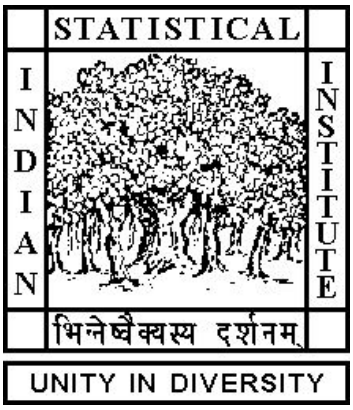
BIVARIATE CURRENT STATUS DATA WITH COMPETING RISKS USING FRAILTY MODELS

A thesis submitted for the degree of Doctor of Philosophy

By

Biswadeep Ghosh
Applied Statistics Unit

Thesis Supervisor: Prof. Anup Dewanji
Applied Statistics Unit



Indian Statistical Institute

203, B.T. Road, Kolkata-700108

June 2026

This thesis is dedicated to my supervisor, Prof. Anup Dewanji; my parents, Mr. Biswanath Ghosh and Mrs. Sandhya Ghosh; and all of my teachers.

Acknowledgements

It is my immense pleasure to submit this thesis after years of meticulous research at Indian Statistical Institute, Kolkata. This gives me an opportunity to express my gratitude to the persons associated with me without whom this task could not have been complete.

Firstly, I would like to thank my supervisor, Prof. Anup Dewanji for his constant support and invaluable guidance. His supervision and mentorship have immensely helped me shaping this thesis. He also stood by me and encouraged me even when my productivity was at its lowest.

This acknowledgement will be incomplete without mentioning my co-author, Dr. Sudipta Das. His assistance and advice also helped shaping this thesis. I can never forget about the moments of our collaboration and his friendly gestures.

I would also like to thank the faculty members and the entire staff of ASU for their continuous support on several matters. The coursework has been extremely beneficial for me. I would also like to thank the Dean and the Director of Indian Statistical Institute. I would also like to express my gratitude to the Indian Statistical Institute, for supporting me through a fellowship and providing me the infrastructure.

I can never forget the companionship and support of my friends, seniors, and juniors during my difficult time. The tea club adda at Lakshman-da was always a refreshing period in this journey. I would always cherish those moments comprising of smiles and laughter.

I will always be thankful to my dear friends Koushik Biswas and Mriganka Aulia, who were ready to help me at any time. Whenever I had some difficult coding issue, I rushed to them to seek their advice. Words will fall short to describe their continuous support to shape this thesis.

I will also be indebted to all of my teachers, elders, family members. Their teaching and guidance always helped me to overcome any obstacle throughout my career.

Finally, I am grateful to the department of speech and hearing, Ali Yavar Jung National Institute of Speech and Hearing Disabilities, Eastern Regional Centre, Kolkata for providing us the data on hearing loss. I am also grateful to my senior, Dr. Tamalika Koley, for facilitating the availability of this data to me. She also helped me to understand the nature of this data which would later be beneficial for my analysis of the same data.

List of Tables

Contents

Abstract

Bivariate current status data with competing risks arises in many application areas, for example, carcinogenicity studies, epidemiological studies, reliability studies, etc.. As par of our knowledge, bivariate current status data with competing risks has not received much attention in literature. We initiate this research focusing on modeling and estimation of quantities of interest in absence of any covariate.

We model bivariate current status data with competing risks through different frailty structures in order to describe different kinds of association. We consider multiplicative frailty effect on different cause-specific baseline hazard functions. One important part of our analysis is to investigate the identifiability of a proposed model. It has been shown that, when both the cause-specific baseline hazard function and frailty distribution are non-parametric, then the model is not identifiable.

Motivated by the previous result, we consider the analysis of following three different combinations of cause-specific baseline hazard and frailty distributions, namely, (a) both the cause-specific baseline hazard and frailty distribution are parametric, (b) the cause-specific baseline hazard functions are parametric and the frailty distribution is non-parametric, (c) the cause-specific baseline hazard functions are non-parametric and the frailty distribution is parametric. For the parametric frailty distribution, we have considered Gamma distribution of different kinds.

We have used the maximum likelihood estimation method to estimate the parameters when both the cause-specific baseline hazard functions and frailty distribution are assumed to be parametric. In case of parametric cause-specific baseline hazard and non-parametric frailty, we have implemented an algorithm to estimate mixing distribution in order to obtain the NPMLE of frailty distribution. We have used the Bernstein polynomials to model the non-parametric cause-specific baseline hazard functions, and using sieve maximum likelihood estimation, we have estimated the non-parametric cause-specific baseline hazard function and relevant frailty parameter(s). In every case, simulation studies have been carried out in order to investigate the finite sample properties of the proposed methodologies. We have illustrated the proposed methods through the analyses of a hearing loss data.

Chapter 1

Introduction

Survival analysis is a discipline within Statistics that focuses on analyzing time-to-event data (See Kalbfleisch and Prentice, 2002, Cox and Oakes, 2018, Klein and Moeschberger, 2006, Klein et al., 2014, Kleinbaum and Klein, 1996, Fleming and Harrington, 2013, Liu, 2012, Guo, 2010 etc). In survival analysis, our main concern is the inference about the positive random variable that represents the times of certain events. It is also known as the analysis of failure time data. By failure, we mean the occurrence of the well-defined event of interest. For example, failure can be death, the onset of some disease, the failure of some machine components, and many others. Failure time data generally arises in medical studies, but there are other areas, such as biological studies, economic and financial studies, reliability, and sociological and psychological studies. The failure time random variable describes the length of time from a suitably chosen starting point to the occurrence of failure at an end point. The random variable corresponding to failure is usually continuous.

In many studies, for example, in clinical trials, some of the individuals might withdraw from the trial before the end of the follow-up time without failure. Hence, there is incomplete information on the failure time. An individual with such incomplete information is known as censored. When there is no censoring, the usual inference methodologies are used to analyze the data. However, these methodologies do not work in the presence of censoring. In our case, the type of censoring we encounter is termed as "interval censoring". When the exact failure time of an individual is not known but is observed within an interval, it is known as interval censoring. See Sun (2006), Sun and Chen (2022).

1.1 Current Status Data

Current status data is an extreme case of interval-censored data (See Sun, 2006). Only the failure status of an individual, i.e., whether the failure has occurred or not at some monitoring time, is observed. Many fields, for example, carcinogenicity studies, epidemiological studies, and reliability studies, give rise to current status data. Let us describe the current status with the help of an example from reliability studies. Suppose our event of interest is the time of occurrence of a non-fatal and non-observable fault in a machine, but one can only observe the failure status at some inspection point.

An ample amount of research work on current status data already exists in statistical literature. A great deal of parametric, non-parametric, and semi-parametric methods on current status data have been studied. It can be noted that the likelihood of current status data is a function of distribution functions computed at monitoring times. If the monitoring times are assumed to be fixed and finite, then the likelihood function boils down to a product binomial likelihood where the binomial probabilities are the distribution functions computed at the monitoring times, which are non-decreasing, therefore satisfying some isotonic constraints. Hence, it is a constrained optimization problem with some isotonic constraints. Ayer et al. (1955) were the first to study nonparametric maximum likelihood estimation with such constraints. In order to compute the maximum likelihood estimates of the binomial probabilities, they proposed the pool-adjacent-violators algorithm. The paper by Keiding et al. (1996) extensively investigates statistical methods for analyzing current status data. Authors specifically focused on the non-parametric maximum likelihood estimation (NPMLE) based on current status data, and they mentioned that even though some methods for computing the NPMLE had been known for a long time, the mathematical properties of the NPMLE had recently been established at the time of their writing. In this work, the authors applied the NPMLE to compute the cumulative probability of immunization by a certain age using data from a rubella seroprevalence survey conducted in Vienna in 1988. They also conducted a comparative study of the fully parametric, the NPMLE, and semi-parametric approaches.

Diamond et al. (1986) used a semi-parametric proportional hazards model in order to study the association between a group of explanatory variables and failure time. The event of interest in this study was the weaning of an individual, and the data was available only in the current status form. The authors fitted a generalized linear model, which was computationally extensive. Therefore, the authors recommended

another approach where they modeled the baseline hazard using splines with a limited number of parameters. Rossini and Tsiatis (1996) introduced a semi-parametric proportional odds model to study current status data. They proposed to approximate the baseline log-odds function with a step function, thereby transforming the infinite-dimensional nuisance parameter to a finite-dimensional one. Then, the authors applied the method of maximum likelihood to estimate the regression coefficients. The consistency, asymptotic normality of the estimators, and semi-parametric efficiency were established. A detailed simulation study was performed to illustrate the finite sample behavior of the proposed estimators. Lin et al. (1998) proposed a semi-parametric additive hazards model to study current status data. This model assesses how different covariates affect the hazard function additively. A notable aspect of their approach is that, under specific assumptions, the analysis of current status data using this additive hazards regression model can be transformed into a framework similar to the proportional hazards model for right-censored data. This enabled the authors to derive the asymptotic properties of the proposed estimators. To demonstrate a practical use of their estimators, the authors applied this methodology to a data from a carcinogenicity experiment. However, the assumptions made on the monitoring time and the method itself may not be appropriate.

All the papers mentioned above are based on the assumption that all individuals under study are susceptible to failure. However, there exist studies where the sample also includes individuals not susceptible to failure, known as the cured group. A substantial amount of research already exists in the area of lifetime data in the presence of cured individuals (See Fang et al., 2005, Li et al., 2007, Othus et al., 2009 among many others). A significant amount of work has been done on current status data in the presence of a cured subgroup. Lam and Xue (2005) considered the analysis of current status data with cured proportion using a mixture model that combines the logistic regression formulation for the probability of cure and a semi-parametric regression model for the time to occurrence of failure. This semi-parametric model also enables one to study the possible nonlinear effect of some covariate on the response. The asymptotic properties of the estimators had been studied, and it was shown that the estimators are strongly consistent under mild assumptions. Thorough simulation studies had been performed to investigate the finite sample behavior of the proposed estimators and the proposed methodologies were applied to a dataset from a cataract study.

The works mentioned are based on the assumption that the failure time of the event and the monitoring time are independent of each other. But in practice, there

may be some scenarios where it doesn't hold true. This is known as dependent censoring. Ma (2009) considered a generalized linear model for the cure probability, and linear and partly linear proportional hazards models were used to model the failure time under dependent censoring. The authors used the penalized maximum likelihood approach for estimation, and large sample properties of the proposed estimators were studied extensively. Weighted bootstrap methods were applied for the inference about the regression coefficients. To illustrate the behavior of the proposed estimators in practical cases, a detailed simulation study and a real data application had been performed. The regression analysis for current status data with the cured subgroup in the presence of dependent censoring had also been received attention. Liu et al. (2017) introduced a latent variable to model the dependence between failure time and inspection time. A sieve maximum likelihood approach is introduced along with the use of latent variable and Bernstein polynomials. An EM algorithm was proposed to estimate the model parameters. Asymptotic properties of the estimators were established and a detailed simulation study, including a real data analysis, was performed. In most of the dependent censoring works, it is assumed that the dependence structure is known (See Ma, 2009,Zhao et al., 2015 etc.), which may not be so in reality. Cui et al. (2018) introduced a new copula-based approach that doesn't require any association assumption. They proposed a new two-step sieve estimator where the marginal proportional hazards model for monitoring time is first estimated, and then the regression and association parameters are estimated.

1.2 Current Status Data with Competing Risks

Competing risks data arises when there is more than one causes for the occurrence of failure but only one of them becomes responsible for the failure. For each failure, the cause of failure is observed. There may be current status data in the presence of competing risks in which case the cause of failure is also observed if the status at the monitoring time is observed to be failure. A substancial amount of research work has been done for analyzing current status data with competing risks.

Hudgens et al. (2001) were the first to study the interval censored and truncated data with competing risks. In order to obtain non-parametric maximum likelihood estimate of the underlying failure time distribution, they proposed an EM algorithm. Jewell et al. (2003) addressed the problem of non-parametric estimation of survival functions using current status data with two competing risks, which can be generalized to more than two competing risks. In this work, they proposed an algorithm for

computing the NPMLE of the sub-distribution functions, in addition to two naive estimators of the same which have some limitations. Subsequently, it has been observed in the simulation studies that NPMLE works better than the other two approaches. The methodologies have been illustrated using the menopause data from Krailo and Pike (1983). Jewell and Kalbfleisch (2004) viewed the sub-distribution functions for current status data with competing risks as ordered multinomial parameters and proposed an iterative algorithm using a modified version of the pool adjacent violators algorithm to compute the NPMLE in this case. In this work, they also established the uniqueness of the NPMLE and the convergence of the proposed algorithm to the unique NPMLE. Maathuis (2006) conducted an extensive study on the asymptotic properties of the NPMLE and the naive estimator proposed by Jewell et al. (2003). She proved consistency, $n^{\frac{1}{3}}$-rate of convergence, and asymptotic distribution of both the naive estimator and the NPMLE. See also Groeneboom et al. (2008a,b) and Maathuis et al. (2011) for more details. Koley and Dewanji (2019) focused on the problem of NPMLE of the sub-distribution function based on current status data with competing risks. This is generally a constrained optimization problem. The authors used a re-parametrization in terms of interval hazards to transform the problem into a non-constrained one. An EM algorithm was proposed to estimate the unknown hazards. The missing cause of failure is a common problem in the case of competing risks data. Koley and Dewanji (2022) discussed a parametric approach for current status data with two competing risks and missing failure types.

1.3 Bivariate Current Status Data

Bivariate current status data arises when a pair of related individuals is inspected at some monitoring time to note the corresponding failure status at that time. For example, consider the cancer onset times of a mother and her daughter. The exact cancer onset times are not observed, but their presence or status can be ascertained, if at all, only at some monitoring time. Bivariate current status data also arises when one records the failure status of two different organs of the same individual at a monitoring time. For example, one may be interested in the hearing loss status in both left and right ears at a monitoring time. While the mother-daughter pair with their cancer onset status constitutes an observational unit in the former example, an individual with hearing loss status in both ears is treated as an observational unit in the latter example.

A significant amount of research exists for analysis of bivariate current status data

with an effort to model association between the two components in an observational unit in different ways. Wang and Ding (2000) modeled the association using a copula model, and a two-stage estimation procedure is proposed to estimate the Kendall's tau like association parameter. For a common observed monitoring time, van der Laan and Jewell (2002) observed that the identifiable part of the joint distribution function is represented by the two marginal cumulative distribution functions and the joint cumulative distribution function evaluated at the diagonal. They proposed an EM algorithm to compute the NPMLEs of these distribution functions. They used these NPMLEs to construct a test of independence of two survival random variables and a goodness of fit test for the copula model used in Wang and Ding (2000). Ding and Wang (2004) introduced a non-parametric test for independence based on bivariate current status data. Unkel and Farrington (2012) proposed a new time-varying measure of association in bivariate current status data. This measure is convenient and particularly helpful due to its interpretability for suggesting the correct frailty model. An estimator of this measure is introduced and compared with another existing association measure for current status data.

One alternative way to model association in bivariate lifetime and current status data is to consider frailty variables. This approach has received considerable attention in research activities. Chen et al. (2009) studied the regression analysis of multivariate current status data. A full likelihood based approach based on the proportional hazard frailty model was introduced. An EM algorithm is provided for the estimation of model parameters. An added benefit of the frailty variable is that it helps to understand individual heterogeneity. Even though the use of shared frailty is very common, Hens et al. (2009) remarked that the use of correlated frailty eliminates the need to assume that unobserved factors within a cluster act similarly. The authors analyzed data from a cross-sectionally collected serological dataset on hepatitis A and hepatitis B to illustrate this phenomenon. Wen and Chen (2011) proposed a semi-parametric maximum likelihood analysis of clustered current status data with the Gamma-frailty Cox model, using an efficient computational algorithm. Wang et al. (2015) introduced a novel EM algorithm assuming the Gamma frailty proportional hazards model to analyze bivariate current status data. The conditional baseline cumulative hazard functions are approximated using monotone splines, which leads to estimation of a finite number of parameters along with model flexibility. The authors argued that this algorithm is easy to implement, robust to initialization, and converges quickly. Li et al. (2020) studied regression analysis of multivariate current status data using the EM algorithm under a class of flexible semi-parametric

transformation models. Tran et al. (2020) proposed a general frailty model employing a copula function between the frailty terms to construct flexible bivariate frailty distributions with applications to current status data. The proposed model enjoys less correlation structure than the other popular frailty models. Different copula functions were explored, and the Gamma distribution was chosen for the marginals because of its mathematical convenience. A detailed numerical study shows that the proposed method outperforms the use of additive correlated Gamma frailty model of Yashin et al. (1995), especially in the case of negative correlation between the frailties. The proposed method is applied to a bivariate serological dataset. There are many other works related to regression analysis of bivariate current status data. See, for example, Wang et al. (2008), Chen et al. (2015), Kim (2014), Liu and Qin (2018), Hu et al. (2017), Tong et al. (2012) among many others.

1.4 Bivariate Current Status Data with Competing Risks

When we have current status data on failure of multiple individuals in a cluster who may be related, or failure of multiple organs or components from a single individual, and each failure is subject to several competing risks, then this is known as multivariate current status data with competing risks. In this thesis, we focus on bivariate current status data with competing risks.

Even though there is an extensive amount of research on univariate current status data, with or without competing risks, and also on bivariate current status data, as per our knowledge, analysis of bivariate current status data with competing risks has not received much attention in the literature. In this thesis, the focus is on the analysis of bivariate current status data with competing risks using different frailty structures. In particular, we focus on parametric and non-parametric estimation of the model parameters, including frailty parameters.

Even though many different models, including the frailty models, have been used for the analysis of bivariate current status data, identifiability issues of the models have not been properly addressed. Model identifiability is an important issue that needs to be addressed in the case of competing risks data (See Tsiatis, 1975). A review of identifiability results for bivariate frailty models can be found in Iachine (2004). It was shown in this work that shared frailty models could sometimes be non-identifiable in case of infinite frailty mean. Therefore, a proper study of identifiability issues is

an important aspect while dealing with bivariate current status data with competing risks. Abbring and Van den Berg (2003) studied the identifiability of a copula based bivariate lifetime regression model in the presence of two competing risks. Hence, for the sake of completeness, we have investigated the identifiability issues of the different models introduced in this thesis.

1.5 Modeling of Bivariate Lifetime Data

Extensive research work for the analysis of bivariate lifetime data exists in the literature. The dependence in bivariate survival model can be induced by a common random effect variable, known as shared frailty. See Clayton (1978) and Oakes (1986). Oakes (1989) also introduced a class of bivariate survival distributions arising from shared frailty, allowing for negative association. Hougaard (1984, 1986),Yashin et al. (1994),Yashin et al. (1995),Wienke et al. (2001),Zdravkovic et al. (2002) and many others used the correlated Gamma frailty to model association in bivariate survival data. Hanagal and Pandey (2015) discussed shared Gamma frailty model by introducing two new hazard functions, namely, generalized log-logistics and generalized Weibull. The authors applied Bayesian estimation procedure to estimate the model parameters. Hanagal and Sharma (2015) also studied bivariate survival data using shared inverse-Gaussian frailty. To overcome the limitations of the shared frailty model, Hanagal et al. (2017) used correlated frailty models with the same two hazards for bivariate survival data.

Another widely used approach to model the association in bivariate survival data is through copula. Shih and Louis (1995) proposed a two-stage parametric and a two-stage semi-parametric estimator to estimate the association parameter of the copula. A new class of bivariate survival models based on Archimedean copula is introduced by dos Reis Miranda Filho and Demarqui (2025), which accommodates many of the well-known copulas. The baseline hazards are assumed to be piecewise constant and also approximated by means of Bernstein polynomials.

1.6 Modeling of Bivariate Lifetime Data with Competing Risks

A summarized study of modeling and analysis of multivariate survival data with competing risks can be found in the book by Crowder (2012). Bakoyannis and Touloumi

(2012) discussed a review of the practical applications of competing risks data. In this paper, the authors described the widely used Fine and Gray model for sub-distribution function to analyze competing risks data under the assumption of administrative censoring. Several numerical studies have been conducted, and it was shown that incorrect assumptions about administrative censoring have little effect on the inference for sub-distribution functions. Finally, the proposed methods were applied with CASCADE data on HIV-1 seropositive patients from a collaborative multicentre study.

1.6.1 Parametric Modeling

Bandeen-Roche and Liang (2002) proposed a measure of association for bivariate survival data with competing risks based on a measure of Oakes (1989). The authors also described the estimation of this association parameter. Wang and Ghosh (2003) observed that the absolutely continuous bivariate Exponential distribution is unsuitable for standard inferential procedures due to identifiability issues. Therefore, the integrated likelihood approach was generally considered in the frequentist paradigm. In this work, the authors presented a Bayesian approach using a non-informative prior such as Jeffrey's prior. Ibrahim and Kudus (2009) developed a decision tree model for prognostic classification of multivariate survival data and competing risks. Shoaee and Khorram (2020) introduced a new bivariate distribution based on the Marshal-Olkin model, which can be applied in the presence of competing risks. The authors have suggested an EM algorithm to compute the maximum likelihood estimate.

1.6.2 Non-parametric Modeling

There have been some efforts in non-parametric modeling of bivariate lifetime data with competing risks in terms of bivariate cause-specific hazard functions (See Sankaran et al., 2006) and subsequent non-parametric estimation of the cause-specific joint sub-distribution functions. Sankaran and Antony (2007) considered modeling in terms of bivariate cause-specific hazards, defined by Johnson and Kotz (1975) and in line with Dabrowska (1988) for bivariate lifetime data, and then non-parametric estimation of the cause-specific sub-distribution functions. Both of these works do not consider any explicit modeling of the association between the two lifetimes, possibly due to different competing risks.

One important aspect of the analysis of bivariate survival data with competing risks is modeling the association and estimating the association parameters. Bandeen-

Roche and Ning (2008) proposed the "conditional cause-specific hazard ratio" measure, a modification of the "conditional hazard ratio" (see Oakes, 1989). The authors provided an intuitive, non-parametric estimator that localizes Kendall's tau. Cheng et al. (2007) provided non-parametric estimators for the bivariate cause-specific hazard functions and joint sub-distribution functions without any explicit assumption on the dependence measure including competing risks.

1.6.3 Semi-parametric Modeling

Ng and McLachlan (2003) studied a semi-parametric mixture regression model, focusing on inference concerning the effect of regressors on both the probability of occurrence and the hazard rates conditional on each of the failure types. The joint estimation of the logistic and the regression coefficients was considered while the baseline hazard functions remained completely unspecified. An Expectation Conditional Maximization algorithm was proposed to estimate the model parameters. Naskar et al. (2005) considered a general class of lifetime distributions to analyze clustered competing risks data. It was assumed that the marginal probabilities of the various failure types are logistic functions of the covariates. The authors proposed a hybrid estimation method under the non-parametric frailty distribution using the Dirichlet process. Mixture regression modeling is a widely used way to model competing risks data (See Lu and Peng, 2008). Scheike et al. (2010) studied the multivariate competing risks data using a semi-parametric random effects model. In this model, dependence between the cause-specific hazard rates was modeled by copula functions that were allowed to depend on cluster-specific regressors. A two-stage estimation procedure was introduced where the marginal models were estimated at the first stage and then followed by estimation of the association parameters. Dixon et al. (2011) studied a proportional hazards model for multivariate competing risks data where the frailty term acts multiplicatively. The baseline cause-specific hazard functions were assumed to be non-parametric, and the Gamma distribution with mean unity was considered for the frailty variable. The authors proposed a profile likelihood approach for parameter estimation.

1.6.4 Copula Modeling

Cheng and Fine (2012) observed that the bivariate cumulative incidence function is non-parametrically identifiable based on bivariate survival data with competing risks because each likelihood contribution can be written in terms of bivariate cumulative

incidence function. They modeled the bivariate incidence function using a copula between univariate incidence functions. Then, parametric and non-parametric estimation of the univariate incidence functions were proposed. In this regard, two estimating equations were proposed to estimate the association parameter.

1.6.5 Frailty Modeling

Gorfine and Hsu (2011) introduced a new frailty-based competing risks model, generalizing the work of Prentice et al. (1978) to accommodate frailty variables to model association in clustered survival data. This was a cause-specific proportional hazards frailty model enjoying many attractive features, for example, the inclusion of covariates, flexible dependence structures within clusters, etc. Lai et al. (2017) modeled clustered survival data with competing risks by introducing a bivariate random effects term in the cause-specific hazard functions.

1.7 Four Types of Frailty

Vaupel et al. (1979) introduced the concept of frailty variable to account for heterogeneity due to the unobserved covariate effect to study the mortality of an individual. Another important use of frailty variable is to model association in clustered survival data (See Wienke, 2010). Thus, in the context of bivariate failure time data, frailty generally describes the between pair heterogeneity and also the within pair association.

The shared frailty model, where individuals in a cluster experience the common effect of a random frailty variable, was introduced in the literature by K.P. Clayton in his 1978 Biometrika paper, although not using the term 'frailty'. In this work, Gamma distribution was considered for the shared frailty variable. Hougaard (1984, 1986, 2000), Duchateau and Janssen (2008) and others elaborated and provided applications of this frailty model. Until now, this model has been widely popular among many practitioners. Murphy (1994, 1995) established consistency and asymptotic normality of maximum likelihood estimator of the shared frailty model without covariates. But as mentioned earlier, Iachine (2004) proved this model can be non-identifiable sometimes. One major criticism of using this model for analyzing clustered survival data is that it does not entirely consider individual heterogeneity. This is another important reason why the correlated frailty model has been developed for the analysis of clustered survival time data. Yashin et al. (1994), Pickles

and Crouchlev (1994), Petersen (1998), Wienke et al. (2003) and others investigated correlated frailty model to overcome this difficulty. The correlated frailty model not only captures the common effect of the frailty variables within a cluster but it also induces individual heterogeneity among the cluster members. Yashin et al. (1995), Yashin and Iachine (1997), Petersen (1998), Wienke et al. (2001, 2003) and Zdravkovic et al. (2002) applied the correlated Gamma frailty model to different settings. Parner (1998) studied the asymptotic properties of the maximum likelihood estimator of correlated Gamma frailty variables with observed covariates.

When one encounters more complicated data, for example, bivariate survival (or bivariate current status) data with competing risks, there can be many kinds of dependence structures. In this regard, Gorfine and Hsu (2011) introduced shared cause-specific frailty (although they did not use the term 'shared cause-specific frailty') models with covariates, in addition to several other kinds of frailty structures, depending on the failure types, and possibly on time as well, to describe different types of associations within clusters and between the failure types. See also Ha et al. (2018). To consider such a wide spectrum of frailty structures leads to complications and lack of interpretation. Therefore, in this thesis, we consider two other types of frailty structures, in addition to shared and correlated frailty, depending on the cause of failure. We call them the 'shared cause-specific frailty' and 'correlated cause-specific frailty'. See Section 2.4 for more details.

1.8 Objectives of the Thesis

There are many scenarios, for example, in carcinogenic studies and reliability experiments, which give rise to current status data. When the failure status (1 for failure, 0 for no failure or censoring) of two related individuals (or two components of a single machine) is observed only at some monitoring or inspection time, then we have bivariate current status data. When each of the failures, if they occur, may be due to one of several causes, we have bivariate current status data with competing risks. As per our knowledge, bivariate current status data with competing risks has not received much attention in the literature. In this thesis, we intend to initiate this research on bivariate current status data with competing risks, in particular, focusing on modeling and estimation of the quantities of interest in the absence of any covariate. The main objectives of the thesis are listed in the following.

1. Modeling of bivariate lifetime data with competing risks through different frailty structures in order to describe different kinds of associations. This is achieved

through multiplicative frailty effect on the different cause-specific baseline hazard functions.

2. Investigating model identifiability for the four types of frailty under situations when the different cause-specific baseline hazard functions and frailty distribution can be parametric or entirely unspecified (non-parametric).

The models in the first objective and the identifiability results of the second objective are useful in the analysis of bivariate lifetime data with competing risks. It is proven that, when both the cause-specific baseline hazard functions and frailty distribution are non-parametric, then the model is not identifiable. For parametric frailty distribution, we consider the commonly used Gamma frailty of different kinds.

3. Maximum likelihood estimation when (a) both the cause-specific baseline hazard functions and frailty distribution are parametric, (b) the cause-specific baseline hazard functions are parametric and the frailty distribution is non-parametric, and (c) the cause-specific baseline hazard functions are non-parametric and the frailty distribution is parametric.

4. Carry out extensive simulation studies to investigate the finite sample properties of the estimates and empirically validate the convergence results.

5. Conduct an analysis of a real dataset on hearing loss (See Section 1.12) to illustrate our methodologies.

1.9 Chapter-wise summary of the thesis

In this section, we provide a summary of the thesis work, as presented in the subsequent chapters. This gives a fairly good idea about the content of the thesis and facilitates some understanding of its importance.

In chapter 2, we introduce the modeling of bivariate lifetime data with competing risks using a general frailty distribution. The four different frailty types under consideration, namely, shared, correlated, shared cause-specific, correlated cause-specific are special cases of this general model. It is of interest to measure the association between the failures of the two individuals in a pair, each due to a failure type from a set of competing risks. In this regard, we use the measure "conditional cause-specific hazard ratio" function proposed by Bandeen-Roche and Liang (2002). Then, we also define model identifiability for bivariate current status data with competing risks, which also applies to bivariate lifetime data with competing risks.

In chapter 3, we consider the fully parametric model; that is, both the baseline cause-specific hazard functions and the four different kinds of frailty distributions are

modeled parametrically. For this, we introduce a family of parametric life distributions that includes many commonly used parametric distributions as special cases. We also consider different gamma frailty distributions for the four frailty types. We establish model identifiability under fairly weak conditions. Then, we derive the conditional cause-specific hazard ratio function corresponding to the four frailty types. This conditional cause-specific hazard ratio is shown to be greater than 1 for shared Gamma frailty model indicating a positive association; for the other three frailty types, the conditional cause-specific hazard ratio function is plotted for some parameter values and found to be greater than 1. Maximum likelihood estimation is considered for the model parameters; consistency and asymptotic normality of the estimators are established. We carry out an extensive simulation study to observe the finite sample properties of the estimators. Finally, we fit the four frailty models to a dataset on hearing loss (See Section 1.12) to illustrate the proposed methodology.

In chapter 4, we model bivariate current status data with competing risks with parametric baseline cause-specific hazard rates and non-parametric frailty distribution. We consider the same parametric family for baseline cause-specific hazards introduced in Chapter 3. For each of the four frailty types, we prove model identifiability under some fairly weak conditions. Then, we introduce an iterative algorithm in which the estimates of parameters of baseline cause-specific hazards and those of the non-parametric shared frailty distribution are alternatively updated to increase the overall likelihood. This algorithm adopts a slight modification of that mentioned in Wang (2007) for estimating a mixing distribution. We have carried out a simulation study assuming constant baseline cause-specific hazards to empirically validate the convergence results and study the finite sample properties of the estimators. Convergence of this algorithm to the maximum likelihood estimates is established by a theorem from Lindsay (1983). However, nothing is known about the convergence rates and the limiting distribution of the estimators. Therefore, we suggest a sub-sampling approach of Politis et al. (1999) to estimate the convergence rate and then obtain the standard errors of the estimates along with the sub-sampling based approximate 95% confidence intervals. We have demonstrated our proposed algorithm by analyzing the hearing loss data, assuming constant baseline cause-specific hazards.

In chapter 5, we consider non-parametric cause-specific baseline hazard functions and Gamma frailty distribution. First, we establish model identifiability for the four types of Gamma frailty distributions. Then, we explore the estimation problem of the non-parametric baseline cause-specific hazard functions with shared Gamma frailty distribution. Following Osman and Ghosh (2012), we model the baseline cause-

specific hazard functions using Bernstein polynomials, and then, using the method of sieve maximum likelihood estimation proposed by Geman and Hwang (1982), we estimate the relevant model parameters. We carry out a limited simulation study to investigate error in this estimation method through integrated absolute error (IAE) and integrated mean squared error (IMSE). Finally, we demonstrate the method through analysis of the hearing loss data.

In chapter 6, we present a summary of the results and include some discussion on the limitations, possible improvements and potential applications of the methodologies developed in this thesis with some concluding remarks. We also indicate some future problems that we wish to attempt in near future.

1.10 Mathematical and Statistical Tools for the Thesis

In this thesis, we have used some important mathematical and statistical tools already present in the literature. A brief description of the same is provided below.

1.10.1 Uniqueness Theorem of Laplace Transformation

The Laplace-Stieltjes transform of a random variable X with cumulative distribution function F is given as

$$\mathcal{L}_F(s) = \mathbb{E}[e^{-sX}] = \int_0^\infty e^{-sx} dF(x) = F(0) + \int_{(0,\infty)} e^{-sx} dF(x), \ s > 0.$$

Theorem 1.1. Let $0 \leq X \sim F$ and let the sequence of real numbers $0 < n_1 < n_2 < \cdots$ satisfy $\sum_{k=1}^{\infty} \frac{1}{n_k} = \infty$. Then, the distribution F is uniquely determined by the countable set of values $\{\mathcal{L}_F(n_k), k = 1, 2, \cdots\}$ of its Laplace-Stieltjes transform.

Proof. See Section 3 of Lin and Dou (2021). □

Let us consider two random variables, X and Y, with the joint distribution function H and marginal distributions F and G; that is, $(X, Y) \sim H, X \sim F, Y \sim G$. Denote the Laplace-Stieltjes transformation of H by $\mathcal{L}_{\mathcal{H}}$. Formally,

$$\mathcal{L}_H(s,t) = \mathbb{E}[\exp(-sX - tY)] = \int_0^\infty \int_0^\infty e^{-sx-ty} dH(x,y), \ s,t > 0.$$

Theorem 1.2. Let $0 \leq X \sim F, 0 \leq Y \sim G$ and $(X, Y) \sim H$. Assume further that the two sequences m_k, n_k, for $k = 1, 2, \cdots$, satisfy

1. $0 < m_1 < m_2 < \cdots$ with $\sum_{i=1}^{\infty} \frac{1}{m_i} = \infty$, and
2. $0 < n_1 < n_2 < \cdots$ with $\sum_{j=1}^{\infty} \frac{1}{n_j} = \infty$.

Then the bivariate distribution H can be uniquely determined by the countably many values of its Laplace-Stieltjes transform: $\{\mathcal{L}_H(m_i, n_j) : i, j = 1, 2, \cdots\}$.

Proof. See Section 3 of Lin and Dou (2021). □

Let us consider m random variables satisfying $0 \leq X_i, i = 1, \cdots, m$ with joint distribution function H on $\mathbb{R}_+^m$. For $s_i > 0, i = 1, \cdots, m$, the Laplace-Stieltjes transform of H is given as

$$\begin{aligned}\mathcal{L}_H(s_1, s_2, \cdots, s_m) &= \mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{m} s_i X_i\Big)\Big] \\ &= \int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty} \exp\Big(-\sum_{i=1}^{m} s_i x_i\Big) dH(x_1, x_2, \cdots, x_m).\end{aligned}$$

Theorem 1.3. Let $\{n_{ik}, k = 1, 2, \cdots\}, i = 1, \cdots, m$, be the $m(\geq 3)$ sequences of real numbers satisfying $0 < n_{i1} < n_{i2} < \cdots$ and $\sum_{k=1}^{\infty} \frac{1}{n_{ik}} = \infty$ for all $i = 1, \cdots, m$.
Then, any m-dimensional distribution function H on $\mathbb{R}_+^m$ is uniquely determined by the countably many values of its Laplace-Stieltjes transform.

Proof. See Section 4 of Lin and Dou (2021). □

The Laplace-Stieltjes transformation has various applications in probability and statistics. In case of frailty modeling, the unconditional survival function is the Laplace-Stieltjes transformation of the distribution function of the frailty variable(s) at the baseline cumulative hazard function(s). We have used the above theorems in proving the identifiability results in Chapter 4 to establish uniqueness of the frailty distribution(s).

1.10.2 Asymptotics of the Maximum Likelihood Estimator

Let us restate the asymptotic properties of maximum likelihood estimators in the multiparameter case. We assume that the distribution depends on a parameter vector

$\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p)$. Let $X_1, \cdots, X_n$ be a set of n i.i.d. observations from the probability distribution $f(., \boldsymbol{\theta})$. The likelihood function of parameter vector $\boldsymbol{\theta}$ is given as

$$\mathcal{L}(\boldsymbol{\theta}|X_1, \cdots, X_n) = \prod_{i=1}^{n} f(X_i; \boldsymbol{\theta}).$$

The maximum likelihood estimator, $\hat{\boldsymbol{\theta}}$ is defined as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta} \in \Omega} \mathcal{L}(\boldsymbol{\theta}|X_1, \cdots, X_n) = \arg\max_{\boldsymbol{\theta} \in \Omega} \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{X}),$$

where the set Ω is the parameter space and $\boldsymbol{X} = (X_1, \cdots, X_n)$. It is often advantageous to work with the log-likelihood function $\ell(\boldsymbol{\theta}; \boldsymbol{X}) = \log(\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{X}))$. The estimator will remain same since the log function is a strictly monotone increasing function. Let us make the following assumptions.

- A.1 The distributions $P_{\boldsymbol{\theta}}$ of the observations are distinct for $\boldsymbol{\theta} \in \Omega$. Otherwise, $\boldsymbol{\theta}$ cannot be estimated consistently.

- A.2 The distributions $P_{\boldsymbol{\theta}}$, for $\boldsymbol{\theta} \in \Omega$, have common support. This implies that the support of the distributions $P_{\boldsymbol{\theta}}$ is independent of the parameter vector $\boldsymbol{\theta}$.

- A.3 The observations $\boldsymbol{X} = (X_1, \cdots, X_n)$, where the observations X_i's are i.i.d. with density function $f(x, \boldsymbol{\theta})$ with respect to a σ-finite measure μ.

- A.4 The parameter space Ω contains an open set ω for which the true parameter value $\boldsymbol{\theta}_0$ is an interior point.

- A.5 The first and second derivatives of $\ell = log\mathcal{L}$ satisfy

  $$\mathbb{E}\left[\frac{\partial}{\partial \theta_j} \log f(x; \boldsymbol{\theta})\right] = 0,$$

  and

  $$I_{jk}(\boldsymbol{\theta}) = \mathbb{E}\left[\frac{\partial}{\partial \theta_j} \log f(x; \boldsymbol{\theta}) . \frac{\partial}{\partial \theta_k} \log f(x; \boldsymbol{\theta})\right] = \mathbb{E}\left[-\frac{\partial^2}{\partial \theta_j \partial \theta_k} \log f(x; \boldsymbol{\theta})\right]$$

  for $j, k = 1, 2 \cdots, p$, where p is the dimension of $\boldsymbol{\theta}$.

- A.6 The $p \times p$ matrix $\mathcal{I}(\boldsymbol{\theta})$, comprising of $I_{jk}(\boldsymbol{\theta})$'s, is a dispersion matrix, and

it is positive semidefinite. For all $\boldsymbol{\theta} \in \omega$, the p statistics

$$\left[\frac{\partial}{\partial\theta_1}\log f(x;\boldsymbol{\theta})\right], \cdots, \left[\frac{\partial}{\partial\theta_p}\log f(x;\boldsymbol{\theta})\right]$$

are affinely independent with probability 1.

- A.7 There exist functions M_{jkl} such that

$$|\frac{\partial^3}{\partial\theta_j\partial\theta_k\partial\theta_l}\log f(x;\boldsymbol{\theta})| < M_{jkl}(\boldsymbol{\theta})$$

for all $\omega \in \Omega$ and $m_{jkl} = \mathbb{E}_{\boldsymbol{\theta_0}} M_{jkl}(X) < \infty$.

Theorem 1.10.1. *Let $X_1, \cdots, X_n$ be i.i.d., each with a density $f(x;\boldsymbol{\theta})$ (with respect to measure μ) which satisfies $A.1 - A.7$. Then, with probability tending to 1 as $n \to \infty$, there exists solutions $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(X_1, \cdots, X_n)$ of the likelihood equations such that*

- *(1) the estimator $\hat{\theta}_j$ is consistent for estimating θ_j, for $j = 1, 2, \cdots, p$.*
- *(2) $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \mathcal{I}^{-1}(\boldsymbol{\theta}))$ as $n \to \infty$.*

Proof. The proof can be found in Lehmann and Casella (2006) (pages $463 - 465$) and hence skipped. □

1.10.3 Estimation of Mixing Distribution Function

Note that the frailty variables are essentially independent random effect variables having a common unknown distribution function G. Although Kiefer and Wolfowitz (1956) studied and proved consistency of the maximum likelihood estimators (MLEs) of both the model parameters (associated with the baseline cause-specific hazards) and the unknown mixing (frailty) distribution G, under some regularity conditions including identifiability of the model, they did not suggest any method for obtaining the MLEs or making related inference. Lindsay (1983) investigated the geometric properties of these MLEs by looking at the convex hull of the set of likelihood functions obtained from all singular values of the random effect. By exploiting the results of convex geometry, this paper established that unique MLE of G exists, under quite general conditions, with support size no more than the number of distinct observations in the sample. See also Laird (1978). Finding this MLE of G reduces to a convex optimization problem, utilizing the directional derivatives of the log-likelihood function

with respect to the distribution G. A convergent algorithm, namely, vertex direction method (VDM), has also been proposed. Lindsay and Lesperance (1995) gave a nice review of all this development, along with some application areas. Heckman and Singer (1984b) first proposed modeling and estimation of frailty with arbitrary mixing distribution G using results of Lindsay (1983) in order to explain population heterogeneity.

Several algorithms have been suggested over the years to find the MLE of G, starting from the use of EM algorithm Laird (1978), which initiates with a large number of support points. Other methods start with a smaller number of initial support points and then add one or more new support points at each iteration. Apart from the VDM of Lindsay (1983), we have the vertex exchange method (VEM) of Böhning (1985), the intra-simplex direction method (ISDM) of Lesperance and Kalbfleisch (1992), etc., among some others. However, these methods seem to converge too slowly to be useful, especially in simulation studies when the MLE needs to be calculated many times. See Böhning (1995) for a review of these computational methods. Wang (2007) suggested a fast algorithm using a linear regression formulation of the quadratic programming method, extending the same proposed by Atwood (1976) and Böhning (1985) in the context of finding optimal design. One important characteristic of this algorithm is to avoid the need of keeping track of support points, while collapsing similar ones and discarding ones with very small probability. In the following, we give some details of the algorithmic method proposed by Wang (2007) to estimate the mixing distribution G non-parametrically, for given the updated estimates of the model parameters, and alternately estimate the model parameters, for given the updated estimate of G. From the geometry of mixture likelihood as discussed in Lindsay (1983), it is well-known that $\hat{G}$, the non-parametric MLE of the mixing distribution G, is a discrete distribution with the size of support points not exceeding the number of distinct observations in the sample. Therefore, finding $\hat{G}$ reduces to the problem of finding a proper set of support points along with their mixing proportions, including their common dimension; see Wang (2007). He suggested starting with any finite support and then updating the set with adding all the local maxima of the directional gradient of log-likelihood with respect to the mixing proportions. Then, an estimate of mixing proportion is obtained by a slight modification of a quadratic approximation developed in Atwood (1976). Zero entries of this set are dropped and a new set of support and mixing proportions is found. This procedure is continued until one finds a set of support points and a set of mixing proportions where the maximum value of the directional derivative is less than a given precision. We use this method

in Chapter 4 for our semi-parametric estimation of the parameters associated with the baseline cause-specific hazard rates and the arbitrary frailty distribution G.

1.10.4 Subsampling Bootstrap

In many areas, for example, in time series and spatial data analysis, practitioners often deal with dependent data. Therefore, the standard bootstrap is not applicable or may not perform well in such scenarios. The subsampling bootstrap technique (See Politis et al., 1999) seems to be useful instead of the usual bootstrap. In subsampling bootstrap approach, smaller subsamples are drawn randomly and without replacement from the original sample.

Subsampling bootstrap is often more advantageous than the usual bootstrap in cases of statistical computations involving large datasets because of smaller sample size, thereby making it computationally less burdensome (Moolgavkar et al., 2013). There are some research works that have investigated and discussed the efficiency of subsampling bootstrap in the case of heavy-tailed data. Kokoszka and Parfionovas (2004) examined the applicability of bootstrap-based unit root tests to time series with heavy-tailed error distribution. Also, Cornea-Madeira and Davidson (2015) addressed the inconsistency of Efron's bootstrap for heavy-tailed distributions. Sengupta et al. (2016) proposed a computationally efficient and robust bootstrap technique under mild conditions for massive datasets also including heavy-tailed distributions based on subsampling bootstrap.

The asymptotic distribution of an estimator may be non-normal with or without the knowledge of the convergence rate; in such cases, the subsampling bootstrap has been explored to provide meaningful estimates of the standard error and confidence interval. Romano and Shaikh (2012) investigated conditions under which subsampling and bootstrap methods produced uniformly valid inference for a broad class of distributions. The main focus of this study was to construct confidence regions and tests that maintain nominal coverage probabilities even when usual asymptotic approximations are inadequate. It is also robust to model misspecification since it does not assume the underlying true distribution. Subsampling is helpful even when the convergence rate of a statistic is unknown or intractable. Bertail et al. (1999) proposed a method to estimate this rate using subsampling, enabling the construction of asymptotically valid confidence regions without prior knowledge of convergence rate. Even though it enjoys many attractive properties, one should be a bit careful about the selection of subsample size. There are some data-driven methods, one of

them is the Minimum Volatility method by Politis and White (2004). It should be noted that the performance of the subsampling procedure depends on certain tuning parameters like subsample size, number of subsamples, etc.. In a recent work, Ma et al. (2024) have developed methodologies to select this tuning parameter optimally for large datasets. This paper provides a closed-form solution of optimal subsample sizes and demonstrates the improved performance of these optimized methods by simulation.

We have implemented the subsampling bootstrap technique in Chapter 4 of this thesis to compute the rate of convergence of the proposed estimators, asymptotic standard errors, and asymptotic confidence intervals.

1.10.5 Bernstein Polynomials

Approximation of a continuous real-valued function arises in various fields of Statistics, for example, in non-parametric regression, density estimation, hazard rate estimation, smoothing of time series data, functional data analysis, machine learning and data science, and many others. In this regard, there have been many applications of kernel method (See, for example, Nadaraya, 1964; Watson, 1964; Ramlau-Hansen, 1983; Wand and Jones, 1994) and spline smoothing (See, for example, Wahba, 1990; Green and Silverman, 1993; Hao et al., 2020; Wu et al., 2022) in the literature. Another alternative is to use the Bernstein polynomials, although not as common as kernel method and spline smoothing. Some of the advantages of the use of Bernstein polynomials have been mentioned in Tenbusch (1994, 1997), including spatial adaptability and better boundary behaviour (See also Osman and Ghosh, 2012).

One of the most widely used applications of Bernstein polynomials in statistics is in density estimation, dating back to the work by Vitale (1975). Several years later, many researchers such as Petrone (1999), Babu et al. (2002), Choudhuri et al. (2004) investigated the use of Bernstein polynomials in different statistical applications. In the non-parametric regression setting, Chang et al. (2007) applied Bernstein polynomials. Chang et al. (2005) used Bernstein polynomials for hazard function approximation in the Bayesian framework without considering any covariate for right censored survival data. Considering the semi-parametric setting, Osman and Ghosh (2012) non-parametrically modeled the baseline hazard functions using Bernstein polynomials to analyze right censored survival data with covariates. The authors motivated the use of Bernstein polynomials from a result by Carnicer and Pena (1993) which states that Bernstein polynomials have the best shape-preserving property among

all approximating polynomials. Wang and Ghosh (2012) used Bernstein polynomials to estimate various shape restricted non-parametric regression function, for example, monotonicity, convexity, etc. Guan (2021) used Bernstein polynomials to approximate baseline hazard function in the cox proportional hazard model for interval censored data. He also proved that the proposed estimate of the survival function enjoys almost $n^{\frac{1}{2}}$ convergence rate, faster than the existing estimators with $n^{\frac{1}{3}}$ convergence rate. Dib et al. (2021) estimated joint distribution function of two random variables using Bernstein polynomials, when one of them is right censored establishing strong uniform convergence of this estimator.

In this context of Bernstein polynomials, the Weierstrass approximation theorem (originally introduced in 1885) plays a crucial role. This theorem asserts that, for any continuous function $f(x)$ on a closed interval, there exists a ploynomial $P(x)$ approximating $f(x)$ uniformly in the sense that, for any $\epsilon > 0$, $|f(x) - P(x)| < \epsilon$ for all x. In 1911, Russian mathematician Sergei Bernstein provided another well-known proof of this theorem using Bernstein polynomials (See Lorentz, 2012; Cui and Pillichshammer, 2025). In the following, we first formally introduce Bernstein polynomials.

Definition 1.1. For each positive integer $m \geq 1$, the mth Bernstein polynomial of a function $f \in C[0,1]$ (that is, f is a continuous real-valued function defined on [0,1]) is defined as

$$B_m(x, f) = \sum_{i=0}^{m} f\Big(\frac{i}{m}\Big)\binom{m}{i}x^i(1-x)^{m-i}, \text{ for } x \in [0,1].$$

We have considered approximation of baseline cause-specific hazard rates using Bernstein polynomials in Chapter 5, while analyzing bivariate current status data with competing risks when the baseline cause-specific hazards are unspecified. Use of Bernstein polynomials has some apparent computational advantages in our particular setting (See Sections 1.8 and 1.9), as explained later in Section 5.3. Also, since the baseline cause-specific hazards need to be continuous functions for the sake of model identifiability (See Section 5.2), use of Bernstein polynomials to approximate these hazard functions seems to be more suitable for easy implementation of numerical methods to maximize the complicated likelihood function (See Section 2.6) as that in our setting.

For the purpose of getting smooth estimates of the unspecified baseline cause-specific hazards, we have used sieve maximum likelihood estimator (Grenander, 1981; Geman and Hwang, 1982), as described briefly in the following with reference to

our particular setting. Consider the two classes $\mathcal{H}_1$ and $\mathcal{H}_2$ of continuous baseline cause-specific hazard functions for the two individuals in a pair (See Section 2.5). It is difficult to carry out likelihood maximization over the infinite dimensional space $\mathcal{H}_1 \times \mathcal{H}_2$. So, we consider a subspace $\mathcal{H}_{1m} \times \mathcal{H}_{2m}$ of it for the purpose of maximization, where $\mathcal{H}_{km}$ denotes the class of mth Bernstein polynomials approximating all the baseline cause-specific hazard functions for the kth individual, for $k = 1, 2$, with the typical ith coefficient of the form $f(\frac{i}{m})$ in each polynomial treated as unknown constant γ_i to be estimated. The sequence of subspaces, such as $\mathcal{H}_{1m} \times \mathcal{H}_{2m}$, of the original parameter space $\mathcal{H}_1 \times \mathcal{H}_2$, for $m \geq 1$, is called a sieve.

Under some standard regularity conditions, in the context of right censored survival data with covariates, Osman and Ghosh (2012) proved that there exists an estimator in the corresponding mth subspace which converges to the true hazard function, as $m \to \infty$, with respect to the Hellinger metric. See also Geman and Hwang (1982). This estimator in the mth subspace, obtained by corresponding likelihood maximization, is called the sieve maximum likelihood estimator.

1.11 Computational Tools

1.11.1 Numerical Optimization

Almost every problem in applied statistics boils down to an optimization problem with some objective function. Nowadays, we often deal with complicated objective functions. It is often a challenging task to obtain an analytical solution to such difficult optimization problem. Therefore, we resort to various numerical optimization algorithms (see Nocedal and Wright, 1999) to get an approximate solution to the difficult optimization problem at hand.

In this thesis, we have encountered both linear and non-linear constrained optimization problems in order to compute both parametric and non-parametric MLEs. We have used the R function 'optimParallel' of the 'optimParallel' package to apply quasi-newton methods like limited memory BFGS for box constrained optimization, 'Nelder-Mead' method in 'optim' function for derivative-free optimization of objective function with computationally expensive gradient calculations, and 'solnl' of the 'NlcOptim' package for nonlinear constrained optimization method, respectively.

1.11.2 Parallel Computing

Parallel computing is a handy and effective tool used in the area of computation. It enables the user to perform the task faster by splitting the task into smaller parts and processing them independently and simultaneously on multiple processors instead of performing the same task with the help of one processor. We explain it with the help of the following example.

Suppose that the given task is to build a house with four rooms. If we have four people instead of one and each of them works independently of each other and builds one room simultaneously, the house will be built much faster. Parallel computing also works according to the same principle. Parallel computing is heavily used in scientific computing, which arises in many fields like artificial intelligence, big data analytics, video game graphics, search engines, weather forecasting, etc. Today, statisticians often deal with problems that involve computationally burdensome calculations and processing of large data sets. Also, some statistical methodologies like bootstrapping need multiple replications of a single step. In all of these cases, parallel computing turns out to be highly beneficial for statisticians.

Parallel and doParallel packages of R software are used extensively in this thesis for parallel computing to handle difficult computation problems like numerical maximization of the complicated log-likelihood function.

1.11.3 Integration of Software packages

Nowadays, it is a common practice to integrate different software packages to consolidate the strengths of each software and tackle the complications of a difficult research problem. Different software turns out to be optimal for solving different tasks. It is also helpful for a collaborative work where the collaborators are from other disciplines. For example, R is popular for statistical analysis and C++ is popularly known for its high performance and accurate computing. Therefore, R and C++ integration can be highly useful for complex statistical analysis and computing. In this regard, Dirk Eddelbuettel and Romain Francois created and proposed the use of Rcpp; see Eddelbuettel (2013).

Rcpp, RcppNumerical, and some other packages along with parallel computing have been used to combine strengths of R and C++ to optimize their performance in different simulation studies and data analyses presented in this thesis.

1.12 Data Set

We have illustrated the different methodologies discussed in this thesis using a dataset on hearing loss. A short description of this dataset is given below.

Hearing loss is a common phenomenon that can occur at any age and gradually increases with age. The problem of hearing loss can be broadly classified into two categories according to which part of the ear is affected. **Sensorineural** hearing loss (SNHL) is related to any sensory problem or neural structures in the inner ear. **Conductive** hearing loss occurs due to obstructions in the outer or middle ear so that the sound does not properly enter through the outer ear to the middle ear. There are few observations with both the types of hearing loss being present, called Mixed type. For our analysis, we consider the failure time as the occurrence time of first hearing loss, so that the corresponding cause of failure is uniquely defined. However, due to the nature of current status data, we have some observations with both causes present (the Mixed type). These can be treated as observations with missing failure type, the analysis of which is beyond the scope of this thesis. For illustration, we discard these observations with Mixed type of hearing loss. In this thesis, this reduced dataset has been analyzed using parametric and semi-parametric modeling in chapters 3, 4 and 5.

This dataset has been collected from the department of speech and hearing, Ali Yavar Jung National Institute of Speech and Hearing Disabilities, Eastern Regional Centre, Kolkata. This dataset comprises 880 individuals who appeared with problem of speech impairment in the time span of January to February, 2016. As a related issue, hearing loss diagnosis, if any, is also carried out for these patients and the outcome of such diagnosis is the response of interest in our study. The monitoring time (X) has been considered to be the age (in months) of an individual at the time of diagnosis; see Koley and Dewanji (2022). We consider observations on only those individuals with hearing loss diagnosis in both left and right ears, so that we have bivariate current status data with the two types of hearing loss as competing risks. We have such information for 726 out of 880 individuals. We use the data on this smaller set of individuals for our analyses in this thesis. We use the labels 1 and 2 for the two competing risks, namely, SNHL and Conductive, respectively, for our analyses. The hearing loss data is summarized in Table 1.1, with cause of failure 0 denoting right censored observation. The full data is available from the author upon request.

Table 1.1: The cross-classified frequency table for the hearing loss data.

Left ear	Right ear		
	0	1	2
0	261	0	1
1	0	435	0
2	1	0	28

Chapter 2

Modeling of Bivariate Current Status Data with Competing Risks

2.1 Introduction

As discussed in Section 1.4, bivariate current status data with competing risks arises from two different scenarios. One scenario is when the failure status of two related individuals are observed at some monitoring or inspection time, and the failure of each individual occurs due to only one out of a set of possible causes. For example, the onset times of cancer in mother and daughter, with two types of cancer, namely, breast and ovarian, may be observed at some routine clinical check. Another scenario is when the failure status of two parts of a single individual or a machine are monitored. For example, hearing loss in an individual's left and right ears can occur due to one of three different causes (See Section 1.12). In our subsequent description, we shall refer to the first scenario and call the two failure times as those of the first and the second individual.

2.2 General Modeling using Frailty

Let T_1 and T_2 denote the failure times of the first and the second individual, respectively, with J_1 and J_2 denoting the corresponding causes of failure. It is assumed that the random variable J_k takes values in the set $\{1, \cdots, L_k\}$, for $k = 1, 2$. Therefore, the set of competing risks can be different for the two individuals. In some cases, we need to consider the same set of competing risks.

In the case of bivariate current status data with competing risks, the quadruplet (T_1, T_2, J_1, J_2) is not observed directly. Rather, the failure status of an individual due to a particular risk is observed. If the observed status is failure, then we also observe

the corresponding cause of failure, J_1 or J_2. Let X_1 and X_2 be the monitoring times for the first and the second individual, respectively. This allows for different monitoring times for the two individuals although, in most situations, the two individuals have the same monitoring time denoted by a random variable X. It is assumed that the monitoring time random variables X_1 and X_2 are independent of the failure time random variables, given by (T_1, T_2, J_1, J_2), without having any common parameter. Note that the observed data is of the form (X_1, X_2, J_1, J_2). From now on, we will follow the convention that J_1 or J_2 takes the value 0 when no failure is observed for the corresponding individual at the monitoring time, being a right censored observation. We also assume that the two failure times T_1 and T_2 are absolutely continuous.

The sub-distribution function corresponding to the kth individual due to cause j at time t_k is

$$F_j^{(k)}(t_k) = P\big[T_k \leq t_k, J_k = j\big] = \int_0^{t_k} f_j^{(k)}(u)du,$$

for all $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$, where $f_j^{(k)}(u)$ is the sub-density function for the kth individual due to cause j. The cumulative distribution function corresponding to the kth individual is

$$F^{(k)}(t_k) = P\big[T_k \leq t_k\big] = \sum_{j=1}^{L_k} F_j^{(k)}(t_k),$$

and the survival function for the kth individual is

$$S^{(k)}(t_k) = P\big[T_k > t_k\big] = 1 - F^{(k)}(t_k).$$

The primary quantity of interest in the context of bivariate failure time data with competing risks is the joint sub-distribution function, given by

$$F_{j_1 j_2}(t_1, t_2) = P\big[T_1 \leq t_1, T_2 \leq t_2, J_1 = j_1, J_2 = j_2\big],$$

for $t_k > 0, j_k = 1, \cdots, L_k$ and $k = 1, 2$.

Explicit expressions for all these quantities, including the joint sub-distribution function, require modeling of dependence between the failure times of the two individuals, possibly differently for different failure types. As discussed in Sections 1.5 and 1.6, we model this dependence solely by means of different types of frailty variables, so that the two failure times and types (T_1, J_1), and (T_2, J_2) are conditionally independent, given the frailties. In our development of the model, in addition to the

frailty distributions, we also require modeling of the cause-specific baseline hazard functions, as described below.

The jth cause-specific hazard function of the kth individual at time t_k is defined as

$$\lambda_j^{(k)}(t_k) = \lim_{h \to 0+} \frac{P\big[t_k < T_k < t_k + h, J_k = j | T_k \geq t_k\big]}{h} = \frac{f_j^{(k)}(t_k)}{S^{(k)}(t_k)},$$

for all $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$. In general, let $\boldsymbol{\epsilon^{(k)}} = (\epsilon_1^{(k)}, \epsilon_2^{(k)}, \cdots, \epsilon_{L_k}^{(k)})$ be the vector of frailty variables related to the kth individual. The conditional cause-specific hazard function for the kth individual due to cause j, given the frailty vector $\boldsymbol{\epsilon^{(k)}}$, is modeled multiplicatively as

$$\lambda_j^{(k)}(t_k | \boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k)\epsilon_j^{(k)}, \tag{2.1}$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$, where $h_{0j}^{(k)}(t_k)$ is the cause-specific baseline hazard function for the kth individual due to cause j at time t_k and $H_{0j}^{(k)}(t_k) = \int_0^{t_k} h_{0j}^{(k)}(u)du$ denotes the corresponding cause-specific cumulative baseline hazard function for individual k.

The conditional cause-specific cumulative hazard, given the frailty vector $\boldsymbol{\epsilon^{(k)}}$ is

$$\Lambda_j^{(k)}(t_k | \boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k} \lambda_j^{(k)}(u | \boldsymbol{\epsilon^{(k)}})du = \int_0^{t_k} h_{0j}^{(k)}(u)\epsilon_j^{(k)}du = H_{0j}^{(k)}(t_k)\epsilon_j^{(k)},$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$. The conditional cumulative hazard for the kth individual, given the frailty vector $\boldsymbol{\epsilon^{(k)}}$, is

$$\Lambda^{(k)}(t_k | \boldsymbol{\epsilon^{(k)}}) = \sum_{j=1}^{L_k} \Lambda_j^{(k)}(t_k | \boldsymbol{\epsilon^{(k)}}) = \sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k)\epsilon_j^{(k)},$$

for $t_k > 0$ and $k = 1, 2$. The conditional survival function for the kth individual at time t_k, given $\boldsymbol{\epsilon^{(k)}}$, is

$$S^{(k)}(t_k | \boldsymbol{\epsilon^{(k)}}) = \exp\big[-\{\Lambda^{(k)}(t_k | \boldsymbol{\epsilon^{(k)}})\}\big] = \exp\Big[-\sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k)\epsilon_j^{(k)}\Big],$$

for $t_k > 0$ and $k = 1, 2$. The sub-distribution function due to cause j for the kth

individual, conditional on $\boldsymbol{\epsilon^{(k)}}$, is

$$F_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k} \lambda_j^{(k)}(u|\boldsymbol{\epsilon^{(k)}}) S^{(k)}(u|\boldsymbol{\epsilon^{(k)}}) du = \int_0^{t_k} h_{0j}^{(k)}(u)\epsilon_j^{(k)} \exp\Big[-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(u)\epsilon_{j'}^{(k)}\Big] du,$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$. Using the conditional independence of T_1 and T_2 given the frailty vector $\boldsymbol{\epsilon^{(1)}}$ and $\boldsymbol{\epsilon^{(2)}}$, the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2)$ at time t_1 and t_2 corresponding to causes j_1 and j_2 for the two individuals, respectively, is

$$\begin{aligned} F_{j_1 j_2}(t_1, t_2) &= \mathbb{E}\Big[F_{j_1}^{(1)}(t_1|\boldsymbol{\epsilon^{(1)}}) F_{j_2}^{(2)}(t_2|\boldsymbol{\epsilon^{(2)}})\Big] \\ &= \int_0^{\infty} \cdots \int_0^{\infty} \left[\prod_{k=1}^{2} \int_0^{t_k} h_{0j_k}^{(k)}(u_k)\epsilon_{j_k}^{(k)} \exp\Big(-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(u_k)\epsilon_{j'}^{(k)}\Big) du_k \right] g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) d\boldsymbol{\epsilon^{(1)}} d\boldsymbol{\epsilon^{(2)}}, \end{aligned} \tag{2.2}$$

for $t_k > 0, j_k = 1, 2, \cdots, L_k$ and $k = 1, 2$, where $g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}})$ is the joint probability density function of the frailty vectors $\boldsymbol{\epsilon^{(1)}}$ and $\boldsymbol{\epsilon^{(2)}}$, respectively. The unconditional joint sub-density function is

$$\begin{aligned} f_{j_1 j_2}(t_1, t_2) &= \frac{\partial^2}{\partial t_1 \partial t_2} F_{j_1, j_2}(t_1, t_2) \\ &= \int_0^{\infty} \cdots \int_0^{\infty} \left[\prod_{k=1}^{2} h_{0j_k}^{(k)}(t_k)\epsilon_{j_k}^{(k)} \exp\Big(-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(t_k)\epsilon_{j'}^{(k)}\Big) \right] g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) d\boldsymbol{\epsilon^{(1)}} d\boldsymbol{\epsilon^{(2)}}, \end{aligned} \tag{2.3}$$

for $t_k > 0, j_k = 1, \cdots, L_k$ and $k = 1, 2$. The unconditional joint survival function is

$$\begin{aligned} S(t_1, t_2) &= \mathbb{E}\Big[S^{(1)}(t_1|\boldsymbol{\epsilon^{(1)}}) S^{(2)}(t_2|\boldsymbol{\epsilon^{(2)}})\Big] \\ &= \int_0^{\infty} \cdots \int_0^{\infty} \left[\exp\Big(-\sum_{k=1}^{2} \sum_{j'=1}^{L_k} H_{0j'}^{(k)}(t_k)\epsilon_{j'}^{(k)}\Big) \right] g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) d\boldsymbol{\epsilon^{(1)}} d\boldsymbol{\epsilon^{(2)}}, \end{aligned} \tag{2.4}$$

for $t_1, t_2 > 0$. The unconditional sub-distribution function for the kth individual due

to cause j is

$$F_j^{(k)}(t_k) = \mathbb{E}\big[F_j^{(k)}(t_k|\boldsymbol{\epsilon}^{(\boldsymbol{k})})\big]$$
$$= \int_0^\infty \Bigg[\int_0^{t_k} h_{0j}^{(k)}(u)\epsilon_j^{(k)} \exp\Big(- \sum_{j'=1}^{L_k} H_{0j'}^{(k)}(u)\epsilon_{j'}^{(k)}\Big) du \Bigg] g^{(k)}(\boldsymbol{\epsilon}^{(\boldsymbol{k})}) d\boldsymbol{\epsilon}^{(\boldsymbol{k})}, \qquad (2.5)$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$, where $g^{(k)}(\boldsymbol{\epsilon}^{(\boldsymbol{k})})$ is the marginal density of the frailty vector $\boldsymbol{\epsilon}^{(\boldsymbol{k})}$. The unconditional sub-density function for the kth individual due to cause j is

$$f_j^{(k)}(t_k) = \frac{\partial}{\partial t_k} F_j^{(k)}(t_k)$$
$$= \int_0^\infty h_{0j}^{(k)}(t_k)\epsilon_j^{(k)} \exp\Big(- \sum_{j'=1}^{L_k} H_{0j'}^{(k)}(t_k)\epsilon_{j'}^{(k)}\Big) g^{(k)}(\boldsymbol{\epsilon}^{(\boldsymbol{k})}) d\boldsymbol{\epsilon}^{(\boldsymbol{k})}, \qquad (2.6)$$

for $t_k > 0, j = 1, 2, \cdots, L_k$ and $k = 1, 2$.

2.3 Conditional Cause-specific Hazard Ratio

One important aspect of modeling bivariate failure time data with competing risks is to provide a measure of association of failure times of the two individuals, where each failure occurs due to one cause out of a set of possible causes. For bivariate failure time data, Oakes (1989) proposed the measure of association $\theta(t_1, t_2)$, called 'conditional hazard ratio', as

$$\theta(t_1, t_2) = \frac{\lambda(t_1|T_2 = t_2)}{\lambda(t_1|T_2 > t_2)}$$
$$= \frac{f(t_1, t_2) \Big/ - \frac{\partial}{\partial s_2} S(t_1, s_2)\Big|_{s_2=t_2}}{-\frac{\partial}{\partial s_1} S(s_1, t_2)\Big|_{s_1=t_1} \Big/ S(t_1, t_2)}$$
$$= \frac{f(t_1, t_2) S(t_1, t_2)}{\Big[- \frac{\partial}{\partial s_2} S(t_1, s_2)\Big|_{s_2=t_2}\Big] \times \Big[- \frac{\partial}{\partial s_1} S(s_1, t_2)\Big|_{s_1=t_1}\Big]},$$

for all $t_k > 0, k = 1, 2$, where $\lambda(t_1|T_2 = t_2)$ and $\lambda(t_1|T_2 > t_2)$ are the conditional hazards of T_1, given $T_2 = t_2$ and $T_2 > t_2$, respectively. Note that this measure

$\theta(t_1, t_2)$ is defined to be the proportion by which the first individual's failure risk at time t_1, given that the second individual has failed at time t_2, is increased over the same failure risk, given that the second individual has not yet failed by time t_2.

Motivated by this, Bandeen-Roche and Liang (2002) extended this measure to the case of bivariate failure time data with competing risks. They proposed the following measure, named the 'conditional cause-specific hazard ratio', as given by

$$\begin{aligned} CR_{j_1 j_2}(t_1, t_2) &= \frac{\lambda_{j_1}^{(1)}(t_1 | T_2 = t_2, J_2 = j_2)}{\lambda_{j_1}^{(1)}(t_1 | T_2 > t_2)} \\ &= \frac{S(t_1, t_2) f_{j_1 j_2}(t_1, t_2)}{\Big[\int_{t_2}^{\infty} \sum_{j=1}^{L_2} f_{j_1 j}(t_1, t) dt\Big] \times \Big[\int_{t_1}^{\infty} \sum_{j=1}^{L_1} f_{j j_2}(t, t_2) dt\Big]}, \end{aligned}$$

for $t_k > 0, j_k = 1, 2, \cdots, L_k, k = 1, 2$. It is easy to see that this measure is always positive. The conditional cause-specific hazard ratio $CR_{j_1 j_2}(t_1, t_2)$ is the risk of failure of the first individual at time t_1 due to cause j_1, given the second individual has failed at time t_2 due to cause j_2, over the same failure risk, given that the second individual has not failed at time t_2. According to Clayton (1978) (See also Oakes, 1989, Bandeen-Roche and Liang, 2002, and Gorfine and Hsu, 2011), if $CR_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) = 1$, there is no association between failure of the first individual at time t_1 due to cause j_1 and failure of the second individual at time t_2 due to cause j_2. When $CR_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ is larger (smaller) than 1, we have positive (negative) association between such failures of the two individuals.

2.4 Four Types of Frailty

As described earlier, frailty variables are used to model association in bivariate or multivariate failure time data. The use of shared and correlated frailty is widely popular to model the common and cluster-specific associations. But, for bivariate failure time data with competing risks, many other complicated types of associations exist, which we model through the frailty variables, possibly depending on the cause and time of failure as well (Gorfine and Hsu, 2011).

2.4.1 Shared Frailty

Often, a common random effect is used to describe association between the two failure times regardless of their failure types. This common random effect is the shared

frailty, which induces dependence between the two failure times and types. For example, the association between cancer onset times of a mother and daughter pair may be described by shared frailty because of their common family environment. If we substitute $\epsilon_j^{(k)} = \epsilon$, for all $j = 1, \cdots, L_k$ and $k = 1, 2$, then the general model (2.1) reduces to

$$\lambda_j^{(k)}(t_k|\epsilon) = h_{0j}^{(k)}(t_k)\epsilon,$$

where ϵ is the shared frailty variable. Since the frailty variable is not specific to failure type, the model can accommodate two different sets of competing risks for the two individuals.

2.4.2 Correlated Frailty

As the shared frailty model assumes a common frailty for both the individuals, thereby not being able to capture individual effects, correlated frailty model allows this individual heterogeneity by considering two correlated frailty variables for the two individuals within a pair. For example, the association between the cancer onset times of a mother-daughter pair arises due to the two correlated frailty variables describing their individual heterogeneous effects. Formally, writing $\epsilon_j^{(k)} = \epsilon^{(k)}$, for all $j = 1, 2, \cdots, L_k$, then the general model (2.1) reduces to

$$\lambda_j^{(k)}(t_k|\epsilon^{(k)}) = h_{0j}^{(k)}(t_k)\epsilon^{(k)},$$

where $\epsilon^{(k)}$ is the frailty variable for the kth individual and the pair $(\epsilon^{(1)}, \epsilon^{(2)})$ are correlated. As for the shared frailty, the two correlated frailty variable are not specific to the failure types. Hence, the model also allows two different sets of competing risks for the two individuals.

2.4.3 Shared cause-specific Frailty

There are other complicated dependencies in bivariate failure time data with competing risks, which cannot be modeled through the shared or correlated frailty model. For example, the dependence between the breast cancer onset time of a mother and the same for her daughter, in the presence of other cancer types, requires a frailty variable specific to breast cancer incidence. Similarly, the breast cancer incidence in a mother and ovarian cancer incidence in her daughter may be associated with each other in a complicated way which cannot be captured simply by a shared frailty variable or correlated frailty variables. We need a more complicated frailty struc-

ture; a simple alternative is to introduce frailty variables specific to failure types. Shared cause-specific frailty is one such model in which different frailty variables are considered for the different failure types, each of which is shared between the two individuals. Therefore, we need the set of competing risks to be the same for the two individuals with $L_1 = L_2 = L$, say. Formally, writing $\epsilon_j^{(k)} = \epsilon_j$ for $j = 1, \cdots, L$, the general model (2.1), reduces to

$$\lambda_j^{(k)}(t_k|\epsilon_j) = h_{0j}^{(k)}(t_k)\epsilon_j,$$

where ϵ_j is the shared frailty between two individuals due to cause j, for $j = 1, \cdots, L$.

2.4.4 Correlated cause-specific Frailty

In further generalization of the shared cause-specific frailty model described in the previous subsection, this model incorporates heterogeneity between the two individuals with respect to failure times due to certain failure types, while capturing the dependence through the correlated frailty variables specific to failure types. As a result, this model also requires the same set of competing risks for the two individuals with $L_1 = L_2 = L$, say. Formally, we have $(\epsilon_j^{(1)}, \epsilon_j^{(2)})$ as the correlated pair of frailty variables corresponding to failure type j, for $j = 1, 2, \cdots, L$. Thus, we write the general model (2.1) as

$$\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k)\epsilon_j^{(k)},$$

for $j = 1, \cdots, L$ and $k = 1, 2$.

2.5 Model Identifiability

Clearly, the pivotal quantity in the modeling of bivariate failure time, or current status, data with competing risks is the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2)$ in (2.2), or the unconditional joint sub-density function $f_{j_1 j_2}(t_1, t_2)$ in (2.3). All other quantities introduced in Section 2.2 can be derived from this joint sub-distribution function. That is, specification of this unconditional joint sub-distribution function describes the model for bivariate failure time, or current status, data with competing risks. We, therefore, discuss model identifiaiblity through this

quantity, as in the following. Let us define the classes $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{G}$ as

$$\mathcal{H}_1 = \left\{ \left(h_{01}^{(1)}(x), \cdots, h_{0L_1}^{(1)}(x)\right) : h_{0j}^{(1)}(x) > 0 \,\text{for}\, x > 0, j = 1, \cdots, L_1 \,\text{and} \int_0^\infty \sum_{j=1}^{L_1} h_{0j}^{(1)}(x)dx = \infty \right\},$$

$$\mathcal{H}_2 = \left\{ \left(h_{01}^{(2)}(x), \cdots, h_{0L_2}^{(2)}(x)\right) : h_{0j}^{(2)}(x) > 0 \,\text{for}\, x > 0, j = 1, \cdots, L_2 \,\text{and} \int_0^\infty \sum_{j=1}^{L_2} h_{0j}^{(2)}(x)dx = \infty \right\},$$

$$\mathcal{G} = \left\{ G(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) : G(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) \,\text{is the joint CDF of}\, L_1 + L_2 \,\text{variate frailty vector.} \right\}.$$

Definition 2.1. The general model (2.1) is identifiable within families $\mathcal{H}_1, \mathcal{H}_2, \mathcal{E}$, if for some $\boldsymbol{h_0^{(1)}(x)} = \left\{h_{01}^{(1)}(x), \cdots, h_{0L_1}^{(1)}(x)\right\}, \boldsymbol{\tilde{h}_0^{(1)}(x)} = \left\{\tilde{h}_{01}^{(1)}(x), \cdots, \tilde{h}_{0L_1}^{(1)}(x)\right\}$, where $\boldsymbol{h_0^{(1)}(x)}, \boldsymbol{\tilde{h}_0^{(1)}(x)} \in \mathcal{H}_1, \boldsymbol{h_0^{(2)}(x)} = \left\{h_{01}^{(2)}(x), \cdots, h_{0L_1}^{(2)}(x)\right\}, \boldsymbol{\tilde{h}_0^{(2)}(x)} = \left\{\tilde{h}_{01}^{(2)}(x), \cdots, \tilde{h}_{0L_2}^{(2)}(x)\right\}$, where $\boldsymbol{h_0^{(2)}(x)}, \boldsymbol{\tilde{h}_0^{(2)}(x)} \in \mathcal{H}_2$, and $G(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}), \tilde{G}(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) \in \mathcal{G}$, leading to expressions of $F_{j_1j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G)$ and $\tilde{F}_{j_1j_2}(t_1, t_2; \boldsymbol{\tilde{h}_0^{(1)}}, \boldsymbol{\tilde{h}_0^{(2)}}, \tilde{G})$, respectively, using (2.2), the equality

$$F_{j_1j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) = \tilde{F}_{j_1j_2}(t_1, t_2; \boldsymbol{\tilde{h}_0^{(1)}}, \boldsymbol{\tilde{h}_0^{(2)}}, \tilde{G}), \text{ for all } t_k > 0, j_k = 1, \cdots, L_k \text{ and } k = 1, 2,$$

implies that $h_{0j}^{(k)}(x) = \tilde{h}_{0j}^{(k)}(x)$, for all $x > 0$, $j = 1, \cdots, L_k$, $k = 1, 2$, and $G(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}) = \tilde{G}(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}})$ almost everywhere.

As described above, the classes $\mathcal{H}_1$,$\mathcal{H}_2$ and $\mathcal{G}$ are non-parametric in the sense that the baseline cause-specific hazards in $\mathcal{H}_1$,$\mathcal{H}_2$ and the frailty distributions in $\mathcal{G}$ do not belong to any specific parametric families. In the following, we prove that, in this generality, the model (2.1) is not identifiable.

From Section 2.2, note that the expression for joint sub-density function is

$$f_{j_1j_2}(t_1, t_2) = \int_0^\infty \cdots \int_0^\infty \prod_{k=1}^{2} \left(h_{0j_k}^{(k)}(t_k)\epsilon_{j_k}^{(k)}\right) \exp\left[-\sum_{k=1}^{2} \sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k)\epsilon_j^{(k)} \right] \times dG(\epsilon_1^{(1)}, \cdots, \epsilon_{L_1}^{(1)}, \epsilon_1^{(2)}, \cdots, \epsilon_{L_2}^{(2)}),$$

where $h_{0j_k}^{(k)}(t_k) \in \mathcal{H}_k$, for $t_k > 0$, $j_k = 1, \cdots, L_k$ and $k = 1, 2$. Let us write

$$\tilde{h}_{0j_1}^{(1)}(t_1) = c_1 h_{0j_1}^{(1)}(t_1) \quad \text{and} \quad \tilde{h}_{0j_2}^{(2)}(t_2) = \frac{1}{c_2} h_{0j_2}^{(2)}(t_2),$$

for all $t_1, t_2 > 0$, $j_k = 1, \cdots, L_k$ and $k = 1, 2$, where c_1 and c_2 are some positive constants. It can be easily verified that the $\tilde{h}^{(k)}_{0j_k}(t_k)$, for $k = 1, 2$, are also valid baseline cause-specific hazard functions, thus belonging to $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively.

Also, let us consider the frailty distribution given by

$$\tilde{G}\Big(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}\Big) = \frac{c_1^{L_1}}{c_2^{L_2}} G\Big(c_1\boldsymbol{\epsilon^{(1)}}, \frac{1}{c_2}\boldsymbol{\epsilon^{(2)}}\Big).$$

Since $G \in \mathcal{G}$, it is an easy exercise to check that $\tilde{G}$ also belongs to $\mathcal{G}$. Using the transformation of variables, $\boldsymbol{\delta^{(1)}} = c_1\boldsymbol{\epsilon^{(1)}}$ and $\boldsymbol{\delta^{(2)}} = \frac{1}{c_2}\boldsymbol{\epsilon^{(2)}}$, one can check that $\int_0^\infty \cdots \int_0^\infty d\tilde{G}\Big(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}\Big) = 1$.

With this choice of $\tilde{h}^{(k)}_{0j_k}(t_k)$, for $k = 1, 2$, and $\tilde{G}\Big(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}\Big)$, we have the corresponding joint sub-density function given by

$$\begin{aligned}
\tilde{f}_{j_1 j_2}(t_1, t_2) &= \int_0^\infty \cdots \int_0^\infty c_1 h^{(1)}_{0j_1}(t_1) \frac{1}{c_2} h^{(2)}_{0j_2}(t_2) \epsilon^{(1)}_{j_1} \epsilon^{(2)}_{j_2} \exp\Big[- \sum_{j=1}^{L_1} c_1 H^{(1)}_{0j}(t_1)\epsilon^{(1)}_j - \sum_{j=1}^{L_2} \frac{1}{c_2} H^{(2)}_{0j}(t_2)\epsilon^{(2)}_j \Big] \\
&\quad \times \frac{c_1^{L_1}}{c_2^{L_2}} dG(c_1\epsilon^{(1)}_1, \cdots, c_1\epsilon^{(1)}_{L_1}, \frac{1}{c_2}\epsilon^{(2)}_1, \cdots, \frac{1}{c_2}\epsilon^{(2)}_{L_2}) \\
&= \int_0^\infty \cdots \int_0^\infty \prod_{k=1}^{2} \Big(h^{(k)}_{0j_k}(t_k)\delta^{(k)}_{j_k} \Big) \exp\Big[- \sum_{k=1}^{2} \sum_{j=1}^{L_k} H^{(k)}_{0j}(t_k)\delta^{(k)}_j \Big] \\
&\quad \times dG(\delta^{(1)}_1, \cdots, \delta^{(1)}_{L_1}, \delta^{(2)}_1, \cdots, \delta^{(2)}_{L_2}),
\end{aligned}$$

which is same as $f_{j_1,j_2}(t_1, t_2)$, for $t_k > 0$, $j_k = 1, \cdots, L_k$ and $k = 1, 2$. Therefore, when both the baseline cause-specific hazards and the frailty distribution are unspecified, the model (2.1) is not identifiable.

2.6 The Likelihood

From Section 2.2, we know that the observed data is of the form (X_1, X_2, J_1, J_2). Suppose $(x_{i1}, x_{i2}, j_{i1}, j_{i2})$ denotes the realization of (X_1, X_2, J_1, J_2) for the ith pair in the sample, for $i = 1, \cdots, n$. In order to construct the likelihood function based on the above observations, let us consider a typical observation from a pair of individuals as (x_1, x_2, j_1, j_2), where $x_1 > 0$ and $x_2 > 0$ denote the monitoring times of the

two individuals with j_1 and j_2 being the corresponding causes of failure, for $j_k = 0, 1, \cdots, L_k$ and $k = 1, 2$. Recall that $j_k = 0$ means that the corresponding observation is right censored. We consider likelihood contributions for different combinations of (j_1, j_2), as given below.

Let $h(x_1, x_2)$ denote the joint density of the two monitoring times (X_1, X_2), which is not of interest and considered as nuisance. Then, for $j_k = 1, \cdots, L_k,\ k = 1, 2$, the likelihood contribution is proportional to
$P\Big[T_1 \leq X_1 = x_1, T_2 \leq X_2 = x_2, J_1 = j_1, J_2 = j_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G\Big] h(x_1, x_2)$, which is given by

$$\begin{aligned} f^*(x_1, x_2, j_1, j_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) &= \left[\int_0^{x_1} \int_0^{x_2} f_{j_1 j_2}(u_1, u_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) du_1 du_2 \right] h(x_1, x_2) \\ &\propto F_{j_1 j_2}(x_1, x_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G). \end{aligned}$$

Also, for $j_1 = 0$ and $j_2 \in \{1, \cdots, L_2\}$, we have $\{T_1 > X_1 = x_1,\ T_2 \leq X_2 = x_2, J_2 = j_2\}$. Then, the likelihood contribution is proportional to

$$\begin{aligned} f^*(x_1, x_2, j_1, j_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) &= \left[\int_0^{x_2} f_{j_2}^{(2)}(u_2; \boldsymbol{h_0^{(2)}}, G) du_2 - \right. \\ &\quad \left. \int_0^{x_1} \int_0^{x_2} \sum_{j=1}^{L_1} f_{j j_2}(u_1, u_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) du_1 du_2 \right] \times h(x_1, x_2) \\ &\propto \left[F_{j_2}^{(2)}(x_2; \boldsymbol{h_0^{(2)}}, G) - \sum_{j=1}^{L_1} F_{j j_2}(x_1, x_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \right]. \end{aligned}$$

Similarly, for $j_1 \in \{1, \cdots, L_1\}$ and $j_2 = 0$, we have $\{T_1 \leq X_1 = x_1, J_1 = j_1,\ T_2 > X_2 = x_2\}$ and the likelihood contribution is proportional to

$$\begin{aligned} f^*(x_1, x_2, j_1, j_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) &= \left[\int_0^{x_1} f_{j_1}^{(1)}(u_1; \boldsymbol{h_0^{(1)}}, G) du_1 - \right. \\ &\quad \left. \int_0^{x_1} \int_0^{x_2} \sum_{j=1}^{L_2} f_{j_1 j}(u_1, u_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) du_1 du_2 \right] \times h(x_1, x_2) \\ &\propto \left[F_{j_1}^{(1)}(x_1; \boldsymbol{h_0^{(1)}}, G) - \sum_{j=1}^{L_2} F_{j_1 j}(x_1, x_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \right]. \end{aligned}$$

Finally, when failure times of both the individuals are censored (that is, $j_1 = j_2 = 0$), we have $\{T_1 > X_1 = x_1,\ T_2 > X_2 = x_2\}$ and the likelihood contribution is proportional to

$$\begin{aligned} f^*(x_1, x_2, j_1, j_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) &= \left[\int_{x_1}^{\infty} \int_{x_2}^{\infty} f(u_1, u_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) du_1 du_2 \right] h(x_1, x_2) \\ &\propto S(x_1, x_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G). \end{aligned}$$

Suppose we have data from n independent pairs of individuals giving observations $(x_{i1}, x_{i2}, j_{i1}, j_{i2})$, for $i = 1, 2, \cdots, n$. Then, the likelihood function is given by

$$L(\boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \propto \prod_{i=1}^{n} f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G).$$

Note that this likelihood function is proportional to the product of n independent and identically distributed (iid) density functions like $f^*(x_1, x_2, j_1, j_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G)$ of the random vector (X_1, X_2, J_1, J_2). The corresponding dominating measure $\boldsymbol{\mu}$ is given by $\mathcal{L} \times \mathcal{L} \times \mathcal{C}_1 \times \mathcal{C}_2$, where $\mathcal{L}$ is the Lebesgue measure and $\mathcal{C}_k$ is the count measure on $\{0, 1, \cdots, L_k\}$ for $k = 1, 2$. This likelihood function can be rewritten as

$$\begin{aligned} L(\boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \propto \prod_{i=1}^{n} \prod_{j_{i1}=1}^{L_1} \prod_{j_{i2}=1}^{L_2} & \left\{ F_{j_{i1}j_{i2}}(x_{i1}, x_{i2}; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \right\}^{\delta_{j_{i1}}^{(1)} \delta_{j_{i2}}^{(2)}} \\ & \times \left\{ F_{j_{i1}}^{(1)}(x_{i1}; \boldsymbol{h_0^{(1)}}, G) - \sum_{j_{i2}=1}^{L_2} F_{j_{i1}j_{i2}}(x_{i1}, x_{i2}; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \right\}^{\delta_{j_{i1}}^{(1)}(1-\delta_i^{(2)})} \\ & \times \left\{ F_{j_{i2}}^{(2)}(x_{i2}; \boldsymbol{h_0^{(2)}}, G) - \sum_{j_{i1}=1}^{L_1} F_{j_{i1}j_{i2}}(x_{i1}, x_{i2}; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \right\}^{(1-\delta_i^{(1)})\delta_{j_{i2}}^{(2)}} \\ & \times \left\{ S(x_{i1}, x_{i2}; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G) \right\}^{(1-\delta_i^{(1)})(1-\delta_i^{(2)})}, \end{aligned}$$

where $\delta_{j_{ik}}^{(k)} = \mathbb{I}(T_{ik} \leq x_{ik}, J_k = j_{ik}), \delta_i^{(k)} = \sum_{j_{ik}=1}^{L_k} \delta_{j_{ik}}^{(k)}$ for $x_{ik} > 0, j_{ik} = 1, \cdots, L_k, k = 1, 2, i = 1, \cdots, n$. It is easy to see that all the terms (or, the contributions) in this likelihood function can be derived from the joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, G)$.

2.7 Concluding Remarks

In this chapter, we have discussed modeling of bivariate failure time (or, current status) data with competing risks in terms of baseline cause-specific hazard rates and general frailty distribution(s). When both the baseline cause-specific hazards and the frailty distribution(s) are unspecified, this model is proven to be non-identifiable in Chapter 5. Thus, in Chapters 3-5, we consider the other three configurations, namely, (1) both being in some parametric families, (2) the baseline cause-specific hazards being in a parametric family, while the frailty distribution is unspecified, and (4) the baseline cause-specific hazards are unspecified, while the frailty distribution being in a parametric family, respectively.

Although we suggest only four types of frailty in Section 2.5, there is possibility of more general frailty structure, possibly depending on failure times and failure types (See Gorfine and Hsu, 2011). Working with such general frailty structure may be challenging in terms of model interpretatbility, investigating identifiability, etc.

In this thesis, we consider bivariate current status data with competing risks. In general, one can consider multivariate, or clustered, current status data with competing risks with more than two, or even unequal, number of individuals within a cluster. Although the shared frailty model can be easily applied to such general data structure, one needs some suitable generalization of the correlated frailty model, possibly using some copula structure, as in Tran et al. (2020), in particular with unequal number of individuals within a cluster. Similarly, generalizing the shared cause-specific frailty for multivariate current status data is straightforward with the same set of competing risks for all the individuals within a cluster. Generalizing the correlated cause-specific frailty for multivariate current status data, on the other hand, may be challenging.

Chapter 3

Parametric Analysis of Bivariate Current Status Data with Competing Risks

3.1 Introduction

As discussed in Chapter 1, if we have current status data on failure of a pair of individuals or failure of two organs or components from a single individual, and each failure is subject to several competing risks, then we have bivariate current status data with competing risks. For example, we may have current status data on incidence times of first major disease from a couple (husband and wife) while the different major diseases may be treated as the competing risks. On the other hand, we have current status data on hearing loss in both right and left ears in which the hearing loss may be due to three different primary reasons, namely, SNHL, Conductive and Mixed (See Koley and Dewanji, 2022). Dependence between the two failure times is modeled through introduction of random frailty (Duchateau and Janssen, 2008 and Wienke, 2010). Modeling multivariate failure time with competing risks using frailty seems to be natural (See Gorfine and Hsu, 2011) because of its simple and appealing interpretation.

In this chapter, we consider a parametric class of cause-specific baseline hazard functions, as described in the next section, while different kinds of Gamma frailty distributions are assumed for the four types of frailty (See Sections 1.7 and 2.4). That is, we consider a fully parametric analysis of bivariate current status data with competing risks. We also investigate identifiability of the four models in terms of the corresponding joint sub-distribution function (See Section 2.5) for the four different frailty types. For each model, we derive the conditional cause-specific hazard ratio

measure (Bandeen-Roche and Liang, 2002) to study the nature of association (See Section 2.3). Two of our proposed frailty models allow for dependence at the level of competing risks, unlike the commonly used shared and correlated frailty models, incorporating a variety of dependence structures.

In Section 3.2, we describe the data and the four models with proofs of the corresponding identifiability results. We compute the conditional cause-specific hazard ratio as a function of the bivariate failure time (t_1, t_2) for all possible pairs of (j_1, j_2) in Section 3.3. In Section 3.4, we discuss maximum likelihood estimation of the associated parameters. We also study the asymptotic properties of maximum likelihood estimators (MLEs) of the parameters under some regularity conditions. Section 3.5 presents a detailed simulation study to investigate finite sample properties of the MLEs. Section 3.6 considers analysis of the hearing loss data using the four models and compares them using the AIC criterion. Section 3.7 ends with some concluding remarks.

3.2 Modeling and Identifiability

Recall from Section 2.2 that T_1 and T_2 denote the failure times of the first and the second individual of a pair, respectively, with J_1 and J_2 denoting the corresponding causes of failure. Also, X_1 and X_2 denote the monitoring times for the first and the second individual, respectively. For bivariate current status data with competing risks, our observed data consists of n independent realizations of the quadruplet (X_1, X_2, J_1, J_2).

As discussed in Section 3.1, we model dependence between (T_1, J_1) and (T_2, J_2) by means of different types of frailty variables, in particular, having some Gamma distributions. In our development of the model, in addition to the frailty distributions, we also require modeling of the cause-specific baseline hazard functions, as described below.

The jth cause-specific hazard function of the kth individual at time t_k is defined as

$$\lambda_j^{(k)}\big(t_k;\boldsymbol{\eta}\big) = \lim_{h\to 0+} \frac{P\Big[t_k < T_k < t_k + h, J_k = j | T_k \geq t_k;\boldsymbol{\eta}\Big]}{h} = \frac{f_j^{(k)}\big(t_k;\boldsymbol{\eta}\big)}{S^{(k)}\big(t_k;\boldsymbol{\eta}\big)},$$

where $f_j^{(k)}(t_k;\boldsymbol{\eta})$ is the sub-density function for the kth individual due to cause j and $S^{(k)}(t_k;\boldsymbol{\eta})$ is the survival function for the kth individual, as in Section 2.2, with associated parameter vector $\boldsymbol{\eta}$, for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$.

In general, let $\boldsymbol{\epsilon^{(k)}} = (\epsilon_1^{(k)}, \cdots, \epsilon_{L_k}^{(k)})$ denote the vector of frailty variables corresponding to different failure types for the kth individual, for $k = 1, 2$. Note that the parameter vector $\boldsymbol{\eta}$ consists of two sets of parameters, one for the baseline cause-specific hazard functions, denoted by $\boldsymbol{\xi}$, and the other for the frailty models, denoted by $\boldsymbol{\theta}$, so that $\boldsymbol{\eta} = (\boldsymbol{\xi}, \boldsymbol{\theta})$. Following Section 2.2, we consider the general multiplicative model for the conditional cause-specific hazard function of the kth individual, given the frailty vector $\boldsymbol{\epsilon^{(k)}}$, as

$$\lambda_j^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k;\boldsymbol{\xi})\epsilon_j^{(k)}, \tag{3.1}$$

for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$, where $h_{0j}^{(k)}(t_k;\boldsymbol{\xi})$ is the jth cause-specific baseline hazard function for the kth individual. The corresponding cumulative cause-specific baseline hazard function for the kth individual is $H_{0j}^{(k)}(t_k;\boldsymbol{\xi}) = \int_0^{t_k} h_{0j}^{(k)}(u;\boldsymbol{\xi})du$.

As already mentioned, we consider parametric models for the cause-specific baseline hazard functions $h_{0j}^{(k)}(t_k;\boldsymbol{\xi})$'s with associated parameter vector $\boldsymbol{\xi} \in \boldsymbol{\Xi}$. We assume these parametric cause-specific baseline hazard functions $h_{0j}^{(k)}(t_k;\boldsymbol{\xi})$, for $j = 1, \cdots, L_k,\ k = 1, 2$, to be identifiable within $\boldsymbol{\Xi}$ in the sense that if, for some $\boldsymbol{\xi}$ and $\tilde{\boldsymbol{\xi}}$ in $\boldsymbol{\Xi}$, the equality $h_{0j}^{(k)}(t_k;\boldsymbol{\xi}) = h_{0j}^{(k)}(t_k;\tilde{\boldsymbol{\xi}})$, for all $t_k > 0,\ j = 1, \cdots, L_k$, and $k = 1, 2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$.

For our purpose, we introduce a new parametric class of cause-specific baseline hazard functions $h_{0j}^{(k)}(t_k;\boldsymbol{\xi})$'s which includes hazards of $Exponential$, $Weibull$, $Gamma$, $Log-logistic$ and $Weibull-Gompertz$ distributions as special cases. This parametric class is particularly helpful as it contains various types of cause-specific baseline hazard functions such as increasing, decreasing and bathtub-shaped. We define this parametric class of $h_{0j}^{(k)}(t_k;\boldsymbol{\xi})$, suppressing the dependence on k and j for simplicity, as

$$h(t;\gamma,\alpha) = a(\gamma,\alpha)t^{\gamma-1}b(t;\gamma,\alpha), \tag{3.2}$$

for all $t > 0$, where $\alpha > 0$ and $\gamma > 0$ are scale and shape parameters, respectively, $a(\gamma,\alpha)$ is a positive-valued function of γ and α; it is also assumed that $a(\gamma,\alpha)$ is an one-one function in α for fixed γ. The function $b(t;\gamma,\alpha)$ is positive and such that

$\lim_{t\to 0+} b(t;\gamma,\alpha) = 1$. It can be easily proved, letting $\lim_{t\to 0+}$, that this parametric class (3.2) is identifiable in α and γ in the above-mentioned sense. Note that the parameters α and γ depend on k and j in general. Therefore, the parameter $\boldsymbol{\xi}$ is the vector of all these α's and γ's. The special cases are given below.

1. With $\gamma = 1, a(1,\alpha) = \alpha$ and $b(t,1,\alpha) = 1$ for all $t > 0$, we have the *Exponential* hazard with $h(t;\gamma,\alpha) = \alpha$, a constant.

2. With $a(\gamma,\alpha) = \gamma\alpha^{\gamma}$ and $b(t,\gamma,\alpha) = 1$, for all $t > 0$, we have $h(t;\gamma,\alpha) = \alpha^{\gamma}\gamma t^{\gamma-1}$, the *Weibull* hazard.

3. With $a(\gamma,\alpha) = \frac{\alpha^{\gamma}}{\Gamma(\gamma)}$ and $b(t,\gamma,\alpha) = \frac{\Gamma(\gamma)\exp(-t\alpha)}{\int_{t\alpha}^{\infty} e^{-y}y^{\gamma-1}dy}$, for all $t > 0$, we have the *Gamma* hazard of the form $h(t;\gamma,\alpha) = \alpha^{\gamma}\frac{t^{\gamma-1}\exp(-t\alpha)}{\int_{t\alpha}^{\infty} e^{-y}y^{\gamma-1}dy}$.

4. With $a(\gamma,\alpha) = \gamma\alpha^{\gamma}$ and $b(t,\gamma,\alpha) = \frac{1}{1+(t\alpha)^{\gamma}}$, for all $t > 0$, we have the $Log-logistic$ hazard given by $h(t;\gamma,\alpha) = \frac{\alpha^{\gamma}\gamma t^{\gamma-1}}{1+(t\alpha)^{\gamma}}$, which is a non-monotonic (upside bathtub shaped) function.

5. With $a(\gamma,\alpha) = \gamma\alpha^{\gamma}$ and $b(t,\gamma,\alpha) = e^{\alpha t}$ for all $t > 0$, we have the expression similar to that of *Weibull-Gompertz* hazard (dos Santos et al., 1995) given by $h(t;\gamma,\alpha) = \gamma\alpha^{\gamma}t^{\gamma-1}e^{\alpha t}$, which is a non-monotonic function.

The conditional survival function of kth individual, given $\boldsymbol{\epsilon^{(k)}}$, is

$$S^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon^{(k)}}) = \exp\left[-\sum_{j=1}^{L_k}\int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_j^{(k)}du_k\right]$$
$$= \exp\left[-\sum_{j=1}^{L_k}\epsilon_j^{(k)}H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\right]$$

for $t_k > 0, j = 1,\cdots,L_k$ and $k = 1,2$. The conditional jth sub-distribution function of the kth individual, given $\boldsymbol{\epsilon^{(k)}}$, is

$$F_j^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_j^{(k)}\exp\left[-\sum_{i=1}^{L_k}H_{0i}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_i^{(k)}\right]du_k,$$

for $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1,2$.

Using the conditional independence between T_1 and T_2, given the frailty vectors $\boldsymbol{\epsilon^{(1)}}$ and $\boldsymbol{\epsilon^{(2)}}$, the unconditional joint sub-distribution function $F_{j_1j_2}\big(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}\big)$ at time t_1 and t_2 due to causes j_1 and j_2, respectively, turns out to be

$$\int_0^\infty \cdots \int_0^\infty \left[\int_0^{t_1}\int_0^{t_2} \prod_{k=1}^{2} \left(h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_{j_k}^{(k)} \exp\left(-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_{j'}^{(k)} \right) \right) du_k \right] \times g(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}};\boldsymbol{\theta}) d\boldsymbol{\epsilon^{(1)}} d\boldsymbol{\epsilon^{(2)}},$$

for $t_k > 0,\ j_k = 1,\cdots,L_k$ and $k = 1,2$, where $g(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}};\boldsymbol{\theta})$ denotes the joint density of the frailty vectors $\boldsymbol{\epsilon^{(1)}}$ and $\boldsymbol{\epsilon^{(2)}}$ with the associated parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. In the following subsections, we shall consider different types of Gamma density for this frailty vector $(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}})$.

The corresponding unconditional joint sub-density function $f_{j_1j_2}\big(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}\big)$ is

$$\int_0^\infty \cdots \int_0^\infty \left[\prod_{k=1}^{2} \left(h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})\epsilon_{j_k}^{(k)} \exp\left(-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(t_k;\boldsymbol{\xi})\epsilon_{j'}^{(k)} \right) \right) \right] g(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}};\boldsymbol{\theta}) d\boldsymbol{\epsilon^{(1)}} d\boldsymbol{\epsilon^{(2)}}.$$

It is of interest to investigate identifiability of the model for bivariate failure time with competing risks (T_1,T_2,J_1,J_2) given by the joint sub-distribution function $F_{j_1j_2}\big(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}\big)$ with different Gamma frailty distributions. Note that these identifiability results are important while analyzing bivariate current status data with competing risks (X_1,X_2,J_1,J_2) since the likelihood contributions (See Section 2.6) can be obtained from the joint sub-distribution function.

Definition 3.1. The model (3.1) for bivariate failure time with competing risks is identifiable within $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$ if, for some $\boldsymbol{\xi},\ \tilde{\boldsymbol{\xi}} \in\ \boldsymbol{\Xi}$ and $\boldsymbol{\theta},\ \tilde{\boldsymbol{\theta}} \in\ \boldsymbol{\Theta}$, the equality

$$F_{j_1j_2}\big(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}\big) = F_{j_1j_2}\big(t_1,t_2;\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{\theta}}\big),$$

for all $t_k > 0,\ j_k = 1,\cdots,L_k$ and $k = 1,2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}$.

In Sections 3.2.1-3.2.4, we investigate identifiability of the model for the four different types of frailty, discussed in Section 2.4.

The conditional cause-specific hazard ratio function, denoted by $CR_{j_1j_2}(t_1,t_2)$, between failure of the first individual at time t_1 due to cause j_1 and failure of the

second individual at time t_2 due to cause j_2 is defined as (See Section 2.3)

$$CR_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) = \frac{S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})f_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})}{\left(\int_{t_1}^{\infty}\sum_{j=1}^{L_1} f_{jj_2}(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})du_1\right)\times\left(\int_{t_2}^{\infty}\sum_{j=1}^{L_2} f_{j_1j}(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})du_2\right)}$$

for $t_k > 0,\ j_k = 1,\cdots,L_k$ and $k = 1,2$. In the following subsections, in addition to investigating model identifiability, we also derive this cross-ratio function for the models arising from the four different Gamma frailty distributions.

3.2.1 Shared Gamma Frailty

The model is given by

$$\lambda_j^{(k)}\big(t_k;\boldsymbol{\xi}|\epsilon\big) = h_{0j}^{(k)}\big(t_k;\boldsymbol{\xi}\big)\epsilon, \tag{3.3}$$

for $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1,2$, where ϵ is the shared frailty variable.

We assume the shared frailty ϵ to follow Gamma$(\frac{1}{\sigma^2},\frac{1}{\sigma^2})$ distribution, for some $\sigma > 0$, with density given by

$$g(\epsilon;\sigma) = \frac{1}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}e^{-\frac{\epsilon}{\sigma^2}}\epsilon^{\frac{1}{\sigma^2}-1}, \tag{3.4}$$

having mean 1 and variance σ^2. The frailty parameter space is, therefore, given by $\boldsymbol{\Theta} = \{\sigma : \sigma > 0\}$.

Following the discussion in the beginning of this section, for this shared Gamma frailty model, conditional on the shared frailty ϵ, we have

$$S^{(k)}(t_k;\boldsymbol{\xi}|\epsilon) = \exp\Big[-\sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\epsilon\Big] = \exp\Big[-\epsilon H_0^{(k)}(t_k;\boldsymbol{\xi})\Big],$$

where $H_0^{(k)}(t_k;\boldsymbol{\xi}) = \sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k;\boldsymbol{\xi})$, and

$$F_j^{(k)}(t_k;\boldsymbol{\xi}|\epsilon) = \int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\epsilon\exp\Big[-\epsilon H_0^{(k)}(u_k;\boldsymbol{\xi})\Big]du_k,$$

for $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1,2$. Therefore, the unconditional joint sub-distribution function can be obtained, using conditional independence and then inte-

grating with respect to the Gamma density of ϵ, as

$$\begin{aligned}
F_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\sigma) &= \mathbb{E}\Big[F_{j_1}^{(1)}(t_1;\boldsymbol{\xi}|\epsilon)\times F_{j_2}^{(2)}(t_2;\boldsymbol{\xi}|\epsilon)\Big] \\
&= \int_0^{t_1}\int_0^{t_2}\prod_{k=1}^{2} h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi})\Big\{\int_0^{\infty}\exp\Big[-\epsilon\Big(\sum_{k=1}^{2}H_0^{(k)}(u_k;\boldsymbol{\xi})\Big)\Big]\epsilon^2 g(\epsilon;\sigma)d\epsilon\Big\} \\
&\qquad\qquad \times du_2du_1 \\
&= \int_0^{t_1}\int_0^{t_2}\frac{\prod_{k=1}^{2}h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi})}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\Big\{\int_0^{\infty}\exp\Big[-\epsilon\Big(\sum_{k=1}^{2}H_0^{(k)}(u_k;\boldsymbol{\xi})+\frac{1}{\sigma^2}\Big)\Big]\epsilon^{2+\frac{1}{\sigma^2}-1}d\epsilon\Big\} \\
&\qquad\qquad \times du_2du_1 \\
&= \int_0^{t_1}\int_0^{t_2}\frac{(1+\sigma^2)\Big(\prod_{k=1}^{2}h_{0,j_k}^{(k)}(u_k;\boldsymbol{\xi})\Big)du_2du_1}{\Big[1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(u_k;\boldsymbol{\xi})\Big]^{2+\frac{1}{\sigma^2}}},
\end{aligned}$$

with the corresponding unconditional joint sub-density function as

$$f_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\sigma)=\frac{\partial^2}{\partial t_1\partial t_2}F_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\sigma)=\frac{(1+\sigma^2)\prod_{k=1}^{2}h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})}{\Big[1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k;\boldsymbol{\xi})\Big]^{2+\frac{1}{\sigma^2}}},$$

for all $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$. The unconditional joint survival function is obtained as

$$S(t_1,t_2;\boldsymbol{\xi},\sigma)=\int_0^{\infty}\left(\prod_{k=1}^{2}S^{(k)}(t_k;\boldsymbol{\xi}|\epsilon)\right)g(\epsilon;\sigma)d\epsilon=\Big[1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k;\boldsymbol{\xi})\Big]^{-\frac{1}{\sigma^2}},$$

for all $t_1,t_2>0$.

For the shared frailty model, as per **Definition 3.1**, we have the following definition of identifiability.

Definition 3.2. The shared Gamma frailty model (3.3) is identifiable within $\boldsymbol{\Xi}\times\boldsymbol{\Theta}$ with $\boldsymbol{\Theta}=\{\sigma:\sigma>0\}$ if, for some $\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}\in\boldsymbol{\Xi}$ and $\sigma,\tilde{\sigma}\in\boldsymbol{\Theta}$, the equality of $F_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\sigma)$ and $F_{j_1j_2}(t_1,t_2;\tilde{\boldsymbol{\xi}},\tilde{\sigma}^2)$ with the above expressions, for all $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$, implies $\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}$ and $\sigma=\tilde{\sigma}$.

Note that, for the parametric family of cause-specific baseline hazard functions described in (3.2), the parameter vector $\boldsymbol{\xi}$ can be written as

$$\boldsymbol{\xi}=\{(\gamma_j^{(k)},\alpha_j^{(k)}),\ j=1,\cdots,L_k,\ k=1,2\}.$$

We can also write $\boldsymbol{\xi} = (\boldsymbol{\xi^{(1)}}, \boldsymbol{\xi^{(2)}})$, where $\boldsymbol{\xi^{(k)}} = \{(\gamma_j^{(k)}, \alpha_j^{(k)}),\ j = 1, \cdots, L_k\}$, for $k = 1, 2$. With this break-up of notation, it is to be noted that the unconditional jth sub-distribution function for the kth individual depends only on $\boldsymbol{\xi^{(k)}}$ and σ and is given by

$$\begin{aligned}
F_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \sigma) &= \int_0^\infty \int_0^{t_k} h_{0j}^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \epsilon \exp\Big[-\epsilon H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \Big] g(\epsilon; \sigma) du_k d\epsilon \\
&= \int_0^{t_k} \frac{h_{0j}^{(k)}(u_k; \boldsymbol{\xi^{(k)}})}{\sigma^{\frac{2}{\sigma^2}}.\Gamma(\frac{1}{\sigma^2})} \left\{ \int_0^\infty \exp\Big[-\epsilon\Big(H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) + \frac{1}{\sigma^2} \Big) \Big] \epsilon^{1+\frac{1}{\sigma^2}-1} d\epsilon \right\} du_k \\
&= \int_0^{t_k} h_{0j}^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \Big[1 + \sigma^2 H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \Big]^{-1-\frac{1}{\sigma^2}} du_k \\
&= \int_0^{t_k} \frac{a(\gamma_j^{(k)}, \alpha_j^{(k)}) u_k^{\gamma_j^{(k)}-1} b(u_k; \gamma_j^{(k)}, \alpha_j^{(k)}) du_k}{\big[1 + \sigma^2 H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \big]^{1+\frac{1}{\sigma^2}}}, \qquad (3.5)
\end{aligned}$$

for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$.

Theorem 3.1. The shared Gamma frailty model (3.3), with the cause-specific baseline hazard functions belonging to the family defined in (3.2), is identifiable within $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$.

Proof. The equality of the joint unconditional sub-distribution functions with $(\boldsymbol{\xi}, \sigma)$ and $(\tilde{\boldsymbol{\xi}}, \tilde{\sigma})$ implies equality of the corresponding sub-distribution functions $F_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \sigma)$ and $F_j^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{\sigma})$, for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Differentiating both sides with respect to t_k, for fixed k, we get equality of the corresponding unconditional sub-density functions to have

$$\frac{a(\gamma_j^{(k)}, \alpha_j^{(k)}) t_k^{\gamma_j^{(k)}-1} b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)})}{\Big[1 + \sigma^2 H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) \Big]^{1+\frac{1}{\sigma^2}}} = \frac{a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) t_k^{\tilde{\gamma}_j^{(k)}-1} b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)})}{\Big[1 + \tilde{\sigma}^2 H_0^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}) \Big]^{1+\frac{1}{\tilde{\sigma}^2}}},$$

for $k = 1, 2$. After rearranging the terms, we get

$$1 = t_k^{\tilde{\gamma}_j^{(k)} - \gamma_j^{(k)}} \frac{a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)})}{a(\gamma_j^{(k)}, \alpha_j^{(k)})} \times \frac{b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) \Big[1 + \sigma^2 H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) \Big]^{1+\frac{1}{\sigma^2}}}{b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) \Big[1 + \tilde{\sigma}^2 H_0^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}) \Big]^{1+\frac{1}{\tilde{\sigma}^2}}}. \qquad (3.6)$$

Now, letting $t_k \to 0+$, the right hand side of (3.6) tends to 0 or ∞ depending on

whether $\tilde{\gamma}_j^{(k)}$ is larger or smaller than $\gamma_j^{(k)}$, while the left hand side remains at 1, resulting in a contradiction. Hence, $\tilde{\gamma}_j^{(k)} = \gamma_j^{(k)}$ for all $j = 1, \cdots, L_k$ and $k = 1, 2$. Heckman and Singer (1984a) used a similar technique for univariate failure time model without competing risks. Now, putting this equality in (3.6) and again letting $t_k \to 0+$, we get $\tilde{\alpha}_j^{(k)} = \alpha_j^{(k)}$ for all $j = 1, \cdots, L_k$ and $k = 1, 2$, since $a(\gamma, \alpha)$ is a one-one function in α for fixed γ. So, we have $\tilde{\xi}^{(k)} = \xi^{(k)}$, for $k = 1, 2$. Therefore, from (3.6), we get

$$\left[1 + \sigma^2 H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\right]^{1+\frac{1}{\sigma^2}} = \left[1 + \tilde{\sigma}^2 H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\right]^{1+\frac{1}{\tilde{\sigma}^2}},$$

for all $t_k > 0$ and $k = 1, 2$.

Note that, for fixed k, $H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})$ is the cumulative hazard function of a continuous random variable. Therefore, there exists a unique value t_k^* for which $H_0^{(k)}(t_k^*; \boldsymbol{\xi^{(k)}}) = 1$. Taking limit as $t_k \to t_k^*$, we get

$$\left[1 + \sigma^2\right]^{1+\frac{1}{\sigma^2}} = \left[1 + \tilde{\sigma}^2\right]^{1+\frac{1}{\tilde{\sigma}^2}}.$$

It can be checked that the function $g(x) = (1+x)^{1+\frac{1}{x}}$, for $x > 0$, is a monotonically increasing function with $\lim_{x\to 0+} g(x) = e$ and hence is a one-to-one function. Therefore, the above equality implies $\sigma = \tilde{\sigma}$. Hence the shared Gamma frailty model (3.3) is identifiable.

□

3.2.2 Correlated Gamma Frailty

The correlated Gamma frailty model is given by

$$\lambda_j^{(k)}\big(t_k; \boldsymbol{\xi} | \epsilon^{(k)}\big) = h_{0j}^{(k)}\big(t_k; \boldsymbol{\xi}\big)\epsilon^{(k)}, \tag{3.7}$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$. The two correlated frailty variables $\epsilon^{(k)}$, $k = 1, 2$, are constructed in such a way that they have a common part accounting for the correlation between them and the other parts accounting for the individual effects (see Yashin et al., 1995). We write $\epsilon^{(k)} = \frac{\mu_0}{\mu_k} Y_0 + Y_k$, for $k = 1, 2$, where Y_i follows a Gamma(μ_i, κ_i) distribution, for $i = 0, 1, 2$. We also assume that the Y_i's are independent random variables with $\mu_0, \mu_1, \mu_2, \kappa_0, \kappa_1, \kappa_2 > 0$ satisfying $\mu_k = \kappa_0 + \kappa_k$ for $k = 1, 2$. These two relationships between the parameters ensure that the frailty

means are equal to unity. The association between $\epsilon^{(1)}$ and $\epsilon^{(2)}$ is modeled through the common random variable Y_0. Note that there are three independent parameters in this model. We consider the re-parametrization of these parameters given by

$$\sigma_1 = \frac{1}{\sqrt{\kappa_0 + \kappa_1}},\ \sigma_2 = \frac{1}{\sqrt{\kappa_0 + \kappa_2}},\ \rho = \frac{\kappa_0}{\sqrt{\kappa_0 + \kappa_1}\sqrt{\kappa_0 + \kappa_2}},$$

such that σ_k is the standard deviation of $\epsilon^{(k)}$ and ρ is the correlation coefficient between $\epsilon^{(1)}$ and $\epsilon^{(2)}$ satisfying the condition $\rho < \min\left(\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right)$. The frailty parameter space is, therefore, given by

$$\boldsymbol{\Theta} = \{\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho) : \sigma_1 > 0, \sigma_2 > 0, 0 < \rho < \min\big(\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\big)\}.$$

The joint density $g(\epsilon^{(1)}, \epsilon^{(2)}; \boldsymbol{\theta})$ of $(\epsilon^{(1)}, \epsilon^{(2)})$ can be written in terms of the three independent Gamma densities of $Y_0,\ Y_1$ and Y_2.

As before, given the frailty variable $\epsilon^{(k)}$, we have

$$S^{(k)}(t_k; \boldsymbol{\xi}|\epsilon^{(k)}) = \exp\Big[- \epsilon^{(k)} H_0^{(k)}(t_k; \boldsymbol{\xi})\Big] = \exp\Big[- \Big(\frac{\mu_0}{\mu_k} Y_0 + Y_k\Big) H_0^{(k)}(t_k; \boldsymbol{\xi})\Big]$$

and

$$F_j^{(k)}(t_k; \boldsymbol{\xi}|\epsilon^{(k)}) = \int_0^{t_k} h_{0j}^{(k)}(u_k; \boldsymbol{\xi}) \epsilon^{(k)} \exp\Big[- \epsilon^{(k)} H_0^{(k)}(u_k; \boldsymbol{\xi})\Big] du_k,$$

for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Then, the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ for the correlated Gamma frailty model is obtained as, using conditional independence and integrating with respect to the densities of $Y_0,\ Y_1$ and Y_2,

$$\int_0^{t_1}\int_0^{t_2} \frac{\prod_{k=1}^2 \sigma_k^2 h_{0j_k}^{(k)}(u_k; \boldsymbol{\xi})}{\Big(A(u_1, u_2; \boldsymbol{\xi}, \sigma_1, \sigma_2)\Big)^{\frac{\rho}{\sigma_1\sigma_2}} \prod_{k=1}^2 \Big(B_k(u_k; \boldsymbol{\xi}, \sigma_k)\Big)^{\left(\frac{1}{\sigma_k^2} - \frac{\rho}{\sigma_1\sigma_2}\right)}} \times \Bigg[\frac{\frac{\rho}{\sigma_1\sigma_2}(\frac{\rho}{\sigma_1\sigma_2} + 1)}{\Big(A(u_1, u_2; \boldsymbol{\xi}, \sigma_1, \sigma_2)\Big)^2} +$$
$$\sum_{k=1}^2 \frac{\frac{\rho}{\sigma_1\sigma_2}\left(\frac{1}{\sigma_k^2} - \frac{\rho}{\sigma_1\sigma_2}\right)}{A(u_1, u_2; \boldsymbol{\xi}, \sigma_1, \sigma_2) B_k(u_k; \boldsymbol{\xi}, \sigma_k)} + \prod_{k=1}^2 \frac{\left(\frac{1}{\sigma_k^2} - \frac{\rho}{\sigma_1\sigma_2}\right)}{B_k(u_k; \boldsymbol{\xi}, \sigma_k)} \Bigg] du_2 du_1,$$

where $A(u_1, u_2; \boldsymbol{\xi}, \sigma_1, \sigma_2) = 1 + \sum_{k=1}^2 \sigma_k^2 H_0^{(k)}(u_k; \boldsymbol{\xi})$, $B_k(u_k; \boldsymbol{\xi}, \sigma_k) = 1 + \sigma_k^2 H_0^{(k)}(u_k; \boldsymbol{\xi})$ for $u_k > 0$ and $k = 1, 2$.

The unconditional joint survival function is given by, using conditional independence and integrating with respect to the densities of $Y_0,\ Y_1$ and Y_2 as before,

$$S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) = (A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2))^{-\frac{\rho}{\sigma_1\sigma_2}} \prod_{k=1}^{2}\Big(B_k(t_k;\boldsymbol{\xi},\sigma_k)\Big)^{-\left(\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2}\right)},$$

for all $t_1,t_2>0$. The unconditional joint sub-density function $f_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})$ is given by

$$\prod_{k=1}^{2}\Big(\sigma_k^2 h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})\Big)\, S(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2)\Bigg[\frac{\frac{\rho}{\sigma_1\sigma_2}\left(\frac{\rho}{\sigma_1\sigma_2}+1\right)}{\Big(A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2)\Big)^2}+ \sum_{k=1}^{2}\frac{\frac{\rho}{\sigma_1\sigma_2}\left(\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}{A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2)B_k(t_k;\boldsymbol{\xi},\sigma_k)}+\prod_{k=1}^{2}\frac{\left(\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}{B_k(t_k;\boldsymbol{\xi},\sigma_k)}\Bigg],$$

for all $t_k,\ j_k=1,\cdots,L_k$ and $k=1,2$.

Definition 3.3. The correlated Gamma frailty model (3.7) is identifiable within families $\boldsymbol{\Xi}\times\boldsymbol{\Theta}$ with $\boldsymbol{\Theta}$ as described above in this subsection if, for some $\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}\in\boldsymbol{\Xi}$ and $(\sigma_1,\sigma_2,\rho),(\tilde{\sigma_1},\tilde{\sigma_2},\tilde{\rho})\in\boldsymbol{\Theta}$, the equality of $F_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2,\rho)$ and $F_{j_1j_2}(t_1,t_2;\tilde{\boldsymbol{\xi}},\tilde{\sigma_1},\tilde{\sigma_2},\tilde{\rho})$ with the above expressions, for all $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$, implies $\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}$ and $\sigma_k=\tilde{\sigma}_k$, for $k=1,2$ and $\rho=\tilde{\rho}$.

For $k=1,2$, integrating $F_j^{(k)}(t_k;\boldsymbol{\xi}|\epsilon^{(k)})$ with respect to the distribution of $\epsilon^{(k)}$, or Y_0 and Y_k, we get the corresponding unconditional sub-distribution function $F_j^{(k)}(t_k;\boldsymbol{\xi},\sigma_1,\sigma_2,\rho)$ as

$$\begin{aligned}
&\int_0^\infty\int_0^\infty\Bigg(\int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\Big(\frac{\mu_0}{\mu_k}y_0+y_k\Big)\exp\Bigg[-\Big(\frac{\mu_0}{\mu_k}y_0+y_k\Big)H_0^{(k)}(u_k;\boldsymbol{\xi})\Bigg]du_k\Bigg)\\
&\qquad\qquad\times\frac{\mu_0^{\kappa_0}}{\Gamma(\kappa_0)}\frac{\mu_k^{\kappa_k}}{\Gamma(\kappa_k)}\exp\left[-\mu_0y_0-\mu_ky_k\right]y_0^{\kappa_0-1}y_k^{\kappa_k-1}dy_0dy_k\\
&=\int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})[I_{1j}(u_k)+I_{2j}(u_k)]du_k,
\end{aligned}$$

say, where

$$I_{1j}(u_k) = \frac{\mu_0}{\mu_k}\frac{\mu_0^{\kappa_0}}{\Gamma(\kappa_0)}\frac{\mu_k^{\kappa_k}}{\Gamma(\kappa_k)}\int_0^\infty\int_0^\infty \exp\Big[-\mu_0 y_0\Big(\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big)-\mu_k y_k\Big(\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big)\Big] \times y_0^{\kappa_0+1-1}y_k^{\kappa_k-1}dy_0dy_k$$
$$= \frac{\kappa_0}{\mu_k}\Big[\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big]^{-\kappa_0-\kappa_k-1} = \frac{\kappa_0}{\mu_k}\Big[\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big]^{-\mu_k-1}$$

and

$$I_{2j}(u_k) = \frac{\mu_0^{\kappa_0}}{\Gamma(\kappa_0)}\frac{\mu_k^{\kappa_k}}{\Gamma(\kappa_k)}\int_0^\infty\int_0^\infty \exp\Big[-\mu_0 y_0\Big(\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big)-\mu_k y_k\Big(\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big)\Big] \times y_0^{\kappa_0-1}y_k^{\kappa_k+1-1}dy_0dy_k$$
$$= \frac{\kappa_k}{\mu_k}\Big[\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big]^{-\kappa_0-\kappa_k-1} = \frac{\kappa_k}{\mu_k}\Big[\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big]^{-\mu_k-1}.$$

Note that this unconditional sub-distribution function depends only on $\xi^{(k)}$ and σ_k and is given by

$$F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},\sigma_k) = \int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\left(\frac{\kappa_0+\kappa_k}{\mu_k}\Big[\frac{H_0^{(k)}(u_k;\boldsymbol{\xi})}{\mu_k}+1\Big]^{-\mu_k-1}\right)du_k$$
$$= \int_0^{t_k}\frac{a(\gamma_j^{(k)},\alpha_j^{(k)})u_k^{\gamma_j^{(k)}-1}b(u_k;\gamma_j^{(k)},\alpha_j^{(k)})du_k}{\big[1+\sigma_k^2 H_0^{(k)}(u_k;\boldsymbol{\xi^{(k)}})\big]^{1+\frac{1}{\sigma_k^2}}},$$

for $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1, 2$. This expression is similar to (3.5), the unconditional sub-distribution function for the shared Gamma frailty model (3.3) with σ replaced by σ_k.

Theorem 3.2. The correlated Gamma frailty model (3.7), with the cause-specific baseline hazard functions belonging to the family defined in (3.2), is identifiable within $\boldsymbol{\Xi}\times\boldsymbol{\Theta}$.

Proof. The equality of unconditional joint sub-distribution functions also implies the equality of unconditional marginal sub-distribution functions evaluated at $(\boldsymbol{\xi},\boldsymbol{\theta})$ and $(\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{\theta}})$. Since the unconditional marginal sub-distribution functions have the same form as those of the shared Gamma frailty model, following the same technique as

that was used to prove **Theorem 3.1**, we have

$$\xi_j^{(k)} = \tilde{\xi}_j^{(k)} \quad \text{and} \quad \sigma_k = \tilde{\sigma}_k,$$

for $j = 1, \cdots, L_k,\ k = 1, 2$.

Next, using the equality of unconditional joint survival functions $S(t_1, t_2; \boldsymbol{\xi}, \sigma_1, \sigma_2, \rho)$ and $S(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\rho})$, along with $\xi_j^{(k)} = \tilde{\xi}_j^{(k)}$ and $\sigma_k = \tilde{\sigma}_k$ for $j = 1, \cdots, L_k,\ k = 1, 2$, we have

$$\begin{aligned}\big[1 + \sum_{k=1}^{2} \sigma_k^2 H_0^{(k)}(t_k; \boldsymbol{\xi})\big]^{-\frac{\rho}{\sigma_1\sigma_2}} \prod_{k=1}^{2} \big[1 + \sigma_k^2 H_0^{(k)}(t_k; \boldsymbol{\xi})\big]^{\left(\frac{\rho}{\sigma_1\sigma_2} - \frac{1}{\sigma_k^2}\right)} \\ = \big[1 + \sum_{k=1}^{2} \sigma_k^2 H_0^{(k)}(t_k; \boldsymbol{\xi})\big]^{-\frac{\tilde{\rho}}{\sigma_1\sigma_2}} \prod_{k=1}^{2} \big[1 + \sigma_k^2 H_0^{(k)}(t_k; \boldsymbol{\xi})\big]^{\left(\frac{\tilde{\rho}}{\sigma_1\sigma_2} - \frac{1}{\sigma_k^2}\right)},\end{aligned}$$

for all $t_1, t_2 > 0$. This implies

$$\left[\frac{\prod_{k=1}^{2}\Big(1 + \sigma_k^2 H_0^{(k)}(t_k; \boldsymbol{\xi})\Big)}{1 + \sum_{k=1}^{2} \sigma_k^2 H_0^{(k)}(t_k; \boldsymbol{\xi})}\right]^{\frac{\rho - \tilde{\rho}}{\sigma_1\sigma_2}} = 1,$$

for all $t_1, t_2 > 0$. Hence, we get $\rho = \tilde{\rho}$. Therefore, the correlated Gamma frailty model (3.7) is identifiable. □

3.2.3 Shared Cause-specific Gamma Frailty

The shared cause-specific Gamma frailty model is given by

$$\lambda_j^{(k)}\big(t_k; \boldsymbol{\xi} | \epsilon_j\big) = h_{0j}^{(k)}\big(t_k; \boldsymbol{\xi}\big)\epsilon_j \tag{3.8}$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. We write $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_L)$ and assume that the ϵ_j's independently follow Gamma$(\frac{1}{\sigma_j^2}, \frac{1}{\sigma_j^2})$ distribution with $\sigma_j > 0$, so that $\mathbb{E}(\epsilon_j)$ is unity with density similar to that given in (3.4), for $j = 1, \cdots, L$. The frailty parameter space is, therefore, given by

$$\boldsymbol{\Theta} = \{\boldsymbol{\theta} = (\sigma_1, \cdots, \sigma_L) : \sigma_j > 0, j = 1, 2, \cdots, L\}.$$

As before, conditional on $\boldsymbol{\epsilon}$, the survival function of kth individual is

$$S^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon}) = \exp\left[-\sum_{j=1}^{L} H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\epsilon_j\right]$$

and the jth sub-distribution function of kth individual is

$$F_j^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon}) = \int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_j \exp\left[-\sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_{j'}\right] du_k,$$

for all $t_k > 0,\ j = 1, 2, \cdots, L$ and $k = 1, 2$. The unconditional joint sub-distribution function, $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$, is given by

$$\int_0^{\infty}\cdots\int_0^{\infty}\left[\int_0^{t_1}\int_0^{t_2}\left(\prod_{k=1}^{2} h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_{j_k}\right)\times\exp\left(-\sum_{j=1}^{L}\epsilon_j\sum_{k=1}^{2} H_{0j}^{(k)}(u_k;\boldsymbol{\xi})\right)du_2du_1\right]g(\boldsymbol{\epsilon};\boldsymbol{\theta})d\boldsymbol{\epsilon},$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, where $g(\boldsymbol{\epsilon};\boldsymbol{\theta})$ is the joint density of L independent Gamma variables $\epsilon_1, \cdots, \epsilon_L$. The expression of $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ simplifies to

$$\begin{aligned} F_{jj}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) = (1+\sigma_j^2)\int_0^{t_1}\int_0^{t_2}\left(\prod_{k=1}^{2} h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\right)\times\left[1+\sigma_j^2\sum_{k=1}^{2} H_{0j}^{(k)}(u_k;\boldsymbol{\xi})\right]^{-2} \\ \times\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2\sum_{k=1}^{2} H_{0j'}^{(k)}(u_k;\boldsymbol{\xi})\right]^{-\frac{1}{\sigma_{j'}^2}} du_2du_1, \end{aligned}$$

for $j_1 = j_2 = j$ (say), and

$$\begin{aligned} F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) = \int_0^{t_1}\int_0^{t_2}\prod_{k=1}^{2}\left(h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi})\right)\times\prod_{k=1}^{2}\left[1+\sigma_{j_k}^2\left(H_{0j_k}^{(1)}(u_1;\boldsymbol{\xi})+H_{0j_k}^{(2)}(u_2;\boldsymbol{\xi})\right)\right]^{-1} \\ \times\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2\sum_{k=1}^{2} H_{0j'}^{(k)}(u_k;\boldsymbol{\xi})\right]^{-\frac{1}{\sigma_{j'}^2}} du_2du_1, \end{aligned}$$

for $j_1 \neq j_2$ and $t_k > 0,\ j_k = 1, 2, \cdots, L,\ k = 1, 2$. Then, the joint sub-density function

$f_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ is given by

$$f_{jj}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) = (1+\sigma_j^2)\left(\prod_{k=1}^{2} h_{0j}^{(k)}(t_k; \boldsymbol{\xi})\right) \times \left[1 + \sigma_j^2 \sum_{k=1}^{2} H_{0j}^{(k)}(t_k; \boldsymbol{\xi})\right]^{-2} \times \prod_{j'=1}^{L} \left[1 + \sigma_{j'}^2 \sum_{k=1}^{2} H_{0j'}^{(k)}(t_k; \boldsymbol{\xi})\right]^{-\frac{1}{\sigma_{j'}^2}},$$

for $j_1 = j_2 = j$ (say), and

$$f_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) = \left(\prod_{k=1}^{2} h_{0j_k}^{(k)}(t_k; \boldsymbol{\xi})\right) \times \prod_{k=1}^{2} \left[1 + \sigma_{j_k}^2 \left(H_{0j_k}^{(1)}(t_1; \boldsymbol{\xi}) + H_{0j_k}^{(2)}(t_2; \boldsymbol{\xi})\right)\right]^{-1} \times \prod_{j'=1}^{L} \left[1 + \sigma_{j'}^2 \sum_{k=1}^{2} H_{0j'}^{(k)}(t_k; \boldsymbol{\xi})\right]^{-\frac{1}{\sigma_{j'}^2}},$$

for $j_1 \neq j_2$ and $t_k > 0,\ j_k = 1, \cdots, L,\ k = 1, 2$. The unconditional joint survival function is given by

$$\begin{aligned} S(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) &= \int_0^\infty \cdots \int_0^\infty \exp\left[-\sum_{j=1}^{L} \epsilon_j \sum_{k=1}^{2} H_{0j}^{(k)}(t_k; \boldsymbol{\xi})\right] g(\boldsymbol{\epsilon}; \boldsymbol{\theta}) d\boldsymbol{\epsilon} \\ &= \prod_{j=1}^{L} \left[1 + \sigma_j^2 \sum_{k=1}^{2} H_{0j}^{(k)}(t_k; \boldsymbol{\xi})\right]^{-\frac{1}{\sigma_j^2}}, \end{aligned}$$

for all $t_1, t_2 > 0$. Now, it is easy to check that the unconditional survival function of the kth individual, by letting $t_{k'} \to 0$ for $k' \neq k$, is given by

$$S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \boldsymbol{\theta}) = \prod_{j=1}^{L} \left[1 + \sigma_j^2 H_{0j}^{(k)}(t_k; \gamma_j^{(k)}, \alpha_j^{(k)})\right]^{-\frac{1}{\sigma_j^2}},$$

depending only on $\boldsymbol{\xi^{(k)}}$ and $\boldsymbol{\theta}$.

Definition 3.4. The shared cause-specific Gamma frailty model (3.8) is identifiable within families $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$ with $\boldsymbol{\Theta}$ as described above in this subsection if, for some $\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}} \in \boldsymbol{\Xi}$ and $\boldsymbol{\theta} = (\sigma_1, \cdots, \sigma_L), \tilde{\boldsymbol{\theta}} = (\tilde{\sigma_1}, \cdots, \tilde{\sigma_L}) \in \boldsymbol{\Theta}$, the equality of $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \sigma_1, \cdots, \sigma_L)$ and $F_{j_1 j_2}(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{\sigma_1}, \cdots, \tilde{\sigma_L})$ with the above expressions, for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $\sigma_k = \tilde{\sigma}_k$, for $k = 1, \cdots, L$.

Note that, as in Section 3.2.1, the unconditional jth sub-distribution function $F_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \sigma_j)$ for the kth individual depends only on $\boldsymbol{\xi^{(k)}}$ and σ_j and is given by

the same expression as that in (3.5), with σ replaced by σ_j, for $j = 1, \cdots, L$ and $k = 1, 2$.

Theorem 3.3. The shared cause-specific Gamma frailty model (3.8), with the cause-specific baseline hazard functions belonging to the family defined in (3.2), is identifiable within $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$.

Proof. The equality of the joint unconditional sub-distribution functions with $(\boldsymbol{\xi}, \boldsymbol{\theta})$ and $(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\theta}})$ implies equality of the corresponding unconditional sub-density functions (See Section 3.2.1) $f_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \boldsymbol{\theta})$ and $f_j^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{\boldsymbol{\theta}})$ for $t_k > 0, j = 1, \cdots, L$ and $k = 1, 2$. This gives, after some rearrangements as in (3.6),

$$t_k^{\gamma_j^{(k)} - \tilde{\gamma}_j^{(k)}} = \frac{a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) \left[1 + \sigma_j^2 H_{0j}^{(k)}(t_k, \gamma_j^{(k)}, \alpha_j^{(k)})\right] \prod_{j'=1}^{L} \left[1 + \sigma_{j'}^2 H_{0j'}^{(k)}(t_k, \gamma_{j'}^{(k)}, \alpha_{j'}^{(k)})\right]^{\frac{1}{\sigma_{j'}^2}}}{a(\gamma_j^{(k)}, \alpha_j^{(k)}) b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) \left[1 + \tilde{\sigma}_j^2 H_{0j}^{(k)}(t_k, \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)})\right] \prod_{j'=1}^{L} \left[1 + \tilde{\sigma}_{j'}^2 H_{0j'}^{(k)}(t_k, \tilde{\gamma}_{j'}^{(k)}, \tilde{\alpha}_{j'}^{(k)})\right]^{\frac{1}{\tilde{\sigma}_{j'}^2}}}, \tag{3.9}$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Then, using similar arguments as those in the proof of **Theorem 3.1**, one can show

$$\gamma_j^{(k)} = \tilde{\gamma}_j^{(k)} \text{ and } \alpha_j^{(k)} = \tilde{\alpha}_j^{(k)} \text{ for } j = 1, \cdots, L \text{ and } k = 1, 2.$$

Putting these equalities in (3.9), we get

$$\frac{1 + \sigma_j^2 H_{0j}^{(k)}(t_k, \gamma_j^{(k)}, \alpha_j^{(k)})}{1 + \tilde{\sigma}_j^2 H_{0j}^{(k)}(t_k, \gamma_j^{(k)}, \alpha_j^{(k)})} = \frac{\prod_{j'=1}^{L} \left[1 + \tilde{\sigma}_{j'}^2 H_{0j'}^{(k)}(t_k, \gamma_{j'}^{(k)}, \alpha_{j'}^{(k)})\right]^{\frac{1}{\tilde{\sigma}_{j'}^2}}}{\prod_{j'=1}^{L} \left[1 + \sigma_{j'}^2 H_{0j'}^{(k)}(t_k, \gamma_{j'}^{(k)}, \alpha_{j'}^{(k)})\right]^{\frac{1}{\sigma_{j'}^2}}} = \frac{S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \tilde{\boldsymbol{\theta}})}{S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \boldsymbol{\theta})}, \tag{3.10}$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$.

Now, the equality of the joint unconditional sub-distribution functions with $(\boldsymbol{\xi}, \boldsymbol{\theta})$ and $(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\theta}})$ also implies the equality of the corresponding unconditional survival functions for the kth individual; that is, $S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \tilde{\boldsymbol{\theta}}) = S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \boldsymbol{\theta})$. Using this equality in (3.10) yields $\sigma_j = \tilde{\sigma}_j$, for $j = 1, \cdots, L$. This completes the proof. □

3.2.4 Correlated Cause-specific Gamma Frailty

The correlated cause-specific Gamma frailty model is given by

$$\lambda_j^{(k)}\big(t_k;\boldsymbol{\xi}|\epsilon_j^{(k)}\big) = h_{0j}^{(k)}\big(t_k;\boldsymbol{\xi}\big)\epsilon_j^{(k)}, \tag{3.11}$$

for all $t_k > 0$ and $j = 1, \cdots, L,\ k = 1, 2$. As in Section 3.2.2, we assume that the pair $(\epsilon_j^{(1)}, \epsilon_j^{(2)})$ follows a correlated Gamma distribution, as defined therein but with the corresponding set of parameters $(\sigma_{1j}, \sigma_{2j}, \rho_j)$ depending on j. We also assume that the L pairs of Gamma frailty variables $(\epsilon_j^{(1)}, \epsilon_j^{(2)})$, for $j = 1, \cdots, L$, are independent. Therefore, the frailty parameter space is given by

$$\boldsymbol{\Theta} = \{\boldsymbol{\theta} = ((\sigma_{1j}, \sigma_{2j}, \rho_j),\ j = 1, \cdots, L)\},$$

such that $\sigma_{1j} > 0, \sigma_{2j} > 0, 0 < \rho_j < \min\big(\frac{\sigma_{1j}}{\sigma_{2j}}, \frac{\sigma_{2j}}{\sigma_{1j}}\big)$, for $j = 1, \cdots, L$. As in subsection 3.2.2, the joint density $g(\epsilon_j^{(1)}, \epsilon_j^{(2)}; \boldsymbol{\theta})$ can be written in terms of three independent Gamma densities of $Y_{0j},\ Y_{1j}$ and Y_{2j} with set of parameters $(\sigma_{1j}, \sigma_{2j}, \rho_j)$, for $j = 1, \cdots, L$.

We have, conditional on frailty vector $\boldsymbol{\epsilon^{(k)}} = (\epsilon_1^{(k)}, \cdots, \epsilon_L^{(k)})$,

$$S^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon^{(k)}}) = \exp\Big[-\sum_{j=1}^{L} H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\epsilon_j^{(k)}\Big] = \exp\Big[-\sum_{j=1}^{L} H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\left(\frac{\mu_{0j}}{\mu_{kj}}Y_{0j} + Y_{kj}\right)\Big]$$

and

$$F_j^{(k)}(t_k;\boldsymbol{\xi}|\boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k} h_{0j}^{(k)}\big(u_k;\boldsymbol{\xi}\big)\epsilon_j^{(k)} \exp\Big[-\sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k;\boldsymbol{\xi})\epsilon_{j'}^{(k)}\Big] du_k,$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$.

The unconditional joint survival function $S(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ is given by

$$\begin{aligned}
&\prod_{j=1}^{L}\Big[1 + \sum_{k=1}^{2}\sigma_{kj}^2 H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\Big]^{-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}} \times \prod_{k=1}^{2}\prod_{j=1}^{L}\Big[1 + \sigma_{kj}^2 H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\Big]^{\left(\frac{\rho_j}{\sigma_{1j}\sigma_{2j}} - \frac{1}{\sigma_{kj}^2}\right)} \\
&= \prod_{j=1}^{L}\Big[A_j(t_1, t_2;\boldsymbol{\xi}, \sigma_{1j}, \sigma_{2j})\Big]^{-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}} \prod_{k=1}^{2}\prod_{j=1}^{L}\Big[B_{kj}(t_k;\boldsymbol{\xi}, \sigma_{kj})\Big]^{\left(\frac{\rho_j}{\sigma_{1j}\sigma_{2j}} - \frac{1}{\sigma_{kj}^2}\right)} \quad \text{(say)},
\end{aligned}$$

where $A_j(t_1, t_2;\boldsymbol{\xi}, \sigma_{1j}, \sigma_{2j}) = 1 + \sum_{k=1}^{2}\sigma_{kj}^2 H_{0j}^{(k)}(t_k;\boldsymbol{\xi})$ and $B_{kj}(t_k;\boldsymbol{\xi}, \sigma_{kj}) = 1 + \sigma_{kj}^2 H_{0j}^{(k)}(t_k;\boldsymbol{\xi})$,

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$.

The unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ is obtained as

$$\int_0^{t_1}\int_0^{t_2} \prod_{k=1}^{2} \left(\sigma_{kj}^2 h_{0j}^{(k)}(u_k;\boldsymbol{\xi})\right) S(u_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta}) \Bigg[\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(A_j(u_1,u_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})\right)^2} + \prod_{k=1}^{2} \frac{\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{B_{kj}(u_k;\boldsymbol{\xi},\sigma_{kj})} + \sum_{k=1}^{2} \frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{A_j(u_1,u_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})B_{kj}(u_k;\boldsymbol{\xi},\sigma_{kj})} \Bigg] du_2 du_1,$$

for $j_1 = j_2 = j$ (say), and

$$\int_0^{t_1}\int_0^{t_2} \prod_{k=1}^{2} \left(\sigma_{kj_k}^2 h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi})\right) S(u_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta}) \times \Bigg[\prod_{k=1}^{2} \frac{\frac{\rho_{j_k}}{\sigma_{1j_k}\sigma_{2j_k}}}{A_{j_k}(u_1,u_2;\boldsymbol{\xi},\sigma_{1j_k},\sigma_{2j_k})} + \prod_{k=1}^{2} \frac{\frac{1}{\sigma_{kj_k}^2}-\frac{\rho_{j_k}}{\sigma_{1j_k}\sigma_{2j_k}}}{B_{kj_k}(u_k;\boldsymbol{\xi},\sigma_{kj_k})} + \frac{\frac{\rho_{j_2}}{\sigma_{1j_2}\sigma_{2j_2}}\left(\frac{1}{\sigma_{1j_1}^2}-\frac{\rho_{j_1}}{\sigma_{1j_1}\sigma_{2j_1}}\right)}{\left(A_{j_2}(u_1,u_2;\boldsymbol{\xi},\sigma_{1j_2},\sigma_{2j_2})\right)\left(B_{1j_1}(u_1;\boldsymbol{\xi},\sigma_{1j_1})\right)} + \frac{\frac{\rho_{j_1}}{\sigma_{1j_1}\sigma_{2j_1}}\left(\frac{1}{\sigma_{2j_2}^2}-\frac{\rho_{j_2}}{\sigma_{1j_2}\sigma_{2j_2}}\right)}{\left(A_{j_1}(u_1,u_2;\boldsymbol{\xi},\sigma_{1j_1},\sigma_{2j_1})\right)\left(B_{2j_2}(u_2;\boldsymbol{\xi},\sigma_{2j_2})\right)} \Bigg] du_2 du_1,$$

for $j_1 \neq j_2$. Accordingly, the unconditional joint sub-density function $f_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ is given by

$$\prod_{k=1}^{2} \left(\sigma_{kj}^2 h_{0j}^{(k)}(t_k;\boldsymbol{\xi})\right) S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) \Bigg[\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})\right)^2} + \prod_{k=1}^{2} \frac{\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})} + \sum_{k=1}^{2} \frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})} \Bigg],$$

for $j_1 = j_2 = j$ (say), and

$$\prod_{k=1}^{2} \left(\sigma_{kj_k}^2 h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})\right) S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) \times \Bigg[\prod_{k=1}^{2} \frac{\frac{\rho_{j_k}}{\sigma_{1j_k}\sigma_{2j_k}}}{A_{j_k}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j_k},\sigma_{2j_k})} + \prod_{k=1}^{2} \frac{\frac{1}{\sigma_{kj_k}^2}-\frac{\rho_{j_k}}{\sigma_{1j_k}\sigma_{2j_k}}}{B_{kj_k}(t_k;\boldsymbol{\xi},\sigma_{kj_k})} + \frac{\frac{\rho_{j_2}}{\sigma_{1j_2}\sigma_{2j_2}}\left(\frac{1}{\sigma_{1j_1}^2}-\frac{\rho_{j_1}}{\sigma_{1j_1}\sigma_{2j_1}}\right)}{A_{j_2}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j_2},\sigma_{2j_2})B_{1j_1}(t_1;\boldsymbol{\xi},\sigma_{1j_1})} + \frac{\frac{\rho_{j_1}}{\sigma_{1j_1}\sigma_{2j_1}}\left(\frac{1}{\sigma_{2j_2}^2}-\frac{\rho_{j_2}}{\sigma_{1j_2}\sigma_{2j_2}}\right)}{A_{j_1}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j_1},\sigma_{2j_1})B_{2j_2}(t_2;\boldsymbol{\xi},\sigma_{2j_2})} \Bigg],$$

for $j_1 \neq j_2$.

Definition 3.5. The correlated cause-specific Gamma frailty model (3.11) is identifiable within the family $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$ with $\boldsymbol{\Theta}$ as described above in this subsection if, for some

$\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}} \in \boldsymbol{\Xi}$ and $\boldsymbol{\theta} = ((\sigma_{1j}, \sigma_{2j}, \rho_j),\ j = 1, \cdots, L)$, $\tilde{\boldsymbol{\theta}} = ((\tilde{\sigma}_{1j}, \tilde{\sigma}_{2j}, \tilde{\rho}_j),\ j = 1, \cdots, L) \in \boldsymbol{\Theta}$, the equality $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta}) = F_{j_1 j_2}(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\theta}})$ with the above expressions, for all $t_k > 0, j_k = 1, \cdots, L$ and $k = 1, 2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $\sigma_{kj} = \tilde{\sigma}_{kj},\ \rho_j = \tilde{\rho}_j$ for $j = 1, \cdots, L$ and $k = 1, 2$.

As in the previous subsetions, the unconditional jth sub-distribution function $F_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \sigma_{kj})$ for the kth individual depends only on $\boldsymbol{\xi^{(k)}}$ and σ_{kj} and is given by the same expression as that in (3.5), with σ replaced by σ_{kj}, for $j = 1, \cdots, L$ and $k = 1, 2$.

Theorem 3.4. The correlated cause-specific Gamma frailty model (3.11), with the cause-specific baseline hazard functions belonging to the family defined in (3.2), is identifiable within $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$.

Proof. Following the same techniques as those in the proof of **Theorem 3.3**, using the equality of the unconditional marginal sub-density functions $f_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, \sigma_{kj})$ and $f_j^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{(\boldsymbol{k})}, \tilde{\sigma}_{kj})$, one can prove that

$$\xi_j^{(k)} = \tilde{\xi}_j^{(k)} \text{ and } \sigma_{kj} = \tilde{\sigma}_{kj},$$

for all $j = 1, \cdots, L$ and $k = 1, 2$. With these equalities and the equality of unconditional joint sub-density functions $f_{jj}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ and $f_{jj}(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\theta}})$, for $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$, as implied by **Definition 3.5**, we get

$$\left[\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})\right)^2} + \sum_{k=1}^{2}\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}-\frac{1}{\sigma_{kj}^2}\right)}{A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})} + \prod_{k=1}^{2}\frac{\left(\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}-\frac{1}{\sigma_{kj}^2}\right)}{B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})}\right]\times$$
$$\prod_{j'=1}^{L}\left[A_{j'}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j'},\sigma_{2j'})\right]^{-\left(\frac{\rho_{j'}}{\sigma_{1j'}\sigma_{2j'}}\right)}\prod_{k=1}^{2}\prod_{j'=1}^{L}\left[B_{kj'}(t_k;\boldsymbol{\xi},\sigma_{1j'})\right]^{\left(\frac{\rho_{j'}}{\sigma_{1j'}\sigma_{2j'}}-\frac{1}{\sigma_{kj'}^2}\right)} =$$
$$\left[\frac{\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})\right)^2} + \sum_{k=1}^{2}\frac{\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\left(\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}-\frac{1}{\sigma_{kj}^2}\right)}{A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})} + \prod_{k=1}^{2}\frac{\left(\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}-\frac{1}{\sigma_{kj}^2}\right)}{B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})}\right]\times$$
$$\prod_{j'=1}^{L}\left[A_{j'}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j'},\sigma_{2j'})\right]^{-\left(\frac{\tilde{\rho}_{j'}}{\sigma_{1j'}\sigma_{2j'}}\right)}\prod_{k=1}^{2}\prod_{j'=1}^{L}\left[B_{kj'}(t_k;\boldsymbol{\xi},\sigma_{1j'})\right]^{\left(\frac{\tilde{\rho}_{j'}}{\sigma_{1j'}\sigma_{2j'}}-\frac{1}{\sigma_{kj'}^2}\right)}$$

for $t_k > 0, j = 1, \cdots, L$ and $k = 1, 2$.

Note that, as $t_k \to 0+$, for $k = 1, 2$, we have $A_j(t_1, t_2; \boldsymbol{\xi}, \sigma_{1j}, \sigma_{2j}) \to 1$, $B_{kj}(t_k; \boldsymbol{\xi}, \sigma_{1j}) \to$

1, for $k = 1, 2$. Hence, letting $t_k \to 0+$, for $k = 1, 2$, in the above equality, we get

$$\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big)+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big(\frac{1}{\sigma_{1j}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big)+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big(\frac{1}{\sigma_{2j}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big)+\Big(\frac{1}{\sigma_{1j}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big)\Big(\frac{1}{\sigma_{2j}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\Big)$$
$$=\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big(1+\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big)+\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big(\frac{1}{\sigma_{1j}^2}-\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big)+\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big(\frac{1}{\sigma_{2j}^2}-\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big)+\Big(\frac{1}{\sigma_{1j}^2}-\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big)\Big(\frac{1}{\sigma_{2j}^2}-\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}\Big)$$

for $j = 1, \cdots, L$. Therefore, both sides of the above equality being the same quadratic functions of in ρ_j and $\tilde{\rho}_j$, respectively, we get $\rho_j = \tilde{\rho}_j$ for $j = 1, \cdots, L$. This completes the proof. □

3.3 Conditional Cause-specific Hazard Ratio

In this section, we investigate the conditional cause-specific hazard ratio $CR_{j_1j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ (See Section 2.3) for the four different frailty distributions.

3.3.1 Shared Gamma Frailty

For the shared Gamma frailty model, the conditional cause-specific hazard ratio $CR_{j_1j_2}(t_1, t_2; \boldsymbol{\xi}, \sigma)$ is obtained as follows. First,

$$\int_{t_2}^{\infty}\sum_{j_2=1}^{L_2} f_{j_1j_2}(t_1, u_2; \boldsymbol{\xi}, \sigma)du_2 = \int_{t_2}^{\infty}\sum_{j_2=1}^{L_2}\frac{(1+\sigma^2)h_{0j_1}^{(1)}(t_1; \boldsymbol{\xi^{(1)}})h_{0j_2}^{(2)}(u_2; \boldsymbol{\xi^{(2)}})}{\Big[1+\sigma^2\Big(H_0^{(1)}(t_1; \boldsymbol{\xi^{(1)}})+H_0^{(2)}(u_2; \boldsymbol{\xi^{(2)}})\Big)\Big]^{2+\frac{1}{\sigma^2}}}du_2$$
$$=\frac{h_{0j_1}^{(1)}(t_1; \boldsymbol{\xi^{(1)}})}{\Big[1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\Big]^{1+\frac{1}{\sigma^2}}},$$

using the transformation $z = 1+\sigma^2\Big(H_0^{(1)}(t_1; \boldsymbol{\xi^{(1)}})+H_0^{(2)}(u_2; \boldsymbol{\xi^{(2)}})\Big)$, where $h_0^{(2)}(u_2; \boldsymbol{\xi^{(2)}}) = \sum_{j=1}^{L_2} h_{0j}^{(2)}(u_2; \boldsymbol{\xi^{(2)}})$, for $j_1 = 1, 2, \cdots, L_1$ and $t_1 > 0$. Similarly, we have

$$\int_{t_1}^{\infty}\sum_{j_1=1}^{L_1} f_{j_1j_2}(u_1, t_2; \boldsymbol{\xi}, \sigma)du_1 = \frac{h_{0j_2}^{(2)}(t_2; \boldsymbol{\xi^{(2)}})}{\Big[1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\Big]^{1+\frac{1}{\sigma^2}}},$$

for $j_2 = 1, 2, \cdots, L_2$ and $t_2 > 0$. Then, we have the expression for $CR_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \sigma)$ as

$$\frac{\frac{(1+\sigma^2)\prod_{k=1}^{2} h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi^{(k)}})}{\left[1+\sigma^2\sum_{k=1}^{2} H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right]^{2+\frac{1}{\sigma^2}}} \times \left[1+\sigma^2\sum_{k=1}^{2} H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right]^{-\frac{1}{\sigma^2}}}{\frac{h_{0j_1}^{(1)}(t_1;\boldsymbol{\xi^{(1)}})}{\left[1+\sigma^2\sum_{k=1}^{2} H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right]^{1+\frac{1}{\sigma^2}}} \times \frac{h_{0j_2}^{(2)}(t_2;\boldsymbol{\xi^{(2)}})}{\left[1+\sigma^2\sum_{k=1}^{2} H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right]^{1+\frac{1}{\sigma^2}}}} = 1+\sigma^2,$$

for $t_k > 0,\ j_k = 1, \cdots, L_k$ and $k = 1, 2$. Therefore, the failure at time t_1 due to cause j_1 for the first individual and the failure at time t_2 due to cause j_2 for the second individual are positively associated and this association measure is constant for any two times t_1, t_2 and two causes j_1, j_2 for the first and the second individual, respectively. Recall that σ^2 is the variance of the shared Gamma frailty variable ϵ. Hence, larger the variance of shared frailty, higher is the value of the association measure.

3.3.2 Correlated Gamma Frailty

In this case, the conditional cause-specific hazard ratio function $CR_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \sigma_1, \sigma_2, \rho)$ between occurrence of failure at time t_1 due to cause j_1 for the first individual and the occurrence of failure at time t_2 due to cause j_2 for the second individual is given by

$$\left[\frac{\left(\prod_{k=1}^{2}\sigma_k^2 h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})\right)\left(S(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2,\rho)\right)^2}{\left(\int_{t_2}^{\infty}\sum_{j=1}^{L_2} f_{j_1 j}(t_1,u_2;\boldsymbol{\xi},\sigma_1,\sigma_2,\rho)du_2\right)\times\left(\int_{t_1}^{\infty}\sum_{j=1}^{L_1} f_{j j_2}(u_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2,\rho)du_1\right)}\right]\times$$
$$\left[\frac{\frac{\rho}{\sigma_1\sigma_2}\left(\frac{\rho}{\sigma_1\sigma_2}+1\right)}{\left(A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2)\right)^2}+\sum_{k=1}^{2}\frac{\frac{\rho}{\sigma_1\sigma_2}\left(\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}{A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2)B_k(t_k;\boldsymbol{\xi},\sigma_k)}+\prod_{k=1}^{2}\frac{\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2}}{B_k(t_k;\boldsymbol{\xi},\sigma_k)}\right]$$

for $t_k > 0,\ j_k = 1, \cdots, L_k$ and $k = 1, 2$.

In case of constant cause-specific baseline hazard (that is, $h_{0j}^{(k)}(t_k;\boldsymbol{\xi}) = \alpha_j^{(k)}$, for $j = 1, \cdots, L_k$ and $k = 1, 2$), the above expression for $CR_{j_1 j_2}(t_1, t_2, \boldsymbol{\xi}, \sigma_1, \sigma_2, \rho)$ simplifies to

$$\frac{1}{\prod_{k=1}^{2}\sigma_k^2\alpha^{(k)}} \times \frac{(S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})^2)\, D(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})}{\int_{t_1}^{\infty} S(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})D(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})du_1 \times \int_{t_2}^{\infty} S(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})D(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})du_2}$$

where

$$D(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) = \left[\frac{\frac{\rho}{\sigma_1\sigma_2}(\frac{\rho}{\sigma_1\sigma_2}+1)}{(A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2))^2} + \prod_{k=1}^{2}\frac{\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2}}{B_k(t_k;\boldsymbol{\xi},\sigma_k)} + \sum_{k=1}^{2}\frac{\frac{\rho}{\sigma_1\sigma_2}(\frac{1}{\sigma_k^2}-\frac{\rho}{\sigma_1\sigma_2})}{A(t_1,t_2;\boldsymbol{\xi},\sigma_1,\sigma_2)B_k(t_k;\boldsymbol{\xi},\sigma_k)}\right]$$

and $\alpha^{(k)} = \sum_{j=1}^{L_k}\alpha_j^{(k)}$. It is difficult to find this function analytically, even to obtain a suitable lower bound. Note that this function is the same for all (j_1, j_2), indicating that the measure of association does not depend on the two causes of failure, although depending on the two failure times. We numerically evaluate this for constant cause-specific baseline hazards and different values of (t_1, t_2), as presented in Figure 3.1. We assume two competing risks for each individual with $\boldsymbol{\xi^{(1)}} = (\alpha_1^{(1)}, \alpha_2^{(1)}) = (0.2, 0.25)$ and $\boldsymbol{\xi^{(2)}} = (\alpha_1^{(2)}, \alpha_2^{(2)}) = (0.15, 0.1)$. The frailty parameter vector $\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho)$ is taken as $(0.95, 0.85, 0.8)$. In Figure 3.1, we give plots of the conditional cause-specific hazard ratio function $CR_{1,1}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ against t_1 for five different values of t_2 as $(0.05, 0.2, 0.5, 0.9, 2)$. For this particular set of model parameters, Figure 3.1 shows that the conditional cause-specific hazard ratio function is always greater than unity. We have computed this function for other parameter values and observed the same pattern (not reported here). Therefore, failure times of the two individuals due to some fixed causes seem to be positively associated.

3.3.3 Shared Cause-specific Gamma Frailty

In this case, the conditional cause-specific hazard ratio function $CR_{j_1j_2}(t_1, t_2; \boldsymbol{\xi}, \sigma_1, \cdots, \sigma_L)$ between occurrence of failure at time t_1 due to cause j_1 for the first individual and occurrence of failure at time t_2 due to cause j_2 for the second individual is given by

$$\frac{(1+\sigma_j^2)\left(\prod_{k=1}^{2} h_{0j}^{(k)}(t_k;\boldsymbol{\xi})\right) S^2(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) \times \left[1+\sigma_j^2\sum_{k=1}^{2}H_{0j}^{(k)}(t_k;\boldsymbol{\xi})\right]^{-2}}{\left[\int_{t_2}^{\infty}\sum_{j'=1}^{L} f_{jj'}(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})du_2\right] \times \left[\int_{t_1}^{\infty}\sum_{j'=1}^{L} f_{j'j}(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})du_1\right]},$$

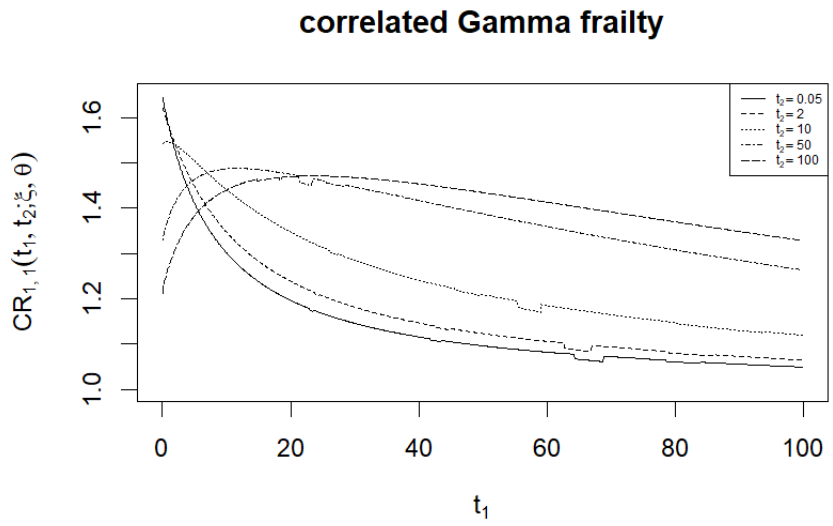


Figure 3.1: The conditional cause-specific hazard ratio function corresponding to the correlated Gamma frailty model.

for $j_1 = j_2 = j$, and

$$\frac{\left(\prod_{k=1}^{2} h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})\right) S^2(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta}) \prod_{k=1}^{2}\left[1+\sigma_{j_k}^2\left(H_{0j_k}^{(1)}(t_1;\boldsymbol{\xi})+H_{0j_k}^{(2)}(t_2;\boldsymbol{\xi})\right)\right]^{-1}}{\left[\int_{t_2}^{\infty}\sum_{j'=1}^{L} f_{j_1 j'}(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})du_2\right] \times \left[\int_{t_1}^{\infty}\sum_{j'=1}^{L} f_{j' j_2}(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})du_1\right]},$$

for $j_1 \neq j_2$.

It is difficult to find a closed form of this conditional cause-specific hazard ratio function, even to obtain a suitable lower bound. Thus, we numerically evaluate this function with two competing risks for different values of (t_1, t_2) and (j_1, j_2), as presented in Figure 3.2. In this case also, we consider the Exponential type cause-specific hazard rates with the same set of parameters as in subsection 3.2.2. We have taken (σ_1, σ_2) as $(1.65, 0.45)$. We again plot the conditional cause-specific hazard ratio function against t_1 for the five different values of t_2, as in Figure 3.1. Note that, in this case, this function depends on (j_1, j_2). The plots for the four different combinations of (j_1, j_2) are presented in the four panels $3.2a - 3.2d$. We see from Figure 3.2 that the conditional cause-specific hazard ratio function is greater than unity in each panel. Therefore, failure times of the two individuals due to any particular combination of (j_1, j_2) seem to be positively associated.

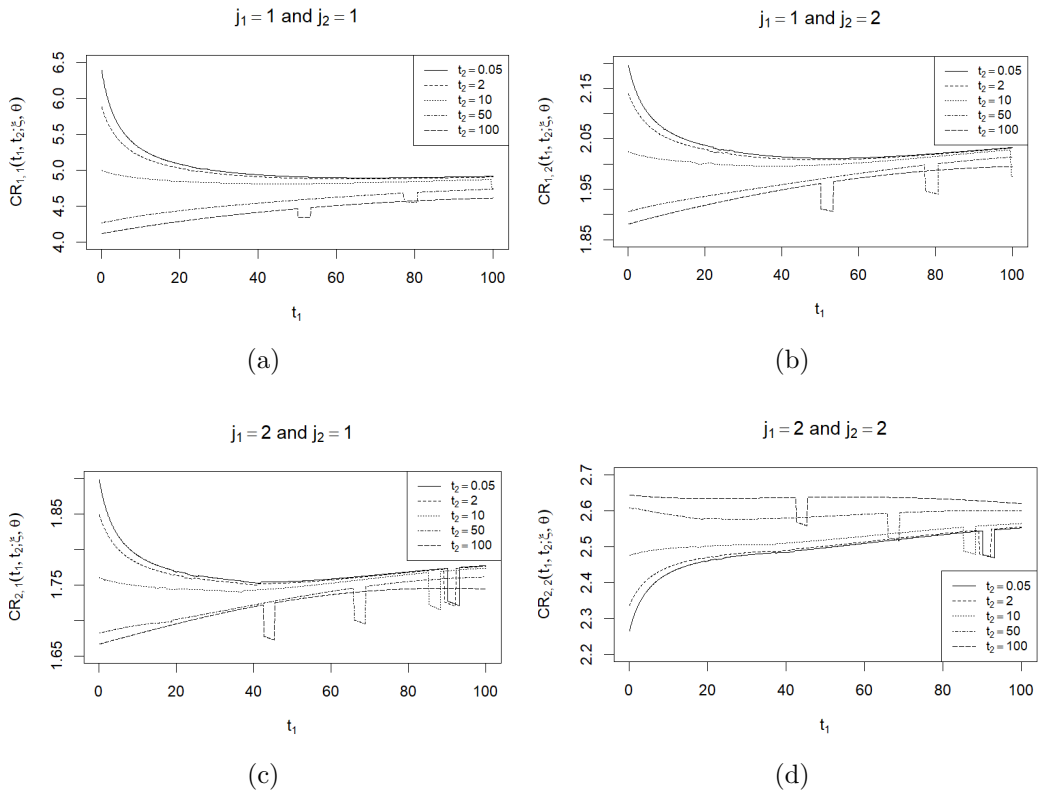


Figure 3.2: The conditional cause-specific hazard ratio function corresponding to the shared cause-specific Gamma frailty model.

3.3.4 Correlated Cause-specific Gamma Frailty

The conditional cause-specific hazard ratio function $CR_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, \boldsymbol{\theta})$ between the occurrence of failure at time t_1 due to cause j_1 for the first individual and occurrence of failure at time t_2 due to cause j_2 for the second individual is

$$\left[\frac{\prod_{k=1}^{2}\left(\sigma_{kj}^2 h_{0j}^{(k)}(t_k;\boldsymbol{\xi})\right)\left(S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})\right)^2}{\left(\int_{t_2}^{\infty}\sum_{j'=1}^{L} f_{jj'}(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})du_2\right)\times\left(\int_{t_1}^{\infty}\sum_{j'=1}^{L} f_{j'j}(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})du_1\right)}\right]\times$$
$$\left[\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})\right)^2}+\sum_{k=1}^{2}\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}-\frac{1}{\sigma_{kj}^2}\right)}{A_j(t_1,t_2;\boldsymbol{\xi},\sigma_{1j},\sigma_{2j})B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})}+\prod_{k=1}^{2}\frac{\left(\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}-\frac{1}{\sigma_{kj}^2}\right)}{B_{kj}(t_k;\boldsymbol{\xi},\sigma_{kj})}\right],$$

for $j_1 = j_2 = j$, say, and

$$\left[\frac{\prod_{k=1}^{2}\left(\sigma_{kj_k}^2 h_{0j_k}^{(k)}(t_k;\boldsymbol{\xi})\right)\left(S(t_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})\right)^2}{\left(\int_{t_2}^{\infty}\sum_{j=1}^{L} f_{j_1 j}(t_1,u_2;\boldsymbol{\xi},\boldsymbol{\theta})du_2\right)\times\left(\int_{t_1}^{\infty}\sum_{j=1}^{L} f_{jj_2}(u_1,t_2;\boldsymbol{\xi},\boldsymbol{\theta})du_1\right)}\right]$$
$$\times\left[\prod_{k=1}^{2}\frac{\frac{\rho_{j_k}}{\sigma_{1,j_k}\sigma_{2,j_k}}}{A_{j_k}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j_k},\sigma_{2j_k})}+\frac{\frac{\rho_{j_2}}{\sigma_{1j_2}\sigma_{2j_2}}\left(\frac{\rho_{j_1}}{\sigma_{1j_1}\sigma_{2j_1}}-\frac{1}{\sigma_{1j_1}^2}\right)}{A_{j_2}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j_2},\sigma_{2j_2})B_{1j_1}(t_1;\boldsymbol{\xi},\sigma_{1j_1})}+\right.$$
$$\left.\frac{\frac{\rho_{j_1}}{\sigma_{1j_1}\sigma_{2j_1}}\left(\frac{\rho_{j_2}}{\sigma_{1j_2}\sigma_{2j_2}}-\frac{1}{\sigma_{2j_2}^2}\right)}{A_{j_1}(t_1,t_2;\boldsymbol{\xi},\sigma_{1j_1},\sigma_{2j_1})B_{2j_2}(t_2;\boldsymbol{\xi},\sigma_{2j_2})}+\prod_{k=1}^{2}\frac{\left(\frac{\rho_{j_k}}{\sigma_{1j_k}\sigma_{2j_k}}-\frac{1}{\sigma_{kj_k}^2}\right)}{B_{kj_k}(t_k;\boldsymbol{\xi},\sigma_{kj_k})}\right],$$

for $j_1 \neq j_2$.

As in the previous subsection, it is also difficult to find a closed form of this conditional cause-speific hazard ratio function, even to obtain a suitable lower bound. Therefore, we numerically evaluate this function for different values of (t_1, t_2) and (j_1, j_2), as presented in Figure 3.3. In this case also, we consider the Exponential type cause-specific hazard rates with the same set of parameters as before. In addition, we take $(\sigma_{11}, \sigma_{12}) = (1.2, 1.8)$ for the first individual and $(\sigma_{21}, \sigma_{22}) = (0.8, 0.4)$ for the second individual. The cause-specific correlation parameters ρ_1 and ρ_2 are taken as 0.7 and 0.25, respectively. As before, we plot the conditional cause-specific hazard ratio function against t_1 for the five different values of t_2 in the four panels $3.3a - 3.3d$ for the four different combinations of (j_1, j_2). Again, we see from Figure 3.3 that this function is greater than unity in each panel. Therefore, failure times of the two individuals due to any particular combination of (j_1, j_2) seem to be positively associated.

3.4 Maximum Likelihood Estimation

Recall, from the beginning of Section 3.2, that the observed data is of the form (X_1, X_2, J_1, J_2). Let us write $J_k = 0$ when the failure time corresponding to the kth individual is right censored at the monitoring time X_k, for $k = 1, 2$. A typical observation from a pair of individuals is, therefore, (x_1, x_2, j_1, j_2), where $x_1 > 0$ and $x_2 > 0$ denote the monitoring times of the two individuals with j_1 and j_2 being the corresponding causes of failure, for $j_k = 0, 1, \cdots, L_k,\ k = 1, 2$. Suppose $(x_{i1}, x_{i2}, j_{i1}, j_{i2})$ denotes the realization of (X_1, X_2, J_1, J_2) for the ith pair in the sam-

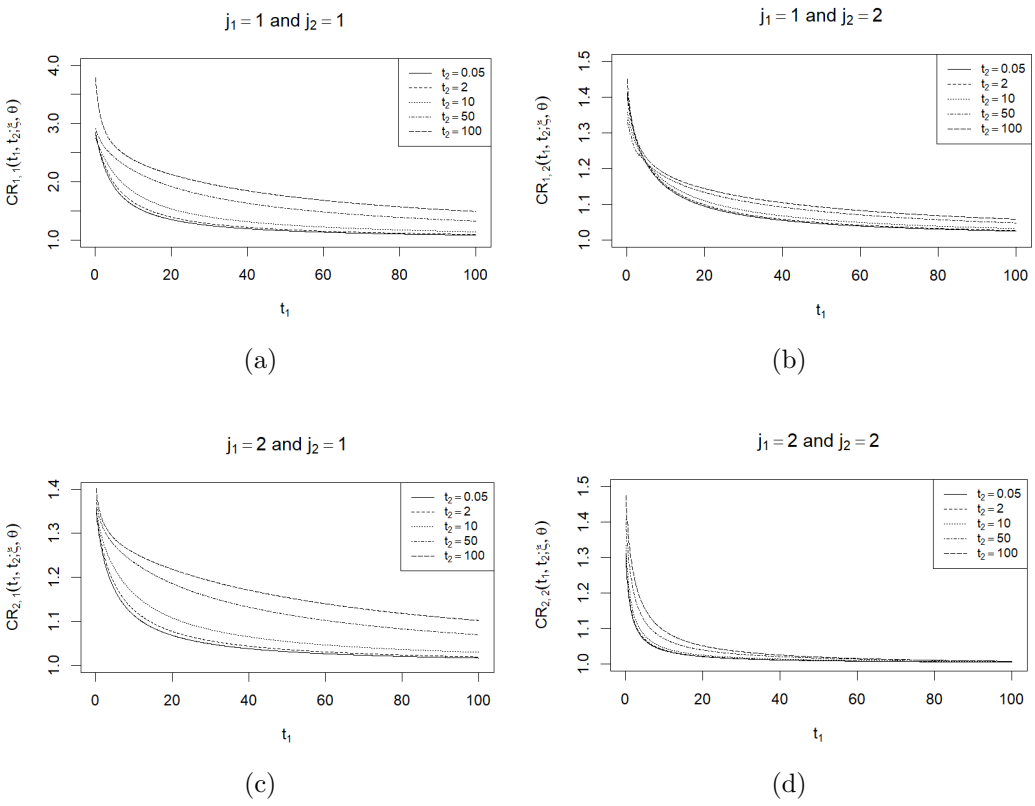


Figure 3.3: The conditional cause-specific hazard ratio function corresponding to the correlated cause-specific Gamma frailty model.

ple, for $i = 1, \cdots, n$. Following Section 2.6, the likelihood function based on the above observations is given by

$$L(\boldsymbol{\xi}, \boldsymbol{\theta}) \propto \prod_{i=1}^{n} f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{\xi}, \boldsymbol{\theta}),$$

where $f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}, \boldsymbol{\theta})$ is the likelihood contribution from the ith pair giving observation $(x_{1i}, x_{2i}, j_{1i}, j_{2i})$ (See Section 2.6) under the assumed fully parametric model for the random vector (X_1, X_2, J_1, J_2) in the present chapter. As argued in Section 2.6, this f^* is the joint density of the random vector (X_1, X_2, J_1, J_2). As a result, the likelihood function can be written as proportional to the product of n independent and identically distributed density functions, which can be useful to study the asymptotic properties later in this section.

For the purpose of maximizing this likelihood function with respect to the associated model parameters, we can rewrite this in terms of the joint sub-distribution

function (See Section 2.6) as

$$
\begin{aligned}
L(\boldsymbol{\xi},\boldsymbol{\theta}) \propto \prod_{i=1}^{n}\prod_{j_{i1}=1}^{L_1}\prod_{j_{i2}=1}^{L_2}&\Big\{F_{j_{i1}j_{i2}}(x_{i1},x_{i2};\boldsymbol{\xi},\boldsymbol{\theta})\Big\}^{\delta_{j_{i1}}^{(1)}\delta_{j_{i2}}^{(2)}}\\
&\times\Big\{F_{j_{i1}}^{(1)}(x_{i1};\boldsymbol{\xi^{(1)}},\boldsymbol{\theta})-\sum_{j_{i2}=1}^{L_2}F_{j_{i1}j_{i2}}(x_{i1},x_{i2};\boldsymbol{\xi},\boldsymbol{\theta})\Big\}^{\delta_{j_{i1}}^{(1)}(1-\delta_i^{(2)})}\\
&\times\Big\{F_{j_{i2}}^{(2)}(x_{i2};\boldsymbol{\xi^{(2)}},\boldsymbol{\theta})-\sum_{j_{i1}=1}^{L_1}F_{j_{i1}j_{i2}}(x_{i1},x_{i2};\boldsymbol{\xi},\boldsymbol{\theta})\Big\}^{(1-\delta_i^{(1)})\delta_{j_{i2}}^{(2)}}\\
&\times\Big\{S(x_{i1},x_{i2};\boldsymbol{\xi},\boldsymbol{\theta})\Big\}^{(1-\delta_i^{(1)})(1-\delta_i^{(2)})},
\end{aligned}
$$

where $\delta_{j_{ik}}^{(k)}=\mathbb{I}(T_{ik}\leq x_{ik},J_k=j_{ik}),\delta_i^{(k)}=\sum_{j_{ik}=1}^{L_k}\delta_{j_{ik}}^{(k)}$ for $x_{ik}>0,j_{ik}=1,\cdots,L_k,k=1,2,i=1,\cdots,n$.

It is important to note that we have a fairly complex likelihood function so that it is not possible to obtain analytic solution of the maximum likelihood estimators (MLEs) of the model parameters. Therefore, we have to make use of some efficient numerical maximization procedure to compute the MLEs. We have performed the task of numerical maximization of the likelihood function with the help of parallel computing. We have used **optimParallel** function from the **optimParallel** package of **R** statistical software for this purpose. The **optimParallel** function generally requires a cluster of $1+2p$ processor cores to solve a minimization problem consisting of p parameters. The computational burden is lesser for the shared and correlated Gamma frailty models, because we have closed form expressions of the joint sub-distribution functions. To speed up the parallel computation in the case of shared and correlated cause-specific Gamma frailty models, we have also used **RcppNumerical** package of **R** software to numerically evaluate the integrals present in the likelihood functions. The asymptotic variance-covariance matrix of the MLEs is estimated by inverting the corresponding hessian matrix computed at the MLEs. See Section 1.11 for details.

let $\boldsymbol{\eta_0}=(\boldsymbol{\xi_0},\boldsymbol{\theta_0})$ be the true parameter vector. Since the likelihood function can be expressed as a product of n independent and identically distributed (iid) densities $f^*(x_{i1},x_{i2},j_{i1},j_{i2};\boldsymbol{\xi},\boldsymbol{\theta})$'s, derivation of the asymptotic results, including consistency and asymptotic normality of maximum likelihood estimator $\hat{\boldsymbol{\eta}}=(\hat{\boldsymbol{\xi}},\hat{\boldsymbol{\theta}})$ of $\boldsymbol{\eta}=(\boldsymbol{\xi},\boldsymbol{\theta})$ follows routinely from the standard likelihood theory under some regularity conditions (See Section 1.10.2). The regularity conditions required for the asymptotic results are:

1. The parameter vector $\boldsymbol{\eta}$ is identifiable with respect to the density function $f^*(.;\boldsymbol{\eta})$.

2. The parameter space $\boldsymbol{\Xi} \times \boldsymbol{\Theta}$ is open.

3. The observations $(x_{i1}, x_{i2}, j_{i1}, j_{i2})$, for $i = 1, 2, \cdots, n$, are iid random vectors with common density $f^*(.;\boldsymbol{\eta})$.

4. The sample space of the random vector (X_1, X_2, J_1, J_2), that is the support of the density $f^*(.;\boldsymbol{\eta})$, is independent of the parameter vector $\boldsymbol{\eta}$.

5. All third-order partial derivatives of $f^*(.;\boldsymbol{\eta})$ exist and are continuous functions of $\boldsymbol{\eta}$.

6. The third-order partial derivatives of $f^*(.;\boldsymbol{\eta})$ are bounded above by functions of (x_1, x_2, j_1, j_2) in the neighbourhood of the true parameter vector $\boldsymbol{\eta_0}$.

7. The Fisher information matrix corresponding to the joint density $f^*(.;\boldsymbol{\eta})$ is positive definite.

Then, we have the following theorem, the proof of which follows from Lehmann and Casella (2006) and hence skipped.

Theorem 3.4.1. *The maximum likelihood estimator of the parameter vector $\boldsymbol{\eta}$, denoted by $\hat{\boldsymbol{\eta}}$, satisfies*

1. $\hat{\boldsymbol{\eta}} \xrightarrow{P} \boldsymbol{\eta_0}$ *and*

2. $\sqrt{n}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta_0})$ *asymptotically follows a mean zero multivariate Normal distribution with variance-covariance matrix estimated by the inverse of hessian matrix computed at $\hat{\boldsymbol{\eta}}$.*

3.5 Simulation Studies

In order to carry out a simulation study to investigate the finite sample behaviour of $\hat{\boldsymbol{\eta}}$, we consider the same monitoring time $X = x$ for both the individuals in a pair, which vary over the different pairs according to a distribution given by the density function $h(x;\mu)$, say, with an associated parameter μ. For our study, we consider Exponential distribution with mean $1/\mu$ for the common monitoring time X. In order to study the effect of censoring on the estimate $\hat{\boldsymbol{\eta}}$, we choose the value of μ so

that the probability of both the failure times T_1 and T_2 being censored equals a fixed value p_{cen}, say. That is,

$$\int_0^\infty P[T_1 > x, T_2 > x; \boldsymbol{\eta_0}] h(x; \mu) dx = p_{cen}.$$

We consider different values of p_{cen} resulting in different values of μ by numerically solving the above equation.

For a sample of n pairs, we simulate n i.i.d. monitoring times $x_1, \cdots, x_n$ from the Exponential distribution with mean $1/\mu$. Note that, for each i, given $X_1 = X_2 = x_i$, the pair (J_1, J_2) follows a multinomial distribution with $(L_1 + 1) \times (L_2 + 1)$ number of cells and the cell probabilities as worked out in Section 2.6 and indicated in the likelihood $L(\boldsymbol{\xi}, \boldsymbol{\theta})$ of Section 3.4. Therefore, given the monitoring time x_i, we simulate the cause of failure $J_k = j_{ik}$, taking values in $\{0, 1, \cdots, L_k\}$, for $k = 1, 2$, using the above multinomial distribution. Repeating this for $i = 1, \cdots, n$, we obtain one simulated sample of bivariate current status data with competing risks, $\{(x_i, x_i, j_{i1}, j_{i2}); i = 1, \cdots, n\}$. The above procedure is repeated 10000 times to get 10000 such simulated samples each of size n.

For our simulation, we consider two competing risks for each individual ($L_1 = L_2 = 2$) with constant or Exponential type baseline cause-specific hazard function for each cause; that is, $h_{0j}(t_k; \boldsymbol{\xi}) = \alpha_j^{(k)}$, for $j = 1, 2$ and $k = 1, 2$, with $\boldsymbol{\xi} = (\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)})$. We consider the four different frailty models discussed in Section 2.4. For the shared frailty model, we take $(\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)}) = (2.4, 5.8, 3.5, 4.5)$ and $\boldsymbol{\theta} = \sigma = 0.95$. For the correlated Gamma frailty model, we take $(\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)}) = (3, 2, 2.5, 0.8)$ and $\boldsymbol{\theta} = (\sigma_1 = 0.9, \sigma_2 = 0.7, \rho = 0.65)$. For the shared cause-specific Gamma frailty model, we take $(\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)}) = (7, 6, 8.5, 10)$ and $\boldsymbol{\theta} = (\sigma_1 = 0.95, \sigma_2 = 0.85)$. While trying with the correlated cause-specific Gamma frailty model, we encountered some issues with the convergence of likelihood maximization by numerical method. From our experience with real data analysis using this frailty model, we find that the numerical method is more complicated than the previous three models possibly because of larger dimension of the parameter space. Therefore, some manipulation with the initial parameter values is required, which could not be replicated for simulation studies. We intuitively feel that very large and informative datasets are required to produce reasonable estimates for the correlated cause-specific Gamma

frailty model. Li et al. (2001) noted that convergence issues of maximum likelihood estimators may arise due to irregular behaviour of the log-likelihood function in the neighbourhood of the maxima, which may be the case here as well.

For every simulated dataset, we obtain the MLEs of the model parameters and their standard errors using the method described in Section 3.4. Then, the estimated bias of each parameter estimate is obtained by subtracting the true value of that parameter from the average over the corresponding estimates from the 10000 simulated samples. For each parameter, the average of the 10000 standard errors is obtained and denoted by ASE. We also compute the square root of the sample variance of these 10000 MLEs to get sample standard error (denoted by SSE). Next, for each parameter, we compute the proportion of times the corresponding 10000 approximate 95% confidence intervals based on normal approximation contain the true value. This gives an estimated coverage probability, denoted by CP, of the approximate 95% confidence interval for that parameter. We carry out the simulation study for sample sizes $n = 50, 150, 300$.

The results for the three frailty models are summarized in Tables 3.1-3.3. In each of the three cases, we see decrease in bias, ASE and SSE with increase in sample size, as expected. The values of ASE and SSE tend to be closer with increasing sample size, as expected. The CP values also tend to 0.95 with increasing sample size giving evidence in favour of asymptotic normality. We have carried out similar simulation studies with different values of the model parameters, but the results are not reported here, as those are qualitatively similar.

To investigate the effect of joint censoring probability p_{cen} on the different MLEs, we consider two values of p_{cen}, 0.1 and 0.2. There is no definite pattern for the bias, ASE, SSE and CP values corresponding to the two values of p_{cen}. As the joint censoring probability p_{cen} is increased from 0.1 to 0.2, it is found that the probabilities of the other possibilities, including those of one individual being censored, are sometimes decreased. There does not seem to be any definite pattern of information loss or gain. As a result, the effect of such increase in p_{cen} is not clear and we observe this indefinite pattern in our simulation result as well.

Table 3.1: Simulation results for shared Gamma frailty model for different joint censoring probability.

p_{cen}	n	Properties	$\hat{\alpha}_1^{(1)}$	$\hat{\alpha}_2^{(1)}$	$\hat{\alpha}_1^{(2)}$	$\hat{\alpha}_2^{(2)}$	$\hat{\sigma}$
0.1	50	Bias	−0.2310	−0.6091	−0.2798	−0.3960	−0.2281
		SSE	1.0839	2.0898	1.5835	1.9169	0.2414
		ASE	1.3336	2.9576	1.9244	2.4028	0.3601
		CP	0.8473	0.8545	0.8470	0.8416	0.8977
	150	Bias	−0.1812	−0.4514	−0.2697	−0.3397	−0.1303
		SSE	0.6717	1.4275	0.9280	1.1754	0.1242
		ASE	0.7963	1.7954	1.1125	1.4081	0.2197
		CP	0.8943	0.8937	0.8993	0.8995	0.9407
	300	Bias	−0.1706	−0.4096	−0.2362	−0.3024	−0.0993
		SSE	0.4745	1.0294	0.6446	0.8076	0.0842
		ASE	0.5662	1.2816	0.7948	1.0046	0.1594
		CP	0.9137	0.9204	0.9216	0.9230	0.9493
	500	Bias	−0.0367	−0.0811	−0.0483	−0.0598	−0.0320
		SSE	0.4052	0.8987	0.5636	0.7116	0.0790
		ASE	0.4560	1.0301	0.6373	0.8051	0.1241
		CP	0.9435	0.9441	0.9433	0.9415	0.9567
	1000	Bias	−0.0226	−0.0547	−0.0333	−0.0427	−0.0171
		SSE	0.2927	0.6529	0.4074	0.5143	0.0679
		ASE	0.3251	0.7339	0.4537	0.5729	0.0892
		CP	0.9489	0.9522	0.9486	0.9499	0.9613
0.2	50	Bias	−0.1665	−0.4312	−0.1979	−0.2601	−0.2108
		SSE	0.9914	1.8407	1.3894	1.6486	0.2257
		ASE	1.1229	2.3651	1.5632	1.9311	0.3718
		CP	0.8718	0.8820	0.8780	0.8785	0.9141
	150	Bias	−0.1410	−0.3311	−0.1862	−0.2502	−0.1273
		SSE	0.5840	1.1851	0.8039	0.9666	0.1225
		ASE	0.6612	1.4116	0.9030	1.1206	0.2272
		CP	0.9141	0.9207	0.9130	0.9153	0.9431
	300	Bias	−0.1210	−0.2986	−0.1790	−0.2322	−0.0893
		SSE	0.4172	0.8445	0.5552	0.6836	0.0848
		ASE	0.4728	1.0082	0.6416	0.7974	0.1645
		CP	0.9284	0.9306	0.9269	0.9271	0.9523
	500	Bias	−0.0287	−0.0639	−0.0413	−0.0543	−0.0324
		SSE	0.3454	0.7131	0.4629	0.5637	0.0912
		ASE	0.3767	0.8012	0.5102	0.6335	0.1284
		CP	0.9499	0.9530	0.9501	0.9513	0.9595
	1000	Bias	−0.0165	−0.0396	−0.0221	−0.0313	−0.0184
		SSE	0.2487	0.5147	0.3357	0.4096	0.0689
		ASE	0.2684	0.5710	0.3641	0.4520	0.0921
		CP	0.9524	0.9573	0.9546	0.9586	0.9660

Table 3.2: Simulation results for correlated Gamma frailty model for different joint censoring probability.

p_{cen}	n	Properties	$\hat{\alpha}_1^{(1)}$	$\hat{\alpha}_2^{(1)}$	$\hat{\alpha}_1^{(2)}$	$\hat{\alpha}_2^{(2)}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\rho}$
0.1	50	Bias	1.5273	1.0162	1.4087	0.4493	−0.0296	0.0950	−0.0961
		SSE	5.1014	3.4022	3.9115	1.3526	0.2782	0.2461	0.3503
		ASE	4.8932	3.2608	3.5827	1.2012	0.3904	0.4091	1.1411
		CP	0.9120	0.9130	0.9801	0.9623	0.9951	0.9926	0.9999
	150	Bias	0.3964	0.2611	0.3618	0.1146	−0.0246	0.0163	−0.0814
		SSE	2.0332	1.3450	1.2447	0.4102	0.1927	0.1768	0.2864
		ASE	1.8929	1.2121	1.1170	0.3839	0.2021	0.2161	0.5177
		CP	0.9176	0.9152	0.9651	0.9595	0.9863	0.9846	0.9994
	300	Bias	0.0234	0.0296	0.1625	0.0507	−0.0145	0.0062	−0.0552
		SSE	1.0641	0.7175	0.6503	0.2284	0.1367	0.1280	0.2244
		ASE	1.0192	0.6896	0.6809	0.2365	0.1382	0.1455	0.3051
		CP	0.9219	0.9223	0.9661	0.9608	0.9621	0.9833	0.9934
	500	Bias	0.0155	0.0172	0.0994	0.0315	−0.0109	0.0036	−0.0345
		SSE	0.7831	0.5289	0.4811	0.1712	0.1062	0.0998	0.1781
		ASE	0.7520	0.5103	0.5046	0.1759	0.1060	0.1111	0.2219
		CP	0.9269	0.9307	0.9667	0.9624	0.9522	0.9754	0.9859
	1000	Bias	0.0084	0.0093	0.0518	0.0152	−0.0051	0.0022	−0.0172
		SSE	0.5206	0.3528	0.3270	0.1155	0.0741	0.0701	0.1260
		ASE	0.5154	0.3499	0.3453	0.1205	0.0746	0.0777	0.1488
		CP	0.9425	0.9413	0.9697	0.9629	0.9539	0.9720	0.9786
0.2	50	Bias	1.2595	0.8514	1.1080	0.3538	0.0283	0.1312	−0.0702
		SSE	4.1206	2.8198	3.0998	1.0841	0.3191	0.2874	0.3315
		ASE	3.7998	2.5914	2.7069	0.9352	0.4383	0.4637	0.9992
		CP	0.9190	0.9126	0.9842	0.9613	0.9944	0.9907	0.9991
	150	Bias	0.2867	0.1916	0.3002	0.0953	−0.0204	0.0336	−0.0589
		SSE	1.4890	1.0073	0.9537	0.3672	0.2196	0.1996	0.2691
		ASE	1.3328	0.9109	0.9262	0.3351	0.2347	0.2554	0.4927
		CP	0.9271	0.9256	0.9760	0.9606	0.9855	0.9819	0.9954
	300	Bias	0.1141	0.0706	0.1624	0.0498	−0.0165	0.0202	−0.0413
		SSE	0.8716	0.5925	0.5571	0.2065	0.1571	0.1478	0.2096
		ASE	0.8411	0.5751	0.5933	0.2157	0.1610	0.1752	0.2914
		CP	0.9295	0.9314	0.9717	0.9600	0.9613	0.9776	0.9850
	500	Bias	0.069	0.0420	0.0960	0.0308	−0.0103	0.0133	−0.0256
		SSE	0.6530	0.4472	0.4051	0.1527	0.1247	0.1142	0.1664
		ASE	0.6323	0.4328	0.4400	0.1609	0.1237	0.1347	0.2145
		CP	0.9358	0.9334	0.9772	0.9696	0.9536	0.9778	0.9807
	1000	Bias	0.0209	0.0120	0.0523	0.0173	−0.0076	0.0078	−0.0117
		SSE	0.4400	0.3017	0.2816	0.1042	0.0866	0.0827	0.1200
		ASE	0.4336	0.2973	0.3030	0.1110	0.0869	0.0949	0.1466
		CP	0.9413	0.9397	0.9750	0.9698	0.9550	0.9737	0.9761

Table 3.3: Simulation results for shared cause-specific Gamma frailty model for different joint censoring probability.

p_{cen}	n	Properties	$\hat{\alpha}_1^{(1)}$	$\hat{\alpha}_2^{(1)}$	$\hat{\alpha}_1^{(2)}$	$\hat{\alpha}_2^{(2)}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
0.1	50	Bias	0.3132	0.3448	0.4420	0.6502	0.1270	0.1475
		SSE	3.4688	3.1404	4.0554	4.9451	0.7937	0.9567
		ASE	3.6174	3.3713	4.2811	5.0905	0.8272	0.9276
		CP	0.9170	0.9137	0.9238	0.9205	0.9880	0.9726
	150	Bias	0.1699	0.1557	0.1678	0.2456	0.0292	0.0364
		SSE	1.9770	1.8282	2.2922	2.7348	0.4105	0.4756
		ASE	1.9402	1.8021	2.2603	2.6988	0.3902	0.4566
		CP	0.9388	0.9320	0.9400	0.9402	0.9213	0.9216
	300	Bias	0.0694	0.0762	0.1000	0.1046	0.0120	0.0198
		SSE	1.3751	1.2829	1.6074	1.9153	0.2807	0.3301
		ASE	1.3458	1.2567	1.5767	1.8795	0.2680	0.3166
		CP	0.9392	0.9369	0.9385	0.9406	0.9236	0.9286
	500	Bias	0.0521	0.0447	0.0694	0.0743	0.0049	0.0110
		SSE	1.0353	0.9796	1.2334	1.4712	0.2105	0.2465
		ASE	1.0408	0.9688	1.2167	1.4481	0.2058	0.2428
		CP	0.9454	0.9406	0.9432	0.9419	0.9364	0.9386
	1000	Bias	0.0197	0.0219	−0.0044	0.0111	0.0017	0.0066
		SSE	0.7984	0.6883	0.9454	1.0236	0.1767	0.1670
		ASE	0.7833	0.6905	0.9210	1.0283	0.1735	0.1623
		CP	0.9380	0.9364	0.9420	0.9430	0.9350	0.9350
0.2	50	Bias	0.3027	0.3050	0.3074	0.6349	0.1537	0.1482
		SSE	3.0609	2.7242	3.4885	4.5166	0.8040	0.9393
		ASE	3.1078	2.8411	3.6258	4.4740	0.8195	0.9233
		CP	0.9303	0.9338	0.9381	0.9383	0.9935	0.9790
	150	Bias	0.1485	0.1033	0.1546	0.2216	0.0271	0.0441
		SSE	1.6827	1.5469	1.9703	2.4492	0.4022	0.4771
		ASE	1.6628	1.5279	1.9442	2.3931	0.3902	0.4680
		CP	0.9465	0.9422	0.9454	0.9428	0.9362	0.9306
	300	Bias	0.0659	0.0499	0.0903	0.0952	0.0119	0.0218
		SSE	1.1745	1.0772	1.3701	1.6915	0.2755	0.3221
		ASE	1.1603	1.0715	1.3609	1.6662	0.2701	0.3171
		CP	0.9456	0.9451	0.9483	0.9453	0.9358	0.9410
	500	Bias	0.0312	0.0397	0.0435	0.0793	0.0155	0.0079
		SSE	0.9578	0.8366	1.1369	1.2973	0.2550	0.2394
		ASE	0.9666	0.8339	1.1340	1.2861	0.2507	0.2347
		CP	0.9476	0.9468	0.9483	0.9452	0.9439	0.9430
	1000	Bias	0.0137	0.0293	0.0069	0.0675	0.0095	−0.0017
		SSE	0.6888	0.6224	0.8185	0.8832	0.1811	0.1666
		ASE	0.6805	0.5903	0.7980	0.9073	0.1762	0.1642
		CP	0.9480	0.9380	0.9450	0.9550	0.9470	0.9440

3.6 Data Analysis

Possibility of hearing loss slowly increases with age, a common phenomenon known as presbycusis. Broadly, there are two causes of hearing loss depending on the part of an ear. **Sensorineural** hearing loss (SNHL) occurs due to any sensory problem or a neural structure related problem in the inner ear; **Conductive** hearing loss is causally related to any hindrance in the outer or middle ear so that sound cannot enter the inner ear properly. In this section, we analyze the dataset collected from the department of speech and hearing, Ali Yavar Jung National institute of Speech and Hearing Disabilities, Eastern Regional centre, in Kolkata, India. This dataset contains information about hearing loss of 880 individuals who appeared for diagnosis in the time span of January to February month of the year 2016. The monitoring time (X) is considered to be the age (in months) of an individual when he(she) appeared for diagnosis. See Section 1.12 for details. We intend to use hearing loss information in both the ears so that we have bivariate data. We have such information from 726 out of 880 individuals. Since hearing loss can occur due to any of the two above-mentioned causes, this gives a bivariate current status dataset with two competing risks, namely SNHL and Conductive with corresponding labels 1 and 2, respectively. A summary of the dataset is presented in Table 3.4 below, with cause of failure 0 meaning censored observation. There is evidence of high positive correlation between the causes of hearing loss in the two ears since only two observations appear in the off-diagonal cells. This is also evident from the results in this section where the MLEs of the correlation parameters are found to be close to 1.

Table 3.4: A summary of the hearing loss dataset.

Cause J_1 for T_1	Cause J_2 for T_2		
	0	1	2
0	261	0	1
1	0	435	0
2	1	0	28

We start by assuming constant (Exponential type) baseline cause-specific hazards $\alpha_j^{(k)}$'s, denoting the jth cause-specific hazard due to cause j for the kth individual, for $j = 1, 2$ and $k = 1, 2$. The MLEs of the model parameters and the corresponding standard errors (SE) along with the AIC value for each of the four frailty distributions are reported in Tables 3.5 (for shared and correlated Gamma frailty) and 3.6 (for shared and correlated cause-specific Gamma frailty), respectively. From Tables

3.5 and 3.6, we observe that SNHL is the dominating cause of hearing loss in both ears.

Table 3.5: Results for the constant baseline cause-specific hazard model with shared and correlated Gamma frailty.

Shared Gamma frailty				Correlated Gamma frailty			
Parameter	MLE	SE	AIC	Parameter	MLE	SE	AIC
$\alpha_1^{(1)}$	8.9551	6.9567	1423.9852	$\alpha_1^{(1)}$	9.2268	5.4607	1427.1321
$\alpha_2^{(1)}$	0.6123	0.4981		$\alpha_2^{(1)}$	0.6597	0.4350	
$\alpha_1^{(2)}$	8.5792	6.3573		$\alpha_1^{(2)}$	9.5207	5.8413	
$\alpha_2^{(2)}$	0.5813	0.4516		$\alpha_2^{(2)}$	0.7043	0.4956	
σ	3.0123	0.1612		σ	3.0816	0.1402	
-	-	-		ρ	0.9998	0.0056	

From our calculation, we observe that the magnitudes of the MLEs of σ_1 and σ_2 for constant cause-specific baseline hazards with correlated Gamma frailty are very close. Therefore, it is of interest to test the hypothesis, $H_0 : \sigma_1 = \sigma_2$ vs $H_1 : \sigma_1 \neq \sigma_2$. The likelihood ratio test statistic comes out to be 0.2784, lesser than 3.84, the upper 0.05 point of χ_1^2 distribution. Hence, we fail to reject the null hypothesis of equality of frailty variances at 0.05 level of significance. We, therefore, assume $\sigma_1 = \sigma_2 = \sigma$, say, for a simplified analysis and the corresponding results are presented in the right panel of Table 3.5. It is to be noted that the equality of σ_1 and σ_2 in the correlated frailty model does not make it the shared frailty model.

We next fit the two cause-specific frailty models with constant cause-specific hazards. For the shared cause-specific Gamma frailty model, we have two frailty parameters σ_1 and σ_2. For the correlated cause-specific Gamma frailty model, we have six frailty parameters $\sigma_{kj},\ \rho_j$, for $j = 1, 2$ and $k = 1, 2$. In view of our previous calculations, it is of interest to test for the equality of σ_{1j} and σ_{2j}, for $j = 1, 2$, so that there may be some reduction in number of parameters. The likelihood ratio test for this hypothesis, for the sample at hand, leads to the observed value of the test statistic as 0.0232 to be compared with χ_2^2 distribution. This is lesser than the corresponding 95th percentile. Therefore, we fail to reject the equality at 5% significance level. So, we assume this equality $\sigma_{1j} = \sigma_{2j} = \sigma_j$ (say), for $j = 1, 2$, for further analysis and the results are reported in Table 3.6. In case of constant baseline cause-specific hazards with shared and correlated cause-specific Gamma frailty models, we have computed the asymptotic standard errors of the MLEs using the bootstrap method.

Table 3.6: Results for the constant baseline cause-specific hazard model with shared and correlated cause-specific Gamma frailty.

Shared cause-specific Gamma frailty				Correlated cause-specific Gamma frailty			
Parameter	MLE	SE	AIC	Parameter	MLE	SE	AIC
$\alpha_1^{(1)}$	42.7027	30.9252	1216.8936	$\alpha_1^{(1)}$	1521.5801	51.0891	1208.6964
$\alpha_2^{(1)}$	0.0578	0.0110		$\alpha_2^{(1)}$	0.1733	0.0498	
$\alpha_1^{(2)}$	44.1087	32.9971		$\alpha_1^{(2)}$	1541.6708	52.1110	
$\alpha_2^{(2)}$	0.0598	0.0142		$\alpha_2^{(2)}$	0.1349	0.0167	
σ_1	11.8831	1.0925		σ_1	3.9893	0.0997	
σ_2	41.5261	18.5925		σ_2	12.8552	0.0735	
-	-	-		ρ_1	0.9991	0.0039	
-	-	-		ρ_2	0.9978	0.0076	

Note that the AIC value for the correlated cause-specific Gamma frailty is the lowest indicating the best fit with constant baseline cause-specific hazards. For this model, both the parameters σ_1 and σ_2 appear to be significant. So, there is evidence of dependence between occurrence of hearing loss in left and right ears due to a particular cause.

Next, we consider Weibull type baseline cause-specific hazards (that is, power functions of time) with the scale parameters $\alpha_j^{(k)}$ and shape parameters $\gamma_j^{(k)}$, for $j = 1, 2$ and $k = 1, 2$. We use four Gamma frailty models as before. The results for shared Gamma frailty and correlated Gamma frailty models are presented in Table 3.7. Similar to the case of constant cause-specific hazards, we test the equality of frailty variances (that is, $\sigma_1^2 = \sigma_2^2$) for the analysis of correlated frailty model. In this case, the likelihood ratio test statistic comes out to be 0.0245. So, we cannot reject the null hypothesis since this value is less than 3.84. So, we assume $\sigma_1 = \sigma_2 = \sigma$, say, for further analysis, the results of which are presented in the right panel of Table 3.7. Note that the AIC values are decreased compared to those for the constant cause-specific hazards in Table 3.5. There is evidence of time dependence on the baseline cause-specific hazards from the estimates of $\gamma^{(1)}$ and $\gamma^{(2)}$, indicating increasing cause-specific hazards with age. Contrary to the results of constant baseline cause-specific hazards, correlated Gamma frailty model gives better result than the shared Gamma frailty model. Again, we have used bootstrap to approximate the asymptotic standard errors of the MLEs for the power function baseline cause-specific hazard with

shared and correlated Gamma frailty models.

Table 3.7: Results for the power function baseline cause-specific hazard model with shared and correlated Gamma frailty.

Shared Gamma frailty				Correlated Gamma frailty			
Parameter	MLE	SE	AIC	Parameter	MLE	SE	AIC
$\alpha_1^{(1)}$	10.6668	1.5778	1291.3132	$\alpha_1^{(1)}$	10.7854	0.5985	1218.7970
$\alpha_2^{(1)}$	0.0593	0.0477		$\alpha_2^{(1)}$	0.0201	0.0276	
$\alpha_1^{(2)}$	13.5783	1.7545		$\alpha_1^{(2)}$	23.4554	0.9932	
$\alpha_2^{(2)}$	0.0291	0.0216		$\alpha_2^{(2)}$	1.4341	0.4965	
$\gamma_1^{(1)}$	1.8712	0.0856		$\gamma_1^{(1)}$	4.2621	0.0323	
$\gamma_2^{(1)}$	0.3872	0.0650		$\gamma_2^{(1)}$	0.3240	0.0687	
$\gamma_1^{(2)}$	1.7982	0.0928		$\gamma_1^{(2)}$	3.8178	0.0347	
$\gamma_2^{(2)}$	0.3481	0.0770		$\gamma_2^{(2)}$	0.7037	0.0876	
σ	3.9339	0.1010		σ	6.0581	0.1202	
-	-	-		ρ	0.9992	0.0036	

Table 3.8: Results for the power function baseline cause-specific hazard model with shared and correlated cause-specific Gamma frailty.

Shared cause-specific Gamma frailty				Correlated cause-specific Gamma frailty			
Parameter	MLE	SE	AIC	Parameter	MLE	SE	AIC
$\alpha_1^{(1)}$	63.5871	4.6621	1175.3721	$\alpha_1^{(1)}$	68.2560	2.6277	1175.5856
$\alpha_2^{(1)}$	0.3556	0.2406		$\alpha_2^{(1)}$	0.4147	0.2518	
$\alpha_1^{(2)}$	34.4121	4.1326		$\alpha_1^{(2)}$	34.2791	2.2893	
$\alpha_2^{(2)}$	0.2570	0.2071		$\alpha_2^{(2)}$	0.3164	0.1347	
$\gamma_1^{(1)}$	3.1908	1.4924		$\gamma_1^{(1)}$	3.3882	0.0993	
$\gamma_2^{(1)}$	3.0352	1.2006		$\gamma_2^{(1)}$	3.1607	0.3195	
$\gamma_1^{(2)}$	3.4541	1.5987		$\gamma_1^{(2)}$	3.6732	0.0953	
$\gamma_2^{(2)}$	3.0345	1.2313		$\gamma_2^{(2)}$	3.4327	0.3245	
σ_1	5.9858	1.1507		σ_1	5.9188	0.0609	
σ_2	15.1334	2.9455		σ_2	15.2319	0.1601	
-	-	-		ρ_1	0.9998	0.0011	
-	-	-		ρ_2	0.9938	0.0094	

Finally, we consider the two cause-specific frailty models with power function (that is, Weibull type) baseline cause-specific hazards having different shape parameters for the different causes. As before, we assume $\sigma_{1j} = \sigma_{2j} = \sigma_j$, for $j = 1, 2$, for the correlated cause-specific Gamma frailty model. The corresponding results are presented in Table 3.8. We observe better fit compared to all other models in terms of AIC value (See Tables 3.5-3.8). Also, the estimates of the shape parameters are significant, indicating time dependence. The model with Weibull type baseline cause-specific hazards and shared cause-specific frailty gives the least value of AIC. However, one should use caution before adopting this model. These are complicated models with many parameters, and numerical stability of the estimates is an issue. Although some models may be identifiable, it does not guaranty efficient estimation, for example, see Li et al. (2001). In summary, there is evidence of strong time dependence on the different cause-specific hazards with the cause SNHL dominating the other cause. Also, there is evidence of some association between occurrence of hearing loss in left and right ears, the nature of which seems to depend on the underlying cause.

3.7 Concluding Remarks

In this work, we consider parametric modeling and analysis of bivariate current status data with competing risks, while association between the two individuals in a pair is modeled through four different Gamma frailty distributions possibly depending on the cause of failure. These four different frailty distributions cover a wide range of dependence as discussed in Section 2.4. The baseline cause-specific hazards are assumed to belong to a parametric family that includes many commonly used parametric hazard functions. However, there are some other parametric hazard functions, for example, Log-normal, Pareto, generalized Gamma, etc., which do not belong to this family. A remaining task is to construct and analyze models consisting of the above-mentioned cause-specific hazard functions with the four different Gamma frailty distributions.

In our data analysis in Section 3.6, we have assumed Exponential type baseline cause-specific hazards and also Weibull type baseline cause-specific hazards with different shape parameters for these two failure types. In practice, the different baseline cause-specific hazards need not have the same parametric model (Exponential or Weibull); in many cases, they may have different shapes suggesting different theoretical models (See Section 5.5 for the analysis with non-parametric baseline cause-specific hazards). Therefore, the parametric analysis described in this chapter (in particular, in Section 3.6) suffers from some lack of flexibility. This is, however, a drawback for any parametric analysis in the absence of a proper goodness of fit test.

Note that we have used different Gamma frailty in this work because of its popularity and mathematical tractability. It is also an important task to consider other frailty distributions, for example, inverse Gaussian, Log-normal (see Gorfine and Hsu, 2011), etc., which also produce tractable expressions in many cases, and consequently analyze the resulting models for bivariate current status data with competing risks. From our data analysis in Section 3.6, it is seen that the fit of the model substantially depends on the kind of frailty distribution. In view of that, it is of interest to investigate different types of frailty distributions subject to identifiability of the resulting bivariate model. This applies to routine analysis of bivariate failure time data as well using some specific frailty distribution.

As evident from Sections 3.5-3.6, our analysis involves heavy computation and some manipulation with initial values possibly because of some irregular behaviour of the likelihood in the neighbourhood of the true maxima. This numerical problem gets further compounded with larger number of competing risks. Even if the model is identifiable, one needs large number of observations to tackle this problem of nu-

merical instability. With a smaller dataset, one may be advised to consider a simpler frailty model and possibly combine several causes to have fewer number of competing risks.

Extending this work to multivariate current status data with competing risks seems, in principle, straightforward. But it is associated with quite an amount of complications in terms of notation, mathematical tractability to derive identifiability results, etc., and also computational issues for analysis.

Chapter 4

Bivariate Current status Data with Competing Risks: Parametric Baseline Hazards and Non-parametric Frailty

4.1 Introduction

In Section 2.5, the general model (2.1) is shown to be non-identifiable when both the baseline cause-specific hazards and the frailty distribution are unspecified. In Chapter 3, we have considered a fully parametric model for both along with an investigation of the corresponding model identifiability. In this chapter, we consider parametric models for the baseline cause-specific hazards, but the frailty distribution remains unspecified, resulting in a semi-parametric setup. Nevertheless, it is relevant to study identifiability of models with such parametric hazards but with non-parametric frailty distribution.

We consider the same parametric class of baseline cause-specific hazard functions (3.2) introduced in Section 3.2. This parametric class is particularly helpful as it contains various shapes of hazard functions. Hence, it enables us to choose appropriate baseline cause-specific hazard functions according to the nature of the observed data. As described in Section 2.2, we model the dependence between the failure times using frailty variables. We also consider the four different frailty structures mentioned in Section 2.4. The common assumption of these models to ensure identifiability is that the frailty mean(s) is/are equal to unity.

In Section 4.2, we describe the data and the four frailty models with proofs of corresponding identifiability results. In Section 4.3, we discuss the semi-parametric

estimation procedure under the non-parametric shared frailty model by means of estimation of mixing distribution. Section 4.4 discusses some computational issues related to this semi-parametric estimation method. In Section 4.5, we first estimate the convergence rates of the proposed estimators using subsampling bootstrap method. We also propose bootstrap based confidence intervals for the related parameters and the unknown quartiles of the frailty distribution. Section 4.6 presents a simulation study to investigate the finite sample properties of the proposed estimators. We illustrate the estimation method through analysis of the hearing loss data in Section 4.7 using the non-parametric shared frailty model. We end with some concluding remarks in Section 4.8.

4.2 Modeling and Identifiability

Recall from Section 2.2 that T_1 and T_2 denote the failure time of the first and the second individual of a pair, respectively, with J_1 and J_2 denoting the cause of failure for the first and the second individual, respectively. We assume that J_k has support $\{1, 2, \cdots, L_k\}$, out of which only one is responsible for the failure of the kth individual, for $k = 1, 2$. The cause-specific hazard function $\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}})$ for the failure of the kth individual at time $T_k = t_k$ due to cause j, conditional on the frailty vector $\boldsymbol{\epsilon^{(k)}}$, is defined as

$$\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k)\epsilon_j^{(k)}, \tag{4.1}$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$, where $h_{0j}^{(k)}(t_k)$ is the jth baseline cause-specific hazard at time t_k for the kth individual. We consider the same class of parametric baseline cause-specific hazard functions $h_{0j}^{(k)}(t_k; \boldsymbol{\xi})$ (See Section 3.2), where $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ is the associated parameter vector.

As in Section 2.2, the conditional survival function of the kth individual, given the frailty vector $\boldsymbol{\epsilon^{(k)}}$, is

$$S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}|\boldsymbol{\epsilon^{(k)}}) = \exp\left[-\sum_{j=1}^{L_k} \epsilon_j^{(k)} H_{0j}^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) \right],$$

where $H_{0j}^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) = \int_0^{t_k} h_{0j}^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) du_k$, and the jth sub-distribution function of

the kth individual, conditional on the frailty vector $\boldsymbol{\epsilon}^{(\boldsymbol{k})}$, is

$$F_j^{(k)}(t_k;\boldsymbol{\xi}^{(\boldsymbol{k})}|\boldsymbol{\epsilon}^{(\boldsymbol{k})}) = \int_0^{t_k} h_{0j}^{(k)}(u_k;\boldsymbol{\xi}^{(\boldsymbol{k})})\epsilon_j^{(k)} \exp\left[-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(u_k;\boldsymbol{\xi}^{(\boldsymbol{k})})\epsilon_{j'}^{(k)}\right] du_k,$$

for all $j = 1, \cdots, L_k$ and $k = 1, 2$. Therefore, using conditional independence given frailty, the joint unconditional sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi})$ under the general model (4.1) is

$$F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}) = \int_0^{\infty} \cdots \int_0^{\infty} \Bigg[\int_0^{t_1} \int_0^{t_2} \prod_{k=1}^{2} \left(h_{0j_k}^{(k)}(u_k;\boldsymbol{\xi}^{(\boldsymbol{k})})\epsilon_{j_k}^{(k)}\right) \times \exp\left(-\sum_{k=1}^{2}\sum_{j=1}^{L_k} H_{0j}^{(k)}(u_k;\boldsymbol{\xi}^{(\boldsymbol{k})})\epsilon_j^{(k)}\right) du_2 du_1 \Bigg] dG(\boldsymbol{\epsilon}^{(\boldsymbol{1})}, \boldsymbol{\epsilon}^{(\boldsymbol{2})}), \quad (4.2)$$

for all $j_k = 1, \cdots, L_k$ and $k = 1, 2$, where $G(\boldsymbol{\epsilon}^{(\boldsymbol{1})}, \boldsymbol{\epsilon}^{(\boldsymbol{2})})$ is an arbitrary probability measure on $[0, \infty)^{L_1} \times [0, \infty)^{L_2}$, representing the joint frailty distribution of $(\boldsymbol{\epsilon}^{(\boldsymbol{1})}, \boldsymbol{\epsilon}^{(\boldsymbol{2})})$. In this work, we assume it to belong to a family $\mathcal{G}$ of arbitrary joint distribution functions. It is to be noted that this joint sub-distribution function also depends on the joint frailty distribution $G(\boldsymbol{\epsilon}^{(\boldsymbol{1})}, \boldsymbol{\epsilon}^{(\boldsymbol{2})}) \in \mathcal{G}$; so we write this as $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, G)$ for completeness. The focus of this section is to investigate identifiability of the model for bivariate failure time with competing risks (T_1, T_2, J_1, J_2) given by the joint sub-distribution function $F_{j_1 j_2}\big(t_1, t_2; \boldsymbol{\xi}, G\big)$ with the four different arbitrary frailty distributions, as mentioned earlier. For this, we need the following definition of model identifiability. Note that the family $\mathcal{G}$ of the joint frailty distribution functions differs in nature and dimension for the four different frailty types.

Definition 4.1. The model (4.1) for bivariate failure time with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi} \times \mathcal{G}$ if, for some $\boldsymbol{\xi},\ \tilde{\boldsymbol{\xi}} \in\ \boldsymbol{\Xi}$ and $G,\ \tilde{G} \in\ \mathcal{G}$, the equality

$$F_{j_1 j_2}\big(t_1, t_2; \boldsymbol{\xi}, G\big) = F_{j_1 j_2}\big(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{G}\big),$$

for all $t_k > 0,\ j_k = 1, \cdots, L_k$ and $k = 1, 2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $G(\boldsymbol{\epsilon}^{(\boldsymbol{1})}, \boldsymbol{\epsilon}^{(\boldsymbol{2})}) = \tilde{G}(\boldsymbol{\epsilon}^{(\boldsymbol{1})}, \boldsymbol{\epsilon}^{(\boldsymbol{2})})$ almost surely.

4.2.1 Shared Frailty

The non-parametric shared frailty model is the simplest one with one common frailty variable ϵ that is shared between the two individuals in a pair. So, we have $\epsilon_j^{(k)} = \epsilon$, for all $j = 1, \cdots, L_k$ and $k = 1, 2$. This model enjoys the flexibility of allowing two different sets of competing risks for the two individuals. Formally, the model is given by

$$\lambda_j^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}|\epsilon) = h_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}})\epsilon, \tag{4.3}$$

for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Here the shared frailty $\epsilon > 0$ is assumed to have an arbitrary distribution function G defined on $[0, \infty)$. So, the class $\mathcal{G}$ of the arbitrary frailty distributions is given by

$$\mathcal{G} = \{G(\epsilon) : \epsilon > 0\}.$$

The conditional survival function of kth individual, given the shared frailty ϵ, is

$$S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}|\epsilon) = \exp\Big[-\epsilon \sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) \Big] = \exp\Big[-\epsilon H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) \Big],$$

where $H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}}) = \sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k; \boldsymbol{\xi^{(k)}})$, for all $t_k > 0$ and $k = 1, 2$. Also, the conditional jth sub-distribution function of kth individual, given ϵ, is

$$\begin{aligned} F_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}|\epsilon) &= \int_0^{t_k} \epsilon h_{0j}^{(k)}(u_k; \boldsymbol{\xi_j^{(k)}}) \exp\Big[-\epsilon H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \Big] du_k \\ &= \int_0^{t_k} \epsilon a(\gamma_j^{(k)}, \alpha_j^{(k)}) u_k^{\gamma_j^{(k)}-1} b(u_k; \gamma_j^{(k)}, \alpha_j^{(k)}) \exp\Big[-\epsilon H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \Big] du_k, \end{aligned}$$

using (3.2), for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Following (4.2), the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\xi}, G)$ is given by

$$\int_0^{\infty} \int_0^{t_1} \int_0^{t_2} \epsilon^2 \prod_{k=1}^{2} \left(a(\gamma_{j_k}^{(k)}, \alpha_{j_k}^{(k)}) u_k^{\gamma_{j_k}^{(k)}-1} b(u_k; \gamma_{j_k}^{(k)}, \alpha_{j_k}^{(k)}) \right) \times \exp\Big[-\epsilon \sum_{k=1}^{2} H_0^{(k)}(u_k; \boldsymbol{\xi^{(k)}}) \Big] du_2 du_1 dG(\epsilon).$$

Similarly, the unconditional jth sub-distribution function for kth individual is given

by

$$F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)=\int_0^{\infty}\int_0^{t_k}\epsilon a(\gamma_j^{(k)},\alpha_j^{(k)})u_k^{\gamma_j^{(k)}-1}b(u_k;\gamma_j^{(k)},\alpha_j^{(k)})\times\exp\Big[-\epsilon H_0^{(k)}(u_k;\boldsymbol{\xi^{(k)}})\Big]du_k dG(\epsilon),$$

with the corresponding unconditional sub-density function

$$f_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)=\int_0^{\infty}\epsilon a(\gamma_j^{(k)},\alpha_j^{(k)})t_k^{\gamma_j^{(k)}-1}b(t_k;\gamma_j^{(k)},\alpha_j^{(k)})\times\exp\Big[-\epsilon H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\Big]dG(\epsilon). \tag{4.4}$$

Definition 4.2. The non-parametric shared frailty model (4.3) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi}\times\mathcal{G}$ if, for some $\boldsymbol{\xi},\ \tilde{\boldsymbol{\xi}}\in\ \boldsymbol{\Xi}$ and $G,\ \tilde{G}\in\ \mathcal{G}=\{G(\epsilon):\epsilon>0\}$, as defined in the beginning of this subsection, the equality

$$F_{j_1j_2}\big(t_1,t_2;\boldsymbol{\xi},G\big)=F_{j_1j_2}\big(t_1,t_2;\tilde{\boldsymbol{\xi}},\tilde{G}\big),$$

for all $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$, implies $\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}$ and $G(\epsilon)=\tilde{G}(\epsilon)$ almost surely.

Theorem 4.1. The non-parametric shared frailty model (4.3) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi}\times\mathcal{G}$, provided $\lim_{t\to\infty}H_0^{(k)}(t;\boldsymbol{\xi^{(k)}})=\infty$ for all $\boldsymbol{\xi^{(k)}}$, $k=1,2$, and $\mathbb{E}(\epsilon)=\int_0^{\infty}\epsilon dG(\epsilon)=1$.

Proof. From the equality of the joint unconditional sub-distribution functions as in **Definition 4.2**, for all $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$, we get equality of the unconditional sub-distribution functions $F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ and $F_j^{(k)}(t_k;\tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}},\tilde{G})$; in particular, we get equality of the corresponding unconditional sub-density functions $f_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ and $f_j^{(k)}(t_k;\tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}},\tilde{G})$. From the expression above in (4.4), we have

$$\begin{aligned}&\int_0^{\infty}\epsilon a(\gamma_j^{(k)},\alpha_j^{(k)})t_k^{\gamma_j^{(k)}-1}b(t_k;\gamma_j^{(k)},\alpha_j^{(k)})\exp\left[-\epsilon H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right]dG(\epsilon)\\&=\int_0^{\infty}\epsilon a(\tilde{\gamma}_j^{(k)},\tilde{\alpha}_j^{(k)})t_k^{\tilde{\gamma}_j^{(k)}-1}b(t_k;\tilde{\gamma}_j^{(k)},\tilde{\alpha}_j^{(k)})\exp\left[-\epsilon H_0^{(k)}(t_k;\tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}})\right]d\tilde{G}(\epsilon)\end{aligned}$$

for all $t_k>0,\ j=1,\cdots,L_k$ and $k=1,2$.

Following the technique used by Heckman and Singer (1984a) to show identifiability of shape parameters, we rearrange the above equation to write

$$t_k^{\gamma_j^{(k)}-\tilde{\gamma}_j^{(k)}} \frac{a(\gamma_j^{(k)},\alpha_j^{(k)})b(t_k;\gamma_j^{(k)},\alpha_j^{(k)})A(t_k;\boldsymbol{\xi^{(k)}},G)}{a(\tilde{\gamma}_j^{(k)},\tilde{\alpha}_j^{(k)})b(t_k;\tilde{\gamma}_j^{(k)},\tilde{\alpha}_j^{(k)})A(t_k;\tilde{\boldsymbol{\xi}}^{(k)},\tilde{G})} = 1, \tag{4.5}$$

where $A(t_k;\boldsymbol{\xi^{(k)}},G) = \int_0^\infty \epsilon \exp\left[-\epsilon H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right] dG(\epsilon)$, for $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1,2$. Note that, by dominated convergence theorem, we have $\lim_{t_k\to 0+} A(t_k;\boldsymbol{\xi^{(k)}},G) = \int_0^\infty \epsilon dG(\epsilon) = 1$ by the assumption of the theorem. Similarly, we have $\lim_{t_k\to 0+} A(t_k;\tilde{\boldsymbol{\xi}}^{(k)},\tilde{G}) = \int_0^\infty \epsilon d\tilde{G}(\epsilon) = 1$.

Now, depending on whether $\gamma_j^{(k)} > \tilde{\gamma}_j^{(k)}$ or $\gamma_j^{(k)} < \tilde{\gamma}_j^{(k)}$, the limit of the left hand side of (4.5), as $t_k \to 0+$, is 0 or ∞, while the same limit of the right hand side of (4.5) is 1, which is a contradiction. Therefore, we have $\gamma_j^{(k)} = \tilde{\gamma}_j^{(k)}$, for $j = 1,\cdots,L_k$ and $k = 1,2$. Using this equality in (4.5) and re-arranging the terms again, we get

$$\frac{a(\gamma_j^{(k)},\tilde{\alpha}_j^{(k)})}{a(\gamma_j^{(k)},\alpha_j^{(k)})} = \frac{b(t_k;\gamma_j^{(k)},\alpha_j^{(k)})A(t_k;\boldsymbol{\xi^{(k)}},G)}{b(t_k;\tilde{\gamma}_j^{(k)},\tilde{\alpha}_j^{(k)})A(t_k;\tilde{\boldsymbol{\xi}}^{(k)},\tilde{G})}.$$

Now, as before, letting $t_k \to 0+$ in both sides of the above equation, we have

$$\frac{a(\gamma_j^{(k)},\tilde{\alpha}_j^{(k)})}{a(\gamma_j^{(k)},\alpha_j^{(k)})} = 1.$$

Therefore, using the one-to-one property of the function $a(\cdot,\cdot)$, we get $\alpha_j^{(k)} = \tilde{\alpha}_j^{(k)}$, for $j =,\cdots,L_k$ and $k = 1,2$. Therefore, we now have $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$.

Now, equality of the joint unconditional sub-distribution functions also implies equality of the unconditional survival functions $S^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G) = S^{(k)}(t_k;\tilde{\boldsymbol{\xi}}^{(k)},\tilde{G})$, with $\boldsymbol{\xi^{(k)}} = \tilde{\boldsymbol{\xi}}^{(k)}$, for all $t_k > 0$. Using the expression of the conditional survival function $S^{(k)}(t_k;\boldsymbol{\xi^{(k)}}|\epsilon)$ in the beginning of this subsection, we therefore have

$$\int_0^\infty \exp\Big[-\epsilon H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\Big] dG(\epsilon) = \int_0^\infty \exp\Big[-\epsilon H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\Big] d\tilde{G}(w),$$

for all $t_k > 0$. Since the cumulative baseline hazard functions are monotonically increasing with $H_0^{(k)}(0; \boldsymbol{\xi^{(k)}}) = 0$ and $\lim_{t\to\infty} H_0^{(k)}(t; \boldsymbol{\xi^{(k)}}) = \infty$, there exists exactly one time point t^*_{kn} such that $H_0^{(k)}(t^*_{kn}; \boldsymbol{\xi^{(k)}}) = n$ for each $n \in \mathbb{N}$, by intermediate value property. Now, taking limit as $t_k \to t^*_{kn}$ for each $n \in \mathbb{N}$ in the above equation, we get

$$\int_0^\infty \exp(-n\epsilon) dG(\epsilon) = \int_0^\infty \exp(-n\epsilon) d\tilde{G}(\epsilon) \text{ for each } n \in \mathbb{N}.$$

From the above equation, it follows that $G(\epsilon) = \tilde{G}(\epsilon)$ almost surely for all $\epsilon > 0$ by the uniqueness theorem regarding Laplace-Stieltjes transformation (See Section 1.10.1). Hence, the identifiability of this model is proved.

□

4.2.2 Correlated Frailty

Let $\epsilon^{(1)} > 0$ and $\epsilon^{(2)} > 0$ be two correlated frailty variables associated with failure times of the first and the second individual, respectively, regardless of the causes of failure. Dependence between the two failure times is modeled through the correlation between these two frailty variables. The joint distribution function of $(\epsilon^{(1)}, \epsilon^{(2)})$ is denoted by $G(\epsilon^{(1)}, \epsilon^{(2)})$ and defined on $[0, \infty) \times [0, \infty)$, so that the family $\mathcal{G}$ is now

$$\mathcal{G} = \{G(\epsilon^{(1)}, \epsilon^{(2)}) : \epsilon^{(1)} > 0, \epsilon^{(2)} > 0\}.$$

Let us denote the marginal distribution of $\epsilon^{(k)}$ as $G^{(k)}(\epsilon^{(k)})$, for $k = 1, 2$, which can be obtained as $G^{(1)}(\epsilon^{(1)}) = G(\epsilon^{(1)}, \infty)$ and $G^{(2)}(\epsilon^{(2)}) = G(\infty, \epsilon^{(2)})$. Formally, the non-parametric correlated frailty model is

$$\lambda_j^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}} | \epsilon^{(k)}) = h_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \epsilon^{(k)}, \tag{4.6}$$

for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. As before, the conditional survival function for the kth individual, given $\epsilon^{(k)}$, is

$$S^{(k)}(t_k; \boldsymbol{\xi^{(k)}} | \epsilon^{(k)}) = \exp\left[-\epsilon^{(k)} H_0^{(k)}\big(t_k; \boldsymbol{\xi^{(k)}}\big) \right]$$

and the conditional jth sub-distribution function of the kth individual, given $\epsilon^{(k)}$, is

$$F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}}|\epsilon^{(k)}) = \int_0^{t_k} \epsilon^{(k)} a(\gamma_j^{(k)},\alpha_j^{(k)}) u_k^{\gamma_j^{(k)}-1} b(u_k;\gamma_j^{(k)},\alpha_j^{(k)}) \exp\left[-\epsilon^{(k)} H_0^{(k)}(u_k;\boldsymbol{\xi^{(k)}})\right] du_k,$$

for all $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1,2$. Therefore, the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1,t_2;\boldsymbol{\xi},G)$ is given by

$$\int_0^\infty \int_0^\infty \int_0^{t_1} \int_0^{t_2} \prod_{k=1}^{2} \left(\epsilon^{(k)} a(\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)}) u_k^{\gamma_{j_k}^{(k)}-1} b(u_k;\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)}) \right) \times \exp\left[-\sum_{k=1}^{2} \epsilon^{(k)} H_0^{(k)}(u_k;\boldsymbol{\xi^{(k)}}) \right] du_2 du_1 dG(\epsilon^{(1)},\epsilon^{(2)}),$$

for all $t_k > 0,\ j_k = 1,\cdots,L_k$ and $k = 1,2$. Similarly, the unconditional jth sub-distribution function $F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ of the kth individual is

$$\int_0^\infty \int_0^{t_k} \epsilon^{(k)} a(\gamma_j^{(k)},\alpha_j^{(k)}) u_k^{\gamma_j^{(k)}-1} b(u_k;\gamma_j^{(k)},\alpha_j^{(k)}) \exp\left[-\epsilon^{(k)} H_0^{(k)}(u_k;\boldsymbol{\xi^{(k)}})\right] du_k dG^{(k)}(\epsilon^{(k)})$$

with the corresponding unconditional sub-density function $f_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ given by

$$\int_0^\infty \epsilon^{(k)} a(\gamma_j^{(k)},\alpha_j^{(k)}) t_k^{\gamma_j^{(k)}-1} b(t_k;\gamma_j^{(k)},\alpha_j^{(k)}) \exp\left[-\epsilon^{(k)} H_0^{(k)}(t_k;\boldsymbol{\xi^{(k)}})\right] dG^{(k)}(\epsilon^{(k)}),$$

for all $t_k > 0,\ j = 1,\cdots,L_k$ and $k = 1,2$.

Definition 4.3. The non-parametric correlated frailty model (4.6) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi}\times\mathcal{G}$ if, for some $\boldsymbol{\xi},\ \tilde{\boldsymbol{\xi}} \in\ \boldsymbol{\Xi}$ and $G,\ \tilde{G} \in\ \mathcal{G} = \{G(\epsilon^{(1)},\epsilon^{(2)}) : \epsilon^{(1)} > 0, \epsilon^{(2)} > 0\}$, as defined in the beginning of this subsection, the equality

$$F_{j_1 j_2}\big(t_1,t_2;\boldsymbol{\xi},G\big) = F_{j_1 j_2}\big(t_1,t_2;\tilde{\boldsymbol{\xi}},\tilde{G}\big),$$

for all $t_k > 0,\ j_k = 1,\cdots,L_k$ and $k = 1,2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $G(\epsilon^{(1)},\epsilon^{(2)}) = \tilde{G}(\epsilon^{(1)},\epsilon^{(2)})$ almost surely.

Theorem 4.2. The non-parametric correlated frailty model (4.6) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by

(3.2) is identifiable within $\boldsymbol{\Xi} \times \mathcal{G}$, provided $\lim_{t\to\infty} H_0^{(k)}(t; \boldsymbol{\xi^{(k)}}) = \infty$ for all $\boldsymbol{\xi^{(k)}}$ and $\mathbb{E}(\epsilon^{(k)}) = \int_0^\infty \epsilon^{(k)} dG^{(k)}(\epsilon^{(k)}) = 1$, for $k = 1, 2$.

Proof. As in the previous subsection, equating the unconditional jth sub-density functions $f_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, G)$ and $f_j^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{G})$, we get

$$\begin{aligned}
&\int_0^\infty \epsilon^{(k)} a(\gamma_j^{(k)}, \alpha_j^{(k)}) t_k^{\gamma_j^{(k)}-1} b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) \exp\left[-\epsilon^{(k)} H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\right] dG^{(k)}(\epsilon^{(k)}) \\
&\quad = \int_0^\infty \epsilon^{(k)} a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) t_k^{\tilde{\gamma}_j^{(k)}-1} b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) \exp\left[-\epsilon^{(k)} H_0^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}})\right] d\tilde{G}^{(k)}(\epsilon^{(k)}),
\end{aligned}$$

for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Rearranging the terms as in the previous subsection, we now have

$$t_k^{\gamma_j^{(k)} - \tilde{\gamma}_j^{(k)}} \frac{a(\gamma_j^{(k)}, \alpha_j^{(k)}) b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) A(t_k; \boldsymbol{\xi^{(k)}}, G^{(k)})}{a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) A(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{G}^{(k)})} = 1,$$

for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Note that this is the same identity as that in (4.5) in the previous subsection with G replaced by $G^{(k)}$. Therefore, following the same arguments as those used in the proof of **Theorem 4.1** and using the assumption of **Theorem 4.2**, we can prove that $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$.

Now, the equality of the unconditional joint sub-distribution functions implies equality of the unconditional joint survival functions $S(t_1, t_2; \boldsymbol{\xi}, G)$ and $S(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{G})$ with $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$, for all $t_k > 0$ and $k = 1, 2$. Using the expression of the conditional survival functions $S^{(k)}(t_k; \boldsymbol{\xi^{(k)}} | \epsilon^{(k)})$, for $k = 1, 2$, in the beginning of this subsection and using conditional independence, the above equality of the joint survival functions gives

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \exp\Big[-\sum_{k=1}^{2} \epsilon^{(k)} H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\Big] dG(\epsilon^{(1)}, \epsilon^{(2)}) \\
&\qquad\qquad = \int_0^\infty \int_0^\infty \exp\Big[-\sum_{k=1}^{2} \epsilon^{(k)} H_0^{(k)}(t_k; \boldsymbol{\xi^{(k)}})\Big] d\tilde{G}(\epsilon^{(1)}, \epsilon^{(2)}).
\end{aligned}$$

Since the cumulative baseline hazard functions are monotonically increasing with $H_0^{(k)}(0; \boldsymbol{\xi^{(k)}}) = 0$ and $\lim_{t\to\infty} H_0^{(k)}(t; \boldsymbol{\xi^{(k)}}) = \infty$, there exists exactly one point t_{kn}^*

such that $H_0^{(k)}(t_{kn}^*; \boldsymbol{\xi^{(k)}}) = n$, for every $n \in \mathbb{N}$ and for $k = 1, 2$, by intermediate value property. Therefore, taking limit as $t_k \to t_{kn}^*$ for every $n \in \mathbb{N}$ and for $k = 1, 2$ in the above equation, we get

$$\int_0^\infty \int_0^\infty \exp\left[-\sum_{k=1}^{2} \epsilon^{(k)} n\right] dG(\epsilon^{(1)}, \epsilon^{(2)}) = \int_0^\infty \int_0^\infty \exp\left[-\sum_{k=1}^{2} \epsilon^{(k)} n\right] d\tilde{G}(\epsilon^{(1)}, \epsilon^{(2)}).$$

Therefore, it follows from the uniqueness theorem regarding bivariate Laplace-Stieltjes transformation (See Section 1.10.1) that $G(\epsilon^{(1)}, \epsilon^{(2)}) = \tilde{G}(\epsilon^{(1)}, \epsilon^{(2)})$ almost surely for $\epsilon^{(k)} > 0$, for $k = 1, 2$. Hence, the identifiability of this model is proved. □

4.2.3 Shared Cause-specific Frailty

Unlike the previous two subsections, we now assume the non-parametric frailty to depend on the cause of failure, but shared by the two individuals in a pair (See Ghosh et al., 2024a,b). In this case, we need the set of competing risks to be the same for both the individuals since the shared frailty is cause-specific. So, we assume $L_1 = L_2 = L$, say. Let $\epsilon_j > 0$ denote the shared frailty for the jth cause, for $j = 1, \cdots, L$, and write $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_L)$. Dependence between the two individuals is modeled through the common frailty variables in $\boldsymbol{\epsilon}$. The model is given by

$$\lambda_j^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}} | \boldsymbol{\epsilon}) = h_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \epsilon_j, \tag{4.7}$$

for $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Let us write the joint distribution function of $\boldsymbol{\epsilon}$ as $G(\boldsymbol{\epsilon})$ and the corresponding marginal distribution of ϵ_j as $G_j(\epsilon_j)$, for $j = 1, \cdots, L$, which can be obtained from $G(\boldsymbol{\epsilon})$. In contrast with the shared cause-specific Gamma frailty model of Section 3.2.3, we do not need to assume independence between the ϵ_j's. So, the family of $\mathcal{G}$ of frailty distributions is defined as

$$\mathcal{G} = \{G(\boldsymbol{\epsilon}) : \epsilon_j > 0,\ j = 1, \cdots, L\}.$$

As before, the conditional survival function for the kth individual, given $\boldsymbol{\epsilon}$, is

$$S^{(k)}(t_k; \boldsymbol{\xi^{(k)}} | \boldsymbol{\epsilon}) = \exp\left[-\sum_{j=1}^{L} \epsilon_j H_{0j}^{(k)}\left(t_k; \boldsymbol{\xi_j^{(k)}}\right) \right]$$

and the conditional jth sub-distribution function of the kth individual, given $\boldsymbol{\epsilon}$, is

$$F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}}|\boldsymbol{\epsilon}) = \int_0^{t_k} \epsilon_j a(\gamma_j^{(k)},\alpha_j^{(k)}) u_k^{\gamma_j^{(k)}-1} b(u_k;\gamma_j^{(k)},\alpha_j^{(k)}) \exp\,[-\sum_{j'=1}^{L} \epsilon_{j'} H_{0j'}^{(k)}\big(t_k;\boldsymbol{\xi_{j'}^{(k)}}\big)] du_k,$$

for all $t_k > 0,\ j = 1,\cdots,L$ and $k = 1,2$. Therefore, the unconditional jth sub-distribution function $F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ of the kth individual is

$$\int_0^{\infty}\cdots\int_0^{\infty}\int_0^{t_k} \epsilon_j a(\gamma_j^{(k)},\alpha_j^{(k)}) u_k^{\gamma_j^{(k)}-1} b(u_k;\gamma_j^{(k)},\alpha_j^{(k)}) \exp\,[-\sum_{j'=1}^{L} \epsilon_{j'} H_{0j'}^{(k)}\big(u_k;\boldsymbol{\xi_{j'}^{(k)}}\big)] du_k dG(\boldsymbol{\epsilon})$$

with the corresponding unconditional sub-density function $f_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ given by

$$\int_0^{\infty}\cdots\int_0^{\infty} \epsilon_j a(\gamma_j^{(k)},\alpha_j^{(k)}) t_k^{\gamma_j^{(k)}-1} b(t_k;\gamma_j^{(k)},\alpha_j^{(k)}) \exp\,[-\sum_{j'=1}^{L} \epsilon_{j'} H_{0j'}^{(k)}\big(t_k;\boldsymbol{\xi_{j'}^{(k)}}\big)] dG(\boldsymbol{\epsilon})$$

for all $t_k > 0,\ j = 1,\cdots,L$ and $k = 1,2$. Similarly, the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1,t_2;\boldsymbol{\xi},G)$ is

$$\int_0^{\infty}\cdots\int_0^{\infty}\int_0^{t_1}\int_0^{t_2}\prod_{k=1}^{2}\left(\epsilon_{j_k} a(\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)}) u_k^{\gamma_{j_k}^{(k)}-1} b(u_k;\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)})\right) \times \exp\left[-\sum_{k=1}^{2}\sum_{j=1}^{L} \epsilon_j H_{0j}^{(k)}(u_k;\boldsymbol{\xi_j^{(k)}})\right] du_2 du_1 dG(\boldsymbol{\epsilon}),$$

and the corresponding unconditional joint sub-density function $f_{j_1,j_2}(t_1,t_2;\boldsymbol{\xi},G)$ is

$$\int_0^{\infty}\cdots\int_0^{\infty}\prod_{k=1}^{2}\left(\epsilon_{j_k} a(\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)}) t_k^{\gamma_{j_k}^{(k)}-1} b(t_k;\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)})\right) \exp\left[-\sum_{k=1}^{2}\sum_{j=1}^{L} \epsilon_j H_{0j}^{(k)}(t_k;\boldsymbol{\xi_j^{(k)}})\right] dG(\boldsymbol{\epsilon}),$$

for $t_k > 0,\ j_k = 1,\cdots,L$ and $k = 1,2$.

Definition 4.4. The non-parametric shared cause-specific frailty model (4.7) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi}\times\mathcal{G}$ if, for some $\boldsymbol{\xi},\ \tilde{\boldsymbol{\xi}} \in \boldsymbol{\Xi}$ and $G,\ \tilde{G} \in \mathcal{G} = \{G(\boldsymbol{\epsilon}) : \epsilon_j > 0, j = 1,\cdots,L\}$, as defined in the beginning of this

subsection, the equality

$$F_{j_1 j_2}\big(t_1, t_2; \boldsymbol{\xi}, G\big) = F_{j_1 j_2}\big(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{G}\big),$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $G(\boldsymbol{\epsilon}) = \tilde{G}(\boldsymbol{\epsilon})$ almost surely.

Theorem 4.3. The non-parametric shared cause-specific frailty model (4.7) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi} \times \mathcal{G}$, provided $\lim_{t\to\infty} H_{0j}^{(k)}(t; \boldsymbol{\xi_j^{(k)}}) = \infty$, for all $\boldsymbol{\xi_j^{(k)}}$, and $\mathbb{E}(\epsilon_j) = \int_0^\infty \epsilon_j dG_j(\epsilon_j) = 1$, for $j = 1, \cdots, L$ and $k = 1, 2$.

Proof. As before, equating the unconditional jth sub-density functions $f_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, G)$ and $f_j^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{(\boldsymbol{k})}, \tilde{G})$, we get

$$\begin{aligned}
&\int_0^\infty \cdots \int_0^\infty \epsilon_j a(\gamma_j^{(k)}, \alpha_j^{(k)}) t_k^{\gamma_j^{(k)}-1} b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) \exp\big[-\sum_{j'=1}^{L} \epsilon_{j'} H_{0j'}^{(k)}\big(t_k; \boldsymbol{\xi_{j'}^{(k)}}\big)\big] dG(\boldsymbol{\epsilon}) = \\
&\qquad \int_0^\infty \cdots \int_0^\infty \epsilon_j a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) t_k^{\tilde{\gamma}_j^{(k)}-1} b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) \exp\big[-\sum_{j'=1}^{L} \epsilon_{j'} H_{0j'}^{(k)}\big(t_k; \tilde{\boldsymbol{\xi}}_{\boldsymbol{j'}}^{(\boldsymbol{k})}\big)\big] d\tilde{G}(\boldsymbol{\epsilon}),
\end{aligned}$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Re-arranging the terms as in the previous subsection, we now have

$$t_k^{\gamma_j^{(k)} - \tilde{\gamma}_j^{(k)}} \frac{a(\gamma_j^{(k)}, \alpha_j^{(k)}) b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) B_j(t_k; \boldsymbol{\xi^{(k)}}, G)}{a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) B_j(t_k; \tilde{\boldsymbol{\xi}}^{(\boldsymbol{k})}, \tilde{G})} = 1,$$

where $B_j(t_k; \boldsymbol{\xi^{(k)}}, G) = \int_0^\infty \cdots \int_0^\infty \epsilon_j \exp\big[-\sum_{j'=1}^{L} \epsilon_{j'} H_{0j'}^{(k)}\big(t_k; \boldsymbol{\xi_{j'}^{(k)}}\big)\big] dG(\boldsymbol{\epsilon})$, for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Note that this is the same identity as that in (4.5) in Section 4.2.1 with the univariate G replaced by the L-variate G and $A(t_k; \boldsymbol{\xi^{(k)}}, G)$ replaced by $B_j(t_k; \boldsymbol{\xi^{(k)}}, G)$. However, as for $A(t_k; \boldsymbol{\xi^{(k)}}, G)$, one can prove, by using the dominated convergence theorem and the assumption of **Theorem 4.3**, that

$$\lim_{t_k \to 0+} B_j(t_k; \boldsymbol{\xi^{(k)}}, G) = \lim_{t_k \to 0+} B_j(t_k; \tilde{\boldsymbol{\xi}}^{(\boldsymbol{k})}, \tilde{G}) = 1.$$

Therefore, following the same arguments as those used in the proof of **Theorem 4.1** and using the assumption of **Theorem 4.3**, we can prove that $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$.

Now, the equality of the unconditional joint sub-distribution functions implies

equality of the unconditional joint survival functions $S(t_1, t_2; \boldsymbol{\xi}, G)$ and $S(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{G})$ with $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$, for all $t_k > 0$ and $k = 1, 2$. Using the expression of the conditional survival functions $S^{(k)}(t_k; \boldsymbol{\xi^{(k)}}|\boldsymbol{\epsilon})$, for $k = 1, 2$, in the beginning of this subsection and using conditional independence, the above equality of the unconditional joint survival functions gives

$$\int_0^\infty \cdots \int_0^\infty \exp\Big[- \sum_{j=1}^{L} \epsilon_j \sum_{k=1}^{2} H_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \Big] dG(\boldsymbol{\epsilon}) = \int_0^\infty \cdots \int_0^\infty \exp\Big[- \sum_{j=1}^{L} \epsilon_j \sum_{k=1}^{2} H_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \Big] d\tilde{G}(\boldsymbol{\epsilon}).$$

Now, let $\{m_{1n}\}$ be a sequence of positive, strictly increasing, real numbers and bounded above strictly by 1, for $n \in \mathbb{N}$. Since the cumulative baseline cause-specific hazard functions are monotonically increasing with $H_{0j}^{(k)}(0; \boldsymbol{\xi_j^{(k)}}) = 0$ and $\lim_{t\to\infty} H_{0j}^{(k)}(t; \boldsymbol{\xi_j^{(k)}}) = \infty$, for all $\boldsymbol{\xi_j^{(k)}}$ and for $j = 1, \cdots, L$, there exists exactly one point t_{kn}^* such that $H_{01}^{(k)}(t_{kn}^*; \boldsymbol{\xi_1^{(k)}}) = m_{1n}$, for every $n \in \mathbb{N}$ and for $k = 1, 2$, by intermediate value property. So, we can write $t_{kn}^* = H_{01}^{-(k)}(m_{1n}; \boldsymbol{\xi_1^{(k)}})$, where $H_{01}^{-(k)}(\cdot; \boldsymbol{\xi_1^{(k)}})$ denotes the inverse function of $H_{01}^{(k)}(\cdot; \boldsymbol{\xi_1^{(k)}})$, for $k = 1, 2$, which exist by the definition in (3.2). Let us define, for $j = 2, \cdots, L$ and for every $n \in \mathbb{N}$,

$$m_{jn} = \sum_{k=1}^{2} H_{0j}^{(k)}(t_{kn}^*; \boldsymbol{\xi_j^{(k)}}) = \sum_{k=1}^{2} H_{0j}^{(k)}(H_{01}^{-(k)}(m_{1n}; \boldsymbol{\xi_1^{(k)}}); \boldsymbol{\xi_j^{(k)}}).$$

These m_{jn}'s, for fixed j, are clearly a sequence of positive and strictly increasing real numbers. Also, since $m_{1n} < 1$ for every $n \in \mathbb{N}$, we have

$$m_{jn} < \sum_{k=1}^{2} H_{0j}^{(k)}(H_{01}^{-(k)}(1; \boldsymbol{\xi_1^{(k)}}); \boldsymbol{\xi_j^{(k)}}) = c_j, \quad \text{say,}$$

for $j = 2, \cdots, L$ and for every $n \in \mathbb{N}$. Therefore, we have

$$\sum_{n=1}^{\infty} \frac{1}{m_{jn}} = \infty,$$

for $j = 1, \cdots, L$. Then, taking limit as $t_k \to t_{kn}^*$ for every $n \in \mathbb{N}$ and for $k = 1, 2$ in

the above equality of the joint survival functions, we get

$$\int_0^\infty \cdots \int_0^\infty \exp\Big[-\sum_{j=1}^{L} \epsilon_j m_{jn}\Big] dG(\boldsymbol{\epsilon}) = \int_0^\infty \cdots \int_0^\infty \exp\Big[-\sum_{j=1}^{L} \epsilon_j m_{jn}\Big] d\tilde{G}(\boldsymbol{\epsilon}),$$

for every $n \in \mathbb{N}$. Hence $G(\boldsymbol{\epsilon}) = \tilde{G}(\boldsymbol{\epsilon})$ almost surely for $\boldsymbol{\epsilon} > \mathbf{0}$, by the uniqueness theorem regarding multivariate Laplace-Stieltjes transformation (See Section 1.10.1). Therefore, the identifiability of this model is proved.

□

4.2.4 Correlated Cause-specific Frailty

This is a further generalization of the frailty models we considered in the previous three subsections. As in Section 4.2.2, we have two correlated frailty variables for the two individuals; however, as in Section 4.2.3, these two frailties depend on the specific cause of failure. That is, for each cause, say the jth, there is a pair of frailty variables $\boldsymbol{\epsilon_j} = (\epsilon_j^{(1)}, \epsilon_j^{(2)})$ for the two individuals which are correlated as in Section 4.2.2. In this case also, we need the set of competing risks to be the same for both the individuals since the pair of frailties is cause-specific. So, we have $L_1 = L_2 = L$. Let us write $\boldsymbol{\epsilon^{(k)}} = (\epsilon_1^{(k)}, \cdots, \epsilon_L^{(k)})$ denoting the frailties for the kth individual, for $k = 1, 2$. Dependence between the two individuals is modeled through the pairs $\boldsymbol{\epsilon_j}$'s for $j = 1, \cdots, L$. Formally, the model is given by

$$\lambda_j^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}} | \boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \epsilon_j^{(k)}, \tag{4.8}$$

for $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Let us write the joint distribution function of $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}})$ as $G(\boldsymbol{\epsilon})$ and the corresponding marginal distribution of the frailty vector $\boldsymbol{\epsilon^{(k)}}$ as $G^{(k)}(\boldsymbol{\epsilon^{(k)}})$, for $k = 1, 2$, which can be obtained from $G(\boldsymbol{\epsilon})$. For that matter, let $G_j^{(k)}(\epsilon_j^{(k)})$ denote the marginal distribution of the frailty variable $\epsilon_j^{(k)}$, for all j and k. Here also, we do not need to assume independence between the $\epsilon_j^{(k)}$'s for different j and k. So, the family of $\mathcal{G}$ of frailty distributions is defined as

$$\mathcal{G} = \{G(\boldsymbol{\epsilon}) : \epsilon_j^{(k)} > 0,\ j = 1, \cdots, L,\ k = 1, 2\}.$$

The conditional survival function for the kth individual, given the frailty vector

$\boldsymbol{\epsilon^{(k)}}$, is

$$S^{(k)}(t_k;\boldsymbol{\xi^{(k)}}|\boldsymbol{\epsilon^{(k)}}) = \exp\left[-\sum_{j=1}^{L}\epsilon_j^{(k)}H_{0j}^{(k)}\big(t_k;\boldsymbol{\xi_j^{(k)}}\big)\right]$$

and the conditional jth sub-distribution function of the kth individual, given $\boldsymbol{\epsilon^{(k)}}$, is

$$F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}}|\boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k}\epsilon_j^{(k)}a(\gamma_j^{(k)},\alpha_j^{(k)})u_k^{\gamma_j^{(k)}-1}b(u_k;\gamma_j^{(k)},\alpha_j^{(k)})\exp\big[-\sum_{j'=1}^{L}\epsilon_{j'}^{(k)}H_{0j'}^{(k)}\big(t_k;\boldsymbol{\xi_{j'}^{(k)}}\big)\big]du_k,$$

for all $t_k > 0,\ j = 1,\cdots,L$ and $k = 1,2$. Therefore, the unconditional jth sub-distribution function $F_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ of the kth individual is

$$\int_0^{\infty}\cdots\int_0^{\infty}\int_0^{t_k}\epsilon_j^{(k)}a(\gamma_j^{(k)},\alpha_j^{(k)})u_k^{\gamma_j^{(k)}-1}b(u_k;\gamma_j^{(k)},\alpha_j^{(k)})\exp\big[-\sum_{j'=1}^{L}\epsilon_{j'}^{(k)}H_{0j'}^{(k)}\big(u_k;\boldsymbol{\xi_{j'}^{(k)}}\big)\big]du_k dG^{(k)}(\boldsymbol{\epsilon^{(k)}})$$

with the corresponding unconditional sub-density function $f_j^{(k)}(t_k;\boldsymbol{\xi^{(k)}},G)$ given by

$$\int_0^{\infty}\cdots\int_0^{\infty}\epsilon_j^{(k)}a(\gamma_j^{(k)},\alpha_j^{(k)})t_k^{\gamma_j^{(k)}-1}b(t_k;\gamma_j^{(k)},\alpha_j^{(k)})\exp\big[-\sum_{j'=1}^{L}\epsilon_{j'}^{(k)}H_{0j'}^{(k)}\big(t_k;\boldsymbol{\xi_{j'}^{(k)}}\big)\big]dG^{(k)}(\boldsymbol{\epsilon^{(k)}}),$$

for all $t_k > 0,\ j = 1,\cdots,L$ and $k = 1,2$. Similarly, the unconditional joint sub-distribution function $F_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},G)$ is

$$\int_0^{\infty}\cdots\int_0^{\infty}\int_0^{t_1}\int_0^{t_2}\prod_{k=1}^{2}\left(\epsilon_{j_k}^{(k)}a(\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)})u_k^{\gamma_{j_k}^{(k)}-1}b(u_k;\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)})\right)$$
$$\times\exp\left[-\sum_{k=1}^{2}\sum_{j=1}^{L}\epsilon_j^{(k)}H_{0j}^{(k)}(u_k;\boldsymbol{\xi_j^{(k)}})\right]du_2du_1dG(\boldsymbol{\epsilon}),$$

and the corresponding unconditional joint sub-density function $f_{j_1j_2}(t_1,t_2;\boldsymbol{\xi},G)$ is

$$\int_0^{\infty}\cdots\int_0^{\infty}\prod_{k=1}^{2}\left(\epsilon_{j_k}^{(k)}a(\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)})t_k^{\gamma_{j_k}^{(k)}-1}b(t_k;\gamma_{j_k}^{(k)},\alpha_{j_k}^{(k)})\right)\exp\left[-\sum_{k=1}^{2}\sum_{j=1}^{L}\epsilon_j^{(k)}H_{0j}^{(k)}(t_k;\boldsymbol{\xi_j^{(k)}})\right]dG(\boldsymbol{\epsilon}),$$

for $t_k > 0,\ j_k = 1,\cdots,L$ and $k = 1,2$.

Definition 4.5. The non-parametric correlated cause-specific frailty model (4.8) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi}\times\mathcal{G}$ if, for some $\boldsymbol{\xi},\ \tilde{\boldsymbol{\xi}}\in\ \boldsymbol{\Xi}$ and

$G,\ \tilde{G} \in\ \mathcal{G} = \{G(\boldsymbol{\epsilon}) : \epsilon_j^{(k)} > 0, j = 1, \cdots, L, k = 1, 2\}$, as defined in the beginning of this subsection, the equality

$$F_{j_1 j_2}\big(t_1, t_2; \boldsymbol{\xi}, G\big) = F_{j_1 j_2}\big(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{G}\big),$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, implies $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$ and $G(\boldsymbol{\epsilon}) = \tilde{G}(\boldsymbol{\epsilon})$ almost surely.

Theorem 4.4. The non-parametric correlated cause-specific frailty model (4.8) for bivariate failure time data with competing risks with the baseline cause-specific hazards given by (3.2) is identifiable within $\boldsymbol{\Xi} \times \mathcal{G}$, provided $\lim_{t\to\infty} H_{0j}^{(k)}(t; \boldsymbol{\xi_j^{(k)}}) = \infty$, for all $\boldsymbol{\xi_j^{(k)}}$, and $\mathbb{E}(\epsilon_j^{(k)}) = \int_0^\infty \epsilon_j^{(k)} dG_j^{(k)}(\epsilon_j^{(k)}) = 1$, for $j = 1, \cdots, L$ and $k = 1, 2$.

Proof. As in the previous sections, equating the unconditional jth sub-density functions $f_j^{(k)}(t_k; \boldsymbol{\xi^{(k)}}, G)$ and $f_j^{(k)}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{G})$, we get

$$\int_0^\infty \cdots \int_0^\infty \epsilon_j^{(k)} a(\gamma_j^{(k)}, \alpha_j^{(k)}) t_k^{\gamma_j^{(k)}-1} b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) \exp\big[-\sum_{j'=1}^{L} \epsilon_{j'}^{(k)} H_{0j'}^{(k)}\big(t_k; \boldsymbol{\xi_{j'}^{(k)}}\big)\big] dG^{(k)}(\boldsymbol{\epsilon^{(k)}})$$
$$= \int_0^\infty \cdots \int_0^\infty \epsilon_j^{(k)} a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) t_k^{\tilde{\gamma}_j^{(k)}-1} b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) \exp\big[-\sum_{j'=1}^{L} \epsilon_{j'}^{(k)} H_{0j'}^{(k)}\big(t_k; \tilde{\boldsymbol{\xi}}_{\boldsymbol{j'}}^{\boldsymbol{(k)}}\big)\big] d\tilde{G}^{(k)}(\boldsymbol{\epsilon^{(k)}}),$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Rearranging the terms as in the previous subsection, we can write

$$t_k^{\gamma_j^{(k)} - \tilde{\gamma}_j^{(k)}} \frac{a(\gamma_j^{(k)}, \alpha_j^{(k)}) b(t_k; \gamma_j^{(k)}, \alpha_j^{(k)}) C_{jk}(t_k; \boldsymbol{\xi^{(k)}}, G^{(k)})}{a(\tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) b(t_k; \tilde{\gamma}_j^{(k)}, \tilde{\alpha}_j^{(k)}) C_{jk}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{G}^{(k)})} = 1,$$

where $C_{jk}(t_k; \boldsymbol{\xi^{(k)}}, G^{(k)}) = \int_0^\infty \cdots \int_0^\infty \epsilon_j^{(k)} \exp\big[-\sum_{j'=1}^{L} \epsilon_{j'}^{(k)} H_{0j'}^{(k)}\big(t_k; \boldsymbol{\xi_{j'}^{(k)}}\big)\big] dG^{(k)}(\boldsymbol{\epsilon^{(k)}})$, for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. This is again the same identity as that in (4.5) in Section 4.2.1 with the univariate G replaced by the L-variate $G^{(k)}$ and $A(t_k; \boldsymbol{\xi^{(k)}}, G)$ replaced by $C_{jk}(t_k; \boldsymbol{\xi^{(k)}}, G^{(k)})$. However, as for $A(t_k; \boldsymbol{\xi^{(k)}}, G)$, one can prove, by using the dominated convergence theorem and the assumption of **Theorem 4.5**, that

$$\lim_{t_k \to 0+} C_{jk}(t_k; \boldsymbol{\xi^{(k)}}, G^{(k)}) = \lim_{t_k \to 0+} C_{jk}(t_k; \tilde{\boldsymbol{\xi}}^{\boldsymbol{(k)}}, \tilde{G}^{(k)}) = 1.$$

Therefore, following the same arguments as those used in the proof of **Theorem 4.1** and using the assumption of **Theorem 4.5**, we can prove that $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$.

Equality of the unconditional joint sub-distribution functions implies equality of the unconditional joint survival functions $S(t_1, t_2; \boldsymbol{\xi}, G)$ and $S(t_1, t_2; \tilde{\boldsymbol{\xi}}, \tilde{G})$ with $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}$, for all $t_k > 0$ and $k = 1, 2$. Using the expression of the conditional survival functions $S^{(k)}(t_k; \boldsymbol{\xi^{(k)}} | \boldsymbol{\epsilon^{(k)}})$, for $k = 1, 2$, in the beginning of this subsection and using conditional independence, the above equality of the joint survival functions gives

$$\int_0^\infty \cdots \int_0^\infty \exp\Big[- \sum_{j=1}^{L} \sum_{k=1}^{2} \epsilon_j^{(k)} H_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \Big] dG(\boldsymbol{\epsilon}) = \int_0^\infty \cdots \int_0^\infty \exp\Big[- \sum_{j=1}^{L} \sum_{k=1}^{2} \epsilon_j^{(k)} H_{0j}^{(k)}(t_k; \boldsymbol{\xi_j^{(k)}}) \Big] d\tilde{G}(\boldsymbol{\epsilon}).$$

As in the previous subsection, let $\{m_{1n}^{(k)}\}$ be a sequence of positive, strictly increasing, real numbers and bounded above strictly by 1, for $k = 1, 2$ and $n \in \mathbb{N}$. Since the cumulative baseline cause-specific hazard functions are monotonically increasing with $H_{0j}^{(k)}(0; \boldsymbol{\xi_j^{(k)}}) = 0$ and $\lim_{t \to \infty} H_{0j}^{(k)}(t; \boldsymbol{\xi_j^{(k)}}) = \infty$, for $j = 1, \cdots, L$, there exists exactly one point t_{kn}^* such that $H_{01}^{(k)}(t_{kn}^*; \boldsymbol{\xi_1^{(k)}}) = m_{1n}^{(k)}$, for every $n \in \mathbb{N}$ and for $k = 1, 2$, by intermediate value property. For $k = 1, 2$, let us define

$$m_{jn}^{(k)} = H_{0j}^{(k)}(t_{kn}^*; \boldsymbol{\xi_j^{(k)}}) = H_{0j}^{(k)}(H_{01}^{-(k)}(m_{1n}^{(k)}; \boldsymbol{\xi_1^{(k)}}); \boldsymbol{\xi_j^{(k)}}),$$

for $j = 2, \cdots, L$ and $n \in \mathbb{N}$. These $m_{jn}^{(k)}$'s, for fixed j and k, are clearly a sequence of positive and strictly increasing real numbers. Also, since $m_{1n}^{(k)} < 1$ for every $n \in \mathbb{N}$, we have

$$m_{jn}^{(k)} < H_{0j}^{(k)}(H_{01}^{-(k)}(1; \boldsymbol{\xi_1^{(k)}}); \boldsymbol{\xi_j^{(k)}}) = d_j^{(k)}, \quad \text{say,}$$

for $j = 2, \cdots, L$ and for every $n \in \mathbb{N}$. Therefore, we have

$$\sum_{n=1}^{\infty} \frac{1}{m_{jn}^{(k)}} = \infty,$$

for $j = 1, \cdots, L$ and $k = 1, 2$. Then, taking limit as $t_k \to t_{kn}^*$ for every $n \in \mathbb{N}$ and for $k = 1, 2$ in the above equality of the joint survival functions, we get

$$\int_0^\infty \cdots \int_0^\infty \exp\Big[- \sum_{j=1}^{L} \sum_{k=1}^{2} \epsilon_j^{(k)} m_{jn}^{(k)} \Big] dG(\boldsymbol{\epsilon}) = \int_0^\infty \cdots \int_0^\infty \exp\Big[- \sum_{j=1}^{L} \sum_{k=1}^{2} \epsilon_j^{(k)} m_{jn}^{(k)} \Big] d\tilde{G}(\boldsymbol{\epsilon}),$$

for every $n \in \mathbb{N}$. Hence $G(\boldsymbol{\epsilon}) = \tilde{G}(\boldsymbol{\epsilon})$ almost surely for $\boldsymbol{\epsilon} > \mathbf{0}$, by the uniqueness

theorem regarding multivariate Laplace-Stieltjes transformation (See Section 1.10.1). Therefore, the identifiability of this model is proved.

□

4.3 Method of Estimation

In this section, we propose a method of estimating the parameter vector $\boldsymbol{\xi}$ associated with the parametric baseline cause-specific hazard functions and the non-parametric frailty distribution G by means of maximizing the corresponding likelihood function. From the construction of the likelihood function for bivariate current status data with competing risks in Section 2.6 (and also in Chapter 3), it is evident that each likelihood contribution arises from the density of a mixture distribution in which the frailty distribution G plays the role of the mixing distribution. Therefore, we utilize the technique of estimating mixture distribution G, as described in Section 1.10.3.

In this context of semi-parametric modeling, in particular, we need to estimate both $\boldsymbol{\xi}$ and G. Following Wang (2010) (See also Lindsay and Lesperance, 1995), we estimate G for given $\boldsymbol{\xi}$ and then estimate $\boldsymbol{\xi}$ for given G alternatively to increase the overall likelihood at each iteration. In order to keep the description of the estimation method simple, we consider only the shared frailty model in the following discussion. Some indication of how to extend the methodology to other three more general frailty models is given in Section 4.8.

Recall from Section 2.2 that the observed data is of the form (X_1, X_2, J_1, J_2). The observed data for the ith pair is denoted by $(x_{i1}, x_{i2}, j_{i1}, j_{i2})$ for $i = 1, \cdots, n$. The parameter vector that is needed to be estimated is $\boldsymbol{\eta} = \{\boldsymbol{\xi}, G\}$, where $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ and G is the distribution function of the shared frailty variable ϵ belonging to $\mathcal{G} = \{G(\epsilon) : \epsilon > 0, \int_0^\infty \epsilon dG(\epsilon) = 1\}$, in view of the identifiability result in 4.2.1. Following the likelihood construction in Section 2.6, we can write the likelihood function in this setup as

$$L(\boldsymbol{\xi}, G) \propto \prod_{i=1}^{n} f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{\xi}, G),$$

where each likelihood contribution can be written as

$$f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{\xi}, G) = \int_0^\infty f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{\xi}|\epsilon) dG(\epsilon),$$

for $i = 1, \cdots, n$. The expressions for these integrands $f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{\xi}|\epsilon)$, the conditional likelihood contributions given the frailty variable ϵ, can be easily obtained

from the individual contributions worked out in Section 2.6 (See also Section 4.2.1). For example, with $j_k = 1, \cdots, L_k$ and $k = 1, 2$, we have

$$f^*(x_1, x_2, j_1, j_2; \boldsymbol{\xi}|\epsilon) \propto F_{j_1 j_2}(x_1, x_2; \boldsymbol{\xi}|\epsilon),$$

for a typical observation (x_1, x_2, j_1, j_2). This brings us to the semi-parametric mixture framework of Lindsay and Lesperance (1995) and Wang (2010) to estimate $\boldsymbol{\xi}$ and G alternatively.

Note that the log-likelihood function is

$$\ell(\boldsymbol{\xi}, G) \propto \sum_{i=1}^{n} \log \left[f^*(x_{i1}, x_{i2}, j_{i1}, j_{i2}; \boldsymbol{\xi}, G) \right].$$

We need to find the semi-parametric maximum likelihood estimator $(\hat{\boldsymbol{\xi}}, \hat{G})$ of $(\boldsymbol{\xi}, G)$ which maximizes this log-likelihood function $\ell(\boldsymbol{\xi}, G)$. From Section 1.10.3, for every fixed $\boldsymbol{\xi}$, the non-parametric MLE of the mixing distribution G is a discrete distribution with number of mass points not exceeding the number of distinct observations in the observed data. For such a G, let us suppose that there are K mass points $\theta_1, \cdots, \theta_K$ with masses $\pi_1, \cdots, \pi_K$, respectively, where $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_K)$ and $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_K)$ and also the dimension $K (\leq n)$ are all unknown and to be estimated. Then, we can write the density $f^*(x_1, x_2, j_1, j_2; \boldsymbol{\xi}, G)$ for a typical observation (x_1, x_2, j_1, j_2) as

$$f^*(x_1, x_2, j_1, j_2; \boldsymbol{\xi}, G) = \sum_{k=1}^{K} f^*(x_1, x_2, j_1, j_2; \boldsymbol{\xi}|\epsilon = \theta_k)\pi_k,$$

and the log-likelihood function can be written as

$$\ell(\boldsymbol{\xi}, G) \propto \sum_{i=1}^{n} \log \left[\sum_{k=1}^{K} f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_k)\pi_k \right].$$

This log-likelihood may now be written as $\ell(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi})$ to reflect the explicit dependence on the parameter vector $\boldsymbol{\eta} = (\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi})$, including the dimension K. The individual conditional likelihood contributions $f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_k)$, given $\epsilon = \theta_k$, worked out from Section 2.6 and 4.2.1.

For a fixed $\boldsymbol{\xi}$, we use a slight modification of the fast algorithm proposed by Wang (2007) based on directional derivative for the computation of the non-parametric MLE of G, or $(\boldsymbol{\theta}, \boldsymbol{\pi})$ including K. For this purpose, we consider two distribution functions $G = (\boldsymbol{\theta}, \boldsymbol{\pi})$ and δ_θ, where δ_θ is the degenerate distribution function having mass 1 at

point θ, $\theta > 0$. The directional derivative of G in the direction of δ_θ (also called the gradient function) is defined as

$$d(\delta_\theta, G, \boldsymbol{\xi}) = \frac{\partial \ell\{\boldsymbol{\xi}, (1-\omega)G + \omega\delta_\theta\}}{\partial \omega}|_{\omega=0} = \sum_{i=1}^{n} \left[\frac{f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta)}{\sum_{k=1}^{K} f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta)\pi_k} - 1 \right].$$

Characterization of semi-parametric MLE (Wang, 2010): If the log-likelihood function $\ell(\boldsymbol{\xi}, G)$ is bounded for all $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ and $G \in \mathcal{G}$, then a semi-parametric MLE $(\hat{\boldsymbol{\xi}}, \hat{G})$ must satisfy the following conditions:

$$\frac{\partial \ell(\boldsymbol{\xi}, \hat{G})}{\partial \boldsymbol{\xi}}|_{\boldsymbol{\xi}=\hat{\boldsymbol{\xi}}} = \mathbf{0}, \text{ and} \tag{4.9}$$

$$\sup_{\theta \in \mathbb{R}^+} d(\delta_\theta, \hat{G}, \hat{\boldsymbol{\xi}}) = 0. \tag{4.10}$$

In order to describe the algorithm to non-parametrically estimate G, or $(\boldsymbol{\theta}, \boldsymbol{\pi})$ including K, for fixed $\boldsymbol{\xi}$, let us write

$$\begin{aligned} s_i(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi}) &= \frac{\partial \log \left\{ \sum_{k=1}^{K} f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_k)\pi_k \right\}}{\partial \boldsymbol{\pi}} \\ &= \left(\frac{f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_1)}{\sum_{k=1}^{K} f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_k)\pi_k}, \cdots, \frac{f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_K)}{\sum_{k=1}^{K} f^*(x_{1i}, x_{2i}, j_{1i}, j_{2i}; \boldsymbol{\xi}|\epsilon = \theta_k)\pi_k} \right), \end{aligned}$$

for $i = 1, \cdots, n$. Let us denote

$$S^T = S^T(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi}) = \Big(s_1(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi}), \cdots, s_n(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi}) \Big)$$

Therefore, the score vector corresponding to the parameter vector $\boldsymbol{\pi}$ is

$$\frac{\partial \ell(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi})}{\partial \boldsymbol{\pi}} = \sum_{i=1}^{n} s_i(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\pi}) = S^T \mathbf{1},$$

where $\mathbf{1} = (1, \cdots, 1)$, a vector of 1's. It can be shown that

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\pi} \partial \boldsymbol{\pi}^T} = -S^T S$$

Following Atwood (1976), we approximate ℓ by Taylor series expansion with respect

to $\boldsymbol{\pi}$ up to the quadratic term to get

$$\begin{aligned}\ell(\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\pi'}) &\approx \ell(\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\pi}) + \Big(\frac{\partial \ell}{\partial \boldsymbol{\pi'}}|_{\boldsymbol{\pi'}=\boldsymbol{\pi}}\Big)^T(\boldsymbol{\pi'}-\boldsymbol{\pi}) + \frac{1}{2}(\boldsymbol{\pi'}-\boldsymbol{\pi})^T \frac{\partial^2 \ell}{\partial \boldsymbol{\pi'}\partial \boldsymbol{\pi'}^T}|_{\boldsymbol{\pi'}=\boldsymbol{\pi}}(\boldsymbol{\pi'}-\boldsymbol{\pi}) \\ &= \ell(\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\pi}) + \mathbf{1}^T S(\boldsymbol{\pi'}-\boldsymbol{\pi}) - \frac{1}{2}(\boldsymbol{\pi'}-\boldsymbol{\pi})^T S^T S(\boldsymbol{\pi'}-\boldsymbol{\pi})\end{aligned}$$

Therefore, writing $Q(\boldsymbol{\pi'}|\boldsymbol{\pi},\boldsymbol{\theta},\boldsymbol{\xi}) = \ell(\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\pi}) - \ell(\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\pi'})$, we have

$$Q(\boldsymbol{\pi'}|\boldsymbol{\pi},\boldsymbol{\theta},\boldsymbol{\xi}) \approx -\mathbf{1}^T S(\boldsymbol{\pi'}-\boldsymbol{\pi}) + \frac{1}{2}(\boldsymbol{\pi'}-\boldsymbol{\pi})^T S^T S(\boldsymbol{\pi'}-\boldsymbol{\pi}).$$

Let us write $\boldsymbol{\psi} = \boldsymbol{\pi'} - \boldsymbol{\pi}$. Then,

$$\begin{aligned}Q(\boldsymbol{\pi'}|\boldsymbol{\pi},\boldsymbol{\theta},\boldsymbol{\xi}) &\approx -\mathbf{1}^T S\boldsymbol{\psi} + \frac{1}{2}\boldsymbol{\psi}^T S^T S\boldsymbol{\psi} \\ &= \frac{1}{2}(S\boldsymbol{\psi})^T S\boldsymbol{\psi} - \mathbf{1}^T S\boldsymbol{\psi} \\ &= \frac{1}{2}\Bigg[\Big\{S\boldsymbol{\psi} - \mathbf{1} + \mathbf{1}\Big\}^T\Big\{S\boldsymbol{\psi} - \mathbf{1} + \mathbf{1}\Big\}\Bigg] - \mathbf{1}^T S\boldsymbol{\psi} \\ &= \frac{1}{2}\Bigg[\Big(S\boldsymbol{\psi} - \mathbf{1}\Big)^T\Big(S\boldsymbol{\psi} - \mathbf{1}\Big) + \Big(S\boldsymbol{\psi} - \mathbf{1}\Big)^T\mathbf{1} + \mathbf{1}^T\Big(S\boldsymbol{\psi} - \mathbf{1}\Big) + \mathbf{1}^T\mathbf{1}\Bigg] - \mathbf{1}^T S\boldsymbol{\psi} \\ &= \frac{1}{2}||S\boldsymbol{\psi} - \mathbf{1}||^2 - \frac{n}{2} \\ &= \frac{1}{2}||S(\boldsymbol{\pi'}-\boldsymbol{\pi}) - \mathbf{1}||^2 - \frac{n}{2}\end{aligned}$$

It can be shown that $S\boldsymbol{\pi} = \mathbf{1}$. Substituting this in the above equation, we get

$$Q(\boldsymbol{\pi'}|\boldsymbol{\pi},\boldsymbol{\theta},\boldsymbol{\xi}) = \frac{1}{2}||S\boldsymbol{\pi'} - \mathbf{2}||^2 - \frac{n}{2}.$$

For a fixed $\boldsymbol{\xi}$ and given $\boldsymbol{\theta}$, in order to obtain the next update $\boldsymbol{\pi'}$ of $\boldsymbol{\pi}$, we need to minimize $Q(\boldsymbol{\pi'}|\boldsymbol{\pi},\boldsymbol{\theta},\boldsymbol{\xi})$, or $||S\boldsymbol{\pi'}-\mathbf{2}||^2$, with respect to $\boldsymbol{\pi'}$ subject to three constraints, namely, (i) $\pi_k \geq 0$, for all $k = 1, \cdots, K$, (ii) $\sum_{k=1}^K \pi_k = 1$ and (iii) $\sum_{k=1}^K \theta_k \pi_k = 1$. In the following, we propose a modification of the Algorithm 2 (CNM-AP) of Wang (2010) to compute the semi-parametric MLE, which applies the idea of the Algorithm 1 (CN10) of Wang (2007) to non-parametrically estimate G for a fixed $\boldsymbol{\xi}$ by adding one additional mass point for G at each iteration of the algorithm. For this reason, we name the modified algorithm as 'CN10-AP', where 'AP' stands for 'alternating the parameters' $\boldsymbol{\xi}$ and $(\boldsymbol{\theta},\boldsymbol{\pi})$ for estimation in the algorithm.

Algorithm (CN1O-AP)

Set $j = 0$. We start with an initial estimate $(\boldsymbol{\xi}_0, G_0) = (\boldsymbol{\xi}_0, \boldsymbol{\theta}_0, \boldsymbol{\pi}_0)$ such that $\ell(\boldsymbol{\theta}_0, \boldsymbol{\pi}_0, \boldsymbol{\xi}_0) > -\infty$

1. Compute $\theta_j^* = \sup_{\theta \in \mathbb{R}^+} d(\delta_\theta; G_j, \boldsymbol{\xi_j})$, where $G_j = (\boldsymbol{\theta}_j, \boldsymbol{\pi}_j)$. If $d(\delta_{\theta_j^*}; G_j, \boldsymbol{\xi_j}) = 0$, stop.

2. Set $\boldsymbol{\theta}_j^+ = (\boldsymbol{\theta}_j, \theta_j^*)^T$ and $\boldsymbol{\pi}_j^+ = (\boldsymbol{\pi}_j, 0)^T$. Let us denote the optimized solution of $Q(\boldsymbol{\pi}' | \boldsymbol{\pi}_j^+, \boldsymbol{\theta}_j^+)$ by $\boldsymbol{\pi_{j+1}^-}$ with the identifiability constraint $\boldsymbol{\pi_j^+}^T \boldsymbol{\theta_j^+} = 1$. Define $G_{j+1}^- = (\boldsymbol{\theta}_j^+, \boldsymbol{\pi_{j+1}^-})$.

3. Find $\varepsilon \in [0, 1]$ which maximizes $\ell(G_j + \varepsilon(G_{j+1}^- - G_j))$ with respect to ε.

4. Set $G_{j+1} = G_j + \varepsilon(G_{j+1}^- - G_j)$.

5. update $\boldsymbol{\xi_j}$ to $\boldsymbol{\xi_{j+1}}$ by maximizing $\ell(\boldsymbol{\xi}, \boldsymbol{\theta_{j+1}}, \boldsymbol{\pi_{j+1}})$ with respect to $\boldsymbol{\xi}$ using a quasi-Newton method, for example BFGS. Set $j = j + 1$.

6. Repeat step 1 to 5 until convergence.

4.4 Computational Issues

There are several computational issues related to the **CN1O-AP** algorithm to find the semi-parametric MLE discussed in Section 4.3. In step 1, we require to compute $\sup_{\theta \in \mathbb{R}^+} d(\delta_\theta, G_j, \boldsymbol{\xi}_j)$. According to Wang (2007), finding the set of local maxima of the gradient function $d(., G_j, \boldsymbol{\xi}_j)$ and appending that to the set of mass points, instead of only the supremum, makes the algorithm faster. We found it to be more time-consuming possibly because the gradient function is complicated in our case. It is numerically difficult to find all local maxima of the gradient function. Therefore, we only compute the supremum numerically and append it with the set of mass points at each iteration. In order to find the maximum, the directional derivatives (or, the gradients) are computed on a set of grid points (say, 100 points) and we compute the supremum from the gradient values at these grid points.

4.5 Asymptotic Standard Error and Confidence Interval

Derivation of asymptotic distribution of the semi-parametric MLE, obtained by the method of Wang (2010) or the previous section, is difficult and has not been attempted. Even in the context of estimating the mixing distribution non-parametrically

by the method of Lindsay (1983) or Wang (2007), asymptotic distribution of the corresponding estimator has not been derived. Therefore, there is no result in the literature on interval estimation of such non-parametric MLE of G, or semi-parametric MLE of $(\boldsymbol{\xi}, G)$. In particular, asymptotic standard errors of the relevant estimates have also not been provided. For that matter, even the rates of convergence of the proposed estimators are not known. See Lindsay and Lesperance (1995) for a discussion on this.

In view of the above, we resort to the subsampling method of Politis et al. (1999) (See Section 1.10.4) to estimate the convergence rates of the proposed estimators, followed by computation of the standard errors and approximate 95% confidence intervals. We first describe the subsampling method to estimate the rate of convergence of an estimator. Let $X_1, \cdots, X_n$ be a sample of size n from a distribution characterized by a single parameter θ. We consider an estimator $\hat{\theta}_n = f(X_1, \cdots, X_n)$ of the parameter θ and assume that there is some convergence rate τ_n such that

$$\tau_n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} \text{Some distribution.} \tag{4.11}$$

There are $\binom{n}{b}$ ways to select a subsample of size $b(< n)$ without replacement from the sample of size n. Denote a subsample of size b by $(X_1^*, X_2^*, \cdots, X_b^*)$ and the corresponding estimator of θ based on the subsample by $\theta_b^* = f(X_1^*, \cdots, X_b^*)$, the subsample bootstrap analogue of $\hat{\theta}_n$. Then, we have

$$\tau_b(\theta_b^* - \theta) \xrightarrow{\mathcal{D}} \text{Some distribution,} \tag{4.12}$$

as $b \to \infty$ where 'Some' refers to the same distribution as in (4.11). The assumptions required for the subsampling method are

$$b \to \infty, \frac{b}{n} \to 0, \tau_b \to \infty, \frac{\tau_b}{\tau_n} \to 0, \tag{4.13}$$

as $n \to \infty$. It easily follows from the set of assumptions (4.13) that

$$\tau_b(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} 0. \tag{4.14}$$

Subtracting (4.14) from (4.12), we get

$$\tau_b(\theta_b^* - \hat{\theta}_n) \xrightarrow{\mathcal{D}} \text{Some distribution,} \tag{4.15}$$

where 'Some' refers to the same asymptotic distribution in (4.11). The above result is useful in order to estimate the convergence rate and subsampling based confidence intervals.

Let us write $\tau_n = n^\beta$, where we need to estimate the unknown constant β. For this, subsamples of different sizes (b) are drawn and we observe the distribution of $\theta_b^* - \hat{\theta}_n$ based on subsampling bootstrap method for different sizes (b). Let $G_b^{*-1}(t)$ be the tth quantile of the distribution of $\theta_b^* - \hat{\theta}_n$, for a fixed b. Then, we have

$$G_b^{*-1}(t) \approx b^{-\beta} F^{-1}(t), \tag{4.16}$$

where F is the CDF of the same asymptotic distribution of $n^\beta(\hat{\theta}_n - \theta)$, as in (4.11). The proof of this result (4.16) can be found in Chapter 8 of Politis et al. (1999). For $s < t$, using 4.16, we have

$$G_b^{*-1}(t) - G_b^{*-1}(s) \approx b^{-\beta}\big[F^{-1}(t) - F^{-1}(s)\big],$$

which gives

$$\log\big[G_b^{*-1}(t) - G_b^{*-1}(s)\big] \approx -\beta \log b + \log[F^{-1}(t) - F^{-1}(s)]. \tag{4.17}$$

Now, the problem reduces to estimation of β by means of linear regression. We choose several subsample sizes b_j and several pair of quantiles $s_i < t_i$, with $s_i < 0.5 < t_i$, as suggested by Politis et al. (1999). Then, we write,

$$y_{ij} = \log\big[G_{b_j}^{*-1}(t_i) - G_{b_j}^{*-1}(s_i)\big]$$

and

$$c_i = \log[F^{-1}(t_i) - F^{-1}(s_i)].$$

Averaging over the index i (pair of the quantiles), we have

$$\bar{y}_j \approx -\beta \log(b_j) + \bar{c},$$

where $\bar{y}_j$ is the mean of y_{ij}'s over i and $\bar{c}$ is the mean of the c_i's. Politis et al. (1999) recommends regressing $\bar{y}_j$ on the $\log b_j$, based on different j's, to estimate β as

$$\hat{\beta} = -\frac{\mathrm{cov}[\bar{y}, \log b]}{\mathrm{var}[\log b]}.$$

At the end of this section, we report the result of this method to estimate the rate of convergence for our estimators of Section 4.3.

This gives the estimated rate of convergence as $\hat{\tau}_n = n^{\hat{\beta}}$, or . $\hat{\tau}_b = b^{\hat{\beta}}$ for the subsample size b. Hence, for a suitable subsample size b, an approximate standard error of $\hat{\theta}_n$, denoted by $SE(\hat{\theta}_n)$, is given by

$$SE(\hat{\theta}_n) = SE(\theta_b^*) \times \frac{b^{\hat{\beta}}}{n^{\hat{\beta}}},$$

where $SE(\theta_b^*)$ is the standard error of θ_b^* based on subsampling bootstrap method with subsample size b. We consider this computation of standard error for our estimates of Section 4.3 in our simulation study in Section 4.6 and real data analysis in Section 4.7.

We estimate the asymptotic distribution function F by its empirical analogue F_b^*, obtained by using the subsampling bootstrap method based on $b^{\hat{\beta}}(\theta_b^* - \hat{\theta}_n)$, for a suitable b. Note that, for large enough n,

$$F^{-1}\Big(\frac{\alpha}{2}\Big) < \tau_n(\hat{\theta}_n - \theta) < F^{-1}\Big(1 - \frac{\alpha}{2}\Big)$$

occurs with probability approximately equal to $1 - \alpha$, where $F^{-1}(\frac{\alpha}{2})$ and $F^{-1}(1 - \frac{\alpha}{2})$ are the $\frac{\alpha}{2}$th and $(1 - \frac{\alpha}{2})$th quantiles of F. The functional form of F is unknown but it can be estimated by F_b^* for large b and n, satisfying 4.13. Therefore,

$$F_b^*\Big(\frac{\alpha}{2}\Big) < \tau_n(\hat{\theta}_n - \theta) < F_b^*\Big(1 - \frac{\alpha}{2}\Big) \tag{4.18}$$

holds with probability approximately equal to $1 - \alpha$. Replacing τ_n by $\hat{\tau}_n$ in equation 4.18 and then rearranging it suitably, we get

$$\hat{\theta}_n - \hat{\tau}_n^{-1} F_b^{*-1}\Big(1 - \frac{\alpha}{2}\Big) < \theta < \hat{\theta}_n - \hat{\tau}_n^{-1} F_b^{*-1}\Big(\frac{\alpha}{2}\Big) \tag{4.19}$$

as the approximate $100(1 - \alpha)\%$ confidence interval of the parameter θ based on subsampling bootstrap method. These approximate confidence intervals have been reported in Sections 4.6 and 4.7 for our simulation study and data analysis, respectively.

In the following, we describe the estimation of the convergence rates for the semi-parametric estimators of $\boldsymbol{\xi}$ and G (See Section 4.3) using the subsampling method, as discussed in the beginning of this section. In particular, for our shared frailty model and constant baseline cause-specific hazards with two competing risks for each

individual, we have four parameters in $\boldsymbol{\xi}$ with $h_{0j}^{(k}(t) = \alpha_j^{(k)}$, for $j, k = 1, 2$. We consider estimating the convergence rates of the corresponding four estimators and also the same for the three quartiles of the estimated G.

In order to apply the subsampling method for this purpose, we consider simulation of bivariate current status data with competing risks of the form (X_1, X_2, J_1, J_2) from the above model having constant baseline cause-specific hazards with two competing risks $(\alpha_1^{(1)} = 0.8, \alpha_2^{(1)} = 1.3, \alpha_1^{(2)} = 0.2, \alpha_2^{(2)} = 1.1)$ and exponential monitoring time distribution with mean 3.9848 for both the individuals. Also, we consider 5 different sample sizes $n = 1000, 2000, 5000, 10000, 20000$, and 7 subsample sizes (b_j's) suitably chosen for each n. Again, we consider 5 pairs (s_i, t_i), namely, $(s_1, t_1) = (0.25, 0.75)$, $(s_2, t_2) = (0.4, 0.6), (s_3, t_3) = (0.3, 0.7), (s_4, t_4) = (0.35, 0.65), (s_5, t_5) = (0.32, 0.68)$.

We estimate β for the convergence rate n^β for each of the four parameter estimates and the three estimated quartiles of G. We expect and assume that the convergence rates for the four parameter estimates are the same, and those of the three estimated quartiles are also the same but may be different from those of the parameter estimates. We, therefore, take average of the estimated convergence rates of the four parameter estimates and over the five sample sizes (n) as the estimated common convergence rate for the estimators of the parametric components of the model. Similarly, we obtain the estimated common convergence rate for the estimators of the non-parametric components (three quartiles) of the model.

From our simulation based subsampling method, we find the estimated convergence rate for the parameter estimates as $n^{0.3118}$, while the same for the estimated quartiles as $n^{0.2812}$. This aligns with the finding of Chen (1995) which states that, while a $n^{\frac{1}{2}}$ convergence rate is achievable if the number of components (mass points) in the mixing distribution G is known, the best possible rate is only $n^{\frac{1}{4}}$ if it is unknown (See Chen, 1995).

4.6 Simulation Studies

In order to investigate finite sample properties of the semi-parametric MLE $\hat{\boldsymbol{\eta}} = (\hat{\boldsymbol{\xi}}, \hat{G}) = (\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\pi}})$, we consider a simulation study with the similar setup as that in Section 3.5. We consider the same monitoring time X for both the individuals in a pair, where X follows an Exponential distribution with mean $\frac{1}{\mu}$. We solve the value of μ from the joint censoring probability $P\big[T_1 > X, T_2 > X; \mu\big] = p_{cen}$, for a fixed value of p_{cen} (say, 0.1), as in the Section 3.5.

We take $L_1 = L_2 = 2$ and consider constant or Exponential type baseline cause-

specific hazards; that is, $h_{0j}^{(k)}(t_k) = \alpha_j^{(k)}$, for $k = 1, 2$ and $j = 1, 2$, with $\boldsymbol{\xi} = (\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)}) = (0.8, 1.3, 0.2, 1.1)$. We also consider shared Gamma frailty model with $\sigma = 0.632$. We carry out simulation of a sample of size n with each observation of the form (x, x, j_1, j_2) following the same method as that in Section 3.5. This procedure is repeated 1000 times to get 1000 simulated data sets each of the form $\{(x_i, x_i, j_{i1}, j_{i2}); i = 1, \cdots, n\}$.

For each simulated dataset, we calculate the semi-parametric MLE $(\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\pi}})$ using the **CN1O-AP** algorithm described in Section 4.3. We save the value of $\hat{\boldsymbol{\xi}}$ and the three estimated quartiles of G (denoted by $\hat{\zeta}_{0.25}, \hat{\zeta}_{0.5}$ and $\hat{\zeta}_{0.75}$, respectively) in each case. For each estimator, the corresponding bias is estimated by subtracting the average of 1000 estimates from its true value. The sample standard error (SSE) is computed by taking the square root of the sample variance of these 1000 estimates. We also compute the standard error of each estimate in a simulated sample by the subsampling bootstrap method described in Section 4.5. Averaging these 1000 standard errors produces the average standard error (ASE). Then, for each parameter, we obtain the corresponding subsampling bootstrap based approximate 95% confidence interval and compute the proportion (CP) of times these 1000 confidence intervals contain the corresponding true value, as an estimate of the corresponding coverage probability. We carry out this simulation study for $n = 500, 1000, 2000$.

The results of this simulation study are summarized in Table 4.1. Corresponding to each parameter, we observe that Bias, SSE, and ASE decrease with the sample size, as expected. Additionally, the value of SSE and ASE become closer as the sample size increases, validating the computation of asymptotic standard error by subsampling bootstrap method. Note that the CP value for each of the four parameter estimates tends to 0.95 with the increase in sample size. We also observe that the convergence of the non-parametric MLE (the three estimated quartiles of G) is slower. We have carried out similar simulation studies for other values of the parameter vectors and found similar results, which are not reported here.

Table 4.1: Simulation results for the semi-parametric MLEs under constant baseline cause-specific hazards and non-parametric shared frailty model with joint censoring probability $p_{cen} = 0.1$.

n	Properties	$\hat{\alpha}_1^{(1)}$	$\hat{\alpha}_2^{(1)}$	$\hat{\alpha}_1^{(2)}$	$\hat{\alpha}_2^{(2)}$	$\hat{\zeta}_{0.25}$	$\hat{\zeta}_{0.5}$	$\hat{\zeta}_{0.75}$
500	Bias	0.0722	0.0930	0.0254	0.0912	0.0619	0.0806	−0.0415
	SSE	0.1854	0.2926	0.0498	0.2261	0.2637	0.3491	0.4936
	ASE	0.2334	0.3645	0.0608	0.2659	0.1965	0.2230	0.3258
	CP	0.9300	0.9000	0.9650	0.8600	0.4900	0.3600	0.5300
1000	Bias	0.0554	0.0894	0.0177	0.0816	0.0473	0.0297	−0.0408
	SSE	0.1514	0.2382	0.0426	0.1951	0.2454	0.3348	0.4546
	ASE	0.1739	0.2660	0.0459	0.2114	0.1757	0.2100	0.3001
	CP	0.9500	0.9300	0.9500	0.9300	0.5400	0.3700	0.4700
2000	Bias	0.0342	0.0541	0.0053	0.0412	0.0303	0.0207	−0.0356
	SSE	0.1264	0.1815	0.0291	0.1582	0.2159	0.3045	0.3585
	ASE	0.1478	0.2265	0.0354	0.1776	0.1651	0.2001	0.2656
	CP	0.9600	0.9600	0.9300	0.9100	0.6500	0.5000	0.5500

4.7 Data Analysis

Recall from Section 1.12 that the hearing loss of both the ears of an individual can occur due to any one of two causes, namely, **SNHL** and **Conductive**. For ease of our use, we label them as 1 and 2, respectively, as in Section 3.6. In this section, we analyze the hearing loss dataset under the Exponential or constant type baseline cause-specific hazards with non-parametric shared frailty, That is, we assume $h_{0j}^{(k)}(t) = \alpha_j^{(k)}$, for $j = 1, 2$ and $k = 1, 2$. We compute the semi-parametric MLEs of the $\alpha_j^{(k)}$'s and the frailty distribution G using the method of Section 4.3 and calculate the corresponding asymptotic standard errors (ASE) using the method of Section 4.5. We also report the approximate 95% confidence intervals (CI) for all the parameters in $\boldsymbol{\xi}$ and also the three quantiles of G based on the subsampling bootstrap method of Section 4.5. The results are summarized in Table 4.2. We observe that **SNHL** is the dominating cause for both left and right ear. The non-parametric estimate of G has 5 mass points at 1.62×10^{-6}, 0.0010, 0.0653, 0.0694, and 3.6915 with corresponding masses 0.3315, 0.0332, 0.3062, 0.0648, and 0.2642. One can verify that the mean of this estimated G is 1. The three quartiles of this estimated G are presented in Table 4.2. Note that the lower confidence limits of the first two quartiles, $\hat{\zeta}_{0.25}$ and $\hat{\zeta}_{0.50}$, are obtained as slightly smaller than zero, so we decided to truncate them at zero.

Table 4.2: Analysis of the hearing loss data under constant baseline cause-specific hazards and non-parametric shared frailty model.

Parameter	MLE	ASE	Approx. 95% C.I
$\alpha_1^{(1)}$	1.4720	0.3639	[0.6728, 2.1788]
$\alpha_2^{(1)}$	0.0981	0.0312	[0.0295, 0.1496]
$\alpha_1^{(2)}$	1.4363	0.5890	[0.2700, 2.4052]
$\alpha_2^{(2)}$	0.0957	0.0481	[0.0057, 0.1550]
$\zeta_{0.25}$	1.62×10^{-6}	0.0003	[0, 4.32×10^{-5}]
$\zeta_{0.50}$	0.0653	0.2654	[0, 0.0761]
$\zeta_{0.75}$	3.6915	1.0291	[3.0145, 5.0983]

4.8 Concluding Remarks

Modeling of bivariate failure time data with competing risks using frailty has been studied by several authors, for example, see Bandeen-Roche and Liang (2002) and Gorfine and Hsu (2011). Even though identifiability study of the resulting models is an essential aspect of statistical inference (See condition 1 in p69 for the asymptotic results to hold), there is no investigation for such identifiability, as far as we know. In this work, we have studied this model identifiability with a particular class of parametric baseline cause-specific hazard functions and four different types of non-parametric frailty distributions. We have proved identifiability results under fairly reasonable assumptions.

In Section 4.7, we have assumed Exponential type baseline cause-specific hazards for simplicity. This is a very restrictive modeling assumption since, in practice, the baseline cause-specific hazards may have different forms depending on time (See Section 5.5). One common alternative is to try Weibull type baseline cause-specific hazards, possibly with different shape parameters. It is of interest to study how the non-parametric estimate of the mixing distribution G changes with such different models for the baseline cause-specific hazards.

In Chapters 1-3, even till Section 4.2, we have mentioned four types of frailty, whereas in Section 4.3, while describing the semi-parametric estimation of $\boldsymbol{\xi}$ and G, we consider only the shared frailty model. This was done for the ease of notation and some simplicity in explaining the related ideas of estimating the frailty distribution G non-parametrically. The generalization to the other three types of frailty is, in principle, similar to that for the shared frailty model by viewing the mixing distribution G as that of the random frailty vector $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}^{(1)}, \boldsymbol{\epsilon}^{(2)})$, while estimating

it non-parametrically. Following Lindsay (1983), the method works as long as the space of all such $\boldsymbol{\epsilon}$ is a sigma field of measurable sets containing all the single points of the form $\boldsymbol{\epsilon}$. This property holds since the concerned space is some manifolds of $\mathbb{R}_+$. Therefore, for a given $\boldsymbol{\xi}$, the non-parametric estimate of G is a discrete distribution with number of mass points in that space of frailty vectors not exceeding n, the sample size. These mass points may be denoted as $\theta_1, \cdots, \theta_K$, with $K \leq n$, as in Section 4.3, and the θ_k's belong to that space of frailty vectors.

Although the notation becomes complicated and the coding may be more difficult, the description of the method follows exactly the same way as in Section 4.3. For example, for the correlated frailty model, we consider G as the joint distribution of $(\epsilon^{(1)}, \epsilon^{(2)})$ defined on $\mathbb{R}_+^2$. The non-parametric estimate of G is discrete with at most n mass points (say, $\theta_1, \cdots, \theta_K$, with $K \leq n$) in that plane. The derivation of directional derivatives (or, gradients) follows similarly having the same expression as that in Section 4.3. The rest follows similarly by writing the finite number of θ_k's and the corresponding masses (the π_k's) as vectors $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$, respectively. There are now two identifiability constraints for the means of ($\epsilon^{(1)}$ and $\epsilon^{(2)}$) being unity which need to be considered while minimizing the corresponding $Q(\boldsymbol{\pi'}|\boldsymbol{\pi}, \boldsymbol{\theta}, \boldsymbol{\xi})$ function.

In this paper, we have only studied the models with a parametric class of baseline cause-specific hazard functions which includes a number of commonly used popular hazard functions as special cases. However, there are several other choices (for example, Gumbel and generalized F) for baseline cause-specific of hazard functions which do not belong to this class. The identifiability study of these models is also of interest. Similarly, extension of the present work to multivariate failure time data with competing risks is of importance. Although the four types of frailty models considered in this work cover a wide range of dependence structures including heterogeneity between individuals, one can possibly think of other forms of frailty, for example, time-dependent frailty. This will certainly mean a much bigger challenge in terms of notation, proof of identifiability and computation as well.

Chapter 5

Bivariate Current Status Data with Competing Risks: Non-parametric Baseline Hazards and Parametric Frailty

5.1 Introduction

In view of the development in Chapters 3 and 4, it is now relevant to study modeling and analysis of bivariate current status data with competing risks using non-parametric baseline cause-specific hazards with parametric frailty models. In particular, we consider the four different Gamma frailty models, as discussed in the previous chapters. However, for simplicity, we describe our estimation procedure, and subsequent simulation study and data analysis, using only the shared Gamma frailty model.

As discussed before, identifiability is an important aspect of model building, especially in the context of bivariate failure time data with competing risks using frailty models. To highlight the issue further in this chapter, we give an example of non-parametric baseline cause-specific hazards and parametric frailty model which is non-identifiable. Subsequently, we consider non-parametric baseline cause-specific hazards in general and Gamma frailty to investigate identifiability of the corresponding models. Then we develop an estimation method for bivariate current status data with competing risks under the above modeling setup, but with shared Gamma frailty. We choose Gamma distribution for frailty, as it produces analytically tractable quantities.

In Section 5.2, we develop the frailty model for bivariate failure time data with competing risks in terms of the joint sub-distribution function. This section also es-

tablishes the corresponding model identifiability. Section 5.2.1 presents an example of a non-identifiable model even when the frailty distribution is assumed to be parametric, while Sections 5.2.2-5.2.5 investigate identifiability of the four different Gamma frailty models, as discussed. Section 5.3 describes the estimation procedure for the shared Gamma frailty model, where we approximate the non-parametric baseline cause-specific hazards using Bernstein polynomials. A limited simulation study using the shared Gamma frailty model is provided in Section 5.4. The hearing loss data has been analyzed in Section 5.5 using the proposed non-parametric cause-specific hazards with the shared Gamma frailty model. Section 5.6 ends with some concluding remarks.

5.2 Modeling and Identifiability

As in Chapters 2-4, we consider the multiplicative frailty model for $\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}})$, the conditional cause-specific hazard for failure of the kth individual due to cause j, given the frailty vector $\boldsymbol{\epsilon^{(k)}}$, as given by

$$\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k)\epsilon_j^{(k)}, \tag{5.1}$$

for $t_k > 0, j = 1, \cdots, L_k$ and $k = 1, 2$, where $h_{0j}^{(k)}(t_k)$ is the jth baseline cause-specific hazard at time t_k for the kth individual. We assume, for each model, the existence of finite frailty mean.

As in Section 2.2, the marginal survival function of the kth individual, conditional on the frailty vector $\boldsymbol{\epsilon^{(k)}}$, is

$$S^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = \exp\left[-\sum_{j=1}^{L_k} H_{0j}^{(k)}(t_k)\epsilon_j^{(k)}\right],$$

and the jth sub-distribution function of the kth individual, conditional on frailty vector $\boldsymbol{\epsilon^{(k)}}$, is

$$F_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k} h_{0j}^{(k)}(u_k)\epsilon_j^{(k)} \exp\left[-\sum_{j'=1}^{L_k} H_{0j'}^{(k)}(u_k)\epsilon_{j'}^{(k)}\right] du_k,$$

for all $j = 1, \cdots, L_k$ and $k = 1, 2$. Also, the joint sub-distribution function $F_{j_1 j_2}(t_1, t_2)$

under the general model (5.1) is

$$F_{j_1j_2}(t_1,t_2)=\int_0^{\infty}\cdots\int_0^{\infty}\left[\int_0^{t_1}\int_0^{t_2}\prod_{k=1}^{2}\left(h_{0j_k}^{(k)}(u_k)\epsilon_{j_k}^{(k)}\right)\exp\left(-\sum_{k=1}^{2}\sum_{j=1}^{L_k}H_{0j}^{(k)}(u_k)\epsilon_j^{(k)}\right)du_2du_1\right]$$
$$\times\, g(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}};\boldsymbol{\theta})d\boldsymbol{\epsilon^{(1)}}d\boldsymbol{\epsilon^{(2)}}, \tag{5.2}$$

for all $j_k=1,\cdots,L_k$ and $k=1,2$, where the non-parametric baseline cause-specific hazards $h_{0j}^{(k)}(t_k)$'s, $j=1,\cdots,L_k$, belong to the family $\mathcal{H}_k$, for $k=1,2$ (See Section 2.5), and $g(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}};\boldsymbol{\theta})$ is the joint density of the random frailty vector $(\boldsymbol{\epsilon^{(1)}},\boldsymbol{\epsilon^{(2)}})$ with the associated parameter vector $\boldsymbol{\theta}\in\boldsymbol{\Theta}$, the frailty parameter space.

In the next subsection, we provide an example of a parametric family of frailty distributions that leads to non-identifiability of the model (5.1) with non-parametric baseline cause-specific hazards to demonstrate that parametric frailty does not always ensure identifiability. The subsequent four subsections consider the four types of Gamma frailty distributions, since Gamma frailty is commonly used by many authors for its convenience (Yashin et al., 1995; Iachine, 2004; Duchateau and Janssen, 2008; Wienke, 2010 etc).

5.2.1 A Non-identifiable Model

To work with a parametric family of frailty distribution, the joint sub-distribution function $F_{j_1j_2}(t_1,t_2)$ in (5.2) is written as $F_{j_1j_2}(t_1,t_2;\boldsymbol{\theta})$ to indicate the dependence on the parameter vector $\boldsymbol{\theta}\in\boldsymbol{\Theta}$.

Definition 5.1. We say that the model $F_{j_1j_2}(t_1,t_2;\boldsymbol{\theta})$ is identifiable within the class $\mathcal{H}_1\times\mathcal{H}_2\times\boldsymbol{\Theta}$ if, for some $\boldsymbol{h_0^{(k)}},\tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(k)}}\in\mathcal{H}_k$, for $k=1,2$, and $\boldsymbol{\theta},\tilde{\boldsymbol{\theta}}\in\boldsymbol{\Theta}$, the equality

$$F_{j_1j_2}(t_1,t_2;\boldsymbol{\theta})=F_{j_1j_2}(t_1,t_2;\tilde{\boldsymbol{\theta}})\,\text{for all } t_k>0, j_k=1,\cdots,L_k, k=1,2,$$

implies

$$\boldsymbol{h_0^{(k)}}(x)=\tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(k)}}(x)\text{ and }\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}\text{ for all } x\geq 0, k=1,2.$$

Let us consider the model

$$\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon}^{(k)})=h_0^{(k)}(t_k)\epsilon_j^{(k)}, \tag{5.3}$$

for $t_k>0,\ j=1,\cdots,L_k$ and $k=1,2$, where $h_0^{(k)}(t_k)$ is the arbitrary jth baseline cause-specific hazard function for the kth individual, being the same for all $j=$

$1, \cdots, L_k$. Therefore, the class $\mathcal{H}_k$ now reduces to

$$\mathcal{H}_k = \left\{ h_0^{(k)}(x) : h_0^{(k)}(x) \geq 0, \int_0^\infty h_0^{(k)}(x)dx = \infty \right\},$$

for $k = 1, 2$. Let us define the frailty variable $\epsilon_j^{(k)}$ in the general model (5.1) as

$$\epsilon_j^{(k)} = a\eta_j^{(k)},$$

for $j = 1, \cdots, L_k,\ k = 1, 2$, where a is a positive random variable, which acts as the shared frailty between the two individuals in a pair, and $\eta_j^{(k)}$ is j-th cause-specific frailty for the k-th individual. We assume that $\boldsymbol{\eta^{(k)}} = (\eta_1^{(k)}, \cdots, \eta_{L_k}^{(k)}) \sim \text{Dir}(\boldsymbol{\alpha^{(k)}})$, for $k = 1, 2$, where $\text{Dir}(\boldsymbol{\alpha^{(k)}})$ denotes a Dirichlet distribution with parameter vector $\boldsymbol{\alpha^{(k)}} = (\alpha_1^{(k)}, \cdots, \alpha_{L_k}^{(k)})$, with $\alpha_j^{(k)} > 0,\ j = 1, 2, \cdots, L_k$, for $k = 1, 2$. While we assume independence between the two frailty vectors $\boldsymbol{\eta^{(1)}}$ and $\boldsymbol{\eta^{(2)}}$, these are also assumed to be independent of the shared frailty a. Also, a is assumed to follow a Gamma$(\frac{1}{\sigma^2}, \frac{1}{\sigma^2})$ distribution, for some $\sigma > 0$, with density given by (3.4) so that $E[a] = 1$. So, the parameter space $\boldsymbol{\Theta}$ is given by

$$\boldsymbol{\Theta} = \left\{ \boldsymbol{\theta} = (\boldsymbol{\alpha^{(1)}}, \boldsymbol{\alpha^{(2)}}, \sigma) : \boldsymbol{\alpha^{(k)}} \in \mathbb{R}_+^{L_k},\ \text{for } k = 1, 2,\ \text{and } \sigma > 0 \right\}.$$

The joint frailty density function $g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}, \boldsymbol{\theta})$ can be derived from the above description. The choice of Dirichlet distribution for cause-specific frailty vector $\boldsymbol{\eta^{(k)}}$ facilitates the derivation because we have $\sum_{j=1}^{L_k} \eta_j^{(k)} = 1$ almost surely, for $k = 1, 2$. Also, the Gamma distribution for the shared frailty variable a makes the computation for joint sub-distribution function easier.

Note that $\eta_j^{(k)}$ marginally follows the Beta$\left(\alpha_j^{(k)}, \sum_{j' \neq j} \alpha_{j'}^{(k)}\right)$ distribution. Therefore, we have

$$\mathbb{E}(\eta_j^{(k)}) = \frac{\alpha_j^{(k)}}{\sum_{j'=1}^{L_k} \alpha_{j'}^{(k)}},$$

for $j = 1, \cdots, L_k$ and $k = 1, 2$. Also, since a is independent of the $\eta_j^{(k)}$'s and $\mathbb{E}(a) = 1$, we have $\mathbb{E}(\epsilon_j^{(k)}) = \mathbb{E}(\eta_j^{(k)})$. The survival function for the kth individual, conditional

on frailty variables $a, \boldsymbol{\eta^{(k)}}$, is given by

$$\begin{aligned} S^{(k)}(t|a,\boldsymbol{\eta^{(k)}}) &= \exp\left[-\int_0^t \sum_{j=1}^{L_k} \lambda_j^{(k)}(u|a,\boldsymbol{\eta^{(k)}})du \right] \\ &= \exp\left[-\int_0^t a h_0^{(k)}(u) \sum_{j=1}^{L_k} \eta_j^{(k)} du \right] \\ &= \exp\left[-\int_0^t a h_0^{(k)}(u) du \right] \\ &= \exp[-a H_0^{(k)}(t)], \end{aligned}$$

where $H_0^{(k)}(t) = \int_0^t h_0^{(k)}(u)du$, for $k = 1, 2$. Hence, the jth sub-distribution function for the kth individual, conditional on the frailty variables $a, \boldsymbol{\eta^{(k)}}$, is

$$\begin{aligned} F_j^{(k)}(t_k|a,\boldsymbol{\eta^{(k)}}) &= \int_0^{t_k} \lambda_j^{(k)}(u_k|a,\boldsymbol{\eta^{(k)}}) S^{(k)}(u_k|a,\boldsymbol{\eta^{(k)}}) du_k \\ &= \int_0^{t_k} h_0^{(k)}(u_k) a \eta_j^{(k)} \exp\{-a H_0^{(k)}(u_k)\} du_k \\ &= \eta_j^{(k)} \int_0^{a H_0^{(k)}(t_k)} e^{-z} dz \\ &= \eta_j^{(k)} \big(1 - \exp\{-a H_0^{(k)}(t_k)\}\big), \end{aligned}$$

for $t_k > 0$, $j = 1, \cdots, L_k$ and $k = 1, 2$. Therefore, the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{\theta})$ is derived, using conditional independence given a,$\boldsymbol{\eta^{(1)}}$ and $\boldsymbol{\eta^{(2)}}$, as

$$\begin{aligned} F_{j_1 j_2}(t_1, t_2; \boldsymbol{\theta}) &= \mathbb{E}\Big[F_{j_1}^{(1)}(t_1|a,\boldsymbol{\eta^{(1)}}) F_{j_2}^{(2)}(t_2|a,\boldsymbol{\eta^{(2)}}) \Big] \\ &= \mathbb{E}\Big[\eta_{j_1}^{(1)} \eta_{j_2}^{(2)} \Big(1 - \exp\big\{ - a H_0^{(1)}(t_1)\big\}\Big)\Big(1 - \exp\big\{ - a H_0^{(2)}(t_2)\big\}\Big) \Big] \\ &= \left[\prod_{k=1}^{2} \mathbb{E}(\eta_{j_k}^{(k)}) \right] \times \mathbb{E}\Big[(1 - \exp\big\{ - a H_0^{(1)}(t_1)\big\})(1 - \exp\big\{ - a H_0^{(2)}(t_2)\big\}) \Big] \\ &= \left[\prod_{k=1}^{2} \mathbb{E}(\eta_{j_k}^{(k)}) \right] \times \left[1 - \sum_{k=1}^{2} \mathbb{E}\Big(\exp\big\{ - a H_0^{(k)}(t_k)\big\}\Big) + \mathbb{E}\Big(\exp\big\{ - a \sum_{k=1}^{2} H_0^{(k)}(t_k)\big\}\Big) \right]. \end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big[-aH_0^{(k)}(t_k)\Big]\Big] &= \int_0^\infty \exp\Big\{-aH_0^{(k)}(t_k)\Big\}\frac{\exp\left(-\frac{a}{\sigma^2}\right)a^{\frac{1}{\sigma^2}-1}}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}da \\ &= \frac{1}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\int_0^\infty \exp\Big[-a\Big(H_0^{(k)}(t_k)+\frac{1}{\sigma^2}\Big)\Big]a^{\frac{1}{\sigma^2}-1}da \\ &= \frac{1}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\times\frac{\Gamma(\frac{1}{\sigma^2})}{\Big(H_0^{(k)}(t_k)+\frac{1}{\sigma^2}\Big)^{\frac{1}{\sigma^2}}} \\ &= \frac{1}{\Big(1+\sigma^2H_0^{(k)}(t_k)\Big)^{\frac{1}{\sigma^2}}},\end{aligned}$$

for $t_k > 0$ and $k = 1, 2$. Hence, we have

$$F_{j_1j_2}(t_1,t_2;\boldsymbol{\theta}) = \left(\prod_{k=1}^{2}\frac{\alpha_{j_k}^{(k)}}{\sum_{j=1}^{L_k}\alpha_j^{(k)}}\right)\times\left[1-\sum_{k=1}^{2}\Big(1+\sigma^2H_0^{(k)}(t_k)\Big)^{-\frac{1}{\sigma^2}}+\Big(1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k)\Big)^{-\frac{1}{\sigma^2}}\right],$$

for all $t_k > 0,\ j_k = 1, \cdots, L_k$ and $k = 1, 2$.

Let us consider

$$\tilde{\boldsymbol{\theta}} = (\tilde{\alpha}_1^{(1)}, \cdots, \tilde{\alpha}_{L_1}^{(1)}, \tilde{\alpha}_1^{(2)}, \cdots, \tilde{\alpha}_{L_2}^{(2)}, \tilde{\sigma}),$$

where $\tilde{\alpha}_j^{(k)} = c_k\alpha_j^{(k)}$, for some $c_k > 0$ and $j = 1, \cdots, L_k,\ k = 1, 2$; also, $\tilde{\sigma} = \sigma$. This implies $\sum_{j=1}^{L_k}\tilde{\alpha}_j^{(k)} = c_k\sum_{j=1}^{L_k}\alpha_j^{(k)}$, for $k = 1, 2$. We also take $\tilde{h}_0^{(k)}(x) = h_0^{(k)}(x)$, for all $x > 0$ and $k = 1, 2$. Then, the joint sub-distribution function $\tilde{F}_{j_1j_2}(t_1,t_2;\tilde{\boldsymbol{\theta}})$ simplifies to

$$\begin{aligned}\tilde{F}_{j_1j_2}(t_1,t_2;\tilde{\boldsymbol{\theta}}) &= \left(\prod_{k=1}^{2}\frac{\tilde{\alpha}_{j_k}^{(k)}}{\sum_{j=1}^{L_k}\tilde{\alpha}_j^{(k)}}\right)\times\left[1-\sum_{k=1}^{2}\Big(1+\sigma^2H_0^{(k)}(t_k)\Big)^{-\frac{1}{\sigma^2}}+\Big(1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k)\Big)^{-\frac{1}{\sigma^2}}\right] \\ &= \left(\prod_{k=1}^{2}\frac{\alpha_{j_k}^{(k)}}{\sum_{j=1}^{L_k}\alpha_j^{(k)}}\right)\times\left[1-\sum_{k=1}^{2}\Big(1+\sigma^2H_0^{(k)}(t_k)\Big)^{-\frac{1}{\sigma^2}}+\Big(1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(t_k)\Big)^{-\frac{1}{\sigma^2}}\right] \\ &= F_{j_1j_2}(t_1,t_2;\boldsymbol{\theta}),\end{aligned}$$

for all $t_k > 0$ and $j_k = 1, \cdots, L_k$ and $k = 1, 2$. Hence, this model (5.3) is non-identifiable.

5.2.2 Shared Gamma Frailty

In this case, the model (5.1) reduces to

$$\lambda_j^{(k)}(t_k|\epsilon) = h_{0j}^{(k)}(t_k)\epsilon \tag{5.4}$$

where $\epsilon_j^{(k)} = \epsilon$ is the common shared frailty variable, for $j = 1, \cdots, L_k$ and $k = 1, 2$. The shared frailty ϵ is assumed to follow a Gamma$(\frac{1}{\sigma^2}, \frac{1}{\sigma^2})$ distribution with density given by (3.4), for some $\sigma > 0$. Thus, the parameter space is $\mathbf{\Theta} = \big\{\sigma : \sigma > 0\big\}$. Following the same technique as before, the marginal survival function of the kth individual, conditional on the shared frailty ϵ, is

$$S^{(k)}(t; \mathbf{h_0^{(k)}}|\epsilon) = \exp\bigg[- \epsilon \sum_{j=1}^{L_k} H_{0j}^{(k)}(t)\bigg] = \exp\bigg[- \epsilon H_0^{(k)}(t)\bigg],$$

where $H_0^{(k)}(t) = \sum_{j=1}^{L_k} H_{0j}^{(k)}(t)$, for $k = 1, 2$, unlike that in Section 5.2.1. Also, the marginal sub-distribution function due to cause j for the kth individual, conditional on the shared frailty ϵ, is

$$F_j^{(k)}(t_k; \mathbf{h_0^{(k)}}|\epsilon) = \int_0^{t_k} h_{0j}^{(k)}(u_k)\epsilon \exp\big\{ - \epsilon H_0^{(k)}(u_k)\big\} du_k,$$

for $j = 1, 2, \cdots, L_k$ and $k = 1, 2$. Therefore, as before, the unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \mathbf{h_0^{(1)}}, \mathbf{h_0^{(2)}}, \sigma^2)$ is written as

$$\int_0^{\infty} \bigg[\int_0^{t_1} \int_0^{t_2} \bigg(\prod_{k=1}^{2} h_{0j_k}^{(k)}(u_k) \bigg) \epsilon^2 \times \exp\big(- \epsilon \sum_{k=1}^{2} H_0^{(k)}(u_k)\big) du_2 du_1 \bigg] g(\epsilon; \sigma^2) d\epsilon,$$

where $g(\epsilon;\sigma^2)$ is the density of the Gamma$(\frac{1}{\sigma^2},\frac{1}{\sigma^2})$ distribution. This simplifies the unconditional joint sub-distribution function $F_{j_1j_2}(t_1,t_2;\mathbf{h_0^{(1)}},\mathbf{h_0^{(2)}},\sigma^2)$ as

$$
\begin{aligned}
&= \int_0^{\infty}\Bigg[\int_0^{t_1}\int_0^{t_2}\Bigg(\prod_{k=1}^{2}h_{0j_k}^{(k)}(u_k)\Bigg)\epsilon^2\times\exp\big(-\epsilon\sum_{k=1}^{2}H_0^{(k)}(u_k)\big)du_2du_1\Bigg]\frac{\exp\left(-\frac{\epsilon}{\sigma^2}\right)\epsilon^{\frac{1}{\sigma^2}-1}}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}d\epsilon\\
&= \int_0^{t_1}\int_0^{t_2}\Bigg[\int_0^{\infty}\frac{1}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\exp\Big(-\epsilon\big(\sum_{k=1}^{2}H_0^{(k)}(u_k)+\frac{1}{\sigma^2}\big)\Big)\epsilon^{2+\frac{1}{\sigma^2}-1}d\epsilon\Bigg]\Bigg(\prod_{k=1}^{2}h_{0j_k}^{(k)}(u_k)\Bigg)du_2du_1\\
&= \frac{1}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\int_0^{t_1}\int_0^{t_2}\Bigg(\prod_{k=1}^{2}h_{0j_k}^{(k)}(u_k)\Bigg)\frac{\Gamma(2+\frac{1}{\sigma^2})}{\big(\sum_{k=1}^{2}H_0^{(k)}(u_k)+\frac{1}{\sigma^2}\big)^{2+\frac{1}{\sigma^2}}}du_2du_1\\
&= \frac{1}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\int_0^{t_1}\int_0^{t_2}\frac{(1+\frac{1}{\sigma^2})\times\frac{1}{\sigma^2}\times\Gamma(\frac{1}{\sigma^2})\left(\prod_{k=1}^{2}h_{0j_k}^{(k)}(u_k)\right)}{\big(\sum_{k=1}^{2}H_0^{(k)}(u_k)+\frac{1}{\sigma^2}\big)^{2+\frac{1}{\sigma^2}}}du_2du_1\\
&= \int_0^{t_1}\int_0^{t_2}\frac{(1+\sigma^2)\left(\prod_{k=1}^{2}h_{0j_k}^{(k)}(u_k)\right)}{\big(1+\sigma^2\sum_{k=1}^{2}H_0^{(k)}(u_k)\big)^{2+\frac{1}{\sigma^2}}}du_2du_1,
\end{aligned}
$$

for $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$. Similarly, the unconditional sub-distribution function is obtained as

$$
\begin{aligned}
F_j^{(k)}(t_k;\mathbf{h_0^{(k)}},\sigma^2) &= \int_0^{\infty}\Bigg[\int_0^{t_k}h_{0j}^{(k)}(u_k)\epsilon\exp\Big\{-\epsilon H_0^{(k)}(u_k)\Big\}du_k\Bigg]\frac{\exp\left(-\frac{\epsilon}{\sigma^2}\right)\epsilon^{\frac{1}{\sigma^2}-1}}{\sigma^{\frac{2}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}d\epsilon\\
&= \int_0^{t_k}\frac{h_{0j}^{(k)}(u_k)}{\big(1+\sigma^2H_0^{(k)}(u_k)\big)^{1+\frac{1}{\sigma^2}}}du_k, \qquad (5.5)
\end{aligned}
$$

for $j=1,\cdots,L_k$ and $k=1,2$. The general definition of identifiability in Section 2.5 now reduces to the following.

Definition 5.2. The shared Gamma frailty model (5.4) is identifiable within the family $\mathcal{H}_1\times\mathcal{H}_2\times\Theta$ if, for some $\boldsymbol{h_0^{(k)}},\tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\boldsymbol{k})}\in\mathcal{H}_k$, for $k=1,2$, and $\sigma,\ \tilde{\sigma}\in\boldsymbol{\Theta}$, the equality

$$F_{j_1j_2}(t_1,t_2;\boldsymbol{h_0^{(1)}},\boldsymbol{h_0^{(2)}},\sigma^2)=F_{j_1j_2}(t_1,t_2;\tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\mathbf{1})},\tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\mathbf{2})},\tilde{\sigma^2}),$$

for all $t_k>0,\ j_k=1,\cdots,L_k$ and $k=1,2$, implies

$$h_{0j}^{(k)}(x)=\tilde{h}_{0j}^{(k)}(x)\text{ for all }x>0,\ j=1,\cdots,L_k,\ k=1,2\text{ and }\sigma=\tilde{\sigma}.$$

Theorem 5.1. The shared Gamma frailty model (5.4) with non-parametric baseline

cause-specific hazard functions is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$, where $\boldsymbol{\Theta} = \{\sigma : \sigma > 0\}$, provided $H_{0j}^{(k)}(\cdot)$ is differentiable and its inverse, denoted by $H_{0j}^{-(k)}(\cdot)$, exists for all $j = 1, \cdots, L_k$ and $k = 1, 2$.

Proof. From the equality of the joint of the sub-distribution functions as in **Definition 5.2.**, for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$, we obtain equality of the marginal sub-distribution functions $F_j^{(k)}(t_k; \mathbf{h_0^{(k)}}, \sigma^2)$ and $F_j^{(k)}(t_k; \tilde{\mathbf{h}}_{\mathbf{0}}^{(\mathbf{k})}, \tilde{\sigma}^2)$, for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. In particular, we have equality of the corresponding marginal sub-density functions obtained from (5.5), as given by

$$\frac{h_{0j}^{(k)}(t_k)}{[1 + \sigma^2 \sum_{j'=1}^{L_k} H_{0j'}^{(k)}(t_k)]^{1+\frac{1}{\sigma^2}}} = \frac{\tilde{h}_{0j}^{(k)}(t_k)}{[1 + \tilde{\sigma}^2 \sum_{j'=1}^{L_k} \tilde{H}_{0j'}^{(k)}(t_k)]^{1+\frac{1}{\tilde{\sigma}^2}}},$$

for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. This equality can be written as

$$\frac{\frac{d}{dt_k} H_{0j}^{(k)}(t_k)}{\frac{d}{dt_k} \tilde{H}_{0j}^{(k)}(t_k)} = \frac{[1 + \sigma^2 \sum_{j'=1}^{L_k} H_{0j'}^{(k)}(t_k)]^{1+\frac{1}{\sigma^2}}}{[1 + \tilde{\sigma}^2 \sum_{j'=1}^{L_k} \tilde{H}_{0j'}^{(k)}(t_k)]^{1+\frac{1}{\tilde{\sigma}^2}}}.$$

Now, writing $y = \sigma^2 H_{0j}^{(k)}(t_k)$ and $z = \tilde{\sigma}^2 \tilde{H}_{0j}^{(k)}(t_k)$, this equality reduces to

$$\frac{\tilde{\sigma}^2}{\sigma^2} \times \frac{dy}{dz} = \frac{[1 + \sigma^2 \sum_{j'(\neq j)=1}^{L_k} H_{0j'}^{(k)}(H_{0j}^{-(k)}(\frac{y}{\sigma^2})) + y]^{1+\frac{1}{\sigma^2}}}{[1 + \tilde{\sigma}^2 \sum_{j'(\neq j)=1}^{L_k} \tilde{H}_{0j'}^{(k)}(\tilde{H}_{0j}^{-(k)}(\frac{z}{\tilde{\sigma}^2})) + z]^{1+\frac{1}{\tilde{\sigma}^2}}}.$$

This is equivalent to, after integration from 0 to some $w > 0$,

$$\frac{\int_0^w [1 + \sigma^2 \sum_{j'(\neq j)=1}^{L_k} H_{0j'}^{(k)}(H_{0j}^{-(k)}(\frac{y}{\sigma^2})) + y]^{-1-\frac{1}{\sigma^2}} dy}{\int_0^w [1 + \tilde{\sigma}^2 \sum_{j'(\neq j)=1}^{L_k} \tilde{H}_{0j'}^{(k)}(\tilde{H}_{0j}^{-(k)}(\frac{z}{\tilde{\sigma}^2})) + z]^{-1-\frac{1}{\tilde{\sigma}^2}} dz} = \frac{\sigma^2}{\tilde{\sigma}^2}.$$

Now, taking the limit as $w \to 0+$ and using the L'hospital rule, we have

$$\frac{\lim_{w \to 0+} \left[1 + \sigma^2 \sum_{j'(\neq j)=1}^{L_k} H_{0j'}^{(k)}(H_{0j}^{-(k)}(\frac{w}{\sigma^2})) + w\right]^{-1-\frac{1}{\sigma^2}}}{\lim_{w \to 0+} \left[1 + \tilde{\sigma}^2 \sum_{j'(\neq j)=1}^{L_k} \tilde{H}_{0j'}^{(k)}(\tilde{H}_{0j}^{-(1)}(\frac{w}{\tilde{\sigma}^2})) + w\right]^{-1-\frac{1}{\tilde{\sigma}^2}}} = \frac{\sigma^2}{\tilde{\sigma}^2}.$$

Since the composite functions appearing in the above equality are also continuous

functions, we have

$$\lim_{w\to 0+} H_{0j'}^{(k)}(H_{0j}^{-(k)}(\frac{w}{\sigma^2})) = \lim_{w\to 0+} \tilde{H}_{0j'}^{(1)}(\tilde{H}_{0j}^{-(k)}(\frac{w}{\tilde{\sigma}^2})) = 0,$$

for $j \neq j'$. Hence, we obtain $\frac{\sigma^2}{\tilde{\sigma}^2} = 1$, or $\sigma = \tilde{\sigma}$.
Now, equality of the joint sub-distribution functions also implies the equality of the marginal survival functions $S^{(k)}(t_k; \mathbf{h_0^{(k)}}, \sigma^2)$ and $S^{(k)}(t_k; \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\boldsymbol{k})}, \tilde{\sigma}^2)$, for all $t_k > 0$ and $k = 1, 2$. This gives, with $\sigma = \tilde{\sigma}$,

$$\left[1 + \sigma^2 H_0^{(k)}(t_k)\right]^{-\frac{1}{\sigma^2}} = \left[1 + \sigma^2 \tilde{H}_0^{(k)}(t_k)\right]^{-\frac{1}{\sigma^2}},$$

for all $t_k > 0$, which implies

$$H_0^{(k)}(t_k) = \tilde{H}_0^{(k)}(t_k),$$

for all $t_k > 0$ and $k = 1, 2$.

Again, from the equality of the sub-density functions $f_j^{(k)}(t_k; \mathbf{h_0^{(k)}}, \sigma^2)$ and $f_j^{(k)}(t_k; \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\boldsymbol{k})}, \tilde{\sigma}^2)$ and using $\sigma = \tilde{\sigma}$ and $H_0^{(k)}(t_k) = \tilde{H}_0^{(k)}(t_k)$, for all $t_k > 0$, we obtain

$$\frac{h_{0j}^{(k)}(t_k)}{\left[1 + \sigma^2 H_0^{(k)}(t_k)\right]^{1+\frac{1}{\sigma^2}}} = \frac{\tilde{h}_{0j}^{(k)}(t_k)}{\left[1 + \sigma^2 H_0^{(k)}(t_k)\right]^{1+\frac{1}{\sigma^2}}}.$$

This implies

$$h_{0j}^{(k)}(t_k) = \tilde{h}_{0j}^{(k)}(t_k),$$

for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$.
Hence, the shared Gamma frailty model (5.4) is identifiable. □

5.2.3 Correlated Gamma Frailty

The correlated frailty model is given by

$$\lambda_j^{(k)}(t_k|\epsilon^{(k)}) = h_{0j}^{(k)}(t_k)\epsilon^{(k)}, \tag{5.6}$$

for $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$, where $\epsilon^{(k)}$ is the frailty variable for the kth individual and the frailty vector $(\epsilon^{(1)}, \epsilon^{(2)})$ has a correlated Gamma distribution (see Yashin et al., 1995; Wienke, 2010), as described in Section 3.2.2. The frailty parameter space $\boldsymbol{\Theta}$ is given by

$$\boldsymbol{\Theta} = \left\{\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho) : \sigma_1 > 0, \sigma_2 > 0,\ 0 < \rho < \min\left(\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right)\right\}.$$

As before, the survival function of the kth individual, conditional on the frailty variable $\epsilon^{(k)}$, is given by

$$S^{(k)}(t_k; \boldsymbol{h_0^{(k)}}|\epsilon^{(k)}) = \exp\Big[- \epsilon^{(k)} H_0^{(k)}(t_k)\Big] = \exp\Big[- H_0^{(k)}(t_k)(\frac{\mu_0}{\mu_k} Y_0 + Y_k)\Big],$$

for $k = 1, 2$. Also, the jth sub-distribution function of the kth individual, conditional on $\epsilon^{(k)}$, is

$$F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}|\epsilon^{(k)}) = \int_0^{t_k} h_{0j}^{(k)}(u_k)(\frac{\mu_0}{\mu_k} Y_0 + Y_k) \exp\Big[- H_0^{(k)}(u_k)(\frac{\mu_0}{\mu_k} Y_0 + Y_k)\Big] du,$$

for $j = 1, \cdots, L_k$ and $k = 1, 2$. Then, $F_{j_1 j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, \boldsymbol{\theta})$, the unconditional joint sub-distribution function, is given by

$$\int_0^\infty \int_0^\infty \left[\int_0^{t_1} \int_0^{t_2} \left(\prod_{k=1}^{2} h_{0j_k}^{(k)}(u_k) \epsilon^{(k)} \right) \times \exp\Big(- \sum_{k=1}^{2} \epsilon^{(k)} H_0^{(k)}(u_k) \Big) du_2 du_1 \right] g(\epsilon^{(1)}, \epsilon^{(2)}; \boldsymbol{\theta}) d\epsilon^{(1)} d\epsilon^{(2)},$$

for all $t_k > 0$, $j_k = 1, \cdots, L_k$ and $k = 1, 2$, where $g(\epsilon^{(1)}, \epsilon^{(2)}; \boldsymbol{\theta})$ denotes the joint density of $(\epsilon^{(1)}, \epsilon^{(2)})$ following a correlated Gamma distribution, as in Section 3.2.2.

From the expression of the jth sub-distribution function of the kth individual, conditional on $\epsilon^{(k)}$, or equivalently Y_0 and Y_k, we get the corresponding unconditional jth sub-distribution function, for $k = 1, 2$, as

$$\begin{aligned} F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \boldsymbol{\theta}) &= \int_0^\infty \int_0^\infty \left[\int_0^{t_k} h_{0j}^{(k)}(u_k) \left(\frac{\mu_0}{\mu_k} y_0 + y_k \right) \exp\left(- H_0^{(k)}(u_k) \left(\frac{\mu_0}{\mu_k} y_0 + y_k \right) \right) du_k \right] \\ &\qquad \times \prod_{l=0,k} \left(\frac{\mu_l^{\kappa_l}}{\Gamma(\kappa_l)} \exp\left[-\mu_l y_l \right] y_l^{\kappa_l - 1} dy_l \right) \\ &= \int_0^{t_k} h_{0j}^{(k)}(u_k) \big[I_1(u_k) + I_2(u_k) \big] du_k \ \text{(say)}, \end{aligned}$$

where

$$\begin{aligned} I_1(u_k) &= \frac{\mu_0^{\kappa_0}}{\Gamma(\kappa_0)} \frac{\mu_k^{\kappa_k}}{\Gamma(\kappa_k)} \frac{\mu_0}{\mu_k} \times \int_0^\infty \int_0^\infty \exp\left[- \left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1 \right) \sum_{l=0,k} \mu_l y_l \right] y_0^{\kappa_0} y_k^{\kappa_k - 1} dy_0 dy_k \\ &= \frac{\kappa_0}{\mu_k} \left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1 \right)^{-\kappa_0 - 1 - \kappa_k} = \frac{\kappa_0}{\mu_k} \left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1 \right)^{-\mu_k - 1} \end{aligned}$$

and

$$I_2(u_k) = \left(\prod_{l=0,k} \frac{\mu_l^{\kappa_l}}{\Gamma(\kappa_l)}\right) \int_0^\infty \int_0^\infty \exp\left[-\left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1\right) \sum_{l=0,k} \mu_l y_l\right] y_0^{\kappa_0 - 1} y_k^{\kappa_k} dy_0 dy_k$$
$$= \frac{\kappa_k}{\mu_k}\left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1\right)^{-\kappa_0 - 1 - \kappa_k} = \frac{\kappa_k}{\mu_k}\left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1\right)^{-\mu_k - 1},$$

for all $u_k > 0$. Hence,

$$F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \boldsymbol{\theta}) = \int_0^{t_k} h_{0j}^{(k)}(u_k) \left[\frac{\kappa_0 + \kappa_k}{\mu_k}\left(\frac{H_0^{(k)}(u_k)}{\mu_k} + 1\right)^{-\mu_k - 1}\right] du_k$$
$$= \int_0^{t_k} \frac{h_{0j}^{(k)}(u_k) du_k}{\left(1 + \sigma_k^2 H_0^{(k)}(u_k)\right)^{1 + \frac{1}{\sigma_k^2}}},$$

for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$. Note that this unconditional jth sub-distribution function $F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \boldsymbol{\theta})$ has the same form as that in (5.5) for the shared Gamma frailty model of the previous subsection with σ_k^2 replacing σ^2. Since this unconditional jth sub-distribution function depends on $\boldsymbol{\theta}$ only through σ_k^2, we write this as $F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \sigma_k^2)$.

Definition 5.3. The correlated Gamma frailty model (5.6) is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$ if, for some $\boldsymbol{h_0^{(k)}}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(k)}} \in \mathcal{H}_k$, for $k = 1, 2$, and $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$, where $\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho), \tilde{\boldsymbol{\theta}} = (\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\rho})$, the equality of the joint sub-distribution functions

$$F_{j_1 j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, \boldsymbol{\theta}) = F_{j_1 j_2}(t_1, t_2; \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(1)}}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(2)}}, \tilde{\boldsymbol{\theta}}),$$

for all $t_k > 0,\ j_k = 1, \cdots, L_k$ and $k = 1, 2$, implies

$$h_{0j}^{(k)}(x) = \tilde{h}_{0j}^{(k)}(x) \text{ for all } x > 0,\ j = 1, \cdots, L_k,\ \sigma_k = \tilde{\sigma}_k, \text{ for } k = 1, 2, \text{ and } \rho = \tilde{\rho}.$$

Theorem 5.2. The correlated Gamma frailty model (5.6) with non-parametric baseline cause-specific hazard functions is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$, where $\boldsymbol{\Theta} = \left\{\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho) : \sigma_1 > 0, \sigma_2 > 0,\ 0 < \rho \leq \min\left(\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right)\right\}$, provided $H_{0j}^{(k)}(\cdot)$ is differentiable and its inverse, denoted by $H_{0j}^{-(k)}(\cdot)$, exists for all $j = 1, \cdots, L_k$ and $k = 1, 2$.

Proof. Note that the equality of joint sub-distribution functions, as per **Definition 5.3**, implies equality of the unconditional jth sub-distribution functions $F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \sigma_k^2)$

and $F_j^{(k)}(t_k; \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\boldsymbol{k})}, \tilde{\sigma}_k^2)$. Therefore, following the same technique as that used in the proof of **Theorem 5.1** for the shared Gamma frailty model in the previous subsection, we have

$$h_{0j}^{(k)}(t_k) = \tilde{h}_{0j}^{(k)}(t_k) \text{ and } \sigma_k = \tilde{\sigma}_k,$$

for all $t_k > 0,\ j = 1, \cdots, L_k$ and $k = 1, 2$.

Now, the equality of joint sub-distribution functions, as per **Definition 5.3**, also implies equality of the unconditional joint survival functions $S(t_1, t_2; \boldsymbol{h}_{\mathbf{0}}^{(\mathbf{1})}, \boldsymbol{h}_{\mathbf{0}}^{(\mathbf{2})}, \boldsymbol{\theta})$ and $S(t_1, t_2; \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\mathbf{1})}, \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\mathbf{2})}, \tilde{\boldsymbol{\theta}})$. Note that the unconditional joint survival function can be obtained as

$$\begin{aligned}
S(t_1, t_2; \boldsymbol{h}_{\mathbf{0}}^{(\mathbf{1})}, \boldsymbol{h}_{\mathbf{0}}^{(\mathbf{2})}, \boldsymbol{\theta}) &= \int_0^\infty \int_0^\infty \int_0^\infty \exp\left[-\mu_0 y_0 \left(1 + \sum_{k=1}^{2} \frac{H_0^{(k)}(t_k)}{\mu_k}\right) - \sum_{k=1}^{2} \mu_k y_k \left(1 + \frac{H_0^{(k)}(t_k)}{\mu_k}\right)\right] \\
&\qquad \times \prod_{l=0,1,2} \left(\frac{\mu_l^{\kappa_l}}{\Gamma(\kappa_l)} y_l^{\kappa_l - 1} dy_l \right) \\
&= \left(1 + \sum_{k=1}^{2} \frac{H_0^{(k)}(t_k)}{\mu_k}\right)^{-\kappa_0} \times \prod_{k=1}^{2} \left(1 + \frac{H_0^{(k)}(t_k)}{\mu_k}\right)^{-\kappa_k} \\
&= \left(1 + \sum_{k=1}^{2} \sigma_k^2 H_0^{(k)}(t_k)\right)^{-\frac{\rho}{\sigma_1\sigma_2}} \times \prod_{k=1}^{2} \left(1 + \sigma_k^2 H_0^{(k)}(t_k)\right)^{-\frac{1}{\sigma_k^2} + \frac{\rho}{\sigma_1\sigma_2}},
\end{aligned}$$

for all $t_1, t_2 > 0$. Now, using the above expression in the equality of $S(t_1, t_2; \boldsymbol{h}_{\mathbf{0}}^{(\mathbf{1})}, \boldsymbol{h}_{\mathbf{0}}^{(\mathbf{2})}, \boldsymbol{\theta})$ and $S(t_1, t_2; \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\mathbf{1})}, \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\mathbf{2})}, \tilde{\boldsymbol{\theta}})$, along with the already obtained equalities $h_{0j}^{(k)}(t_k) = \tilde{h}_{0j}^{(k)}(t_k)$, for all $t_k > 0, j = 1, 2, \cdots, L_k$, and $\sigma_k = \tilde{\sigma}_k$, for $k = 1, 2$, we have

$$\begin{aligned}
&\left(1 + \sum_{k=1}^{2} \sigma_k^2 H_0^{(k)}(t_k)\right)^{-\frac{\rho}{\sigma_1\sigma_2}} \times \prod_{k=1}^{2} \left(1 + \sigma_k^2 H_0^{(k)}(t_k)\right)^{-\frac{1}{\sigma_k^2} + \frac{\rho}{\sigma_1\sigma_2}} \\
&\qquad = \left(1 + \sum_{k=1}^{2} \sigma_k^2 H_0^{(k)}(t_k)\right)^{-\frac{\tilde{\rho}}{\sigma_1\sigma_2}} \times \prod_{k=1}^{2} \left(1 + \sigma_k^2 H_0^{(k)}(t_k)\right)^{-\frac{1}{\sigma_k^2} + \frac{\tilde{\rho}}{\sigma_1\sigma_2}},
\end{aligned}$$

for all $t_1, t_2 > 0$. This implies

$$\left[\frac{\left(\sigma_1^2 H_0^{(1)}(t_1) + 1\right)\left(\sigma_2^2 H_0^{(2)}(t_2) + 1\right)}{\left(\sigma_1^2 H_0^{(1)}(t_1) + \sigma_2^2 H_0^{(2)}(t_2) + 1\right)} \right]^{\frac{\rho - \tilde{\rho}}{\sigma_1\sigma_2}} = 1,$$

for all $t_1, t_2 > 0$. Hence, $\rho = \tilde{\rho}$.

Therefore, the correlated Gamma frailty model (5.6) is identifiable. □

5.2.4 Shared Cause-specific Gamma Frailty

The shared cause-specific frailty model is given by

$$\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon}) = h_{0j}^{(k)}(t_k)\epsilon_j, \tag{5.7}$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$, where ϵ_j is the shared frailty variable between the two individuals for failure due to cause j, $j = 1, \cdots, L$, and $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_L)$. Note that, in this case, $L_1 = L_2 = L$. We assume that the cause-specific frailty variables ϵ_j's are independent with ϵ_j following Gamma$(\frac{1}{\sigma_j^2}, \frac{1}{\sigma_j^2})$ distribution with $\sigma_j > 0$ having density similar to that in (3.4), so that $E(\epsilon_j) = 1$, for $j = 1, \cdots, L$. Therefore, the frailty parameter space $\boldsymbol{\Theta}$ is given by

$$\boldsymbol{\Theta} = \big\{\boldsymbol{\theta} = (\sigma_1, \cdots, \sigma_L) : \sigma_j > 0, j = 1, \cdots, L\big\}.$$

. Then, as before, the survival function of the kth individual, conditional on $\boldsymbol{\epsilon}$, is

$$S^{(k)}(t_k; \boldsymbol{h_0^{(k)}}|\boldsymbol{\epsilon}) = \exp\bigg[- \sum_{j=1}^{L} H_{0j}^{(k)}(t_k)\epsilon_j \bigg],$$

and the jth sub-distribution function of the kth individual, conditional on $\boldsymbol{\epsilon}$, is

$$F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}|\boldsymbol{\epsilon}) = \int_0^{t_k} h_{0j}^{(k)}(u_k)\epsilon_j \exp\bigg[- \sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k)\epsilon_{j'} \bigg] du_k,$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. The unconditional joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}; \boldsymbol{\theta})$ is given by

$$\int_0^{\infty} \cdots \int_0^{\infty} \bigg[\int_0^{t_1} \int_0^{t_2} \Big(h_{0j_k}^{(k)}(u_k)\epsilon_{j_k} \Big) \times \exp\bigg(- \sum_{j=1}^{L} \epsilon_j \sum_{k=1}^{2} H_{0j}^{(k)}(u_k) \bigg) du_2 du_1 \bigg] g(\boldsymbol{\epsilon}; \boldsymbol{\theta}) d\boldsymbol{\epsilon},$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, where $g(\boldsymbol{\epsilon}; \boldsymbol{\theta})$ denotes the product density of L independent Gamma densities for $\epsilon_1, \cdots, \epsilon_L$, as described above. The unconditional sub-distribution function $F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \boldsymbol{\theta})$ is given by

$$= \int_0^{t_k} h_{0j}^{(k)}(u_k) \bigg[\int_0^{\infty} \cdots \int_0^{\infty} \epsilon_j \exp\bigg(- \sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k)\epsilon_{j'} \bigg) \prod_{j'=1}^{L} \frac{e^{-\frac{\epsilon_{j'}}{\sigma_{j'}^2}} \epsilon_{j'}^{\frac{1}{\sigma_{j'}^2} - 1}}{\sigma_{j'}^{\frac{2}{\sigma_{j'}^2}} \Gamma(\frac{1}{\sigma_{j'}^2})} d\epsilon_{j'} \bigg] du_k$$

$$
\begin{aligned}
&= \int_0^{t_k} h_{0j}^{(k)}(u_k) \Bigg[\Bigg(\frac{1}{\sigma_j^{\frac{2}{\sigma_j^2}} \Gamma(\frac{1}{\sigma_j^2})} \int_0^{\infty} \exp\Big(-\epsilon_j \Big(H_{0j}^{(k)}(u_k) + \frac{1}{\sigma_j^2} \Big) \Big) \epsilon_j^{\frac{1}{\sigma_j^2}} d\epsilon_j \Bigg) \times \\
&\qquad \prod_{j'(\neq j)=1}^{L} \Bigg(\frac{1}{\sigma_{j'}^{\frac{2}{\sigma_{j'}^2}} \Gamma(\frac{1}{\sigma_{j'}^2})} \int_0^{\infty} \exp\Big(-\epsilon_{j'} \Big(H_{0j'}^{(k)}(u) + \frac{1}{\sigma_{j'}^2} \Big) \Big) \epsilon_{j'}^{\frac{1}{\sigma_{j'}^2}-1} d\epsilon_{j'} \Bigg) \Bigg] du_k \\
&= \int_0^{t_k} h_{0j}^{(k)}(u_k) \Bigg[\Bigg(\frac{1}{\sigma_j^{\frac{2}{\sigma_j^2}} \Gamma(\frac{1}{\sigma_j^2})} \frac{\Gamma(\frac{1}{\sigma_j^2}+1)}{\Big(H_{0j}^{(k)}(u_k) + \frac{1}{\sigma_j^2} \Big)^{\frac{1}{\sigma_j^2}+1}} \Bigg) \times \\
&\qquad\qquad \prod_{j'(\neq j)=1}^{L} \Bigg(\frac{1}{\sigma_{j'}^{\frac{2}{\sigma_{j'}^2}} \Gamma(\frac{1}{\sigma_{j'}^2})} \times \frac{\Gamma(\frac{1}{\sigma_{j'}^2})}{\Big(H_{0j'}^{(k)}(u_k) + \frac{1}{\sigma_{j'}^2} \Big)^{\frac{1}{\sigma_{j'}^2}}} \Bigg) \Bigg] du_k \\
&= \int_0^{t_k} h_{0j}^{(k)}(u_k) \Bigg[\Bigg(\frac{1}{\Big(1 + \sigma_j^2 H_{0j}^{(k)}(u_k) \Big)^{\frac{1}{\sigma_j^2}+1}} \Bigg) \times \prod_{j'(\neq j)=1}^{L} \Bigg(\frac{1}{\Big(1 + \sigma_{j'}^2 H_{0j'}^{(k)}(u_k) \Big)^{\frac{1}{\sigma_{j'}^2}}} \Bigg) \Bigg] du_k \\
&= \int_0^{t_k} h_{0j}^{(k)}(u_k) \Bigg[\frac{1}{1 + \sigma_j^2 H_{0j}^{(k)}(u_k)} \times \prod_{j'=1}^{L} \Bigg(\frac{1}{\Big(1 + \sigma_{j'}^2 H_{0j'}^{(k)}(u_k) \Big)^{\frac{1}{\sigma_{j'}^2}}} \Bigg) \Bigg] du_k, \qquad (5.8)
\end{aligned}
$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$.

Definition 5.4. The shared cause-specific Gamma frailty model (5.7) is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$ if, for some $\boldsymbol{h_0^{(k)}}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{(\boldsymbol{k})} \in \mathcal{H}_k$, for $k = 1, 2$, and $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$, where $\boldsymbol{\theta} = (\sigma_1, \cdots, \sigma_L),\ \tilde{\boldsymbol{\theta}} = (\tilde{\sigma}_1, \cdots, \tilde{\sigma}_L)$, the equality of the joint sub-distribution functions

$$
F_{j_1 j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}; \boldsymbol{\theta}) = F_{j_1 j_2}(t_1, t_2; \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{(\boldsymbol{1})}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{(\boldsymbol{2})}; \tilde{\boldsymbol{\theta}}),
$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, implies

$$
h_{0j}^{(k)}(x) = \tilde{h}_{0j}^{(k)}(x) \text{ for all } x > 0, \quad \sigma_j = \tilde{\sigma}_j, \text{ for } j = 1, \cdots, L \text{ and } k = 1, 2.
$$

Theorem 5.3. The shared cause-specific Gamma frailty model (5.7) with non-parametric baseline cause-specific hazard functions is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$, where $\boldsymbol{\Theta} = \big\{ \boldsymbol{\theta} = (\sigma_1, \cdots, \sigma_L) : \sigma_1 > 0, \cdots, \sigma_L > 0 \big\}$, provided $H_{0j}^{(k)}(\cdot)$ is differentiable and its inverse, denoted by $H_{0j}^{-(k)}(\cdot)$, exists for all $j = 1, \cdots, L$ and $k = 1, 2$.

Proof. As before, the equality of joint sub-distribution functions, as per **Definition 5.4**, implies equality of the unconditional jth sub-distribution functions $F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \boldsymbol{\theta})$

and $F_j^{(k)}(t_k; \tilde{\boldsymbol{h}}_{\mathbf{0}}^{(\boldsymbol{k})}, \tilde{\boldsymbol{\theta}})$, as given by (5.8), for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Hence, we have equality of the corresponding sub-density functions, as given by

$$h_{0j}^{(k)}(t_k)\big\{1+\sigma_j^2 H_{0j}^{(k)}(t_k)\big\}^{-1}\prod_{j'=1}^{L}\big\{1+\sigma_{j'}^2 H_{0j'}^{(k)}(t_k)\big\}^{-\frac{1}{\sigma_{j'}^2}} = \tilde{h}_{0j}^{(k)}(t_k)\big\{1+\tilde{\sigma}_j^2 \tilde{H}_{0j}^{(k)}(t_k)\big\}^{-1}\prod_{j'=1}^{L}\big\{1+\tilde{\sigma}_{j'}^2 \tilde{H}_{0j'}^{(k)}(t_k)\big\}^{-\frac{1}{\tilde{\sigma}_{j'}^2}}, \quad (5.9)$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. This implies

$$\frac{\frac{d}{dt_k}\Big[\frac{1}{\sigma_j^2}\log(1+\sigma_j^2 H_{0j}^{(k)}(t_k))\Big]}{\frac{d}{dt_k}\Big[\frac{1}{\tilde{\sigma}_j^2}\log(1+\tilde{\sigma}_j^2 \tilde{H}_{0j}^{(k)}(t_k))\Big]} = \frac{\big\{1+\tilde{\sigma}_j^2 \tilde{H}_{0j}^{(k)}(t_k)\big\}^{-\frac{1}{\tilde{\sigma}_j^2}}\prod_{j'(\neq j)=1}^{L}\big\{1+\tilde{\sigma}_{j'}^2 \tilde{H}_{0j'}^{(k)}(t_k)\big\}^{-\frac{1}{\tilde{\sigma}_{j'}^2}}}{\big\{1+\sigma_j^2 H_{0j}^{(k)}(t_k)\big\}^{-\frac{1}{\sigma_j^2}}\prod_{j'(\neq j)=1}^{L}\big\{1+\sigma_{j'}^2 H_{0j'}^{(k)}(t_k)\big\}^{-\frac{1}{\sigma_{j'}^2}}},$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Let us now write $y = \log[1+\sigma_j^2 H_{0j}^{(k)}(t_k)]$ and $z = \log[1+\tilde{\sigma}_j^2 \tilde{H}_{0j}^{(k)}(t_k)]$. Then, from the above equality, we have

$$\frac{\tilde{\sigma}_j^2}{\sigma_j^2}\times\frac{dy}{dz} = \frac{e^{-\frac{z}{\tilde{\sigma}_j^2}}\prod_{j'(\neq j)=1}^{L}\Big[1+\tilde{\sigma}_{j'}^2 \tilde{H}_{0j'}^{(k)}\Big(\tilde{H}_{0j}^{-(k)}\big(\frac{e^z-1}{\tilde{\sigma}_j^2}\big)\Big)\Big]^{-\frac{1}{\tilde{\sigma}_{j'}^2}}}{e^{-\frac{y}{\sigma_j^2}}\prod_{j'(\neq j)=1}^{L}\Big[1+\sigma_{j'}^2 H_{0j'}^{(k)}\Big(H_{0j}^{-(k)}\big(\frac{e^y-1}{\sigma_j^2}\big)\Big)\Big]^{-\frac{1}{\sigma_{j'}^2}}}.$$

This is equivalent to, after integrating from 0 to some $w > 0$,

$$\frac{\int_0^w e^{-\frac{y}{\sigma_j^2}}\prod_{j'(\neq j)=1}^{L}\Big[1+\sigma_{j'}^2 H_{0j'}^{(k)}\Big(H_{0j}^{-(k)}\big(\frac{e^y-1}{\sigma_j^2}\big)\Big)\Big]^{-\frac{1}{\sigma_{j'}^2}}dy}{\int_0^w e^{-\frac{z}{\tilde{\sigma}_j^2}}\prod_{j'(\neq j)=1}^{L}\Big[1+\tilde{\sigma}_{j'}^2 \tilde{H}_{0j'}^{(k)}\Big(\tilde{H}_{0j}^{-(k)}\big(\frac{e^z-1}{\tilde{\sigma}_j^2}\big)\Big)\Big]^{-\frac{1}{\tilde{\sigma}_{j'}^2}}dz} = \frac{\sigma_j^2}{\tilde{\sigma}_j^2}.$$

Now, taking limit as $w \to 0+$ and applyling L'opital rule, we get

$$\lim_{w\to 0+}\frac{e^{-\frac{w}{\sigma_j^2}}\prod_{j'(\neq j)=1}^{L}\Big[1+\sigma_{j'}^2 H_{0j'}^{(k)}\Big(H_{0j}^{-(k)}\big(\frac{e^w-1}{\sigma_j^2}\big)\Big)\Big]^{-\frac{1}{\sigma_{j'}^2}}}{e^{-\frac{w}{\tilde{\sigma}_j^2}}\prod_{j'(\neq j)=1}^{L}\Big[1+\tilde{\sigma}_{j'}^2 \tilde{H}_{0j'}^{(k)}\Big(\tilde{H}_{0j}^{-(k)}\big(\frac{e^w-1}{\tilde{\sigma}_j^2}\big)\Big)\Big]^{-\frac{1}{\tilde{\sigma}_{j'}^2}}} = \frac{\sigma_j^2}{\tilde{\sigma}_j^2}.$$

Note that, as in Section 5.2.2, the composite functions appearing in the above equality

are continuous functions. Hence, for $j' \neq j$,

$$\lim_{w\to 0+} H_{0j'}^{(k)}\left(H_{0j}^{-(k)}(\frac{e^w-1}{\sigma_j^2})\right) = \lim_{w\to 0+} \tilde{H}_{0j'}^{(k)}\left(\tilde{H}_{0j}^{-(k)}(\frac{e^w-1}{\tilde{\sigma}_j^2})\right) = 0.$$

Hence, since $\lim_{w\to 0+} e^{-\frac{w}{\sigma_j^2}} = \lim_{w\to 0+} e^{-\frac{w}{\tilde{\sigma}_j^2}} = 1$, we get $\frac{\sigma_j^2}{\tilde{\sigma}_j^2} = 1$, or $\sigma_j^2 = \tilde{\sigma}_j^2$, for $j = 1, \cdots, L$. Using this equality in (5.9) and then summing over $j = 1, \cdots, L$, we get

$$\left(\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2 H_{0j'}^{(k)}(t_k)\right]^{-\frac{1}{\sigma_{j'}^2}}\right) \times \left(\sum_{j=1}^{L}\frac{h_{0j}^{(k)}(t_k)}{1+\sigma_j^2 H_{0j}^{(k)}(t_k)}\right)$$
$$= \left(\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2 \tilde{H}_{0j'}^{(k)}(t_k)\right]^{-\frac{1}{\sigma_{j'}^2}}\right) \times \left(\sum_{j=1}^{L}\frac{\tilde{h}_{0j}^{(k)}(t_k)}{1+\sigma_j^2 \tilde{H}_{0j}^{(k)}(t_k)}\right),$$

for all $t_k > 0$ and $k = 1, 2$. Writing

$$c(t_k) = \log\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2 H_{0j'}^{(k)}(t_k)\right]^{\frac{1}{\sigma_{j'}^2}} \quad \text{and} \quad \tilde{c}(t_k) = \log\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2 \tilde{H}_{0j'}^{(k)}(t_k)\right]^{\frac{1}{\sigma_{j'}^2}},$$

the above equality reduces to

$$e^{-c(t_k)}\frac{d}{dt_k}c(t_k) = e^{-\tilde{c}(t_k)}\frac{d}{dt_k}\tilde{c}(t_k),$$

for all $t_k > 0$. By integrating both sides, we get $e^{-c(t_k)} = e^{-\tilde{c}(t_k)}$, since $\lim_{t_k\to 0+} c(t_k) = \lim_{t_k\to 0+} \tilde{c}(t_k) = 0$. This implies $c(t_k) = \tilde{c}(t_k)$ for all $t_k > 0$. That is,

$$\prod_{j'=1}^{L}\left[1+\sigma_{j'}^2 H_{0j'}^{(k)}(t_k)\right]^{\frac{1}{\sigma_{j'}^2}} = \prod_{j'=1}^{L}\left[1+\sigma_{j'}^2 \tilde{H}_{0j'}^{(k)}(t_k)\right]^{\frac{1}{\sigma_{j'}^2}}.$$

Using this equality in (5.9) and then integrating, we get

$$\int_0^{t_k} h_{0j}^{(k)}(u_k)\left[1+\sigma_j^2 H_{0j}^{(k)}(u_k)\right]^{-1} du_k = \int_0^{t_k} \tilde{h}_{0j}^{(k)}(u_k)\left[1+\sigma_j^2 \tilde{H}_{0j}^{(k)}(u_k)\right]^{-1} du_k.$$

This gives $H_{0j}^{(k)}(t_k) = \tilde{H}_{0j}^{(k)}(t_k)$, or $h_{0j}^{(k)}(t_k) = \tilde{h}_{0j}^{(k)}(t_k)$, for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Hence, the theorem is proved. □

5.2.5 Correlated Cause-specific Gamma Frailty

The correlated cause-specific frailty model is given by

$$\lambda_j^{(k)}(t_k|\boldsymbol{\epsilon^{(k)}}) = h_{0j}^{(k)}(t_k)\epsilon_j^{(k)}, \tag{5.10}$$

where $\boldsymbol{\epsilon^{(k)}} = (\epsilon_1^{(k)}, \cdots, \epsilon_L^{(k)})$, for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Here, we also have $L_1 = L_2 = L$. We assume that, while $(\epsilon_j^{(1)}, \epsilon_j^{(2)})$ follows a correlated Gamma frailty distribution for a particular j (See Sections 3.2.2 and 3.2.4 for details), the L pairs of frailty variables $(\epsilon_j^{(1)}, \epsilon_j^{(2)})$, for $j = 1, \cdots, L$, are independent. As in Section 5.2.3, the frailty parameter space for correlated cause-specific Gamma frailty model is given by

$$\boldsymbol{\Theta} = \big\{\boldsymbol{\theta} = (\sigma_{1j}, \sigma_{2j}, \rho_j) : \sigma_{1j} > 0, \sigma_{2j} > 0,\ 0 < \rho_j < \min\big(\frac{\sigma_{1j}}{\sigma_{2j}}, \frac{\sigma_{2j}}{\sigma_{1j}}\big),\ j = 1, \cdots, L\big\}.$$

. The survival function of the kth individual, conditional on the frailty vector $\boldsymbol{\epsilon^{(k)}}$, is

$$S^{(k)}(t_k; \boldsymbol{h_0^{(k)}}|\boldsymbol{\epsilon^{(k)}}) = \exp\bigg[-\sum_{j=1}^{L} H_{0j}^{(k)}(t_k)\epsilon_j^{(k)}\bigg],$$

and the jth sub-distribution function of the kth individual, conditional on frailty vector $\boldsymbol{\epsilon^{(k)}}$, is

$$F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}|\boldsymbol{\epsilon^{(k)}}) = \int_0^{t_k} h_{0j}^{(k)}(u_k)\epsilon_j^{(k)} \exp\bigg[-\sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k)\epsilon_{j'}^{(k)}\bigg] du_k,$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. The unconditional joint sub-distribution $F_{j_1 j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, \boldsymbol{\theta})$ is

$$\int_0^{\infty} \cdots \int_0^{\infty} \bigg[\int_0^{t_1}\int_0^{t_2} \prod_{k=1}^{2} \Big(\epsilon_{j_k}^{(k)} h_{0j_k}^{(k)}(u_k)\Big) \times \exp\bigg(-\sum_{k=1}^{2}\sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k)\epsilon_{j'}^{(k)}\bigg) du_2 du_1\bigg]$$
$$\times\, g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}; \boldsymbol{\theta}) d\boldsymbol{\epsilon^{(1)}} d\boldsymbol{\epsilon^{(2)}},$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, where $g(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}}; \boldsymbol{\theta})$ denotes the joint density of $(\boldsymbol{\epsilon^{(1)}}, \boldsymbol{\epsilon^{(2)}})$ following correlated cause-specific Gamma distribution, as described above.

As in Section 5.2.3, from the expression of the jth sub-distribution function of

the kth individual, conditional on $\boldsymbol{\epsilon}^{(\boldsymbol{k})}$, we get the corresponding unconditional jth sub-distribution function, for $k = 1, 2$, as

$$F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}; \boldsymbol{\theta}) = \int_0^{t_k} \int_0^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} h_{0j}^{(k)}(u_k) \left(\frac{\mu_{0j}}{\mu_{kj}} y_{0j} + y_{kj} \right) \exp \Big[- \sum_{j'=1}^{L} H_{0j'}^{(k)}(u_k) \big(\frac{\mu_{0j'}}{\mu_{kj'}} y_{0j'} + y_{kj'} \big) \Big] \times$$
$$\prod_{l=0,k} \prod_{j'=1}^{L} \left(\frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})} e^{-\mu_{lj'} y_{lj'}} y_{lj'}^{\kappa_{lj'}-1} dy_{lj'} \right) du_k$$
$$= \int_0^{t_k} h_{0j}^{(k)}(u_k) \big(I_{1j}(u_k) + I_{2j}(u_k) \big) du_k \qquad \text{(say)},$$

where $I_{1j}(u_k)$ is given by

$$\left(\frac{\mu_{0j}}{\mu_{kj}} \prod_{l=0,k} \prod_{j'=1}^{L} \frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})} \right) \int_0^{\infty} \cdots \int_0^{\infty} \exp \Big[- \sum_{j'=1}^{L} \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big) \big(\sum_{l=0,k} y_{lj'} \mu_{lj'} \big) \Big]$$
$$\times y_{0j} \prod_{j'=1}^{L} \prod_{l=0,k} \left(y_{lj'}^{\kappa_{lj'}-1} dy_{lj'} \right)$$
$$= \left(\frac{\mu_{0j}}{\mu_{kj}} \frac{\mu_{0j}^{\kappa_{0j}}}{\Gamma(\kappa_{0j})} \int_0^{\infty} \exp \Big[- \Big(1 + \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}} \Big) y_{0j} \mu_{0j} \Big] y_{0j}^{\kappa_{0j}} dy_{0j} \right)$$
$$\times \left(\prod_{j'(\neq j)=1}^{L} \frac{\mu_{0j'}^{\kappa_{0j'}}}{\Gamma(\kappa_{0j'})} \int_0^{\infty} \exp \Big[- \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big) y_{0j'} \mu_{0j'} \Big] y_{0j'}^{\kappa_{0j'}-1} dy_{0j'} \right)$$
$$\times \left(\prod_{j'=1}^{L} \frac{\mu_{kj'}^{\kappa_{kj'}}}{\Gamma(\kappa_{kj'})} \int_0^{\infty} \exp \Big[- \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big) y_{kj'} \mu_{kj'} \Big] y_{kj'}^{\kappa_{kj'}-1} dy_{kj'} \right)$$
$$= \left(\frac{\mu_{0j}}{\mu_{kj}} \frac{\kappa_{0j}}{\mu_{0j}} \Big(1 + \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}} \Big)^{-\kappa_{0j}-1} \right) \times \left(\prod_{j'(\neq j)=1}^{L} \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big)^{-\kappa_{0j'}} \right)$$
$$\times \left(\prod_{j'=1}^{L} \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big)^{-\kappa_{kj'}} \right)$$
$$= \left(\frac{\kappa_{0j}}{\mu_{kj}} \prod_{j'=1}^{L} \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big)^{-\kappa_{0j'}-\kappa_{kj'}} \right) \times \Big(1 + \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}} \Big)^{-1}$$
$$= \frac{\kappa_{0j}}{\mu_{kj}} \Big(1 + \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}} \Big)^{-1} \prod_{j'=1}^{L} \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big)^{-\mu_{kj'}}$$

and $I_{2j}(u_k)$ is given by

$$\left(\prod_{l=0,k}\prod_{j'=1}^{L}\frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})}\right)\int_0^\infty\cdots\int_0^\infty\exp\Big[-\sum_{j'=1}^{L}\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)\sum_{l=0,k}y_{lj'}\mu_{lj'}\Big]y_{kj}\prod_{l=0,k}\prod_{j'=1}^{L}y_{lj'}^{\kappa_{lj'}-1}dy_{lj'}$$

$$\begin{aligned}
&= \left(\frac{\mu_{kj}^{\kappa_{kj}}}{\Gamma(\kappa_{kj})}\int_0^\infty\exp\Big[-\Big(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\Big)y_{kj}\mu_{kj}\Big]y_{kj}^{\kappa_{kj}}dy_{kj}\right)\\
&\times\left(\prod_{j'(\neq j)=1}^{L}\frac{\mu_{kj'}^{\kappa_{kj'}}}{\Gamma(\kappa_{kj'})}\int_0^\infty\exp\Big[-\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)y_{kj'}\mu_{kj'}\Big]y_{kj'}^{\kappa_{kj'}-1}dy_{kj'}\right)\\
&\times\left(\prod_{j'=1}^{L}\frac{\mu_{0j'}^{\kappa_{0j'}}}{\Gamma(\kappa_{0j'})}\int_0^\infty\exp\Big[-\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)y_{0j'}\mu_{0j'}\Big]y_{0j'}^{\kappa_{0j'}-1}dy_{0j'}\right)\\
&= \left(\frac{\kappa_{kj}}{\mu_{kj}}\Big(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\Big)^{-\kappa_{kj}-1}\right)\times\left(\prod_{j'(\neq j)=1}^{L}\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)^{-\kappa_{kj'}}\right)\\
&\quad\times\left(\prod_{j'=1}^{L}\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)^{-\kappa_{0j'}}\right)\\
&= \frac{\kappa_{kj}}{\mu_{kj}}\Big(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\Big)^{-1}\prod_{j'=1}^{L}\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)^{-\kappa_{0j'}-\kappa_{kj'}}\\
&= \frac{\kappa_{kj}}{\mu_{kj}}\Big(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\Big)^{-1}\prod_{j'=1}^{L}\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)^{-\mu_{kj'}}
\end{aligned}$$

for all $u_k > 0$ and $k = 1, 2$. Therefore, we have

$$\begin{aligned}
F_j^{(k)}(t_k;\boldsymbol{h_0^{(k)}},\boldsymbol{\theta}) &= \int_0^{t_k}h_{0j}^{(k)}(u_k)\left[\frac{\kappa_{0j}+\kappa_{kj}}{\mu_{kj}}\Big(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\Big)^{-1}\times\prod_{j'=1}^{L}\Big(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\Big)^{-\mu_{kj'}}\right]du_k\\
&= \int_0^{t_k}h_{0j}^{(k)}(u_k)\left[\frac{1}{1+\sigma_{kj}^2H_{0j}^{(k)}(u_k)}\prod_{j'=1}^{L}\frac{1}{\Big(1+\sigma_{kj'}^2H_{0j'}^{(k)}(u_k)\Big)^{\frac{1}{\sigma_{kj'}^2}}}\right]du_k
\end{aligned}$$

for all $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. Note that this unconditional jth sub-distribution function of the kth individual has the similar form as that in (5.8) for the shared cause-specific Gamma frailty model of the previous subsection with σ_j's

replaced by σ_{kj}'s.

Definition 5.5. The correlated cause-specific Gamma frailty model (5.10) is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$ if, for some $\boldsymbol{h_0^{(k)}}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(k)}} \in \mathcal{H}_k$, for $k = 1, 2$, and $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$, where $\boldsymbol{\theta} = (\sigma_{1j}, \sigma_{2j}, \rho_j;\ j = 1, \cdots, L)$, $\tilde{\boldsymbol{\theta}} = (\tilde{\sigma}_{1j}, \tilde{\sigma}_{2j}, \tilde{\rho}_j;\ j = 1, \cdots, L)$, the equality of the joint sub-distribution functions

$$F_{j_1,j_2}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}; \boldsymbol{\theta}) = F_{j_1,j_2}(t_1, t_2; \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(1)}}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(2)}}; \tilde{\boldsymbol{\theta}}),$$

for all $t_k > 0,\ j_k = 1, \cdots, L$ and $k = 1, 2$, implies

$h_{0j}^{(k)}(x) = \tilde{h}_{0j}^{(k)}(x)$ for all $x > 0,\ k = 1, 2,$ and $\sigma_{1j} = \tilde{\sigma}_{1j},\ \sigma_{2j} = \tilde{\sigma}_{2j},\ \rho_j = \tilde{\rho}_j,$ for $j = 1, \cdots, L$.

Theorem 5.4. The correlated cause-specific Gamma frailty model (5.10) with non-parametric baseline cause-specific hazard functions is identifiable within the family $\mathcal{H}_1 \times \mathcal{H}_2 \times \boldsymbol{\Theta}$, where $\boldsymbol{\Theta} = \big\{\boldsymbol{\theta} = (\sigma_{1j}, \sigma_{2j}, \rho_j) : \sigma_{1j} > 0, \sigma_{2j} > 0,\ 0 < \rho_j < \min\left(\frac{\sigma_{1j}}{\sigma_{2j}}, \frac{\sigma_{2j}}{\sigma_{1j}}\right),\ j = 1, \cdots, L\big\}$, provided $H_{0j}^{(k)}(\cdot)$ is differentiable and its inverse, denoted by $H_{0j}^{-(k)}(\cdot)$, exists for all $j = 1, \cdots, L$ and $k = 1, 2$.

Proof. Equality of the unconditional sub-distribution functions $F_j^{(k)}(t_k; \boldsymbol{h_0^{(k)}}, \boldsymbol{\theta})$ and $F_j^{(k)}(t_k; \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(k)}}, \tilde{\boldsymbol{\theta}})$ follows from the condition of equality of the corresponding joint sub-distribution functions in **Definition 5.5**. Therefore, following similar techniques as those used in the proof of **Theorem 5.3** for the shared cause-specific Gamma frailty model in the previous section, we get

$$h_{0j}^{(k)}(t_k) = \tilde{h}_{0j}^{(k)}(t_k) \text{ and } \sigma_j = \tilde{\sigma}_j$$

for $t_k > 0,\ j = 1, \cdots, L$ and $k = 1, 2$. In order to prove $\rho_j = \tilde{\rho}_j$, for $j = 1, \cdots, L$, we need to consider equality of the unconditional joint sub-distribution functions $F_{j,j}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, \boldsymbol{\theta})$ and $F_{j,j}(t_1, t_2; \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(1)}}, \tilde{\boldsymbol{h}}_{\boldsymbol{0}}^{\boldsymbol{(2)}}, \tilde{\boldsymbol{\theta}})$, for all $j = 1, 2, \cdots, L$, which also follows from **Definition 5.5**. Now, the unconditional joint sub-distribution function $F_{j,j}(t_1, t_2; \boldsymbol{h_0^{(1)}}, \boldsymbol{h_0^{(2)}}, \boldsymbol{\theta})$ can be obtained as

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{t_1}\int_0^{t_2}\left(\prod_{K=1}^{2} h_{0j}^{(k)}(u_k)\epsilon_j^{(k)}\right) \times \exp\left(-\sum_{j'=1}^{L}\sum_{k=1}^{2} H_{0j'}^{(k)}(u_k)\epsilon_{j'}^{(k)}\right) du_2 du_1\right] \\
&= \int_0^{t_1}\int_0^{t_2}\int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty}\left(\prod_{K=1}^{2} h_{0j}^{(k)}(u_k)\right) \times \left(\frac{\mu_{0j}^2}{\mu_{1j}\mu_{2j}} y_{0j}^2 + \frac{\mu_{0j}}{\mu_{1j}} y_{0j}y_{2j} + \frac{\mu_{0j}}{\mu_{2j}} y_{0j}y_{1j} + y_{1j}y_{2j}\right)
\end{aligned}$$

$$\times \exp\Big[- \sum_{j'(\neq j)=1}^{L} \sum_{k=1}^{2} H_{0j'}^{(k)}(u_k) \Big(\frac{\mu_{0j'}}{\mu_{kj'}} y_{0j'} + y_{kj'} \Big) \Big] \times \exp\Big[- \sum_{k=1}^{2} H_{0j}^{(k)}(u_k) \Big(\frac{\mu_{0j}}{\mu_{kj}} y_{0j} + y_{kj} \Big) \Big]$$
$$\times \left(\prod_{l=0}^{2} \prod_{j'=1}^{L} \frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})} e^{-\mu_{lj'} y_{lj'}} y_{lj'}^{\kappa_{lj'}-1} dy_{lj'} \right) du_2 du_1$$
$$= \int_0^{t_1} \int_0^{t_2} \left(\prod_{k=1}^{2} h_{0j}^{(k)}(u_k) \right) \times \big[I_{1j}(u_1,u_2) + I_{2j}(u_1,u_2) + I_{3j}(u_1,u_2) + I_{4j}(u_1,u_2) \big] du_2 du_1, \text{ (say)},$$

where

$$I_{1j}(u_1,u_2) = \frac{\mu_{0j}^2}{\mu_{1j}\mu_{2j}} \int_0^{\infty} \cdots \int_0^{\infty} \exp\Big[- \sum_{j'(\neq j)=1}^{L} \sum_{k=1}^{2} H_{0j'}^{(k)}(u_k) \Big(\frac{\mu_{0j'}}{\mu_{kj'}} y_{0j'} + y_{kj'} \Big) \Big]$$
$$\times \exp\Big[- \sum_{k=1}^{2} H_{0j}^{(k)}(u_k) \Big(\frac{\mu_{0j}}{\mu_{kj}} y_{0j} + y_{kj} \Big) \Big] y_{0j}^2$$
$$\times \prod_{l=0}^{2} \prod_{j'=1}^{L} \left(\frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})} e^{-\mu_{lj'} y_{lj'}} y_{lj'}^{\kappa_{lj'}-1} dy_{lj'} \right)$$
$$= \frac{\mu_{0j}^2}{\mu_{1j}\mu_{2j}} \prod_{j'(\neq j)=1}^{L} \frac{\mu_{0j'}^{\kappa_{0j'}}}{\Gamma(\kappa_{0j'})} \int_0^{\infty} \exp\Big[- \Big(1 + \sum_{k=1}^{2} \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big) \mu_{0j'} y_{0j'} \Big] y_{0j'}^{\kappa_{0j'}-1} dy_{0j'}$$
$$\times \frac{\mu_{0j}^{\kappa_{0j}}}{\Gamma(\kappa_{0j})} \int_0^{\infty} \exp\Big[- \Big(1 + \sum_{k=1}^{2} \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}} \Big) \mu_{0j} y_{0j} \Big] y_{0j}^{\kappa_{0j}+1} dy_{0j}$$
$$\times \prod_{j'=1}^{L} \prod_{k=1}^{2} \left(\frac{\mu_{kj'}^{\kappa_{kj'}}}{\Gamma(\kappa_{kj'})} \int_0^{\infty} \exp\Big[- \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big) \mu_{kj'} y_{kj'} \Big] y_{kj'}^{\kappa_{kj'}-1} dy_{kj'} \right)$$
$$= \frac{\kappa_{0j}(\kappa_{0j}+1)}{\mu_{1j}\mu_{2j}} \Big(1 + \sum_{k=1}^{2} \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}} \Big)^{-2} \times$$
$$\prod_{j'=1}^{L} \Big[\Big(1 + \sum_{k=1}^{2} \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big)^{-\kappa_{0j'}} \times \prod_{k=1}^{2} \Big(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}} \Big)^{-\kappa_{kj'}} \Big],$$
$$I_{2j}(u_1,u_2) = \frac{\mu_{0j}}{\mu_{1j}} \int_0^{\infty} \cdots \int_0^{\infty} \exp\Big[- \sum_{j'(\neq j)=1}^{L} \sum_{k=1}^{2} H_{0j'}^{(k)}(u_k) \Big(\frac{\mu_{0j'}}{\mu_{kj'}} y_{0j'} + y_{kj'} \Big) \Big]$$
$$\times \exp\Big[- \sum_{k=1}^{2} H_{0j}^{(k)}(u_k) \Big(\frac{\mu_{0j}}{\mu_{kj}} y_{0j} + y_{kj} \Big) \Big] y_{0j} y_{2j}$$
$$\times \left(\prod_{l=0}^{2} \prod_{j'=1}^{L} \frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})} e^{-\mu_{lj'} y_{lj'}} y_{lj'}^{\kappa_{lj'}-1} dy_{lj'} \right)$$

$$= \frac{\mu_{0j}}{\mu_{1j}} \prod_{j'(\neq j)=1}^{L} \frac{\mu_{0j'}^{\kappa_{0j'}}}{\Gamma(\kappa_{0j'})} \int_0^\infty \exp\left[-\left(1 + \sum_{k=1}^{2} \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)\mu_{0j'} y_{0j'}\right] y_{0j'}^{\kappa_{0j'}-1} dy_{0j'}$$
$$\times \frac{\mu_{0j}^{\kappa_{0j}}}{\Gamma(\kappa_{0j})} \int_0^\infty \exp\left[-\left(1 + \sum_{k=1}^{2} \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)\mu_{0j} y_{0j}\right] y_{0j}^{\kappa_{0j}} dy_{0j}$$
$$\times \prod_{j'=1}^{L} \frac{\mu_{1j'}^{\kappa_{1j'}}}{\Gamma(\kappa_{1j'})} \int_0^\infty \exp\left[-\left(1 + \frac{H_{0j'}^{(1)}(u_1)}{\mu_{1j'}}\right)\mu_{1j'} y_{1j'}\right] y_{1j'}^{\kappa_{1j'}-1} dy_{1j'}$$
$$\times \prod_{j'(\neq j)=1}^{L} \frac{\mu_{2j'}^{\kappa_{2j'}}}{\Gamma(\kappa_{2j'})} \int_0^\infty \exp\left[-\left(1 + \frac{H_{0j'}^{(2)}(u_2)}{\mu_{2j'}}\right)\mu_{2j'} y_{2j'}\right] y_{2j'}^{\kappa_{2j'}-1} dy_{2j'}$$
$$\times \frac{\mu_{2j}^{\kappa_{2j}}}{\Gamma(\kappa_{2j})} \int_0^\infty \exp\left[-\left(1 + \frac{H_{0j}^{(2)}(u_2)}{\mu_{2j}}\right)\mu_{2j} y_{2j}\right] y_{2j}^{\kappa_{2j}} dy_{2j}$$
$$= \frac{\kappa_{0j}\kappa_{2j}}{\mu_{1j}\mu_{2j}} \left(1 + \sum_{k=1}^{2} \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-1} \times \left(1 + \frac{H_{0j}^{(2)}(u_2)}{\mu_{2j}}\right)^{-1}$$
$$\times \prod_{j'=1}^{L} \left[\left(1 + \sum_{k=1}^{2} \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{0j'}} \prod_{k=1}^{2} \left(1 + \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{kj'}} \right],$$

$$I_{3j}(u_1, u_2) = \frac{\mu_{0j}}{\mu_{2j}} \int_0^\infty \cdots \int_0^\infty \exp\left[-\sum_{j'(\neq j)=1}^{L} \sum_{k=1}^{2} H_{0j'}^{(k)}(u_k) \left(\frac{\mu_{0j'}}{\mu_{kj'}} y_{0j'} + y_{kj'}\right)\right]$$
$$\times \exp\left[-\sum_{k=1}^{2} H_{0j}^{(k)}(u_k) \left(\frac{\mu_{0j}}{\mu_{kj}} y_{0j} + y_{kj}\right)\right] y_{0j} y_{1j}$$
$$\times \left(\prod_{l=0}^{2} \prod_{j'=1}^{L} \frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})} e^{-\mu_{lj'} y_{lj'}} y_{lj'}^{\kappa_{lj'}-1} dy_{lj'} \right)$$
$$= \frac{\mu_{0j}}{\mu_{2j}} \prod_{j'(\neq j)=1}^{L} \frac{\mu_{0j'}^{\kappa_{0j'}}}{\Gamma(\kappa_{0j'})} \int_0^\infty \exp\left[-\left(1 + \sum_{k=1}^{2} \frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)\mu_{0j'} y_{0j'}\right] y_{0j'}^{\kappa_{0j'}-1} dy_{0j'}$$
$$\times \frac{\mu_{0j}^{\kappa_{0j}}}{\Gamma(\kappa_{0j})} \int_0^\infty \exp\left[-\left(1 + \sum_{k=1}^{2} \frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)\mu_{0j} y_{0j}\right] y_{0j}^{\kappa_{0j}} dy_{0j}$$
$$\times \prod_{j'(\neq j)=1}^{L} \frac{\mu_{1j'}^{\kappa_{1j'}}}{\Gamma(\kappa_{1j'})} \int_0^\infty \exp\left[-\left(1 + \frac{H_{0j'}^{(1)}(u_1)}{\mu_{1j'}}\right)\mu_{1j'} y_{1j'}\right] y_{1j'}^{\kappa_{1j'}-1} dy_{1j'}$$
$$\times \frac{\mu_{1j}^{\kappa_{1j}}}{\Gamma(\kappa_{1j})} \int_0^\infty \exp\left[-\left(1 + \frac{H_{0j}^{(1)}(u_1)}{\mu_{1j}}\right)\mu_{1j} y_{1j}\right] y_{1j}^{\kappa_{1j}} dy_{1j}$$
$$\times \prod_{j'=1}^{L} \frac{\mu_{2j'}^{\kappa_{2j'}}}{\Gamma(\kappa_{2j'})} \int_0^\infty \exp\left[-\left(1 + \frac{H_{0j'}^{(2)}(u_2)}{\mu_{2j'}}\right)\mu_{2j'} y_{2j'}\right] y_{2j'}^{\kappa_{2j'}-1} dy_{2j'}$$

$$= \frac{\kappa_{0j}\kappa_{1j}}{\mu_{1j}\mu_{2j}}\left(1+\sum_{k=1}^{2}\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-1}\left(1+\frac{H_{0j}^{(1)}(u_1)}{\mu_{1j}}\right)^{-1}$$
$$\times\prod_{j'=1}^{L}\left[\left(1+\sum_{k=1}^{2}\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{0j'}}\prod_{k=1}^{2}\left(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{kj'}}\right]$$

and

$$I_{4j}(u_1,u_2) = \int_0^\infty\cdots\int_0^\infty \exp\left[-\sum_{j'(\neq j)=1}^{L}\sum_{k=1}^{2}H_{0j'}^{(k)}(u_k)\left(\frac{\mu_{0j'}}{\mu_{kj'}}y_{0j'}+y_{kj'}\right)\right]$$
$$\times\exp\left[-\sum_{k=1}^{2}H_{0j}^{(k)}(u_k)\left(\frac{\mu_{0j}}{\mu_{kj}}y_{0j}+y_{kj}\right)\right]y_{1j}y_{2j}$$
$$\times\left(\prod_{l=0}^{2}\prod_{j'=1}^{L}\frac{\mu_{lj'}^{\kappa_{lj'}}}{\Gamma(\kappa_{lj'})}e^{-\mu_{lj'}y_{lj'}}y_{lj'}^{\kappa_{lj'}-1}dy_{li}\right)$$
$$= \prod_{j'=1}^{L}\frac{\mu_{0j'}^{\kappa_{0j'}}}{\Gamma(\kappa_{0j'})}\int_0^\infty\exp\left[-\left(1+\sum_{k=1}^{2}\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)\mu_{0j'}y_{0j'}\right]y_{0j'}^{\kappa_{0j'}-1}dy_{0j'}$$
$$\times\prod_{j'(\neq j)=1}^{L}\frac{\mu_{1j'}^{\kappa_{1j'}}}{\Gamma(\kappa_{1j'})}\int_0^\infty\exp\left[-\left(1+\frac{H_{0j'}^{(1)}(u_1)}{\mu_{1j'}}\right)\mu_{1j'}y_{1j'}\right]y_{1j'}^{\kappa_{1j'}-1}dy_{1j'}$$
$$\times\frac{\mu_{1j}^{\kappa_{1j}}}{\Gamma(\kappa_{1j})}\int_0^\infty\exp\left[-\left(1+\frac{H_{0j}^{(1)}(u_1)}{\mu_{1j}}\right)\mu_{1j}y_{1j}\right]y_{1j}^{\kappa_{1j}}dy_{1j}$$
$$\times\prod_{j'(\neq j)=1}^{L}\frac{\mu_{2j'}^{\kappa_{2j'}}}{\Gamma(\kappa_{2j'})}\int_0^\infty\exp\left[-\left(1+\frac{H_{0j'}^{(2)}(u_2)}{\mu_{2j'}}\right)\mu_{2j'}y_{2j'}\right]y_{2j'}^{\kappa_{2j'}-1}dy_{2j'}$$
$$\times\frac{\mu_{2j}^{\kappa_{2j}}}{\Gamma(\kappa_{2j})}\int_0^\infty\exp\left[-\left(1+\frac{H_{0j}^{(2)}(u_2)}{\mu_{2j}}\right)\mu_{2j}y_{2j}\right]y_{2j}^{\kappa_{2j}}dy_{2j}$$
$$= \frac{\kappa_{1j}\kappa_{2j}}{\mu_{1j}\mu_{2j}}\prod_{k=1}^{2}\left(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-1}$$
$$\times\prod_{j'=1}^{L}\left[\left(1+\sum_{k=1}^{2}\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{0j'}}\prod_{k=1}^{2}\left(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{kj'}}\right],$$

for $j=1,\cdots,L$. Therefore, the joint sub-distribution function $F_{j,j}(t_1,t_2;\boldsymbol{h}_{\mathbf{0}}^{(\mathbf{1})},\boldsymbol{h}_{\mathbf{0}}^{(\mathbf{2})},\boldsymbol{\theta})$ is given by

$$\int_0^{t_1}\int_0^{t_2}\left(\prod_{k=1}^{2}\frac{h_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)\prod_{j'=1}^{L}\left[\left(1+\sum_{k=1}^{2}\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{0j'}}\prod_{k=1}^{2}\left(1+\frac{H_{0j'}^{(k)}(u_k)}{\mu_{kj'}}\right)^{-\kappa_{kj'}}\right]$$
$$\times\left[\kappa_{0j}(1+\kappa_{0j})\left(1+\sum_{k=1}^{2}\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-2}+\left(1+\sum_{k=1}^{2}\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-1}\times\sum_{k=1}^{2}\kappa_{0j}\kappa_{kj}\left(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-1}\right.$$
$$\left.+\prod_{k=1}^{2}\kappa_{kj}\left(1+\frac{H_{0j}^{(k)}(u_k)}{\mu_{kj}}\right)^{-1}\right]du_2du_1$$
$$=\int_0^{t_1}\int_0^{t_2}\left(\prod_{k=1}^{2}\sigma_{kj}^2h_{0j}^{(k)}(u_k)\right)\times\prod_{j'=1}^{L}\left(1+\sum_{k=1}^{2}\sigma_{kj'}^2H_{0j'}^{(k)}(u_k)\right)^{-\frac{\rho_{j'}}{\sigma_{1j'}\sigma_{2j'}}}$$
$$\times\prod_{k=1}^{2}\left(1+\sigma_{kj'}^2H_{0j'}^{(k)}(u_k)\right)^{-\left(\frac{1}{\sigma_{kj'}^2}-\frac{\rho_{j'}}{\sigma_{1j'}\sigma_{2j'}}\right)}$$
$$\times\left[\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(1+\sum_{k=1}^{2}\sigma_{kj}^2H_{0j}^{(k)}(u_k)\right)^2}+\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{\left(1+\sum_{k=1}^{2}\sigma_{kj}^2H_{0j}^{(k)}(u_k)\right)}\sum_{k=1}^{2}\frac{\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{\left(1+\sigma_{kj}^2H_{0j}^{(k)}(u_k)\right)}\right.$$
$$\left.+\prod_{k=1}^{2}\left(\frac{\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{1+\sigma_{kj}^2H_{0j}^{(k)}(u_k)}\right)\right]du_2du_1.$$

Note that the equality of the two joint sub-distribution functions, as mentioned above, also implies the equality of the corresponding joint sub-density functions $f_{j,j}(t_1,t_2;\boldsymbol{h_0^{(1)}},\boldsymbol{h_0^{(2)}},\boldsymbol{\theta})$ and $f_{j,j}(t_1,t_2;\boldsymbol{h_0^{(1)}},\boldsymbol{h_0^{(2)}},\tilde{\boldsymbol{\theta}})$, for $j=1,2,\cdots,L$, which can be obtained by differentiating the joint sub-distribution functions with respect to t_1 and t_2. Utilizing the already obtained results $h_{0,j}^{(k)}(t_k)=\tilde{h}_{0,j}^{(k)}(t_k)$, for all $t_k>0$, and $\sigma_{kj}=\tilde{\sigma}_{kj}$, for $j=1,\cdots,L$ and $k=1,2$, in the equality of the corresponding joint sub-density functions, we get

$$\prod_{j'=1}^{L}\left[\left(1+\sum_{k=1}^{2}\sigma_{kj'}^2H_{0j'}^{(k)}(t_k)\right)^{-\frac{\rho_{j'}}{\sigma_{1j'}\sigma_{2j'}}}\prod_{k=1}^{2}\left(1+\sigma_{kj'}^2H_{0j'}^{(k)}(t_k)\right)^{-\left(\frac{1}{\sigma_{kj'}^2}-\frac{\rho_{j'}}{\sigma_{1j'}\sigma_{2j'}}\right)}\right]$$
$$\times\left[\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\left(1+\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}\right)}{\left(1+\sum_{k=1}^{2}\sigma_{kj}^2H_{0j}^{(k)}(t_k)\right)^2}+\left(\frac{\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{1+\sum_{k=1}^{2}\sigma_{kj}^2H_{0j}^{(k)}(t_k)}\right)\sum_{k=1}^{2}\frac{\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{1+\sigma_{kj}^2H_{0j}^{(k)}(t_k)}+\right.$$
$$\left.\prod_{k=1}^{2}\left(\frac{\frac{1}{\sigma_{kj}^2}-\frac{\rho_j}{\sigma_{1j}\sigma_{2j}}}{1+\sigma_{kj}^2H_{0j}^{(k)}(t_k)}\right)\right]$$

$$= \prod_{j'=1}^{L} \left[\left(1 + \sum_{k=1}^{2} \sigma_{kj'}^2 H_{0j'}^{(k)}(t_k) \right)^{-\frac{\tilde{\rho}_{j'}}{\sigma_{1j'}\sigma_{2j'}}} \prod_{k=1}^{2} \left(1 + \sigma_{kj'}^2 H_{0j'}^{(k)}(t_k) \right)^{-\left(\frac{1}{\sigma_{kj'}^2} - \frac{\tilde{\rho}_{j'}}{\sigma_{1j'}\sigma_{2j'}}\right)} \right]$$
$$\times \left[\frac{\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}} \left(1 + \frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}} \right)}{\left(1 + \sum_{k=1}^{2} \sigma_{kj}^2 H_{0j}^{(k)}(t_k) \right)^2} + \left(\frac{\frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}}{1 + \sum_{k=1}^{2} \sigma_{kj}^2 H_{0j}^{(k)}(t_k)} \right) \sum_{k=1}^{2} \frac{\frac{1}{\sigma_{kj}^2} - \frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}}{1 + \sigma_{kj}^2 H_{0j}^{(k)}(t_k)} + \right.$$
$$\left. \prod_{k=1}^{2} \left(\frac{\frac{1}{\sigma_{kj}^2} - \frac{\tilde{\rho}_j}{\sigma_{1j}\sigma_{2j}}}{1 + \sigma_{kj}^2 H_{0j}^{(k)}(t_k)} \right) \right],$$

for $t_k > 0,\ j = 1, 2, \cdots, L$ and $k = 1, 2$. Letting limit as $t_1 \to 0+$ and $t_2 \to 0+$ on both sides of the above equation, we get $\rho_j = \tilde{\rho}_j$, for $j = 1, 2, \cdots, L$. This completes the proof. □

5.3 Estimation

In this section, our objective is to estimate the arbitrary baseline cause-specific hazard functions $h_{0j}^{(k)}(\cdot)$, for $j = 1, \cdots, L_k$ and $k = 1, 2$, non-parametrically along with the parameters of the relevant Gamma frailty distribution(s). Recall from the identifiability conditions of the models in Section 5.2, the cumulative baseline cause-specific hazards $H_{0j}^{(k)}(\cdot)$'s need to be continuous so that the corresponding inverses exist. With this objective in mind, we approximate the hazards by Bernstein polynomials (See Section 1.10.5) with some suitable degree m and then carry out maximization of the likelihood function with this approximation. This is known as the method of sieve maximum likelihood estimation (Geman and Hwang, 1982) in the sense that, as $m \to \infty$, the estimators converge to the true hazards with respect to the Hellinger metric.

Following the description of Bernstein polynomials in Section 1.10.5, we define the model for the baseline cause-specific hazard function due to cause j for the kth individual as

$$h_{0j}^{(k)}(t; \boldsymbol{c}_{m_j^{(k)}}) = \sum_{l=1}^{m_j^{(k)}} c_{m_j^{(k)} l} g_{m_j^{(k)} l}(t) = (\boldsymbol{c}_{m_j^{(k)}})^T \boldsymbol{g}_{m_j^{(k)}}(t), \tag{5.11}$$

for a suitably chosen degree $m_j^{(k)}$, as an approximation for the true baseline cause-specific hazard $h_{0j}^{(k)}(t)$, for $j = 1, \cdots, L_k$ and $k = 1, 2$. Here, $\boldsymbol{c}_{m_j^{(k)}}^T = \left(c_{m_j^{(k)} 1}, \cdots, c_{m_j^{(k)} m_j^{(k)}} \right)$ is a vector of unknown constants satisfying $c_{m_j^{(k)} l} \geq 0$, for all $l = 1, \cdots, m_j^{(k)}$, and

$\boldsymbol{g}_{m_j^{(k)}}(t) = \left(g_{m_j^{(k)}1}(t), \cdots, g_{m_j^{(k)}m_j^{(k)}}(t)\right)^T$ is a vector of some known real 'base' functions satisfying $g_{m_j^{(k)}l}(\cdot) \geq 0$ and $\int_0^\infty g_{m_j^{(k)}l}(u)du < \infty$, for $l = 1, \cdots, m_j^{(k)}$. Then, the corresponding model for baseline cumulative cause-specific hazard is

$$H_{0j}^{(k)}(t; \boldsymbol{c}_{m_j^{(k)}}) = \sum_{l=1}^{m_j^{(k)}} c_{m_j^{(k)}l} G_{m_j^{(k)}l}(t), \tag{5.12}$$

where $G_{m_j^{(k)}l}(t) = \int_0^t g_{m_j^{(k)}l}(u)du$ for all $t > 0, j = 1, \cdots, L_k$, for $k = 1, 2$. The total baseline hazard function for the kth individual is

$$h_0^{(k)}(t; \boldsymbol{c^{(k)}}) = \sum_{j=1}^{L_k} h_{0j}^{(k)}(t; \boldsymbol{c}_{m_j^{(k)}}) = \sum_{j=1}^{L_k} \sum_{l=1}^{m_j^{(k)}} c_{m_j^{(k)}l} g_{m_j^{(k)}l}(t),$$

where $\boldsymbol{c^{(k)}} = \left(\boldsymbol{c}_{m_1^{(k)}}, \boldsymbol{c}_{m_2^{(k)}}, \cdots, \boldsymbol{c}_{m_{L_k}^{(k)}}\right)$, for $k = 1, 2$. The total baseline cumulative hazard for the kth individual is

$$H_0^{(k)}(t; \boldsymbol{c^{(k)}}) = \int_0^t h_0^{(k)}(u; \boldsymbol{c^{(k)}})du = \sum_{j=1}^{L_k} \sum_{l=1}^{m_j^{(k)}} c_{m_j^{(k)}l} G_{m_j^{(k)}l}(t),$$

for $t > 0$ and $k = 1, 2$. It is important to note that the monotonicity property of $H_0^{(k)}(t; \boldsymbol{c^{(k)}})$ is maintained because each element of the vector of coefficients $\boldsymbol{c^{(k)}}$ is non-negative and the base functions $g_{m_j^{(k)}l}(t) \geq 0$, for $l = 1, \cdots, m_j^{(k)}, j = 1, 2, \cdots, L_k$ and $k = 1, 2$.

We construct the base functions $g_{m_j^{(k)}l}(.)$ and $G_{m_j^{(k)}l}(.)$, for all $l = 1, \cdots, m_j^{(k)}, j = 1, \cdots, L_k$ and $k = 1, 2$, using the Bernstein polynomials (See Osman and Ghosh, 2012, for similar approximation model for hazard function). For the true continuous baseline cumulative cause-specific hazard function $H_{0j}^{(k)}(.)$ on the interval $[0, \tau]$, the approximating Bernstein polynomial of order $m_j^{(k)}$ is given by

$$H_{0j}^{(k)}(t; \boldsymbol{c}_{m_j^{(k)}}) = \sum_{l=0}^{m_j^{(k)}} c_{m_j^{(k)}l} \binom{m_j^{(k)}}{l} \left(\frac{t}{\tau}\right)^l \left(1 - \frac{t}{\tau}\right)^{m_j^{(k)} - l}.$$

The time limit τ may be, in theory, defined as $\inf\{t : S(t, t; \boldsymbol{h}_0^{(1)}, \boldsymbol{h}_0^{(2)}, \boldsymbol{\theta}) = 0\}$. The Weierstrass theorem ensures that $H_{0j}^{(k)}(.; c_{m_j^{(k)}}) \to H_{0j}^{(k)}(.)$, the corresponding true baseline cumulative hazard function, uniformly over the interval $[0, \tau]$, as $m_j^{(k)} \to \infty$,

for $j = 1, \cdots, L_k$ and $k = 1, 2$. See Lorentz (2012). The derivative of this approximating polynomial gives the corresponding baseline cause-specific hazard function and can be written as

$$h_j^{(k)}(t; \boldsymbol{c}_{m_j^{(k)}}) = \sum_{l=1}^{m_j^{(k)}} \Big\{ H_{0j}^{(k)}\Big(\frac{l}{m_j^{(k)}}\tau\Big) - H_{0j}^{(k)}\Big(\frac{l-1}{m_j^{(k)}}\tau\Big) \Big\} \frac{1}{\tau} f_\beta\Big(\frac{t}{\tau}; l, m_j^{(k)} - l + 1\Big),$$

where $f_\beta(\cdot; l, m_j^{(k)} - l + 1)$ is the probability density function of the beta distribution with parameters l and $m_j^{(k)} - l + 1$. From the theory of approximating Bernstein polynomial, we also know that $h_{0j}^{(k)}(t; \boldsymbol{c}_{m_j^{(k)}}) \to h_{0j}^{(k)}(.)$, the corresponding true baseline hazard function, as $m_j^{(k)} \to \infty$. See Lorentz (2012).

We model the base function $g_{m_j^{(k)}l}(t)$ as $1/\tau \times f_\beta(t/\tau; l, m_j^{(k)} - l + 1)$ in (5.11), so that

$$G_{m_j^{(k)}l}(t) = \int_0^t g_{m_j^{(k)}l}(u)du = \int_0^t f_\beta(u/\tau; l, m_j^{(k)} - l + 1)d(u/\tau).$$

Also, the unknown coefficient $c_{m_j^{(k)}l}$ is equated with the unknown quantity $H_{0j}^{(k)}(\frac{l}{m_j^{(k)}}\tau) - H_{0j}^{(k)}(\frac{l-1}{m_j^{(k)}}\tau)$, for $l = 1, \cdots, m_j^{(k)}$ and for all $j = 1, \cdots, L_k, k = 1, 2$. This gives the complete description of the approximate model (5.11) for the baseline cause-specific hazards, $h_{0j}^{(k)}(\cdot, \boldsymbol{c}_{m_j^{(k)}})$, in which the coefficients $c_{m_j^{(k)}l}$'s are the unknown parameters.

In the subsequent description of the methodology, for ease of notation and coding afterwards, we consider $m_j^{(k)} = m$, for all $j = 1, \cdots, L_k, k = 1, 2$. This simplification allows the base functions $g_{ml}(t)$'s and $G_{ml}(t)$'s to be the same for all j, k; but the unknown coefficients $c_{m_j^{(k)}l}$'s will depend on j and k; so, henceforth, we denote these coefficients with a common m by $c_{j,ml}^{(k)}$. Also, we write $\boldsymbol{c}_{j,m}^{(k)} = (c_{j,m1}^{(k)}, \cdots, c_{j,mm}^{(k)})^T$, so that the baseline cause-specific hazards in (5.11) are written as $h_{0j}^{(k)}(\cdot, \boldsymbol{c}_{j,m}^{(k)})$, and $\boldsymbol{c}_m^{(k)} = \{\boldsymbol{c}_{j,m}^{(k)}, j = 1, \cdots, L_k\}$, for $k = 1, 2$.

As discussed in Section 1.10.5, instead of attempting to estimate the unspecified baseline cause-specific hazards in the infinite dimensional space $\mathcal{H}_1 \times \mathcal{H}_2$, we restrict our search in the subspace of models defined by $h_{0j}^{(k)}(\cdot, \boldsymbol{c}_{j,m}^{(k)})$ in (5.11) with the common m and the base functions $g_{ml}(t)$'s. We write this space as $\mathcal{H}_{1m} \times \mathcal{H}_{2m}$, where

$$\mathcal{H}_{km} = \{h_{0j}^{(k)}(\cdot, \boldsymbol{c}_{j,m}^{(k)}), \quad \text{for} \quad t > 0, j = 1, \cdots, L_k\},$$

for $k = 1, 2$. The sequence of subspaces $\mathcal{H}_{1m} \times \mathcal{H}_{2m}$ of the original parameter space $\mathcal{H}_1 \times \mathcal{H}_2$, for $m \geq 1$, is called a sieve. In the following, we consider the likelihood

function for bivariate current status data with competing risks under the model defined by (5.11). The estimates obtained by maximizing this likelihood with respect to the baseline cause-specific hazards in the space $\mathcal{H}_{1m} \times \mathcal{H}_{2m}$ and $\sigma > 0$ are called the sieve maximum likelihood estimators (Geman and Hwang, 1982).

For the sake of simplicity in describing the method and also the related coding for maximization of likelihood, we consider the shared Gamma frailty model (See Section 5.2.2) for further discussion and analysis, as in Chapter 3. Also, for practical consideration, while deriving the likelihood contributions, we suitably fix τ as a time larger than the largest monitoring time. Using the expression 5.5 and the above-mentioned approximate model (5.11) for the baseline cause-specific hazards, the jth sub-distribution function for the kth individual is

$$\begin{aligned} F_j^{(k)}(t; \boldsymbol{c}_m^{(k)}, \sigma^2) &= \int_0^t h_{0j}^{(k)}(u_k; \boldsymbol{c}_{j,m}^{(k)}) \Big\{ 1 + \sigma^2 H_0^{(k)}(u_k; \boldsymbol{c}_m^{(k)}) \Big\}^{-1-\frac{1}{\sigma^2}} du_k \\ &= \int_0^t \sum_{l=1}^{m} c_{j,ml}^{(k)} g_{ml}(u_k) \left\{ 1 + \sigma^2 \sum_{j'=1}^{L_k} \sum_{l'=1}^{m} c_{j,ml'}^{(k)} G_{ml'}(u_k) \right\}^{-1-\frac{1}{\sigma^2}} du_k, \end{aligned} \tag{5.13}$$

for all $t > 0, j = 1, \cdots, L_k$ and $k = 1, 2$. The joint survival function is given by

$$\begin{aligned} S(t_1, t_2; \boldsymbol{c}_m^{(1)}, \boldsymbol{c}_m^{(2)}, \sigma^2) &= \Big[1 + \sigma^2 \Big\{ H_0^{(1)}(t_1; \boldsymbol{c}_m^{(1)}) + H_0^{(2)}(t_2; \boldsymbol{c}_m^{(2)}) \Big\} \Big]^{-\frac{1}{\sigma^2}} \\ &= \left[1 + \sigma^2 \left\{ \sum_{j=1}^{L_1} \sum_{l=1}^{m} c_{j,ml}^{(1)} G_{ml}(t_1) + \sum_{j=1}^{L_2} \sum_{l=1}^{m} c_{j,ml}^{(2)} G_{ml}(t_2) \right\} \right]^{-\frac{1}{\sigma^2}}, \end{aligned} \tag{5.14}$$

for all $t_1, t_2 > 0$. The joint sub-distribution function $F_{j_1 j_2}(t_1, t_2; \boldsymbol{c}_m^{(1)}, \boldsymbol{c}_m^{(2)}, \sigma^2)$ is given

by

$$
\begin{aligned}
&(1+\sigma^2)\int_0^{t_1}\int_0^{t_2} h_{0j_1}^{(1)}(u_1;\boldsymbol{c}_{j_1,m}^{(1)})h_{0j_2}^{(2)}(u_2;\boldsymbol{c}_{j_2,m}^{(2)}) \\
&\qquad\times\Big[1+\sigma^2\Big\{H_0^{(1)}(u_1;\boldsymbol{c}_m^{(1)})+H_0^{(2)}(u_2;\boldsymbol{c}_m^{(2)})\Big\}\Big]^{-2-\frac{1}{\sigma^2}}du_2du_1 \\
&=(1+\sigma^2)\int_0^{t_1}\int_0^{t_2}\sum_{l=1}^{m}c_{j_1,ml}^{(1)}g_{ml}(u_1)\sum_{l=1}^{m}c_{j_2,ml}^{(2)}g_{ml}(u_2) \\
&\qquad\times\left[1+\sigma^2\left\{\sum_{j_1=1}^{L_1}\sum_{l=1}^{m}c_{j_1,ml}^{(1)}G_{ml}(u_1)+\sum_{j_2=1}^{L_2}\sum_{l=1}^{m}c_{j_2,ml}^{(2)}G_{ml}(u_2)\right\}\right]^{-2-\frac{1}{\sigma^2}}du_2du_1.
\end{aligned}
\tag{5.15}
$$

Now, following Section 2.6, the likelihood function $\mathcal{L}$, for some suitably chosen m, is

$$\mathcal{L}(\boldsymbol{c}_m^{(1)},\boldsymbol{c}_m^{(2)},\sigma)\propto\prod_{i=1}^{n}f^*(x_{i1},x_{i2},j_{i1},j_{i2};\boldsymbol{c}_m^{(1)},\boldsymbol{c}_m^{(2)},\sigma),$$

where each likelihood contribution can be obtained by using (5.13)-(5.15) (See also Section 5.2.2), as in Section 2.6.

The corresponding log-likelihood function is

$$\ell(\boldsymbol{c}_m^{(1)},\boldsymbol{c}_m^{(2)},\sigma)\propto\sum_{i=1}^{n}\log\big[f^*(x_{i1},x_{i2},j_{i1},j_{i2};\boldsymbol{c}_m^{(1)},\boldsymbol{c}_m^{(2)},\sigma)\big],$$

which is maximized with respect to $\boldsymbol{\eta}=(\boldsymbol{c}_m^{(1)},\boldsymbol{c}_m^{(2)},\sigma)$ using some numerical method to obtain the corresponding sieve maximum likelihood estimate $\hat{\boldsymbol{\eta}}=(\hat{\boldsymbol{c}}_m^{(1)},\hat{\boldsymbol{c}}_m^{(2)},\hat{\sigma})$.

5.4 Simulation Studies

Now, we investigate the finite sample behaviour of the semi-parametric sieve MLE $\hat{\boldsymbol{\eta}}=(\hat{\boldsymbol{c}}_m^{(1)},\hat{\boldsymbol{c}}_m^{(2)},\hat{\sigma})$ of the parameter vector $\boldsymbol{\eta}=(\boldsymbol{c}_m^{(1)},\boldsymbol{c}_m^{(2)},\sigma)$ using a simulation study. Again, we consider the similar setup, as in Section 3.5. We consider same monitoring time X for both individuals and we assume that X follows an Exponential distribution with mean $\frac{1}{\mu}$. The unknown parameter μ of the monitoring time X is solved by equating the joint censoring probability to a fixed value $p_{cen}=0.1$(say).

We consider two competing risks corresponding to both individuals; that is, $L_1=L_2=2$. We consider constant or Exponential type true baseline cause-specific haz-

ards; that is, $h_j^{(k)}(t) = \alpha_j^{(k)}$, for $j = 1, 2$, and $k = 1, 2$ with $(\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)}) = (2.4, 5.8, 3.5, 4.5)$. We consider the value of σ as 0.75. In this case, we fix τ as the 99th percentile of the monitoring time distribution and we ignore any monitoring time that turns out to be greater than τ. Considering this setup, we draw a sample of size n with each observation of the form (x, x, j_1, j_2) following the same method as that in Section 3.5. This procedure is repeated 500 times to get 500 such simulated datasets.

Table 5.1: Simulation results on integrated absolute error (IAE) under constant baseline cause-specific hazards and shared Gamma frailty with parameter σ for different values of n and m with $p_{cen} = 0.1$.

	$n = 300, m = 7$		$n = 300, m = 10$		$n = 300, m = 15$	
	$j = 1$	$j = 2$	$j = 1$	$j = 2$	$j = 1$	$j = 2$
$k = 1$	1.937	2.280	0.673	0.763	0.413	0.457
	(1.332)	(1.311)	(0.313)	(0.386)	(0.215)	(0.225)
$k = 2$	1.961	2.120	0.688	0.658	0.448	0.441
	(1.198)	(1.241)	(0.338)	(0.351)	(0.229)	(0.219)
	$n = 500, m = 7$		$n = 500, m = 10$		$n = 500, m = 15$	
	$j = 1$	$j = 2$	$j = 1$	$j = 2$	$j = 1$	$j = 2$
$k = 1$	1.812	1.955	0.645	0.739	0.421	0.431
	(1.031)	(1.044)	(0.291)	(0.324)	(0.197)	(0.191)
$k = 2$	1.958	1.910	0.683	0.676	0.424	0.438
	(1.108)	(1.087)	(0.322)	(0.310)	(0.192)	(0.211)

For each simulated dataset, we compute the semiparametric sieve MLE $(\hat{\boldsymbol{c}}_m^{(1)}, \hat{\boldsymbol{c}}_m^{(2)}, \hat{\sigma})$ using the method described in Section 5.3. We denote the estimated baseline cause-specific hazard function as $\hat{h}_{0j}^{(k)}(t; \hat{\boldsymbol{c}}_{j,m}^{(k)})$, for $j = 1, 2$ and $k = 1, 2$. For each dataset, we compute Integrated Absolute Error (IAE) (Osman and Ghosh, 2012) and Integrated Mean Squared Error (IMSE) (Wang and Ghosh, 2012) corresponding to each estimated baseline cause-specific hazard defined as

$$IAE(\hat{h}_{0j}^{(k)}) = \int_0^\tau |\hat{h}_j^{(k)}(u; \hat{\boldsymbol{c}}_{j,m}^{(k)}) - \alpha_j^{(k)}| du$$

and

$$IMSE(\hat{h}_{0j}^{(k)}) = \int_0^\tau (\hat{h}_{0j}^{(k)}(u; \hat{\boldsymbol{c}}_{j,m}^{(k)}) - \alpha_j^{(k)})^2 du, \text{ respectively.}$$

For a fixed value of m and n, we compute 500 values of IAE and IMSE from 500

simulated datasets. For each combination of j and k, we report the sample mean and standard deviation (in parentheses) of IAE and IMSE, based on the 500 simulated samples, in Tables 5.1 and 5.2, respectively. We carry out this simulation for $m = 7, 10, 15$ and $n = 300, 500$. The corresponding standard deviation of the 500 values of $\hat{\sigma}$ are presented in Table 5.3.

Table 5.2: Simulation results on integrated mean square error (IMSE) under constant baseline cause-specific hazards and shared Gamma frailty with parameter σ for different values of n and m with $p_{cen} = 0.1$

	$n = 300, m = 7$		$n = 300, m = 10$		$n = 300, m = 15$	
	$j = 1$	$j = 2$	$j = 1$	$j = 2$	$j = 1$	$j = 2$
$k = 1$	1.463	1.876	0.172	0.257	0.066	0.085
	(2.438)	(2.083)	(0.165)	(0.238)	(0.061)	(0.081)
$k = 2$	1.446	1.663	0.187	0.194	0.078	0.081
	(1.664)	(2.212)	(0.169)	(0.176)	(0.070)	(0.075)
	$n = 500, m = 7$		$n = 500, m = 10$		$n = 500, m = 15$	
	$j = 1$	$j = 2$	$j = 1$	$j = 2$	$j = 1$	$j = 2$
$k = 1$	1.092	1.346	0.141	0.205	0.057	0.069
	(1.078)	(1.249)	(0.132)	(0.192)	(0.051)	(0.062)
$k = 2$	1.325	1.240	0.159	0.166	0.061	0.071
	(1.255)	(1.125)	(0.151)	(0.153)	(0.053)	(0.064)

From Table 5.1 and 5.2, we observe that means of both IAE and IMSE decrease if either one of m and n increases, as expected. From Table 5.3, it can be seen that mean absolute deviation of estimated σ does not seem to depend on the value of m, as expected, since m is involved only in the approximation model (5.11) for the unspecified baseline cause-specific hazards. However, it decreases with increase in n.

Table 5.3: Standard deviation of the estimated σ based on 500 simulated samples from constant baseline cause-specific hazards and shared Gamma frailty.

n	m		
	7	10	15
300	0.057	0.053	0.054
500	0.049	0.047	0.042

In order to get an idea about the fitted curves, we plot the averaged (over 200 replications of datasets with sample size 500) fitted baseline cause-specific hazard functions along with the true baseline cause-specific hazard functions $h_{0j}^{(k)}(\cdot)$'s (shown in red line) in Figure 5.1(a)-(d). These estimated curves $\hat{h}_{0j}^{(k)}(t; \hat{\boldsymbol{c}}_{j,m}^{(k)})$, for $j, k = 1, 2$,

are obtained by fitting the approximate model (5.11) with $m = 10, 15$ and 20 based on simulated data of sample size $n = 500$ under the constant baseline cause-specific hazards model with $h_{01}^{(1)}(\cdot) = 2.4, h_{02}^{(1)}(\cdot) = 5.8, h_{01}^{(2)}(\cdot) = 3.5, h_{02}^{(2)}(\cdot) = 4.5$ and $\sigma = 0.75$. While the graphs in Figure 5.1(a)-(b) are for the fitted averaged curves $\hat{h}_{0j}^{(1)}(t; \hat{\boldsymbol{c}}_{j,m}^{(1)})$, for $j = 1, 2$, respectively, those in Figure 5.1(c)-(d) are for the fitted averaged curves $\hat{h}_{0j}^{(2)}(t; \hat{\boldsymbol{c}}_{j,m}^{(2)})$, for $j = 1, 2$, respectively. The "dashed", "dottdash" and "dotted" curves represent the fitted averaged curves corresponding to m = 10, 15 and 20, respectively. It is evident from the graphs that the fitted curves produce reasonably good estimates of the true curves at least in this example. Also, although there seems to be some improvement for higher m, the degree of Bernstein polynomial, the difference does not seem to be that notable.

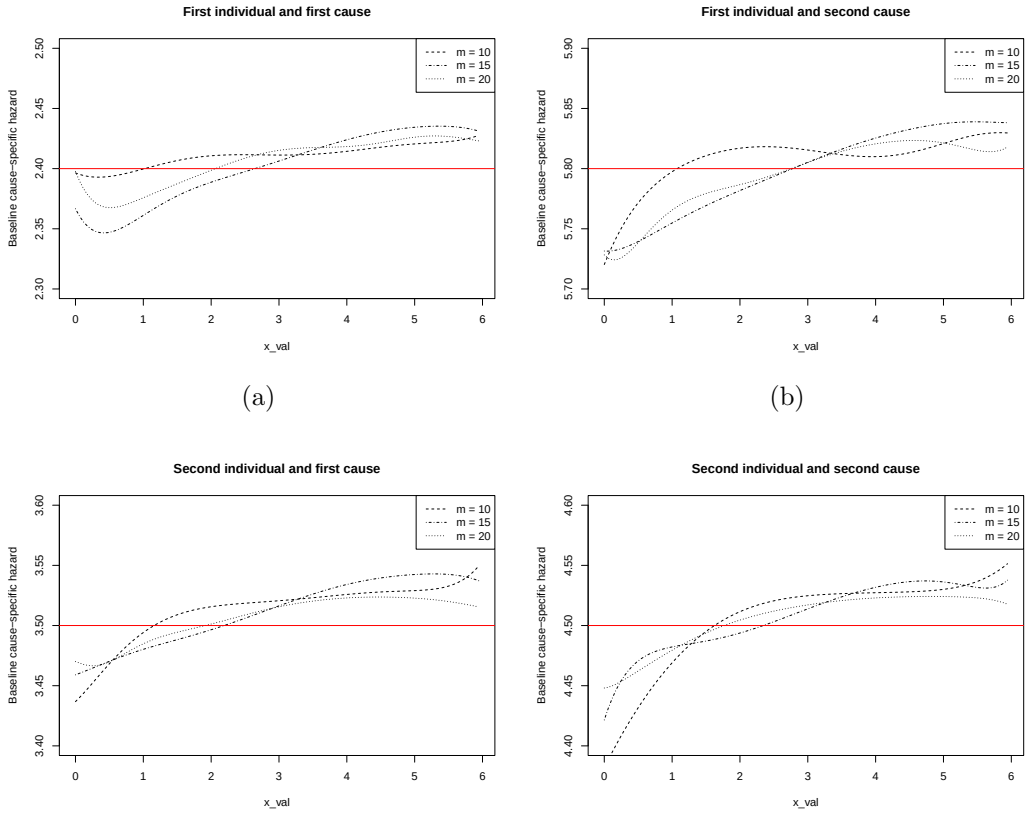


Figure 5.1: Fitted baseline cause-specific hazard functions, averaged over 200 replications of datasets with sample size $n = 500$ under constant baseline cause-specific hazards, for $m = 10, 15$ and 20 with dashed, dottdash and dotted lines, respectively.

5.5 Data Analysis

As discussed in Section 1.12, any one of the two causes, namely, **SNHL** and **Conductive**, is responsible for the hearing loss in each of the ears of an individual. Following the same notation as in Sections 3.6 and 4.7, we label the two causes as 1 and 2, respectively. In this section, we analyze the hearing loss dataset using non-parametric baseline cause-specific hazards with shared Gamma frailty using the method of Section 5.3. We consider the approximate model (5.11) for the unspecified baseline cause-specific hazard functions using Bernstein polynomials with three different values of m, namely $m = 10, 15$ and 20. There are four unknown baseline cause-specific hazards $h_{0j}^{(k)}(\cdot, \boldsymbol{c}_{j,m}^{(k)})$, for $j = 1, 2$ and $k = 1, 2$, to be estimated for each m. These four estimated curves are shown in Figure 5.2(a)-(d), each panel having three graphs for the three m values, showing the effect of increasing m. In each panel, these three graphs for $m = 10, 15$ and 20 are shown using three different line types, namely, 'solid', 'dottdash' and 'dashed', respectively. The curves in Figure 5.2(a)-(b) are for the fitted curves $h_{01}^{(k)}(t; \hat{\boldsymbol{c}}_{1,m}^{(k)})$, for $k = 1, 2$, and those in Figure 5.2(c)-(d) are for the fitted curves $h_{02}^{(k)}(t; \hat{\boldsymbol{c}}_{2,m}^{(k)})$, for $k = 1, 2$, respectively.

Figure 5.2 clearly shows (See the 'dashed' line corresponding to $m = 20$) that the baseline cause-specific hazard for the first cause (that is, SNHL) is dominating in early ages, and then decreases with age in both the ears. The same for the other cause (that is, Conductive hearing loss) is smaller in the early ages and then increases with age, having a peak at the age around 200 months (about 16 years), and finally decreases at later ages. The same phenomenon is observed in both the ears.

This opposing patterns of baseline cause-specific hazards due to SNHL and Conductive hearing loss, as observed in this semi-parametric analysis with unspecified baseline cause-specific hazards, have gone unnoticed in the analyses of Chapters 3 and 4 in which we have assumed parametric models for the baseline cause-specific hazards. These parametric models (Exponential and Weibull type) have forced the methods to give constant or monotonic estimates for the hazards. The semi-parametric analysis in this section reveals that all the hazards do not behave in the same increasing or decreasing manner.

The estimates of σ, the standard deviation of the shared frailty, corresponding to the three m values $10, 15$ and 20, are obtained as 2.8816, 2.8169 and 2.6134, respectively. Note that these estimates do not change much with m, as expected, since the degree m of Bernstein polynomials is relevant to the unspecified baseline cause-specific hazards only. Interestingly, in the parametric analysis of the same data in

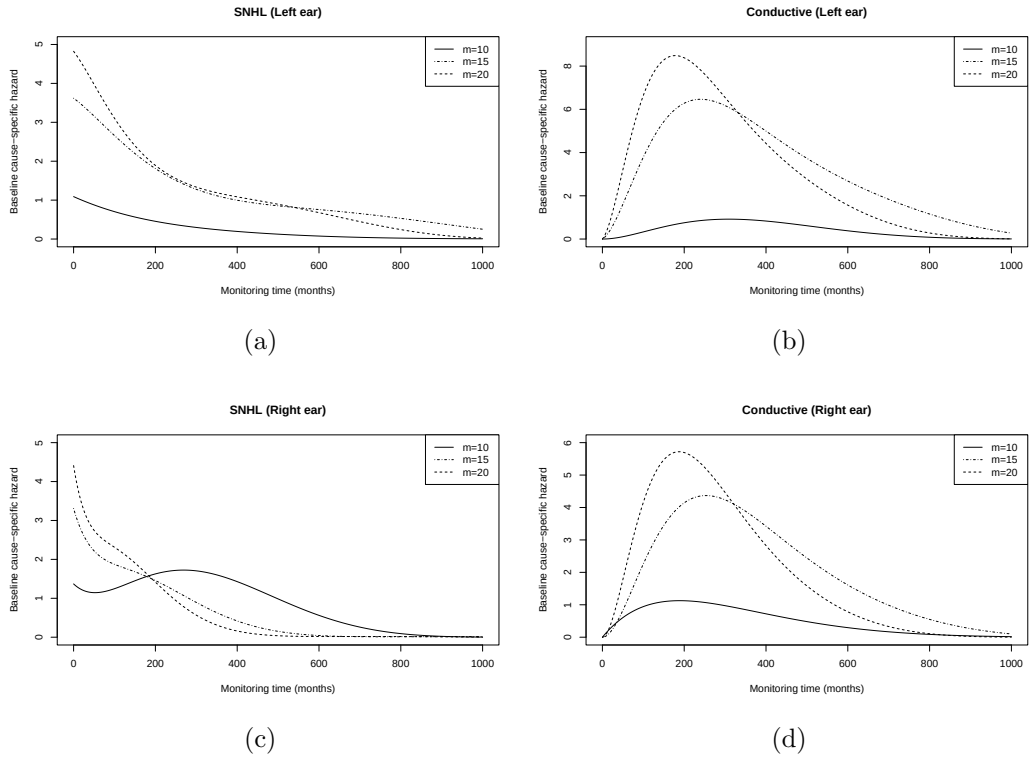


Figure 5.2: Fitted baseline cause-specific hazard functions corresponding to **SNHL** and **Conductive** hearing loss in both ears for $m = 10, 15$ and 20 with solid, dottdash and dashed lines, respectively.

Section 3.6 assuming Weibull (Exponential) type baseline cause-specific hazards with common shape parameter for all the causes and shared frailty, the estimate of σ came out to be 18.3432(6.3803), indicating much larger between pair heterogeneity in the cause-specific hazards. We have observed opposite patterns of baseline cause-specific hazards in our semi-parametric analysis of the present section. These heterogeneous patterns are not reflected in the assumed parametric models of Section 3.6. As a result, in those parametric analyses, the inherent heterogeneity in the baseline cause-specific hazards are absorbed in the frailty distribution, thereby inflating the estimate of σ. In the semi-parametric analysis of the present chapter, as the heterogeneous patterns of the baseline cause-specific hazards are incorporated through the model (5.11), the estimate of σ of the shared frailty model comes down to a reasonable level.

5.6 Concluding Remarks

In this chapter, we consider non-parametric baseline cause-specific hazard functions and Gamma frailty. We establish identifiability of the models corresponding to the four different Gamma frailty distributions under fairly reasonable conditions. We have used only shared Gamma frailty to describe the estimation procedure in Section 5.3 for simplicity, but this can be easily extended to the other three frailty models with considerably increased computational burden.

The computational burden increases with m and also with the number of competing risks. With L number of competing risks for both the individuals, the number of coefficients $c_{j,ml}^{(k)}$ is $2mL$; in addition, we also have the frailty parameters, which is 1 for the shared Gamma model. This large number of parameters not only makes the computation an enormous job (that is why, we consider only 500 replications in our simulation study), the convergence of likelihood maximization is also slow and difficult. This also prohibits investigating the convergence rate, etc., by simulation through the method of subsampling, as in Section 4.5.

As discussed in Section 1.10.5, one could consider an approximate model for the unspecified baseline cause-specific hazards by means of kernel or spline smoothing, both of which involve similar computational challenges. We believe the Bernstein polynomial approximation has certain advantages. Nevertheless, it is worth exploring the other two approximations.

Chapter 6

Discussion

6.1 Summative Remarks

As mentioned in Section 1.3, bivariate current status data with competing risks arises when a pair of related individuals are inspected at some monitoring time(s) to note the corresponding failure status at that time, when the failure may be due to one of several causes. In order to model dependence between the bivariate failure time data, we consider frailty modeling with four different types of frailty, possibly depending on the failure type as well. We also consider general multiplicative frailty models in the sense that, given the frailty, the conditional cause-specific hazard for an individual is the corresponding baseline cause-specific hazard multiplied by the associated frailty. We also assume that the dependence is solely through these frailty variable(s) so that, conditional on the frailty, the failure times and types of the two individuals are independent. For an association measure between the two failure times and types, we consider the "conditional cause-specific hazard ratio" introduced by Bandeen-Roche and Liang (2002). In this context, we have also defined model identifiability and shown that, when both the baseline cause-specific hazard functions and the frailty distribution are arbitrary, this model is not identifiable.

In view of this non-identifiability result, we have discussed parametric modeling and analysis of bivariate current status data with competing risks in Chapter 3. Here, we have proposed a parametric class of hazard functions which includes many well-known hazards of different shapes. This class enables choosing a suitable baseline cause-specific hazards for a given data, possibly based on some additional information on the incidence of different failure types. We have established the corresponding identifiability results for each of the four different Gamma frailty distributions. Then, we have considered maximum likelihood estimation method to estimate the unknown

model parameters. We have carried out an extensive simulation study and analyzed the hearing loss data considering Exponential and Weibull type baseline cause-specific hazard functions with four different Gamma frailty models. It has been observed that the Weibull distribution is a better choice as baseline cause-specific hazard than the Exponential distribution. The cause-specific frailty models have performed better than the shared and correlated Gamma frailty models in terms of AIC. It has been found that **SNHL** is the dominant cause for the hearing loss for both ears of an individual. Nevertheless, some inflexibility in the shape of parametric baseline cause-specific hazard functions may affect the conclusions of such a parametric analysis, especially in the absence of a suitable goodness of fit test. We have not been able to develop such a test for bivariate current status data with competing risks.

In view of such limitation of a fully parametric analysis, as an alternative, we have considered two kinds of semi-parametric analyses. In one, we consider parametric models for the baseline cause-specific hazards from the same class and arbitrary frailty distribution G; then suggest semi-parametric estimation of the model parameters and G as well. Although we have considered only the shared frailty model for G for our description of the methodology, it can be extended to the other three types of frailty distributions. In the second semi-parametric analysis, we consider unspecified baseline cause-specific hazards and parametric Gamma frailty. We suggest an approximating model for the unspecified baseline cause-specific hazards using Bernstein polynomials and then carry out what is known as sieve maximum likelihood estimation. In both the semi-parametric analyses, we have carried out some simulation studies and then analysis of the hearing loss data. Some interesting features regarding the incidence of the three types of hearing loss have been revealed through these analyses.

As described in Section 1.12, this hearing loss data has been collected on a group of patients who appeared at the mentioned health institute with some problem of speech impairment. For these patients, in addition, hearing loss diagnosis has also been carried out. Our response variable of interest in this thesis has been the outcome of such diagnosis in both left and right ears. Therefore, all the conclusions drawn from the different analyses in this thesis refer to the population from which this group of patients appears; that is, the population of individuals with some problem of speech impairment and seeking treatment in this particular institute. Since speech impairment can usually be detected in the early ages, a majority of these patients in the sample are in the early age group of 0-10 years, many of them developing hearing loss due to SNHL. However, there are some elderly patients as well; some of those patients are with some hearing loss problem, mostly with Conductive or Mixed type.

Nevertheless, one should use some caution before generalizing the conclusions to the general population. Even for general group of patients, one needs a multi-institute study to draw any such conclusion on the incidence of the three types of hearing loss.

6.2 Future Works

In this section, we identify some important topics related to this thesis work which may be of practical interest and taken up in future.

6.2.1 Regression analysis of bivariate current status data with competing risks

In this thesis, we have not considered any covariate. In many real life situations, covariates play an important role to describe the behaviour and pattern of the different cause-specific hazard rates representing failure time and type. Andrews et al. (2005), Wang et al. (2008) and many others studied regression analysis of current status data. Gorfine and Hsu (2011) also investigated clustered survival data with competing risks using covariates. Some works on interval-censored data with competing risks in presence of covariates are present in the literature (See Mao et al., 2017, Kim, 2017 and many others). Regression analysis of bivariate current status data with competing risks to incorporate effects of covariates is considered an important topic for future work. Presence of covariates may also facilitate identifiability of models, as in Heckman and Singer (1984a).

6.2.2 Multivariate current status data with competing risks

The focus of this thesis has been modeling and analysis of bivariate (paired) current status data with competing risks. One can also think of multivariate (clustered) current status data with competing risks, possibly with different cluster sizes. It may still be convenient to model the dependence using different frailty distributions, as studied in this thesis. Although multivariate interval-censored data has been considered by many authors (See, for example, Chen et al., 2009, Wen and Chen, 2011, Li et al., 2020), we do not know of any work related to multivariate current status data with competing risks.

The shared Gamma frailty model and the shared cause-specific Gamma frailty model can be easily generalized for modeling multivariate failure time with competing risks. More generally, these two models can be used for clustered failure time data

with unequal number of individuals in different clusters. While the shared Gamma frailty model can allow different sets of competing risks for different individuals, the shared cause-specific Gamma frailty model requires same set of competing risks for all individuals. Extension of the correlated Gamma frailty model for modeling multivariate failure time requires similar manipulation to induce dependence in the joint distribution of the multivariate frailty $(\epsilon^{(1)}, \cdots, \epsilon^{(k)})$, where k denotes the number of individuals in a cluster by introducing different independent Gamma variables, the Y_i's (See Section 3.2.2). Here also the modeling can be used for clustered failure time data with unequal number of individuals possibly with different sets of competing risks. The correlated cause-specific Gamma frailty model may be similarly generalized to multivariate or clustered failure time data with some complicated manipulation to induce dependence, but requiring the same set of competing risks for all the individuals.

6.2.3 Bivariate current status data with competing risks and dependent monitoring time

Throughout this thesis, we have assumed that the failure time and the monitoring time are independent random variables. However, there can exist situations where these two random variables are not independent. For example, in case of the hearing loss data, a particular individual goes for a check-up only if he/she has some speech or hearing issue. We know about some works in literature dealing with current status data and dependent monitoring time, for example, see Ma et al. (2015), Xu et al. (2019) etc. There are also some works on dependent interval-censored data (See Finkelstein et al., 2002, Chen and Shen, 2017 etc.). There does not seem to exist any work on bivariate current status data with competing risks and dependent monitoring time. One can consider a copula function between the failure time and monitoring time to model the dependence.

6.2.4 Bivariate current status data with competing risks and missing failure types

Analysis of bivariate current status data is a well-studied topic in the literature (See Jewell et al., 2005, Wang et al., 2008, Groeneboom, 2013, Hu et al., 2017, Liu and Qin, 2018, among many others). Current status data with competing risks has also been studied by several researchers (See Jewell et al., 2003, Maathuis, 2006, Groeneboom

et al., 2008b, Maathuis et al., 2011 and many others). There are some recent works on univariate current status data with competing risks and missing failure types (See Koley and Dewanji, 2019, 2022). But, bivariate current status data with competing risks and missing failure type has not been addressed in the literature. Therefore, it is an important topic for future work. This also can be generalized to the multivariate case.

6.2.5 Cure rate model in the presence of bivariate current status data with competing risks

There is some research work on the analysis of current status data in the presence of a cured subgroup (See Lam and Xue, 2005, Ma, 2011, Liu et al., 2017, Diao and Yuan, 2019, Lam et al., 2021 and many others). Also, there are some works on regression analysis of cure rate model in case of bivariate interval-censored data; for example, see Yang (2021). There have been works on competing risks data in presence of cure rate models (See Austin et al., 2016, Choi et al., 2018, Nicolaie et al., 2019, among many others). Hudgens et al. (2014) studied parametric analysis of interval censored data with competing risks using cure type model. There has not been any work on analysis of cure rate models in case of bivariate current status data with competing risks. It is an important topic for future work.

Bibliography

Abbring, J. H. and Van den Berg, G. J. (2003). The identifiability of the mixed proportional hazards competing risks model. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(3):701–710.

Andrews, C., van der Laan, M., and Robins, J. (2005). Locally efficient estimation of regression parameters using current status data. *Journal of Multivariate Analysis*, 96(2):332–351.

Atwood, C. L. (1976). Convergent design sequences, for sufficiently regular optimality criteria. *The Annals of Statistics*, 4(6):1124–1138.

Austin, P. C., Lee, D. S., and Fine, J. P. (2016). Introduction to the analysis of survival data in the presence of competing risks. *Circulation*, 133(6):601–609.

Ayer, M., Brunk, H., Ewing, G., Reid, W., and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *The Annals of Mathematical Statistics*, 26(4):641–647.

Babu, G. J., Canty, A. J., and Chaubey, Y. P. (2002). Application of bernstein polynomials for smooth estimation of a distribution and density function. *Journal of Statistical Planning and Inference*, 105(2):377–392.

Bakoyannis, G. and Touloumi, G. (2012). Practical methods for competing risks data: a review. *Statistical Methods in Medical Research*, 21(3):257–272.

Bandeen-Roche, K. and Liang, K. (2002). Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika*, 89(2):299–314.

Bandeen-Roche, K. and Ning, J. (2008). Nonparametric estimation of bivariate failure time associations in the presence of a competing risk. *Biometrika*, 95(1):221–232.

Bertail, P., Politis, D. N., and Romano, J. P. (1999). On subsampling estimators with unknown rate of convergence. *Journal of the American Statistical Association*, 94(446):569–579.

Böhning, D. (1985). Numerical estimation of a probability measure. *Journal of Statistical Planning and Inference*, 11(1):57–69.

Böhning, D. (1995). A review of reliable maximum likelihood algorithms for semiparametric mixture models. *Journal of Statistical Planning and Inference*, 47(1-2):5–28.

Carnicer, J. M. and Pena, J. M. (1993). Shape preserving representations and optimality of the bernstein basis. *Advances in Computational Mathematics*, 1(1-4):173–196.

Chang, I.-S., Chien, L.-C., Hsiung, C. A., Wen, C.-C., and Wu, Y.-J. (2007). Shape restricted regression with random Bernstein polynomials. In *Complex Datasets and Inverse Problems*, volume 54 of *IMS Lecture Notes–Monograph Series*, pages 187–202. Institute of Mathematical Statistics.

Chang, I.-S., Hsiung, C. A., Wu, Y.-j., and Yang, C.-c. (2005). Bayesian survival analysis using bernstein polynomials. *Scandinavian Journal of Statistics*, 32(3):447–466.

Chen, C.-M. and Shen, P.-s. (2017). Semiparametric regression analysis of failure time data with dependent interval censoring. *Statistics in Medicine*, 36(21):3398–3411.

Chen, J. (1995). Optimal rate of convergence for finite mixture models. *The Annals of Statistics*, 23(1):221–233.

Chen, M.-H., Tong, X., and Sun, J. (2009). A frailty model approach for regression analysis of multivariate current status data. *Statistics in Medicine*, 28(27):3424–3436.

Chen, Y., Feng, Y., and Sun, J. (2015). Regression analysis of multivariate current status data with auxiliary covariates under the additive hazards model. *Computational Statistics & Data Analysis*, 87:34–45.

Cheng, Y. and Fine, J. P. (2012). Cumulative incidence association models for bivariate competing risks data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 74(2):183–202.

Cheng, Y., Fine, J. P., and Kosorok, M. R. (2007). Nonparametric association analysis of bivariate competing-risks data. *Journal of the American Statistical Association*, 102(480):1407–1415.

Choi, S., Zhu, L., and Huang, X. (2018). Semiparametric accelerated failure time cure rate mixture models with competing risks. *Statistics in Medicine*, 37(1):48–59.

Choudhuri, N., Ghosal, S., and Roy, A. (2004). Bayesian estimation of the spectral density of a time series. *Journal of the American Statistical Association*, 99(468):1050–1059.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1):141–151.

Cornea-Madeira, A. and Davidson, R. (2015). A parametric bootstrap for heavy-tailed distributions. *Econometric Theory*, 31(3):449–470.

Cox, D. R. and Oakes, D. (2018). *Analysis of Survival Data*. Taylor & Francis.

Crowder, M. J. (2012). *Multivariate survival analysis and competing risks*. CRC Press.

Cui, Q., Zhao, H., and Sun, J. (2018). A new copula model-based method for regression analysis of dependent current status data. *Statistics and Its Interface*, 11(3):463–471.

Cui, T. and Pillichshammer, F. (2025). Bernstein approximation and beyond: proofs by means of elementary probability theory. *Elemente der Mathematik*, 81:1–7.

Dabrowska, D. M. (1988). Kaplan–meier estimate on the plane. *The Annals of Statistics*, 16(4):1475–1489.

Diamond, I. D., McDonald, J. W., and Shah, I. H. (1986). Proportional hazards models for current status data: Application to the study of differentials in age at weaning in Pakistan. *Demography*, 23(4):607–620.

Diao, G. and Yuan, A. (2019). A class of semiparametric cure models with current status data. *Lifetime Data Analysis*, 25(1):26–51.

Dib, K., Bouezmarni, T., Belalia, M., and Kitouni, A. (2021). Nonparametric bivariate distribution estimation using bernstein polynomials under right censoring. *Communications in Statistics-Theory and Methods*, 50(23):5574–5584.

Ding, A. A. and Wang, W. (2004). Testing independence for bivariate current status data. *Journal of the American Statistical Association*, 99(465):145–155.

Dixon, S. N., Darlington, G. A., and Desmond, A. F. (2011). A competing risks model for correlated data based on the subdistribution hazard. *Lifetime Data Analysis*, 17(4):473–495.

dos Reis Miranda Filho, W. and Demarqui, F. N. (2025). A class of semiparametric models for bivariate survival data. *Lifetime Data Analysis*, 31(1):102–125.

dos Santos, D. M., Davies, R. B., and Francis, B. (1995). Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. *Journal of Statistical Planning and Inference*, 47(1-2):111–127.

Duchateau, L. and Janssen, P. (2008). *The frailty model*. Springer.

Eddelbuettel, D. (2013). *Seamless R and C++ integration with Rcpp*. Springer.

Fang, H.-b., Li, G., and Sun, J. (2005). Maximum likelihood estimation in a semiparametric logistic/proportional-hazards mixture model. *Scandinavian Journal of Statistics*, 32(1):59–75.

Finkelstein, D. M., Goggins, W. B., and Schoenfeld, D. A. (2002). Analysis of failure time data with dependent interval censoring. *Biometrics*, 58(2):298–304.

Fleming, T. R. and Harrington, D. P. (2013). *Counting processes and survival analysis*. John Wiley & Sons.

Geman, S. and Hwang, C.-R. (1982). Nonparametric maximum likelihood estimation by the method of sieves. *The Annals of Statistics*, 10(2):401–414.

Ghosh, B., Dewanji, A., and Das, S. (2024a). Model identifiability for bivariate failure time data with competing risk: Non-parametric cause-specific hazards and gamma frailty. *arXiv preprint arXiv:2405.07722*.

Ghosh, B., Dewanji, A., and Das, S. (2024b). Parametric analysis of bivariate current status data with competing risks using frailty model. *arXiv preprint arXiv:2405.05773*.

Gorfine, M. and Hsu, L. (2011). Frailty-based competing risks model for multivariate survival data. *Biometrics*, 67(2):415–426.

Green, P. J. and Silverman, B. W. (1993). *Nonparametric regression and generalized linear models: a roughness penalty approach*. Crc Press.

Grenander, U. (1981). *Abstract Inference*. Wiley, New York.

Groeneboom, P. (2013). The bivariate current status model. *Electronic Journal of Statistics*, 7:1783–1805.

Groeneboom, P., Kim, Y., and Wellner, J. A. (2008a). Current status data with competing risks: Consistency and rates of convergence of the mle. *The Annals of Statistics*, 36(4):1660–1693.

Groeneboom, P., Maathuis, M. H., and Wellner, J. A. (2008b). Current status data with competing risks: Limiting distribution of the mle. *The Annals of Statistics*, 36(3):1031–1063.

Guan, Z. (2021). Maximum approximate bernstein likelihood estimation in proportional hazard model for interval-censored data. *Statistics in Medicine*, 40(3):758–778.

Guo, S. (2010). *Survival analysis*. Oxford University Press.

Ha, I. D., Jeong, J.-H., and Lee, Y. (2018). *Statistical modelling of survival data with random effects: h-likelihood approach*. Springer Publishing Company, Incorporated.

Hanagal, D. D. and Pandey, A. (2015). Gamma frailty models for bivariate survival data. *Journal of Statistical Computation and Simulation*, 85(15):3172–3189.

Hanagal, D. D., Pandey, A., and Ganguly, A. (2017). Correlated gamma frailty models for bivariate survival data. *Communications in Statistics-Simulation and Computation*, 46(5):3627–3644.

Hanagal, D. D. and Sharma, R. (2015). Analysis of bivariate survival data using shared inverse gaussian frailty model. *Communications in Statistics-Theory and Methods*, 44(7):1351–1380.

Hao, M., Lin, Y., and Zhao, X. (2020). Nonparametric inference for right-censored data using smoothing splines. *Statistica Sinica*, 30(1):153–173.

Heckman, J. and Singer, B. (1984a). The identifiability of the proportional hazard model. *The Review of Economic Studies*, 51(2):231–241.

Heckman, J. and Singer, B. (1984b). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, 52(2):271–320.

Hens, N., Wienke, A., Aerts, M., and Molenberghs, G. (2009). The correlated and shared gamma frailty model for bivariate current status data: an illustration for cross-sectional serological data. *Statistics in Medicine*, 28(22):2785–2800.

Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika*, 71(1):75–83.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, 73(2):387–396.

Hougaard, P. (2000). *Analysis of multivariate survival data*. Springer.

Hu, T., Zhou, Q., and Sun, J. (2017). Regression analysis of bivariate current status data under the proportional hazards model. *Canadian Journal of Statistics*, 45(4):410–424.

Hudgens, M. G., Li, C., and Fine, J. P. (2014). Parametric likelihood inference for interval censored competing risks data. *Biometrics*, 70(1):1–9.

Hudgens, M. G., Satten, G. A., and Longini Jr, I. M. (2001). Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. *Biometrics*, 57(1):74–80.

Iachine, I. (2004). Identifiability of bivariate frailty models. *Preprint*, 5.

Ibrahim, N. and Kudus, A. (2009). *Decision tree for prognostic classification of multivariate survival data and competing risks*. INTECH Open Access Publisher London.

Jewell, N. P. and Kalbfleisch, J. D. (2004). Maximum likelihood estimation of ordered multinomial parameters. *Biostatistics*, 5(2):291–306.

Jewell, N. P., van der Laan, M., and Henneman, T. (2003). Nonparametric estimation from current status data with competing risks. *Biometrika*, 90(1):183–197.

Jewell, N. P., Van Der Laan, M., and Lei, X. (2005). Bivariate current status data with univariate monitoring times. *Biometrika*, 92(4):847–862.

Johnson, N. L. and Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis*, 5(1):53–66.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The statistical analysis of failure time data*. John Wiley & Sons.

Keiding, N., Begtrup, K., Scheike, T. H., and Hasibeder, G. (1996). Estimation from current-status data in continuous time. *Lifetime Data Analysis*, 2:119–129.

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics*, 27(4):887–906.

Kim, Y.-J. (2014). Regression analysis of bivariate current status data using a multistate model. *Communications in Statistics-Simulation and Computation*, 43(3):462–475.

Kim, Y.-J. (2017). Analysis of interval censored failure time data with competing risk. *Journal of Applied Statistics*, 44(15):2778–2790.

Klein, J. P. and Moeschberger, M. L. (2006). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.

Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G., and Scheike, T. H. (2014). *Handbook of survival analysis*. CRC Press Boca Raton, FL:.

Kleinbaum, D. G. and Klein, M. (1996). *Survival analysis a self-learning text*. Springer.

Kokoszka, P. and Parfionovas, A. (2004). *Bootstrap Unit Root Tests for Heavy-Tailed Time Series*, pages 175–195. Birkhäuser Boston, Boston, MA.

Koley, T. and Dewanji, A. (2019). Revisiting non-parametric maximum likelihood estimation of current status data with competing risks. *Sankhya B*, 81:39–59.

Koley, T. and Dewanji, A. (2022). Current status data with two competing risks and missing failure types: a parametric approach. *Journal of Applied Statistics*, 49(7):1769–1783.

Krailo, M. and Pike, M. (1983). Estimation of the distribution of age at natural menopause from prevalence data. *American Journal of Epidemiology*, 117(3):356–361.

Lai, X., Yau, K. K., and Liu, L. (2017). Competing risk model with bivariate random effects for clustered survival data. *Computational Statistics & Data Analysis*, 112:215–223.

Laird, N. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association*, 73(364):805–811.

Lam, K. and Xue, H. (2005). A semiparametric regression cure model with current status data. *Biometrika*, 92(3):573–586.

Lam, K. F., Lee, C. Y., Wong, K. Y., and Bandyopadhyay, D. (2021). Marginal analysis of current status data with informative cluster size using a class of semiparametric transformation cure models. *Statistics in Medicine*, 40(10):2400–2412.

Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.

Lesperance, M. L. and Kalbfleisch, J. D. (1992). An algorithm for computing the nonparametric mle of a mixing distribution. *Journal of the American Statistical Association*, 87(417):120–126.

Li, C., Taylor, J. M., and Sy, J. (2001). Identifiability of cure models. *Statistics & Probability Letters*, 54(4):389–395.

Li, S., Hu, T., Zhao, S., and Sun, J. (2020). Regression analysis of multivariate current status data with semiparametric transformation frailty models. *Statistica Sinica*, 30(2):1117–1134.

Li, Y., Tiwari, R. C., and Guha, S. (2007). Mixture cure survival models with dependent censoring. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(3):285–306.

Lin, D., Oakes, D., and Ying, Z. (1998). Additive hazards regression with current status data. *Biometrika*, 85(2):289–298.

Lin, G. D. and Dou, X. (2021). An identity for two integral transforms applied to the uniqueness of a distribution via its laplace–stieltjes transform. *Statistics*, 55(2):367–385.

Lindsay, B. G. (1983). The geometry of mixture likelihoods: A general theory. *The Annals of Statistics*, 11(1):86–94.

Lindsay, B. G. and Lesperance, M. L. (1995). A review of semiparametric mixture models. *Journal of Statistical Planning and Inference*, 47(1-2):29–39.

Liu, H. and Qin, J. (2018). Semiparametric probit models with univariate and bivariate current-status data. *Biometrics*, 74(1):68–76.

Liu, X. (2012). *Survival analysis: models and applications*. John Wiley & Sons.

Liu, Y., Hu, T., and Sun, J. (2017). Regression analysis of current status data in the presence of a cured subgroup and dependent censoring. *Lifetime Data Analysis*, 23:626–650.

Lorentz, G. G. (2012). *Bernstein polynomials*. American Mathematical Society.

Lu, W. and Peng, L. (2008). Semiparametric analysis of mixture regression models with competing risks data. *Lifetime Data Analysis*, 14:231–252.

Ma, L., Hu, T., and Sun, J. (2015). Sieve maximum likelihood regression analysis of dependent current status data. *Biometrika*, 102(3):731–738.

Ma, S. (2009). Cure model with current status data. *Statistica Sinica*, 19(1):233–249.

Ma, S. (2011). Additive risk model for current status data with a cured subgroup. *Annals of the Institute of Statistical Mathematics*, 63(1):117–134.

Ma, Y., Leng, C., and Wang, H. (2024). Optimal subsampling bootstrap for massive data. *Journal of Business & Economic Statistics*, 42(1):174–186.

Maathuis, M. H. (2006). *Nonparametric estimation for current status data with competing risks*. PhD thesis, University of Washington.

Maathuis, M. H., Xu, L., et al. (2011). Nonparametric inference for competing risks current status data with continuous, discrete or grouped observation times. *Biometrika*, 98(2):325–340.

Mao, L., Lin, D.-Y., and Zeng, D. (2017). Semiparametric regression analysis of interval-censored competing risks data. *Biometrics*, 73(3):857–865.

Moolgavkar, S. H., McClellan, R. O., Dewanji, A., Turim, J., Luebeck, E. G., and Edwards, M. (2013). Time-series analyses of air pollution and mortality in the united states: a subsampling approach. *Environmental health perspectives*, 121(1):73–78.

Murphy, S. (1994). Consistency in a proportional hazards model incorporating a random effect. *The Annals of Statistics*, 22(2):712–731.

Murphy, S. A. (1995). Asymptotic theory for the frailty model. *The Annals of Statistics*, 23(1):182–198.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142.

Naskar, M., Das, K., and Ibrahim, J. G. (2005). A semiparametric mixture model for analyzing clustered competing risks data. *Biometrics*, 61(3):729–737.

Ng, S. and McLachlan, G. (2003). An em-based semi-parametric mixture model approach to the regression analysis of competing-risks data. *Statistics in Medicine*, 22(7):1097–1111.

Nicolaie, M. A., Taylor, J. M., and Legrand, C. (2019). Vertical modeling: analysis of competing risks data with a cure fraction. *Lifetime Data Analysis*, 25(1):1–25.

Nocedal, J. and Wright, S. J. (1999). *Numerical optimization*. Springer.

Oakes, D. (1986). Semiparametric inference in a model for association in bivanate survival data. *Biometrika*, 73(2):353–361.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493.

Osman, M. and Ghosh, S. K. (2012). Nonparametric regression models for right-censored data using bernstein polynomials. *Computational Statistics & Data Analysis*, 56(3):559–573.

Othus, M., Li, Y., and Tiwari, R. C. (2009). A class of semiparametric mixture cure survival models with dependent censoring. *Journal of the American Statistical Association*, 104(487):1241–1250.

Parner, E. (1998). Asymptotic theory for the correlated gamma-frailty model. *The Annals of Statistics*, 26(1):183–214.

Petersen, J. H. (1998). An additive frailty model for correlated life times. *Biometrics*, pages 646–661.

Petrone, S. (1999). Bayesian density estimation using bernstein polynomials. *Canadian Journal of Statistics*, 27(1):105–126.

Pickles, A. and Crouchlev, R. (1994). Generalizations and applications of frailty models for survival and event data. *Statistical Methods in Medical Research*, 3(3):263–278.

Politis, D. N., Romano, J. P., Wolf, M., Politis, D. N., Romano, J. P., and Wolf, M. (1999). *Subsampling in the IID Case*. Springer.

Politis, D. N. and White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53–70.

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, pages 541–554.

Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *The Annals of Statistics*, 11(2):453–466.

Romano, J. P. and Shaikh, A. M. (2012). On the uniform asymptotic validity of subsampling and the bootstrap. *The Annals of Statistics*, 40(6):2798–2822.

Rossini, A. and Tsiatis, A. (1996). A semiparametric proportional odds regression model for the analysis of current status data. *Journal of the American Statistical Association*, 91(434):713–721.

Sankaran, P. and Antony, A. A. (2007). Bivariate competing risks models under random left truncation and right censoring. *Sankhya A*, 69(3):425–447.

Sankaran, P., Lawless, J., Abraham, B., and Antony, A. (2006). Estimation of distribution function in bivariate competing risk models. *Biometrical Journal*, 48(3):399–410.

Scheike, T. H., Sun, Y., Zhang, M.-J., and Jensen, T. K. (2010). A semiparametric random effects model for multivariate competing risks data. *Biometrika*, 97(1):133–145.

Sengupta, S., Volgushev, S., and Shao, X. (2016). A subsampled double bootstrap for massive data. *Journal of the American Statistical Association*, 111(515):1222–1232.

Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, 51(4):1384–1399.

Shoaee, S. and Khorram, E. (2020). Survival analysis for a new compounded bivariate failure time distribution in shock and competing risk models via an em algorithm. *Communications in Statistics-Theory and Methods*, 49(21):5123–5153.

Sun, J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data*. Statistics for Biology and Health. Springer Science & Business Media, New York.

Sun, J. and Chen, D.-G. (2022). *Emerging topics in modeling interval-censored survival data*. Springer (Editors).

Tenbusch, A. (1994). Two-dimensional bernstein polynomial density estimators. *Metrika*, 41(1):233–253.

Tenbusch, A. (1997). Nonparametric curve estimation with bernstein estimates. *Metrika*, 45(1):1–30.

Tong, X., Hu, T., and Sun, J. (2012). Efficient estimation for additive hazards regression with bivariate current status data. *Science China Mathematics*, 55(4):763–774.

Tran, T. M., Abrams, S., and Braekers, R. (2020). A general frailty model to accommodate individual heterogeneity in the acquisition of multiple infections: An application to bivariate current status data. *Statistics in medicine*, 39(12):1695–1714.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22.

Unkel, S. and Farrington, C. P. (2012). A new measure of time-varying association for shared frailty models with bivariate current status data. *Biostatistics*, 13(4):665–679.

van der Laan, M. J. and Jewell, N. P. (2002). Bivariate current status data. *U.C. Berkeley Divison Of Biostatistics Working Paper Series*.

Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.

Vitale, R. A. (1975). A bernstein polynomial approach to density function estimation. In *Statistical inference and related topics*, volume 2, pages 87–99. Elsevier.

Wahba, G. (1990). *Spline models for observational data*. SIAM.

Wand, M. P. and Jones, M. C. (1994). *Kernel smoothing*. CRC press.

Wang, C.-P. and Ghosh, M. (2003). Bayesian analysis of bivariate competing risks models with covariates. *Journal of Statistical Planning and Inference*, 115(2):441–459.

Wang, J. and Ghosh, S. K. (2012). Shape restricted nonparametric regression with bernstein polynomials. *Computational Statistics & Data Analysis*, 56(9):2729–2741.

Wang, L., Sun, J., and Tong, X. (2008). Efficient estimation for the proportional hazards model with bivariate current status data. *Lifetime Data Analysis*, 14:134–153.

Wang, N., Wang, L., and McMahan, C. S. (2015). Regression analysis of bivariate current status data under the gamma-frailty proportional hazards model using the em algorithm. *Computational Statistics & Data Analysis*, 83:140–150.

Wang, W. and Ding, A. A. (2000). On assessing the association for bivariate current status data. *Biometrika*, 87(4):879–893.

Wang, Y. (2007). On fast computation of the non-parametric maximum likelihood estimate of a mixing distribution. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(2):185–198.

Wang, Y. (2010). Maximum likelihood computation for fitting semiparametric mixture models. *Statistics and Computing*, 20(1):67–77.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 26(4):359–372.

Wen, C.-C. and Chen, Y.-H. (2011). Nonparametric maximum likelihood analysis of clustered current status data with the gamma-frailty cox model. *Computational Statistics & Data Analysis*, 55(2):1053–1060.

Wienke, A. (2010). *Frailty models in survival analysis*. Chapman and Hall/CRC.

Wienke, A., Holm, N. V., Skytthe, A., and Yashin, A. I. (2001). The heritability of mortality due to heart diseases: a correlated frailty model applied to danish twins. *Twin Research and Human Genetics*, 4(4):266–274.

Wienke, A., Lichtenstein, P., and Yashin, A. I. (2003). A bivariate frailty model with a cure fraction for modeling familial correlations in diseases. *Biometrics*, 59(4):1178–1183.

Wu, Y., Zhang, Y., and Zhou, J. (2022). A spline-based nonparametric analysis for interval-censored bivariate survival data. *Statistica Sinica*, 32(3):1541.

Xu, D., Zhao, S., Hu, T., Yu, M., and Sun, J. (2019). Regression analysis of informative current status data with the semiparametric linear transformation model. *Journal of Applied Statistics*, 46(2):187–202.

Yang, D. (2021). *Regression Analysis of Correlated Interval-Censored Failure Time Data with a Cured Subgroup*. PhD thesis, University of Missouri-Columbia.

Yashin, A. I. and Iachine, I. A. (1997). How frailty models can be used for evaluating longevity limits: Taking advantage of an interdisciplinary approach. *Demography*, 34(1):31–48.

Yashin, A. I., Vaupel, J. W., and Iachine, I. A. (1994). A duality in aging: the equivalence of mortality models based on radically different concepts. *Mechanisms of Ageing and Development*, 74(1-2):1–14.

Yashin, A. I., Vaupel, J. W., and Iachine, I. A. (1995). Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies*, 5(2):145–159.

Zdravkovic, S., Wienke, A., Pedersen, N., Marenberg, M., Yashin, A., and De Faire, U. (2002). Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 swedish twins. *Journal of Internal Medicine*, 252(3):247–254.

Zhao, S., Hu, T., and Sun, J. (2015). Regression analysis of informative current status data with the additive hazards model. *Lifetime Data Analysis*, 21(2):241–258.