & System Safety 51*, 3 (Mar. 1996), 229–240.

[17] Deng, Y., Barros, A., and Grall, A. The influence of initial uncertainties in stochastic degradation modeling. In *Proceedings of the European Safety and Reliability conference, ESREL 2013* (Amsterdam, Netherlands, 2013).

[18] Deng, Y., Barros, A., and Grall, A. Residual useful life estimation based on a time-dependent ornstein-uhlenbeck process. *Chemical Engineering Transaction 33* (2013), 325–333.

[19] Deng, Y., Barros, A., and Grall, A. Calculation of failure level based on inverse first passage problem. In *Proceedings of Annual Reliability and Maintainability Symposium, IEEE* (Colorado Springs, USA, 2014).

[20] Deng, Y., Barros, A., and Grall, A. Degradation modeling based on a time-dependent ornstein-uhlenbeck process and residual useful lifetime estimation. In revision, 2014.

[21] Deng, Y., Barros, A., and Grall, A. Failure level estimation based on inverse first passage problem. Manuscript, 2014.

[22] Deng, Y., Barros, A., and Grall, A. On newton's method for an inverse first passage problem. In *Proceedings of the Sixth SIAM Workshop on Combinatorial Scientific Computing, SIAM* (Lyon, France, 2014).

[23] Di Nardo, E., Nobile, A., Pirozzi, E., and Ricciardi, L. A computational approach to first-passage-time problems for Gauss-Markov processes. *Advances in Applied Probability 33*, 2 (2001), 453–482.

[24] Downes, A. N. *Boundary Crossing Probabilities for Diffusion Processes and Related Problems*. PhD thesis, University of Melbourne, 2008.

[25] Durbin, J. The first-passage density of a continuous Gaussian process to a general boundary. *Journal of applied probability 22*, 1 (1985), 99–122.

[26] Durbin, J. A reconciliation of two different expressions for the first-passage density of brownian motion to a curved boundary. *Journal of Applied Probability 25*, 4 (1988), pp. 829–832.

[27] Durbin, J., and Williams, D. The first-passage density of the Brownian motion process to a curved boundary. *Journal of Applied Probability 29*, 2 (1992), 291–304.

[28] Einstein, A. On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. *Annalen der Physik*, 17 (1905), 549–560.

[29] Ettinger, B., Evans, S. N., and Hening, A. Killed Brownian motion with a prescribed lifetime distribution and models of default. *The Annals of Applied Probability 24*, 1 (2014), 1–33.

[30] Evans, L. C. An Introduction to Stochastic Differential Equations. Lecture notes at Department of Mathematics, University of California at Berkeley.

[31] Ferebee, B. The tangent approximation to one-sided brownian exit densities. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 61*, 3 (1982), 309–326.

[32] Fortet, R. Les fonctions aleatoires du type de markoff associees a certaines equations lineaires aux derivees partielles du type parabolique. *J. Math. Pures Appl. 22* (1943), 177–243.

[33] Gardiner, C. W. *Handbook of Stochastic Methods*. Springer, Berlin, 2004.

[34] Grall, A., Bérenguer, C., Dieulle, L., and Roussignol, M. Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE Transactions on Reliability 51*, 2 (2002), 141–150.

[35] Hanson, R. A numerical method for solving fredholm integral equations of the first kind using singular values. *SIAM Journal on Numerical Analysis 8*, 3 (1971), 616–622.

[36] Higham, D. J. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review 43*, 3 (2001), 525–546.

[37] Hottovy, S. *The Smoluchowski-Kramers approximation for stochastic differential equations with arbitrary state dependent friction.* Ph.d. thesis, University of Arizona, 2013.

[38] Jackson, K., Kreinin, A., and Zhang, W. Randomization in the first hitting time problem. *Statistics & Probability Letters 79*, 23 (2009), 2422 – 2428.

[39] Jacobs, J., Etman, L., van Keulen, F., and Rooda, J. Framework for sequential approximate optimization. *Structural and Multidisciplinary Optimization 27*, 5 (June 2004), 384–400.

[40] Kalbfleisch, J. D., and Prentice, R. L. *The Statistical Analysis of Failure Time Data (Second Edition).* John Wiley & Sons, Inc., Hoboken, New Jersey, 2002.

[41] Lee, M.-L. T., and Whitmore, G. A. Threshold regression for survival analysis : Modeling event times by a stochastic process reaching a boundary. *Statistical Science 21*, 4 (2006), 501–513.

[42] Lemons, D. S., and Gythiel, A. Paul langevin's 1908 paper "on the theory of brownian motion", (sur la théorie du mouvement brownien, c. r. acad. sci. (paris) 146, 530-533 (1908)). *American Journal of Physics 65*, 11 (1997), 1079–1081.

[43] Lerche, H. H. *Boundary Crossing of Brownian Motion.* Springer Verlag, 1986.

[44] Liao, H., Elsayed, E. A., and Chan, L.-Y. Maintenance of continuously monitored degrading systems. *European Journal of Operational Research 175*, 2 (2006), 821–835.

[45] Liao, H., and Tian, Z. A framework for predicting the remaining useful life of a single unit under time-varying operating conditions. *IIE Transactions 45*, 9 (2013), 964–980.

[46] Lo, C. Exact solutions of the fokker-planck equations with moving boundaries. *Annals of Physics 319*, 2 (2005), 326 – 332.

[47] Madec, Y., and Japhet, C. First passage time problem for a drifted ornstein-uhlenbeck process. *Mathematical Biosciences 189*, 2 (2004), 131 – 140.

[48] Nicolai, R. *Maintenance models for systems subject to measurable deterioration.* PhD thesis, Erasmus University Rotterdam, 2008.

[49] Nicolai, R. P., and Dekker, R. A comparison of two non-stationary degradation processes. *Risk, Reliability and Societal Safety, proceedings of ESREL 2007* (2007).

[50] Pagès, G., and Printems, J. Optimal quantization for finance : from random vectors to stochastic processes. *Handbook of Numerical Analysis 15* (2008).

[51] Pandey, M. D., Yuan, X. X., and van Noortwijk, J. M. The influence of temporal uncertainty of deterioration on life-cycle management of structures. *Structure and Infrastructure Engineering 5*, 2 (2009), 145–156.

[52] Paris, P., and Erdogan, F. A critical analysis of crack propagation laws. *Journal of Basic Engineering 85*, 4 (1963), 528–533.

[53] Patie, P., and Winter, C. First exit time probability for multidimensional diffusions : A PDE-based approach. *Journal of Computational and Applied Mathematics 222* (2008), 42–53.

[54] Peng, C.-Y., and Tseng, S.-T. Mis-specification analysis of linear degradation models. *Reliability, IEEE Transactions on 58*, 3 (2009), 444–455.

[55] Peng, C.-Y., and Tseng, S.-T. Progressive-stress accelerated degradation test for highly-reliable products. *Reliability, IEEE Transactions on 59*, 1 (2010), 30–37.

[56] Peskir, G. Limit at zero of the brownian first-passage density. *Probab. Theory Related Fields 124*, 1 (2002), 100–111.

[57] Peskir, G. On integral equations arising in the first-passage problem for brownian motion. *Journal of Integral Equations and Applications 14*, 4 (2002), 397–423.

[58] Peskir, G., and Shiryaev, A. *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics, 2006.

[59] Profeta, C. Conditional hitting time estimation in a nonlinear filtering model by the brownian bridge method. arXiv preprint arXiv :1211.4553, 2012.

[60] Rausand, M. *System Reliability Theory : Models, Statistical Methods, and Applications, 2nd Edition*. Wiley, 2004.

[61] Revuz, D., and Yor, M. *Continuous Martingales and Brownian Motion*. Springer, 1999.

[62] Risken, H. *The Fokker-Planck equation : Methods of solution and applications*, vol. 18. Springer Verlag, 1996.

[63] Roussignol, M. Gamma stochastic process and application to maintenance. Lecture Notes in Université Paris-Est Marne-la-Vallée, 2009.

[64] Sheather, S. J., and Jones, M. C. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological) 53*, 3 (1991), 683–690.

[65] Si, X.-S., Hu, C.-H., Kong, X., and Zhou, D.-H. A residual storage life prediction approach for systems with operation state switches. *Industrial Electronics, IEEE Transactions on 61*, 11 (Nov 2014), 6304–6315.

[66] Si, X.-S., Wang, W., Hu, C.-H., Zhou, D., and Pecht, M. G. Remaining useful life estimation based on a nonlinear diffusion degradation process. *Reliability, IEEE Transactions on 61*, 1 (2012), 50–61.

[67] Si, X.-S., Wang, W., Hu, C.-H., and Zhou, D.-H. Remaining useful life estimation - a review on the statistical data driven approaches. *European Journal of Operational Research 213*, 1 (2011), 1 – 14.

[68] Si, X.-S., Wang, W., Hu, C.-H., and Zhou, D.-H. Estimating remaining useful life with three-source variability in degradation modeling. *Reliability, IEEE Transactions on 63*, 1 (March 2014), 167–190.

[69] Sigman, K., and Wolff, R. W. a review of regenerative processes. *SIAM Review 35*, 2 (1993), 269–288.

[70] Sikorska, J., Hodkiewicz, M., and Ma, L. Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing 25*, 5 (2011), 1803 – 1836.

[71] Skiadas, C., and Skiadas, C. H. Development, simulation, and application of first-exit-time densities to life table data. *Communications in Statistics - Theory and Methods 39*, 3 (2010), 444–451.

[72] Smith, W. L. Regenerative Stochastic Processes, 1955.

[73] Sobczyk, K. Modelling of random fatigue crack growth. *Engineering Fracture Mechanics 24*, 4 (1986), 609 – 623.

[74] Sobczyk, K. Stochastic models for fatigue damage of materials. *Advances in applied probability 19*, 3 (1987), 652–673.

[75] Sobczyk, K., and B.F. Spencer, J. *Random fatigue : from data to theory.* Academic Press, San Diego, 1992.

[76] Sobczyk, K., and Trębicki, J. Fatigue crack growth in random residual stresses. *International Journal of Fatigue 26*, 11 (Nov. 2004), 1179–1187.

[77] Son, K. L. *Modélisation probabilististe du pronostic - application à un cas d'étude et à la prise de désion en maintenance.* PhD thesis, Université de Technologie de Troyes, 2012.

[78] Son, K. L., Fouladirad, M., Barros, A., Levrat, E., and Iung, B. Remaining useful life estimation based on stochastic deterioration models : A comparative study. *Reliability Engineering & System Safety 112*, 0 (2013), 165 – 175.

[79] Song, J.-S., and Zipkin, P. An approximation for the inverse first passage time problem. *Advances in Applied Probability 275*, May 2010 (2011), 264–275.

[80] Steele, J. M. *Stochastic calculus and financial applications.* Springer-Verlag, New York, 2001.

[81] Ting Lee, M.-L., DeGruttola, V., and Schoenfeld, D. A model for markers and latent health status. *Journal of the Royal Statistical Society : Series B (Statistical Methodology) 62*, 4 (Nov. 2000), 747–762.

[82] Tseng, S.-T., and Peng, C.-Y. Stochastic diffusion modeling of degradation data. *Journal of Data Science 5* (2007), 315–333.

[83] Valov, A. *First Passage Times : Integral Equations, Randomization and Analytical Approximations.* Ph.d. thesis, University of Toronto, 2009.

[84] van Noortwijk, J. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety 94*, 1 (2009), 2 – 21.

[85] Vlasiou, M. Regenerative processes. *Wiley Encyclopedia of Operations Research and Management Science* (2011).

[86] Wang, L., and Pötzelberger, K. Boundary crossing probability for Brownian motion and general boundaries. *Journal of Applied Probability*, 34 (1997), 54–65.

[87] Wang, L., and Pötzelberger, K. Crossing probabilities for diffusion processes with piecewise continuous boundaries. *Methodology and Computing in Applied Probability*, 1 (2007), 1–18.

[88] Wang, X., and Xu, D. An inverse gaussian process model for degradation data. *Technometrics 52*, 2 (2010), 188–197.

[89] Whitmore, G. First-passage-time models for duration data : regression structures and competing risks. *The Statistician 35*, 2 (1986), 207–219.

[90] Whitmore, G. a., Crowder, M. J., and Lawless, J. F. Failure inference from a marker process based on a bivariate Wiener model. *Lifetime data analysis 4*, 3 (Jan. 1998), 229–51.

[91] Wolf, F. Lie Algebraic-Solutions of Linear Fokker-Planck Equations. *Journal of Mathematical Physics 29*, 2 (1988), 305–307.

[92] Yashin, A. I., and Manton, K. G. Effects of unobserved and partially observed covariate processes on system failure : A review of models and estimation strategies. *Statistical Science 12*, 1 (1997), pp. 20–34.

[93] Ye, Z.-S., and Chen, N. The inverse gaussian process as a degradation model. *Technometrics 56*, 3 (2014), 302–311.

[94] Zhou, R. R., Serban, N., and Gebraeel, N. Degradation modeling applied to residual lifetime prediction using functional data analysis. *The Annals of Applied Statistics 2B*, 5 (2011), 1586–1610.

[95] Zipkin, P. Linear programming and the inverse method of images. *Annals of Operations Research 208*, 1 (Nov. 2012), 227–243.

[96] Zipkin, P. On the method of images and the asymptotic behavior of first-passage times. *Annals of Operations Research* (2012), 1–24.

[97] Zucca, C., and Sacerdote, L. On the inverse first-passage-time problem for a Wiener process. *The Annals of Applied Probability 19*, 4 (2009), 1319–1346.

Yingjun DENG

Doctorat : Optimisation et Sûreté des Systèmes

Année 2015

Modélisation des dégradations par un processus d'Ornstein-Uhlenbeck et pronostic de défaillances du système

Cette thèse est consacrée à la description, la prédiction et la prévention des défaillances de systèmes. Elle se compose de quatre parties relatives à la modélisation stochastique de dégradation, au pronostic de défaillance du système, à l'estimation du niveau de défaillance et à l'optimisation de maintenance.

Le processus d'Ornstein-Uhlenbeck (OU) dépendant du temps est introduit dans un objectif de modélisation des dégradations. Sur la base de ce processus, le premier instant de passage d'un niveau de défaillance prédéfini est considéré comme l'instant de défaillance du système considéré. Différentes méthodes sont ensuite proposées pour réaliser le pronostic de défaillance. Dans la suite, le niveau de défaillance associé au processus de dégradation est estimé à partir de la distribution de durée de vie en résolvant un problème inverse de premier passage. Cette approche permet d'associer les enregistrements de défaillance et le suivi de dégradation pour améliorer la qualité du pronostic posé comme un problème de premier passage. Le pronostic de défaillances du système permet d'optimiser sa maintenance. Le cas d'un système contrôlé en permanence est considéré. La caractérisation de l'instant de premier passage permet une rationalisation de la prise de décision de maintenance préventive. L'aide à la décision se fait par la recherche d'un niveau virtuel de défaillance dont le calcul est optimisé en fonction de critères proposés.

Mots clés : fiabilité, méthodes statistiques - prévision technologique - durée de vie (ingénierie) - temps entre défaillances, analyse des - maintenance conditionnelle – processus gaussiens.

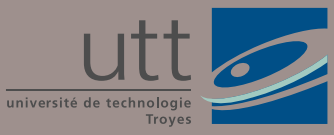
Degradation Modeling Based on a Time-dependent Ornstein-Uhlenbeck Process and Prognosis of System Failures

This thesis is dedicated to describe, predict and prevent system failures. It consists of four issues: i) stochastic degradation modeling, ii) prognosis of system failures, iii) failure level estimation and iv) maintenance optimization.

The time-dependent Ornstein-Uhlenbeck (OU) process is introduced for degradation modeling. The time-dependent OU process is interesting from its statistical properties on controllable mean, variance and correlation. Based on such a process, the first passage time is considered as the system failure time to a pre-set failure level. Different methods are then proposed for the prognosis of system failures, which can be classified into three categories: analytical approximations, numerical algorithms and Monte-Carlo simulation methods. Moreover, the failure level is estimated from the lifetime distribution by solving inverse first passage problems. This is to make up the potential gap between failure and degradation records to reinforce the prognosis process via first passage problems. From the prognosis of system failures, the maintenance optimization for a continuously monitored system is performed. By introducing first passage problems, the arrangement of preventive maintenance is simplified. The maintenance decision rule is based on a virtual failure level, which is solution of an optimization problem for proposed objective functions.

Keywords: reliability (engineering), statistical methods - failure time data analysis - service life (engineering) - technological forecasting - condition-based maintenance – Gaussian processes.

Thèse réalisée en partenariat entre :



Ecole Doctorale "Sciences et Technologies"