CrossMark

Study of the bivariate survival data using frailty models based on Lévy processes

Alexander Begun[1] (ID) · **Anatoli Yashin**[2]

Abstract Frailty models allow us to take into account the non-observable inhomogeneity of individual hazard functions. Although models with time-independent frailty have been intensively studied over the last decades and a wide range of applications in survival analysis have been found, the studies based on the models with time-dependent frailty are relatively rare. In this paper, we formulate and prove two propositions related to the identifiability of the bivariate survival models with frailty given by a nonnegative bivariate Lévy process. We discuss parametric and semiparametric procedures for estimating unknown parameters and baseline hazard functions. Numerical experiments with simulated and real data illustrate these procedures. The statements of the propositions can be easily extended to the multivariate case.

1 Introduction

In survival analysis, frailty is usually defined as a non-observable momentan risk of failure and is included in survival models in the form of an unknown nonnegative random variable or random process characterizing non-homogeneity of population with respect to hazard function. Usually, frailty enters the model multiplicatively to

✉ Alexander Begun
a.begun@uea.ac.uk

Anatoli Yashin
aiy@duke.edu

[1] School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK

[2] Duke University, 2024 W. Main Street, Durham, NC 27705, USA

Springer

the hazard function and allows us to take into account the correlations between failure times. Observed covariates can also be included in the multivariate survival models in the form of Cox-like regression. Identifiability and other properties of the univariate and multivariate survival models with time-independent frailty have been studied in some depth, and these models now are popular and widely used in survival studies and in the search for genes influencing longevity. Several models with time-dependent frailty have also been suggested. Woodbury and Manton (1997) introduced a stochastic process model of human mortality and ageing. They considered hazard as a quadratic function of stochastically changing unobserved frailty. The model was further extended by Yashin and Manton (1997) to consider a partially observed frailty process. Modifications of this model were successfully applied to the analyses of longitudinal data with informative censoring. Ideas on further development and applications of these results to studying ageing and longevity are summarized by Yashin et al. (2012). Gjessing et al. (2003) considered an approach based on the proportional hazard model with time-dependent frailty given by the formula

$$Z(t) = \int_0^t a(u, t - u) \mathrm{d}X(R(u)).$$

Here, $X(t)$ is a nonnegative Lévy process (subordinator) with Laplace exponent $\Phi(c)$, $R(t) = \int_0^t r(u) \mathrm{d}u$ defines the time transformation of subordinator for a nonnegative rate function $r(t)$, and the nonnegative weight function $a(u, s)$ determines the extent to which the previous behavior of transformed subordinator influences the hazard function at time t. The authors derived the expressions for the population survival and hazard functions in a general case:

$$S(t) = \exp\left(-\int_0^t \Phi(b(u, t)) r(u) \mathrm{d}u\right),$$
$$\mu(t) = \lambda(t) \int_0^t \Phi'(b(u, t)) a(u, t - u) r(u) \mathrm{d}u,$$

where $\lambda(t)$ is the baseline hazard function and $b(u, t) = \int_u^t \lambda(s) a(u, s - u) \mathrm{d}s$. Gjessing et al. (2003) showed also that under some conditions, quasi-stationary distributions of survivors can arise. This implies the constant limiting population hazard rate, in spite of the increase of the baseline hazard function. Ata and Özel (2012) considered the proportional hazard model with time-dependent frailty given by the discrete compound Poisson process and applied this model to the earthquake data and to traffic accidents data from Turkey.

The aim of this paper is to study the problem of identifiability for bivariate survival models with/without observed covariates and with time-dependent frailty when this frailty is given by a Lévy process (or Lévy processes). In addition, we demonstrate how these models can be used for longevity datasets based on simulated data.

In Sect. 2, we give the definitions of the univariate survival model under mixed proportional hazard specification and the bivariate correlated frailty model. We then discuss the definitions of the uni- and bivariate survival model with time-dependent

frailty given by the nonnegative Lévy processes. At the end of this section, we discuss known findings related to the identifiability of survival models with time-independent frailty, formulate two new propositions about the identifiability of the bivariate survival models with time-dependent frailties, and give the EM algorithm for estimating unknown baseline hazard functions and parameters for the correlated bivariate model with time-dependent frailty. In Sect. 3, we discuss the results of estimations of the baseline hazard functions and unknown parameters (including the Cox-regression coefficients and parameters characterizing the frailty process) in experiments with simulated and real data for parametric and semiparametric approaches. The real-world example was based on the data from the Danish Twin Registry. Conclusions and outlook are presented in Sect. 4.

2 Survival analysis under a frailty setting

Under mixed proportional hazard specification (Abbring and van den Berg 2003), the hazard rates of the failure times for two related individuals depend multiplicatively on the respective baseline hazards $\lambda_j(t)$, regressor functions $\chi_j(\mathbf{u}_j)$ with observed vector of covariates $\mathbf{u}_j$, and unobserved nonnegative random variable (frailty) Z_j, characterizing the heterogeneity in the population with respect to hazard λ_j

$$\mu_j(t_j|Z_j, \mathbf{u}_j) = Z_j \chi_j(\mathbf{u}_j) \lambda_j(t_j), \qquad j = 1, 2. \tag{1}$$

The function $\chi_j(\mathbf{u}_j)$ is frequently specified as $\chi_j(\mathbf{u}_j) = \exp(\beta_j^* \mathbf{u}_j)$ (the Cox-regression term) for some transposed vector parameter β_j^*, $j = 1, 2$. The univariate population survivals $S_j(t|\chi_j(\mathbf{u}_j)) = P(T_j > t_j)$ for random times of death T_j are

$$S_j(t|\chi_j(\mathbf{u}_j)) = \mathbb{E}e^{-Z_j \chi_j(\mathbf{u}_j) \Lambda_j(t_j)} = L(\chi_j(\mathbf{u}_j) \Lambda_j(t_j)), \ j = 1, 2, \tag{2}$$

with cumulative hazard function $\Lambda_j(t) = \int_0^t \lambda_j(\tau) d\tau$ and Laplace transform $L(c) = \mathbb{E}e^{-cZ}$.

In the bivariate correlated frailty model proposed by Yashin et al. (1995), individual frailties (Z_1, Z_2) were constructed under assumptions

$$
\begin{aligned}
Z_1 &= \frac{m_0}{m_1} Y_0 + Y_1, \\
Z_2 &= \frac{m_0}{m_2} Y_0 + Y_2
\end{aligned}
\tag{3}
$$

for independent nonnegative random variables Y_0, Y_1 and Y_2 and some nonnegative constants m_0, m_1 and m_2. Given frailties (Z_1, Z_2), life spans for both individuals were assumed to be conditionally independent. Since the scale factor common to all subjects in the population can be absorbed into the baseline hazard functions $\lambda_j(.)$, $j = 1, 2$, we can put $\mathbb{E}Z_1 = \mathbb{E}Z_2 = 1$. The heterogeneity in the population can be characterized by the variance of frailties $\mathrm{Var}Z_1 = \sigma_1^2$, $\mathrm{Var}Z_2 = \sigma_2^2$. The correlation between frailties $\mathrm{Corr}(Z_1, Z_2)$ we will denote by ρ.

If Y_j are gamma-distributed random variables, $Y_j \sim G(k_j, m_j)$, $j = 0, 1, 2$, with $k_0 = \rho/\sigma_1\sigma_2$, $k_j = 1/\sigma_j^2 - k_0$, $m_j = 1/\sigma_j^2$, $j = 1, 2$; $0 \leqslant \rho \leqslant \min(\sigma_1/\sigma_2, \sigma_2/\sigma_1)$, then the non-conditional bivariate survival function is given by the formula

$$
\begin{aligned}
S(t_1, t_2 | \chi_1(\mathbf{u}_1), \chi_2(\mathbf{u}_2)) &= \mathbb{E}e^{-Z_1\chi_1(\mathbf{u}_1)\Lambda_1(t_1)}e^{-Z_2\chi_2(\mathbf{u}_2)\Lambda_2(t_2)} \\
&= \frac{\left(1 + \sigma_1^2\chi_1(\mathbf{u}_1)\Lambda_1(t_1)\right)^{-k_1} \left(1 + \sigma_2^2\chi_2(\mathbf{u}_2)\Lambda_2(t_2)\right)^{-k_2}}{\left(1 + \sigma_1^2\chi_1(\mathbf{u}_1)\Lambda_1(t_1) + \sigma_2^2\chi_2(\mathbf{u}_2)\Lambda_2(t_2)\right)^{k_0}}
\end{aligned}
$$

(4)

(Wienke 2010). Note that without loss of generality, we can put $m_0 = m_1$. If left truncation is present at ages (t_{01}, t_{02}), we calculate the conditional survival function by dividing the bivariate survival function by $S(t_{01}, t_{02}|\chi(\mathbf{u}_1), \chi(\mathbf{u}_2))$. To take into account information about censoring, we shall put $\delta_j = 0, 1$, where $\delta_j = 0$ indicates right censoring, $j = 0, 1$.

The assumption that the individual frailty is determined at birth and does not change with age seems to be too strong and unrealistic. To make the approach more flexible, we can weaken this assumption and suppose that the frailty is a random process.

Similarly to Gjessing et al. (2003), we shall consider the frailty process $Z = Z(t) : t \geqslant 0$ defined by a nonnegative Lévy process. In accordance with Lévy-Khinchin formula, such a process can be characterized by its Laplace transform

$$
L(c; t) = \mathbb{E}e^{-cZ(t)} = e^{-t\Phi(c)}
$$

with Laplace exponent $\Phi(c)$, while c is the argument of Laplace transform. Note that

$$
\mathbb{E}Z(t) = t\Phi'(0),
$$
$$
\mathrm{Var}Z(t) = -t\Phi''(0).
$$

Examples of Lévy processes include the standard compound Poisson process, the compound Poisson process with general jump distribution, gamma processes, stable processes, and PVF (power variance function) processes (Gjessing et al. 2003). In this paper, we shall consider the nonnegative Lévy processes (subordinators) with the Laplace exponent given by

$$
\Phi(c) = dc + \int_0^\infty (1 - e^{-cx})\nu(dx),
$$

(5)

nonnegative drift d and the jump measure ν with support $(0, \infty)$ satisfying $\int_0^\infty \min(1, x)\nu(dx) < \infty$. The Laplace exponent is an increasing and concave function. The Laplace exponent and the jump measure for the gamma process are given by the formulas

$$
\begin{aligned}
\Phi_G(c) &= h[\ln(\gamma + c) - \ln\gamma] = h\ln(1 + \gamma^{-1}c), \\
\nu(dx) &= he^{-\gamma x}x^{-1}dx
\end{aligned}
$$

(6)

with the shape ht and the scale parameter γ for the gamma-distributed random variable $Z(t)$. This corresponds to the values of $ht\gamma^{-1}$ and $ht\gamma^{-2}$ for mean and variance of $Z(t)$, respectively. To avoid non-identifiability of the model, we shall standardize the frailty distribution and put $h = \gamma$. In this case, $\mathbb{E}Z(t) = t$, $\mathrm{Var}Z(t) = t\gamma^{-1}$ for $t \geqslant 0$. In "Appendix 1," we can find the formulas for calculating univariate population survivals in this case for a constant and an exponential (Gompertz) baseline hazard functions. We will denote the Laplace exponent for the univariate frailty processes $Z_j(t)$, $j = 1, 2$ by $\Phi_j(.)$.

Let $(Z_1(t), Z_2(t))$ be a bivariate Lévy process with nonnegative components and the Laplace exponent given by

$$\Phi_{\mathrm{biv}}(c) = <d, c> + \int_{R_+^2} (1 - \mathrm{e}^{-<c,x>})\nu(\mathrm{d}x), \tag{7}$$

where $d, c \in \mathbb{R}_+^2$, $<x, y>$ denotes the dot product of vectors $x, y \in \mathbb{R}^2$, the Lévy measure $\nu(A)$ for any Borel set $A \in \mathbb{R}_+^2$ is the expected number of jumps on the time interval $[0,1]$, whose sizes belong to A, and the following integrability conditions for Lévy measure are satisfied:

$$\int_{R_+^2} \min(1, x)\nu(\mathrm{d}x) < \infty,$$

$$\int_{R_+^2} \min(1, x^2)\nu(\mathrm{d}x) < \infty.$$

Note that $\Phi_{\mathrm{biv}}(c_1, 0) = \Phi_1(c_1)$ and $\Phi_{\mathrm{biv}}(0, c_2) = \Phi_2(c_2)$ for marginal Lévy processes.

The bivariate survival function under mixed proportional hazard specification with conditionally independent life spans and time-dependent frailties given by correlated Lévy processes is

$$S(t_1, t_2 | \chi_1(\mathbf{u}_1), \chi_2(\mathbf{u}_2)) = \mathbb{E}\mathrm{e}^{-\chi_1(\mathbf{u}_1)\int_0^{t_1} Z_1(\tau)\lambda_1(\tau)\mathrm{d}\tau} \mathrm{e}^{-\chi_2(\mathbf{u}_2)\int_0^{t_2} Z_2(\tau)\lambda_2(\tau)\mathrm{d}\tau}$$

$$= \exp\left(-\int_0^{t_-} \Phi_{\mathrm{biv}}(\chi_1(\mathbf{u}_1)\Lambda_1(\tau, t_1), \chi_2(\mathbf{u}_2)\Lambda_2(\tau, t_2))\mathrm{d}\tau\right)$$

$$\times \exp\left(-\int_{t_-}^{t_+} \Phi_{\mathrm{uni}}^+(\chi_+(\mathbf{u}_1, \mathbf{u}_2)\Lambda_+(\tau, t_+))\mathrm{d}\tau\right),$$

where $t_- = \min(t_1, t_2)$, $t_+ = \max(t_1, t_2)$, $\Phi_{\mathrm{biv}}(., .)$ is the Laplace exponent for the bivariate frailty process (Z_1, Z_2),

$$\Phi_{\mathrm{uni}}^+(.) = \begin{cases} \Phi_2(.), & \text{if } t_2 > t_1, \\ \Phi_1(.), & \text{otherwise}, \end{cases}$$

$$\Lambda_j(\tau, t) = \int_\tau^t \lambda_j(s)\mathrm{d}s, \qquad j = 1, 2,$$

$$\chi_+(\mathbf{u}_1, \mathbf{u}_2) = \begin{cases} \chi_2(\mathbf{u}_2), & \text{if } t_2 > t_1, \\ \chi_1(\mathbf{u}_1), & \text{otherwise,} \end{cases}$$

and

$$\Lambda_+(\tau, t) = \begin{cases} \int_\tau^t \lambda_2(s)\mathrm{d}s, & \text{if } t_2 > t_1, \\ \int_\tau^t \lambda_1(s)\mathrm{d}s, & \text{otherwise.} \end{cases}$$

Similar to the model with time-independent frailties given by (3), we can construct the bivariate survival function for time-dependent frailties given by the correlated Lévy processes $Z_1(t)$ and $Z_2(t)$ with parameters $\mathbb{E}Z_1(1) = \mathbb{E}Z_2(1) = 1$, $\operatorname{Var}Z_1(1) = \gamma_1^{-1}$, $\operatorname{Var}Z_2(1) = \gamma_2^{-1}$, and $\operatorname{Corr}(Z_1(t), Z_2(t)) = \varrho$, $t > 0$. The formula for calculating bivariate population survival function for time-dependent frailties is given in "Appendix 2."

2.1 Model identifiability

The identifiability of the univariate model with unspecified functional form of frailty distribution and baseline hazard has been studied by Elbers and Ridder (1982). This model is identifiable given information on T for finite $\mathbb{E}Z$ and is not identifiable when frailty has an infinite mean. Identifiability of the correlated frailty models using data on the pair (T_1, T_2) was proved by Honoré (1993) under the assumption of finite means of Z_1 and Z_2. Yashin and Iachine (1999) proved the identifiability of the correlated frailty model without observed covariates assuming that Z_1 and Z_2 are gamma distributed. Abbring and van den Berg (2003) studied the identifiability of the mixed proportional hazards competing risks model. We adopt this method to investigate the identifiability of the mixed bivariate survival model for time-dependent correlated frailties.

Proposition 1 *Let the following assumptions be satisfied.*

Assumption 1. Regressor functions $\chi_i : U_i \to \mathbb{R}_+$ are continuous with supports U_i, $U_i \subset \mathbb{R}^n$, $i = 1, 2$, and $\chi_1(\mathbf{u}_1^*) = \chi_2(\mathbf{u}_2^*) = 1$ for some $\mathbf{u}_1^* \in U_1$, $\mathbf{u}_2^* \in U_2$. Set $\Upsilon = \{(\chi_1(\mathbf{u}_1), \chi_2(\mathbf{u}_2)) | \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2\}$ contains a non-empty open set $\Upsilon_0 \subset \mathbb{R}_+^2$.

Assumption 2. Baseline hazard functions $\lambda_j(.)$ are integrable on $[0, t]$ with $\Lambda_j(t) = \int_0^t \lambda_j(\tau)d\tau < \infty$ for all $t \in \mathbb{R}_+$, and $\Lambda_1(t^*) = \Lambda_2(t^*) = 1$ for some $t^* > 0$, $j = 1, 2$.

Assumption 3. Let μ be a probability measure corresponding to the bivariate frailty variable $(Z_1(1), Z_2(1)) \in \mathbb{R}_+^2$. Then,

$$\int_0^\infty \int_0^\infty e^{b(x_1+x_2)}\mathrm{d}\mu < \infty \tag{8}$$

for some real number $b > 0$.

Then, the mixed bivariate survival model for time-dependent correlated frailties is identified from the bivariate distribution of the failure times $(T_1, T_2 | \mathbf{u}_1, \mathbf{u}_2)$.

Proof Identification of the regressor functions.

From Assumption 3 and Theorem 2.1 in Brychkov et al. (1992), it follows that the Laplace exponent $\Phi_{\text{biv}}(p_1, p_2)$ is an analytic function on $H_b = \{(p_1, p_2 | \text{Re}(p_1) > -b, \text{Re}(p_2) > -b)\}$ and therefore is a real analytic function on $\text{Re}H_b = \{(x_1, x_2) | x_1 > -b, x_2 > -b\}$. Note that $\Phi_1(p) = \Phi_{\text{biv}}(p, 0)$ and $\Phi_2(p) = \Phi_{\text{biv}}(0, p)$. Since $\mathbb{E}Z(1) = \Phi'_1(0)$, $\mathbb{E}Z(2) = \Phi'_2(0)$ and the functions $\Phi_1(x)$ and $\Phi_2(x)$ are analytic in 0, it holds that $\mathbb{E}Z_1(1) < \infty$ and $\mathbb{E}Z_2(1) < \infty$. Gjessing et al. (2003) (Theorem 1) have shown that

$$S_j(t|\chi_j(\mathbf{u}_j)) = \mathbb{E}S_j(t|Z_j, \chi_j(\mathbf{u}_j)) = \exp\left(-\int_0^t \Phi_j(\Lambda_j(\tau, t; \chi_j(\mathbf{u}_j)))\mathrm{d}\tau\right),$$

where

$$\Lambda_j(\tau, t; \chi_j(\mathbf{u}_j)) = \int_\tau^t \chi_j(\mathbf{u_j})\lambda_j(s)\mathrm{d}s = \chi_j(\mathbf{u_j})\Lambda_j(\tau, t), \qquad j = 1, 2.$$

As any subordinator frailty processes, Z_j have increasing and concave Laplace exponents and their derivatives $\Phi'_j(.)$ are decreasing. Moreover,

$$\mathrm{d}S_j(t|\chi_j(\mathbf{u}_j))/\mathrm{d}t = -\chi_j(\mathbf{u}_j)\lambda_j(t)\int_0^t \Phi'_j(\Lambda_j(\tau, t; \chi_j(\mathbf{u}_j)))\mathrm{d}\tau$$

$$\times \exp\left(-\int_0^t \Phi_j(\Lambda_j(\tau, t; \chi_j(\mathbf{u}_j)))\mathrm{d}\tau\right)$$

and

$$(t\lambda_j(t))^{-1}\mathrm{d}S_j(t|\chi_j(\mathbf{u}_j))/\mathrm{d}t \to -\chi_j(\mathbf{u_j})\mathbb{E}Z_j(1) \qquad as \qquad t \downarrow 0.$$

From here and Assumption 1, it follows that

$$\frac{\mathrm{d}S_j(t|\chi_j(\mathbf{u}_j))/\mathrm{d}t}{\mathrm{d}S_j(t|\chi_j(\mathbf{u}_j^*))/\mathrm{d}t} \to \chi_j(\mathbf{u_j}) \qquad as \qquad t \downarrow 0, \qquad j = 1, 2.$$

This formula identifies $\chi_j(.)$ in U_j, since marginal survival functions $S_j(t|\chi_j(\mathbf{u}_j))$ are observed for $t \geqslant 0$ and $\mathbf{u_j}$ are arbitrary for $j = 1, 2$.

Identification of the hazard functions.

For $0 \leqslant \tau \leqslant t \leqslant T < \infty$, it holds that $\Lambda_j(\tau, t) \leqslant C < \infty$, $j = 1, 2$. Therefore, there exists a real $b(T) > 0$ such that the bivariate function $\Phi_{\text{biv}}(\chi_1\Lambda_1(\tau, t), \chi_2\Lambda_2(\tau, t))$ and the univariate functions $\Phi_j(\chi_j\Lambda_j(\tau, t))$ are real analytic functions on the set $\Upsilon_T = \{(\chi_1, \chi_2) | \chi_1 > -b(T), \chi_2 > -b(T)\}$ containing the point (0,0) for fixed (τ, t), $0 \leqslant \tau \leqslant t \leqslant T < \infty$ and $j = 1, 2$. Moreover, the univariate survival functions $S_j(t_j|\chi_j)$, their derivatives $S'_j(t_j|\chi_j)$ in t_j, $j = 1, 2$, and the bivariate survival function $S(t_1, t_2|\chi_1, \chi_2)$ are real analytic functions uniquely defined on Υ_T for $\{(t_1, t_2) | 0 \leqslant t_1, t_2 \leqslant T < \infty\}$.

Differentiating $S_j(t|\chi_j)$ in t, dividing then by χ_j, and setting formally $\chi_j \to 0$, we get the following equations:

$$-\lim_{\chi_j(u_j)\downarrow 0} \frac{S'_j(t|\chi_j)}{\chi_j S_j(t|\chi_j)} = \Phi'_j(0)\lambda_j(t) = \mathbb{E}Z_j(1)\lambda_j(t), \quad j = 1, 2. \quad (9)$$

Since the expression on the right hand of (9) is observed, we find the hazard function in the form

$$\lambda_j(t) = -\Phi'_j(0)^{-1} \lim_{\chi_j\downarrow 0} \frac{S'_j(t|\chi_j)}{\chi_j(u_j)S_j(t|\chi_j)}, \quad j = 1, 2.$$

The constant $\Phi'_j(0)$ can be found from the equation $\Lambda_j(t^*) = 1, j = 1, 2$ (Assumption 2).

Identification of the Laplace exponent.

Since $S(0, 0|\chi_1, \chi_2) = 1$ and the bivariate Laplace exponent is regular in the point $(0,0)$, the following formula holds in some neighborhood $O \subset \mathbb{R}^1_+$ of the point $t = 0$

$$-\ln S(t^{**}, t^{**}|\chi_1, \chi_2) = \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{\chi_1^{n_1}}{n_1!} \frac{\chi_2^{n_2}}{n_2!} \frac{\partial^{n_1+n_2}\Phi_{\text{biv}}(c_1, c_2)}{\partial c_1^{n_1}\partial c_2^{n_2}}\Big|_{c_1=0,c_2=0}$$
$$\times \int_0^{t^{**}} (\Lambda_1(t^{**}) - \Lambda_1(\tau))^{n_1}(\Lambda_2(t^{**}) - \Lambda_2(\tau))^{n_2}d\tau.$$
$$(10)$$

Mixed partial derivatives in (10) can be calculated using the finite difference method as the distance between spaced nodes tends to zero. Note that $\int_0^{t^{**}} (\Lambda_1(t^{**}) - \Lambda_1(\tau))^{n_1}(\Lambda_2(t^{**})-\Lambda_2(\tau))^{n_2}d\tau > 0$ for all $n_1 \in \mathbb{N}$ and $n_2 \in \mathbb{N}$ if $t^{**} > t^*$. The mixed partial derivatives $\partial^{n_1+n_2}\Phi_{\text{biv}}(c_1, c_2)/\partial c_1^{n_1}\partial c_2^{n_2}$ at the point $(0, 0)$ are uniquely defined from (10), and therefore, $\Phi_{\text{biv}}(c_1, c_2)$ can be uniquely defined in some open neighborhood of the point $(0,0)$. That is, the real analytic function $\Phi_{\text{biv}}(., .)$ is uniquely defined in some open set containing $\mathbb{R}^2_+$. Similarly, it can be shown that the real analytic functions $\Phi_j, j = 1, 2$, are uniquely defined in some open set containing R^1_+. □

If individual frailties (Z_1, Z_2) are constructed under assumptions given by

$$\begin{aligned} Z_1 &= Y_0 + Y_1, \\ Z_2 &= \alpha Y_0 + Y_2 \end{aligned} \quad (11)$$

for some positive α, we can weaken the conditions of Proposition 1 and prove the identifiability of the model in the absence of observed covariates.

Proposition 2 *Let the following assumptions be satisfied.*

1. *Assumption 1. Decomposition* (11) *holds for independent positive Lévy processes Y_0, Y_1, Y_2 with zero drift and some $\alpha > 0$.*

Assumption 2. Baseline hazard functions $\lambda_j(.)$ are integrable on $[0, t]$ with $\Lambda_j(t) = \int_0^t \lambda_j(\tau)d\tau < \infty$ for all $t \in \mathbb{R}_+$, $\lim_{t \to \infty} \Lambda_j(t) = \infty$, and $\Lambda_j(t^*) = 1$ for some $t^* > 0$, $j = 1, 2$.

Assumption 3. Jump measures ν_i satisfy $\int_0^\infty x\nu_i(dx) < \infty$, $i = 0, 1, 2$.

Assumption 4. Let μ_i is a probability measure corresponding to the frailty $Y_i(1) \in \mathbb{R}_+^1$, $i = 0, 1, 2$. Then,

$$\int_0^\infty \int_0^\infty e^{bx} d\mu_i < \infty \tag{12}$$

for some real number $b > 0$ and $i = 0, 1, 2$.

Then, the mixed bivariate survival model for time-dependent correlated frailties is identified from the bivariate distribution of the failure times (T_1, T_2).

The proof of Proposition 2 is given in "Appendix 3."

2.2 Model validation

In this section, we will assume that $\chi(\mathbf{u}) = \exp(\beta^*\mathbf{u})$. To validate the regression model, we need to estimate the vectors of Cox-regression parameters β_1 and β_2, vector parameter defining frailty ζ [in the case of correlated frailty model either $(\sigma_1^2, \sigma_2^2, \rho)$ for time-independent gamma-frailty or $(\gamma_1^{-1}, \gamma_2^{-1}, \varrho)$ for gamma-frailty process given by the Laplace exponent (6)], and the baseline hazard function $\lambda_1(t)$ and $\lambda_2(t)$. If the baseline hazard functions follow a parametric form such as the Gompertz or the Weibull function with vector parameter ξ, we can use the classic maximum likelihood method to estimate all unknown parameters. The log-likelihood in this case is given by

$$
\begin{aligned}
\ln L(Data; \theta) = {} & (1 - \delta_{i1})(1 - \delta_{i2}) \sum_i \ln S(t_{i1}, t_{i2}|\mathbf{u}_{i1}, \mathbf{u}_{i2}) \\
& - \delta_{i1}(1 - \delta_{i2}) \sum_i \ln \left(\partial S(t_{i1}, t_{i2}|\mathbf{u}_{i1}, \mathbf{u}_{i2})/\partial t_{i1}\right) \\
& - (1 - \delta_{i1})\delta_{i2} \sum_i \ln \left(\partial S(t_{i1}, t_{i2}|\mathbf{u}_{i1}, \mathbf{u}_{i2})/\partial t_{i2}\right) \\
& + \delta_{i1}\delta_{i2} \sum_i \ln \left(\partial^2 S(t_{i1}, t_{i2}|\mathbf{u}_{i1}, \mathbf{u}_{i2})/\partial t_{i1}\partial t_{i2}\right)
\end{aligned}
\tag{13}
$$

Here, $\theta = (\beta, \zeta, \xi)$ is the vector of unknown parameters, $\beta = (\beta_1, \beta_2)$ stands for Cox-regression coefficients, ζ for frailty parameters, and ξ for parameters defining the baseline hazard functions $\lambda_j(t)$, $j = 1, 2$. The data include information about life spans (t_{i1}, t_{i2}), observed covariates $(\mathbf{u}_{i1}, \mathbf{u}_{i2})$, and censoring $(\delta_{i1}, \delta_{i2})$ for a twin pair i, $i = 1, \ldots, n$. The estimate of the vector parameter θ we can find by maximizing the log-likelihood function (13).

If the form of the baseline hazard functions is not specified, the estimates can be obtained by the EM algorithm. This algorithm combines the maximum partial likelihood estimator of the vector parameter (β, ζ) with the Breslow estimator (Breslow

[1972](#)) of the cumulative baseline hazard function $\Lambda(t) = (\Lambda_1(t), \Lambda_2(t))$. The EM algorithm is an iterative procedure with two steps—E (expectation) and M (maximization) on each iteration. It works as follows: Let $f(z_{i1}, z_{i2}|t_{i1}, t_{i2}, \zeta)$ be the density function of the random variable $(Z_{i1}(t_{i1}), Z_{i2}(t_{i2}))$ given parameter vector ζ, $i = 1, \ldots, n$. Denote the estimates of $\Lambda(t)$, ζ, and β on the lth iteration by $\hat{\Lambda}_l(t)$, $\hat{\zeta}_l$, and $\hat{\beta}_l$, respectively. Similar to Gorfine and Hsu ([2011](#)), we define the failure counting process $N_{ij} = \delta_{ij}\text{Ind}(T_{ij} \leqslant t)$ and the at-risk process $X_{ij} = \text{Ind}(T_{ij} \geqslant t)$, where T_{ij} is the random time to the failure of twin j in twin pair i, $i = 1, \ldots, n$, $j = 1, 2$. Define random processes

$$S_j^{(0)}(\beta, t) = \sum_{i=1}^{n} \sum_{j=1}^{2} X_{ij}(t) Z_{ij}(t) \exp(\beta_j^* \mathbf{u}_{ij}),$$

$$S_j^{(1)}(\beta, t) = \sum_{i=1}^{n} \sum_{j=1}^{2} X_{ij}(t) Z_{ij}(t) u_{ij} \exp(\beta_j^* \mathbf{u}_{ij})$$

and equations

$$\int_0^\infty \sum_{i=1}^{n} \left(\mathbf{u}_{ij} - \frac{\hat{S}_j^{(1)}(\beta, t)}{\hat{S}_j^{(0)}(\beta, t)} \right) dN_{ij}(t) = 0, \tag{14}$$

where the symbol "ˆ" means replacing the unknown frailty $Z_{ij}(t)$ with its conditional expectation given the observed data and the current estimates of $\Lambda(t)$, ζ and β.

To estimate the cumulative baseline hazard functions $\Lambda_j(t)$, we shall use the following estimator

$$\hat{\Lambda}_j(t) = \int_0^t \frac{\sum_{i=1}^{n} dN_{ij}(\tau)}{\hat{S}_j^{(0)}(\beta, t)}, \quad j = 1, 2. \tag{15}$$

This estimator is a step function with jumps at the observed failure times of twins.

Using Bayes' formula, we calculate in the E-step conditional expectation of the log-density function

$$\sum_{i=1}^{n} \hat{E} \ln(f(Z_{i1}(.), Z_{i2}(.)|\zeta)) \tag{16}$$

given observed data and the current estimates of $\Lambda(t)$, ζ and β^h. In the M-step, we update these estimates.

The EM algorithm proceeds as follows:

1. *Initialization.* Set $l = 0$. Put $\hat{\zeta}_0 = (0, \varrho_0)$ for any $0 \leqslant \varrho_0 \leqslant 1$ in the case of time-dependent frailty and $\hat{\zeta}_0 = (0, \rho_0)$ for any $0 \leqslant \rho_0 \leqslant 1$, otherwise. This corresponds to $\hat{Z}_{ij}(t) = t$ and $\hat{Z}_{ij} = 1$, respectively. Calculate $\hat{\beta}_1$ and $\hat{\Lambda}_1(t)$

from (14) and (15), respectively. Given the estimates $\hat{\beta}_1$ and $\hat{\Lambda}_1(t)$, calculate $\hat{\zeta}_1$ by maximizing (16). Set $l = 1$.

2. Given $\hat{\Lambda}_l(t)$ and $\hat{\zeta}_l$, calculate $\hat{\beta}_{l+1}$ from (14).
3. Given the estimates $\hat{\zeta}_l$ and $\hat{\beta}_{l+1}$, calculate $\hat{\Lambda}_{l+1}(t)$ by using formula (15).
4. Given the estimates $\hat{\beta}_{l+1}$, $\hat{\Lambda}_{l+1}(t)$, calculate $\hat{\zeta}_{l+1}$ by maximizing (16).
5. Stop if convergence is reached with respect to estimates of β and ζ. Otherwise, $l \to l + 1$ and repeat steps 2–5.

The convergence of the EM algorithm and the properties of parameter estimates have been discussed elsewhere (Zeng and Lin 2007). For the correlated frailty model with time-independent gamma-distributed frailty, the calculation of expression (16) has been discussed in detail by Iachine (1995). The calculation of this expression in the case of the gamma-frailty process for the same model can be found in "Appendix 4."

In the parametric approach, the choice of the appropriate baseline hazard function plays an important role. The Gompertz function does not guarantee the good fit of the marginal survival function for real longevity data. The following gamma parameterization of the univariate survival function gives better results by fitting the real data in the model with time-independent frailty

$$S(t) = (1 + s^2\tilde{\Lambda}(t))^{-1/s^2} = (1 + \sigma^2\Lambda(t))^{-1/\sigma^2}, \tag{17}$$

where $\tilde{\Lambda}(t) = (\tilde{a}/\tilde{b})(\exp(\tilde{b}t) - 1)$ is the pseudo-baseline cumulative hazard, $t \geqslant 30$, $s^2, \tilde{a}, \tilde{b} > 0$, and $\sigma^2 > 0$ is the variance of the time-independent frailty. Given parameters $s^2, \sigma^2, \tilde{a}$, and $\tilde{b}$, it is not difficult to find the true baseline cumulative hazard $\Lambda(t)$ in the form

$$\Lambda(t) = ((1 + s^2\tilde{\Lambda}(t))^{\sigma^2/s^2} - 1)/\sigma^2.$$

In the case of the time-dependent frailty given by a Lévy process with Laplace exponent $\Phi(c)$, we need to consider the following analog of Eq. (17):

$$S(t) = (1 + s^2\tilde{\Lambda}(t))^{-1/s^2} = \exp\left(-\int_0^t \Phi(\Lambda(\tau, t))d\tau\right). \tag{18}$$

Unfortunately, Eq. (18) does not have a closed-form solution with respect to $\Lambda(t)$. In experiments with real data, we will find the approximative solution to (18) such that the function $\Lambda_{\text{appr}}^{\text{dyn}}(t)$ is a non-decreasing, nonnegative, piecewise-constant function satisfying the following conditions:

$$\Lambda_{\text{appr}}^{\text{dyn}}(0) = 0,$$
$$(1/s^2)\ln(1 + s^2\tilde{\Lambda}(t_i)) = \gamma\int_0^{t_i} \ln\left(1 + \left(\Lambda_{\text{appr}}^{\text{dyn}}(t_i) - \Lambda_{\text{appr}}^{\text{dyn}}(\tau)\right)/\gamma\right)d\tau \tag{19}$$

for times-to-event t_i, $i = 1, \ldots, 2n$, sorted in non-decreasing order, $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{2n-1} \leqslant t_{2n} = \max_{i=1}^{2n} t_i$ (here, n is the number of pairs). The values of $\Lambda_{\text{appr}}^{\text{dyn}}(t_i)$

can be calculated recurrently for $i = 1, 2, \ldots, 2n$ using a simple bisectional procedure. Note that the function $\Lambda_{\text{appr}}^{\text{dyn}}(t)$ converges pointwise to the solution of (19) as $n \to \infty$ and the distance between neighboring moments t_i tends to zero.

To compare two approaches, we assume that in the case of the time-independent frailty, the cumulative hazard is also a non-decreasing, nonnegative, piecewise-constant function $\Lambda_{\text{appr}}^{\text{stat}}(t)$ satisfying the following conditions:

$$
\begin{aligned}
&\Lambda_{\text{appr}}^{\text{stat}}(0) = 0, \\
&\Lambda_{\text{appr}}^{\text{stat}}(t_i) = ((1 + s^2 \tilde{\Lambda}(t_i))^{\sigma^2/s^2} - 1)/\sigma^2.
\end{aligned}
\tag{20}
$$

3 Results

3.1 Experiments with simulated data

In this subsection, we will discuss the results of the consistency test for the correlated frailty models with time-dependent and time-independent frailties (11). It was assumed that $\alpha = 1$, $\text{Var} Z_1 = \text{Var} Z_2 = \sigma^2$, and $\text{Corr}(Z_1, Z_2) = \rho$ in the case of the time-independent frailty or $\text{Var} Z_1(1) = \text{Var} Z_2(1) = \gamma^{-1}$ and $\text{Corr}(Z_1(1), Z_2(1)) = \varrho$ in the case of the time-dependent frailty, baseline hazard functions $\lambda_j(t)$ followed Gompertz (exponential) form $\lambda_j(t) = a \exp(bt)$, and an observed covariate $\mathbf{u}$ influenced longevity so that the conditional hazard function was defined by $\mu_j(t_j|Z_j, \mathbf{u}_j) = Z_j \exp(\beta \mathbf{u}_j)\lambda_j(t)$, $j = 1, 2$. The covariates were randomly generated from the uniform distribution on the interval [0,1] and were independent for the individuals. The (true) values for data generating are given in Tables 1 and 2 and have been chosen so that the mean and the standard deviation of the generated times-to-event were equal to approximately 75 and 12 years, respectively. The bivariate times-to-event have been generated using formula (4) with $\sigma_1 = \sigma_2 = \sigma$ in the case of the time-independent frailty and using formula (25) given in "Appendix 2" in the case of time-dependent frailty. In both cases, it was assumed that $\Lambda_1(t) = \Lambda_2(t) = (a/b)(\exp(bt) - 1)$ and $\chi_1(\mathbf{u}) = \chi_2(\mathbf{u}) = \exp(\beta \mathbf{u})$. We have not truncated or censored the generated data. We estimated unknown parameters and cumulative hazard functions using the classic maximum likelihood estimator (parametric method) and the EM algorithm (semiparametric method). In all cases, we simulated 100 datasets for 500 twin pairs.

Table 1 shows the results of simulation study for the time-independent gamma-frailty model without truncation and censoring. Empirical means and standard deviations of estimates were calculated using the classic maximum likelihood method and the EM algorithm, respectively.

Table 2 shows the results of the simulation study for the time-dependent gamma-frailty model without truncation and censoring. Empirical means and standard deviations of estimates were also calculated using the classic maximum likelihood method and the EM algorithm, respectively. Analysis of estimates in both tables does not indicate the presence of any bias, and estimates calculated using the classic maximum likelihood estimator are generally more efficient. Furthermore, the estimates of

Table 1 Estimates of unknown parameters for the time-independent frailty model

| | True | Estimator | | | |
| | | Classic MLE | | EM algorithm | |
		Mean	SD	Mean	SD
$10^5 \cdot a$	1	1.043	0.369	–	–
$10 \cdot b$	1	1.002	0.049	–	–
β	3	3.012	0.231	2.889	0.244
σ^2	1	1.002	0.125	0.890	0.166
ρ	0.5	0.497	0.086	0.543	0.100

100 simulated datasets for 500 twin pairs

Table 2 Estimates of unknown parameters for the time-dependent frailty model

| | True | Estimator | | | |
| | | Classic MLE | | EM algorithm | |
		Mean	SD	Mean	SD
$10^7 \cdot a$	1	1.059	0.464	–	–
$10 \cdot b$	1	1.005	0.052	–	–
β	3	3.018	0.221	3.006	0.176
$20 \cdot \gamma$	1	1.088	0.579	1.094	0.195
ϱ	0.5	0.465	0.240	0.535	0.088

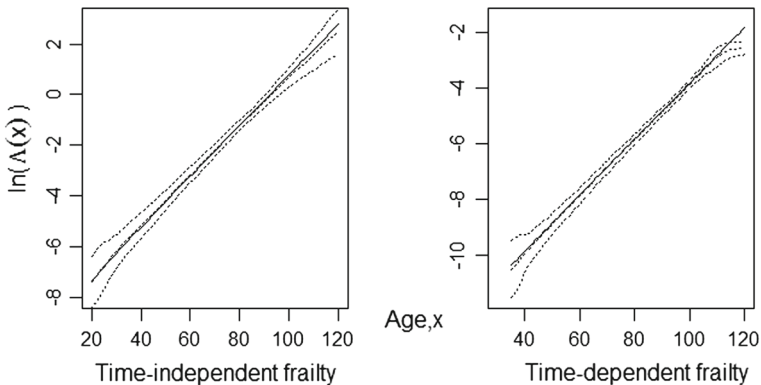100 simulated datasets for 500 twin pairs

Fig. 1 Dependency of $\ln(\Lambda)$ on age. True trajectory (solid line), the empirical mean trajectory, and its lower and upper 95% limit trajectories (dashed lines)

the Cox-regression coefficients and parameters characterizing the frailty distribution are closer to true vales than those for the EM algorithm. One can see in Fig. 1 that in both cases (the time-dependent and the time-independent frailty), the empirical mean log baseline cumulative hazard trajectory calculated using the EM algorithm fits the true log baseline cumulative hazard trajectory very well.

3.2 Experiments with real data

For experiments with real data, we used the datasets from the Danish Twin Registry (DTR). This registry was created in the 1950s. It is one of the oldest population-based registries in the world and contains information about twins born in Denmark since 1870 and who survived to age 6. Multiple births were manually ascertained in birth registers from all 2200 parishes in Denmark. As soon as a twin was traced, a questionnaire was mailed to the twin, to her/his partner or to their closest relatives if neither of the twin partners were alive. The zygosity of twins was assessed on the basis of questions about phenotypic similarities. The reliability of the zygosity diagnosis was validated by comparing laboratory methods based on the blood, serum, and enzyme group determination. In general, the misclassification rates were less than 5%. Other information includes the data on sex, birth, cause of death, health, and lifestyle. An important feature of the Danish twin survival data is their right censoring and left truncation. In our study, we used the longevity data on the like-sex twins with known zygosity born between 1870 and 1900 and who survived until age 30. This non-censored data include 470 male monozygotic (MZ) twin pairs, 475 female MZ twin pairs, 780 male dizygotic (DZ) twin pairs, and 835 male DZ twin pairs. Further details on the Danish Twin Registry can be found in Hauge (1981).

Since the EM algorithm suffers from its slow convergence, we have estimated unknown parameters for the time-independent and the time-dependent frailty models using the classic maximum likelihood method. Table 3 shows these estimates, the logarithm of the maximum value of the likelihood function, and the value of the AKAIKE Information Criterion (AIC) for the data from the Danish Twin Registry described above. The estimates of parameter s^2 were very close to zero in all experiments with real data. We have put this parameter equal to zero to avoid the efficiency loss of estimates. The AIC values for the model with time-dependent frailty are slightly smaller than the ones for the model with time-independent frailty. That is, the model with time-dependent frailty is slightly more informative than the one with time-independent frailty. Figures 2 and 3 show estimated and empirical bivariate probability density functions for the time-independent and the time-dependent frailty models. Note that the shapes of the estimated bivariate probability density functions are very similar.

4 Discussion

Frailty models are a powerful tool for studying non-observable inhomogeneity in a population related to time-to-failure (e.g., death or disease). Models with time-independent frailty have been intensively studied over the last decades and have found a wide range of applications in survival analysis and in searching for genes influencing longevity. However, the studies based on the models with time-dependent frailty are scarce. In this paper, we have attempted to improve the knowledge in this area and to study some properties of multivariate survival models with time-dependent frailty components.

Table 3 Estimates of unknown parameters (standard errors) for the time-independent and [†]the time-dependent frailty models

	Frailty			
	Time-independent		Time-dependent	
	Males	Females	Males	Females
$10^5 \cdot \tilde{a}$	7.004	6.735	7.005	6.710
	(0.782)	(0.743)	(0.835)	(0.997)
$10 \cdot \tilde{b}$	0.916	0.887	0.916	0.888
	(0.015)	(0.014)	(0.016)	(0.020)
s^2	0	0	0	0
	–	–	–	–
σ^2 or $^†10^2 \cdot \gamma$	1.751	1.166	0.626	1.718
	(0.811)	(0.853)	(1.696)	(0.773)
ρ_{MZ} or $^†\varrho_{MZ}$	0.507	0.538	0.477	0.671
	(0.092)	(0.132)	(0.170)	(0.158)
ρ_{DZ} or $^†\varrho_{DZ}$	0.191	0.318	0.176	0.391
	(0.070)	(0.111)	(0.091)	(0.111)
$LogLik$	− 9873.705	− 10,470.3	− 9872.825	− 10,470.14
AIC	19,757.41	20,950.6	19,755.65	20,950.28

Classic MLE estimator. Twin data from the Danish Twin Registry

Proposition 1 we have formulated and proved for the bivariate case. It is not difficult to generalize this result and to prove the identifiability of the frailty model with observed covariates for any number J of related individuals equal or greater than 1 if the time-dependent frailty is a multivariate Lévy process. Similarly, we can generalize Proposition 2 for the case of $J \geqslant 2$. However, the number of frailty components in the multivariate analog of the decomposition (11) will be equal to $2^J − 1$. The shared frailty model where all individuals in a family or cluster share the same non-observable risk of failure does not meet this problem.

In experiments with simulated data, we tested for consistency and used parametric and semiparametric approaches. In the parametric approach, we assumed that the parametric form of the baseline hazard functions is known and follows the Gompertz form. All unknown parameters characterizing frailty distribution, baseline hazard function, and Cox-regression parameters were estimated directly by maximizing the likelihood function. In the semiparametric approach, we used the EM algorithm and estimated the Cox-regression parameters and the parameters of frailty distribution by maximizing the partial likelihood function. The cumulative baseline hazard function was estimated using the Breslow estimator regarding this function as infinite-dimensional parameters. The EM algorithm suffers from its slow convergence. Moreover, in the semiparametric approach, the number of calculations increases with the number of individuals much more rapidly than in the parametric one. It leads to the drastic slowing of the convergence of the EM algorithm and increases substantially the time of
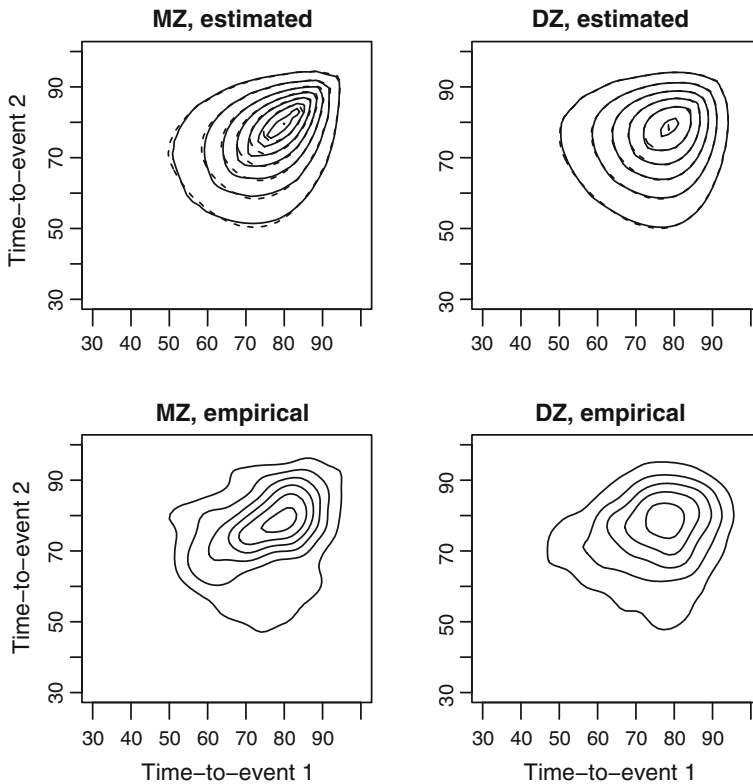
Fig. 2 Estimated and empirical bivariate probability density functions for the time-dependent (solid line) frailty and the time-independent (dashed line) frailty. Male twin pairs from the Danish Twin Registry

estimation. It makes implementing the EM algorithm in the case of the time-dependent frailty for analysis of the real data problematic.

Experiments with real data show that the proposed method and the method with time-independent frailty produce similar shapes of the estimated bivariate probability density functions. The baseline cumulative hazard functions have been chosen so that the estimated marginal survival functions guarantee the best fit to the empirical ones according to Eqs. (19)–(20). A large degree of similarity of the estimated bivariate density functions for the models with time-dependent and time-independent frailties in the range of ages 30–100 years guarantees the similar bivariate fit. The difference between the two approaches can involve the shape of the baseline hazard functions and the asymptotic behavior of the bivariate probability density functions. The models with time-dependent and time-independent frailties are not nested. Therefore, we cannot compare them using the likelihood ratio test. For this purpose, the AKAIKE Information Criterion can be used. In accordance with this criterion, the model with time-dependent frailty is slightly more informative compared to the one with time-independent frailty for the data we considered.
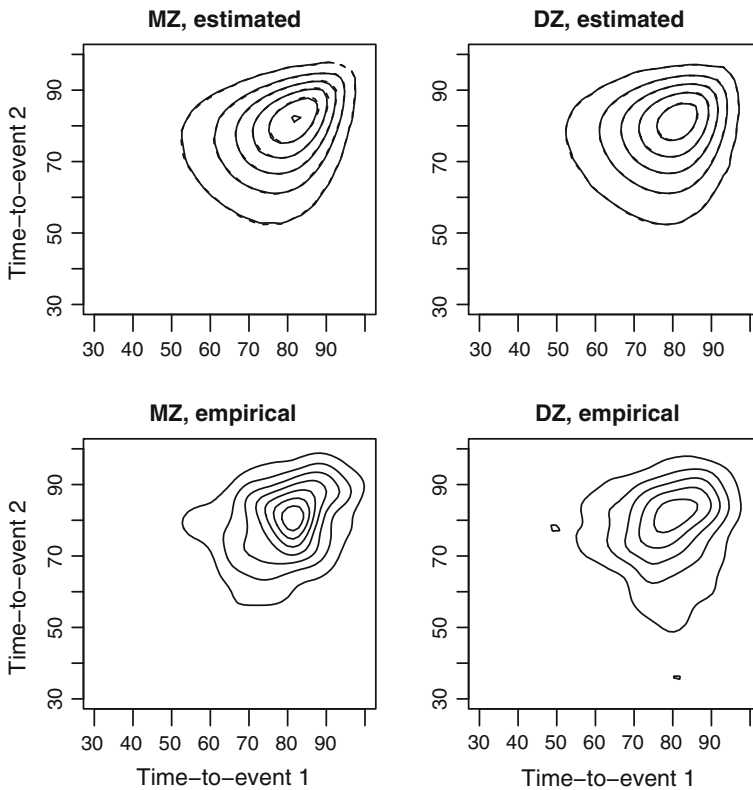
Fig. 3 Estimated and empirical bivariate probability density functions for the time-dependent (solid line) frailty and the time-independent (dashed line) frailty. Female twin pairs from the Danish Twin Registry

Gorfine and Hsu (2011) studied the robustness of the multivariate survival models with frailty components against the violations of the model assumptions. It was found that unnecessary modeling of the dependency between the frailty variates can lead to some efficiency loss of parameter estimates. Misspecification of the frailty distribution can introduce a bias in estimates. Misspecification of the baseline hazard functions can lead to severe bias of all estimates if we use the parametric maximum likelihood estimator, where the baseline hazard functions follow the parametric form. The non-parametric maximum likelihood estimator does not suffer from this drawback. Note that in experiments with real data, we have used a flexible parametrization of the baseline cumulative hazard functions given by formulas (19)–(20). This parameterization does not presume any knowledge about the form of the baseline hazard function. It is sufficient to have a good approximation of the marginal survival function.

An extension of the present study can include the investigation of identifiability of the survival models with competing risks and time-dependent frailty components. The piecewise-constant approximation of the cumulative hazard function has been used in experiments with real data [formulas (19)–(20)]. Other approximative functions such as piecewise linear or piecewise exponential can be used to improve the

bivariate goodness-of-fit. Further, numerical experiments with real data are needed to understand whether the proposed method improves the goodness-of fit on the method with time-independent frailty.

Appendix 1: Univariate survival function for frailty model based on a gamma process

Let $Z = \{Z(t) : t \geqslant 0\}$ is a gamma process with Laplace transform given by the Lévy-Khinchin formula

$$L(c; t) = \mathbb{E}e^{-cZ(t)} = e^{-t\Phi_G(c)},$$

where $\Phi_G(c)$ is the Laplace exponent of the Lévy process Z, the argument for Laplace transform $c \geqslant 0$ and

$$\Phi_G(c) = h[\ln(\gamma + c) - \ln \gamma] = h \ln(1 + \gamma^{-1}c).$$

Here, ht and γ are the shape and the scale parameters, respectively, for gamma-distributed random variable $Z(t)$. Since $Z(t)$ is a.s. increasing in t, we can consider it as subordinator. It is easy to check that

$$
\begin{aligned}
\mathbb{E}Z(t)e^{-cZ(t)} &= t\Phi_G^{(1)}(c)e^{-t\Phi_G(c)} \\
\mathbb{E}Z(t)^2 e^{-cZ(t)} &= \left(t^2\Phi_G^{(1)}(c)^2 - t\Phi_G^{(2)}(c)\right)e^{-t\Phi_G(c)} \\
\mathbb{E}Z(t)^3 e^{-cZ(t)} &= \left(t\Phi_G^{(3)}(c) + t^3\Phi_G^{(1)}(c)^3 - 3t^2\Phi_G^{(1)}(c)\Phi_G^{(2)}(c)\right)e^{-t\Phi_G(c)},
\end{aligned}
\tag{21}
$$

where superscript "(l)" stands for lth-order argument derivative.

In nonhomogeneous population, the process Z (individual for each member of population) can characterize the individual risk of mortality (frailty). In accordance with the definition of the frailty model under mixed proportional hazard specification, we have

$$\mu(t|(Z, \mathbf{u}) = Z(t)\chi(\mathbf{u})\lambda(t)$$

for covariate vector $\mathbf{u}$, parameter vector β, and baseline hazard $\lambda(t)$. In Gjessing et al. (2003) (Theorem 1), it was shown that the population survival in this case can be calculated

$$S(t|\mathbf{u}) = \mathbb{E}S(t|Z, \mathbf{u}) = \exp\left(-\int_0^t \Phi_G(\Lambda(\tau, t; \mathbf{u}))d\tau\right),$$

where

$$\Lambda(\tau, t; \mathbf{u}) = \int_{\tau}^{t} \chi(\mathbf{u})\lambda(s)ds.$$

For constant $\lambda(t) = \lambda_0$, it holds that $\Lambda(\tau, t; \mathbf{u}) = \lambda_0 e^{\beta^* \mathbf{u}}(t - \tau)$,

$$S(t|\mathbf{u}) = \exp\left(h\left(t - \frac{\gamma}{\chi(\mathbf{u})\lambda_0}\left(1 + \frac{\chi(\mathbf{u})\lambda_0 t}{\gamma}\right)\ln\left(1 + \frac{\chi(\mathbf{u})\lambda_0 t}{\gamma}\right)\right)\right),$$

and

$$\mu(t|\mathbf{u}) = -\frac{d\ln S(t|\mathbf{u})}{dt} = h\ln\left(1 + \frac{\chi(\mathbf{u})\lambda_0 t}{\gamma}\right).$$

Using the following formula for indefinite integral

$$\int \ln(1 + \alpha(\exp(\kappa t) - \exp(\kappa \tau)))d\tau = -\kappa^{-1}\text{Li}_2\left(\frac{\alpha \exp(\kappa \tau)}{\alpha \exp(\kappa t) + 1}\right)$$
$$- \tau \ln\left(1 - \frac{\alpha \exp(\kappa t)}{1 + \alpha \exp(\kappa t)}\right)$$

we get for the baseline hazard given by the Gompertz function $\lambda(t) = a\exp(bt)$

$S(t|\mathbf{u})$
$$= \exp\left(-h\int_0^t \ln\left(1 + \gamma^{-1}\Lambda(\tau, t; \mathbf{u})\right)d\tau\right)$$
$$= \exp\left(-h\int_0^t \ln\left(1 + \frac{a\chi(\mathbf{u})}{\gamma b}(\exp(bt) - \exp(b\tau))\right)d\tau\right)$$
$$= \exp\left(-h\left(-t\ln(1 - f(t, \mathbf{u})e^{bt}) - b^{-1}\text{Li}_2(f(t, \mathbf{u})e^{bt}) + b^{-1}\text{Li}_2(f(t, \mathbf{u}))\right)\right).$$
$$(22)$$

Here, $f(t, \mathbf{u}) = (a\chi(\mathbf{u})/(b\gamma(1 + a\chi(\mathbf{u})e^{bt}/(b\gamma)))$ and $\text{Li}_2(z)$ is the dilogarithm function defined by

$$\text{Li}_2(z) = -\int_0^z \frac{\ln(1 - u)}{u}du, z \in \mathbb{C}\setminus[1, \infty).$$

This function is an analytical extension of the infinite series

$$\text{Li}_2 = \sum_{i=1}^{\infty} \frac{z^k}{k^2}, |z| < 1.$$

Appendix 2: Bivariate survival function for frailty model based on gamma processes

Analogously to the univariate case, we can write the bivariate survival function for the bivariate mixed proportional hazard model for the correlated time-dependent frailties given by Lévy processes $Z_1 = \{Z_1(t) : t \geqslant 0\}$ and $Z_2 = \{Z_2(t) : t \geqslant 0\}$ in the form

$$S(t_1, t_2 | \mathbf{u}_1, \mathbf{u}_2) = \mathbb{E}e^{-\chi_1(\mathbf{u}_1) \int_0^{t_1} Z_1(\tau)\lambda_1(\tau)d\tau} e^{-\chi_2(\mathbf{u}_2) \int_0^{t_2} Z_2(\tau)\lambda_2(\tau)d\tau}$$

$$= \exp\left(-\int_0^{t_-} \Phi_{\mathrm{biv}}(\chi_1(\mathbf{u}_1)\Lambda_1(\tau, t_1), \chi_2(\mathbf{u}_2)\Lambda_2(\tau, t_2))d\tau\right)$$

$$\times \exp\left(-\int_{t_-}^{t_+} \Phi_{\mathrm{uni}}^+(\chi_+(\mathbf{u}_1, \mathbf{u}_2)\Lambda_+(\tau, t_+))d\tau\right),$$

where $t_- = \min(t_1, t_2)$, $t_+ = \max(t_1, t_2)$, $\Phi_{\mathrm{biv}}(., .)$ is the Laplace exponent for the bivariate frailty process (Z_1, Z_2),

$$\Lambda_j(\tau, t) = \int_\tau^t \lambda_j(s)ds, \qquad j = 1, 2,$$

$$\chi_+(\mathbf{u}_1, \mathbf{u}_2) = \begin{cases} \chi_2(\mathbf{u}_2), & \text{if } t_2 > t_1, \\ \chi_1(\mathbf{u}_1), & \text{if } t_1 > t_2, \end{cases}$$

$$\Phi_{\mathrm{uni}}^+(.) = \begin{cases} \Phi_2(.), & \text{if } t_2 > t_1, \\ \Phi_1(.), & \text{otherwise}, \end{cases}$$

and

$$\Lambda_+(\tau, t) = \begin{cases} \int_\tau^t \lambda_2(s)ds, & \text{if } t_2 > t_1, \\ \int_\tau^t \lambda_1(s)ds, & \text{if } t_1 > t_2. \end{cases}$$

Assume now that related individuals (e.g., twins) have individual frailty processes $Z_1 = \{Z_1(t) : t \geqslant 0\}$ and $Z_2 = \{Z_2(t) : t \geqslant 0\}$, respectively, so that

$$\begin{aligned} Z_1 &= Y_0 + Y_1, \\ Z_2 &= Y_0 + Y_2, \end{aligned} \tag{23}$$

where $Y_0 = \{Y_0(t) : t \geqslant 0\}$, $Y_1 = \{Y_1(t) : t \geqslant 0\}$, and $Y_2 = \{Y_2(t) : t \geqslant 0\}$ are the gamma processes with Laplace exponents

$$\Phi_{G,j}(c) = h_j[\ln(\gamma + c) - \ln\gamma], \qquad j = 0, 1, 2. \tag{24}$$

with $h_1 = h_2$. Denote $\varrho = h_0/(h_0 + h_1)$. To make the model identifiable, we standardize the gamma-frailty process putting $h = h_0 + h_1 = \gamma$ so that $EZ_j(1) = 1$, $j = 1, 2$. It is easy to check that Z_1 and Z_2 are also gamma processes with Laplace

exponents

$$\Phi_{G,j}(c) = h_j \ln(1 + \gamma^{-1}c), \qquad j = 1, 2.$$

Assume that conditional hazard functions for twins are given by the model with proportional hazard

$$\mu_j(t|(Z_j, \mathbf{u}_j) = Z_j(t)\chi_j(\mathbf{u}_j)\lambda_j(t), \qquad j = 1, 2.$$

Then, the bivariate population survival function has the form

$$
\begin{aligned}
S(t_1, t_2|\mathbf{u}_1, \mathbf{u}_2) = \exp\left(-\int_0^{t_1} \Phi_{G,1}(\Lambda(\tau, t_1; \mathbf{u}_1))\mathrm{d}\tau\right) \\
\times \exp\left(-\int_0^{t_2} \Phi_{G,2}(\Lambda(\tau, t_2; \mathbf{u}_2))\mathrm{d}\tau\right) \\
\times \exp\left(-\int_0^{t_-} \Phi_{G,0}(\Lambda_-(\tau, t_1, t_2; \mathbf{u}_1, \mathbf{u}_2))\mathrm{d}\tau\right) \\
\times \exp\left(-\int_{t_-}^{t_+} \Phi_{G,0}(\Lambda_+(\tau, t_+; \mathbf{u}_1, \mathbf{u}_2))\mathrm{d}\tau\right)
\end{aligned}
\tag{25}
$$

with

$$\Lambda_-(\tau, t_1, t_2; \mathbf{u}_1, \mathbf{u}_2) = \chi_1(\mathbf{u}_1)\Lambda_1(\tau, t_1) + \chi_2(\mathbf{u}_2)\Lambda_2(\tau, t_2)).$$

If the baseline hazard function is given by the Gompertz function, we can use formula (22) to calculate the bivariate population survival.

Appendix 3: The proof of Proposition 2

To prove Proposition 2, we need to check several auxiliary statements.

Lemma 1 *Let $c_1, c_2 \geqslant 0$. Then,*

$$\left|e^{-c_2} - e^{-c_1}\right| \leqslant |c_2 - c_1|.$$
$$\tag{26}$$

Proof Assume that $c_1 \geqslant c_2 \geqslant 0$ and set $\Delta c = c_1 - c_2$. Then,

$$\left|e^{-c_2} - e^{-c_1}\right| = e^{-c_2} - e^{-c_1} \leqslant \frac{e^{\Delta c} - 1}{e^{c_2 + \Delta c}} \leqslant 1 - e^{-\Delta c}.$$

It is easy to check that

$$\max_{\Delta c \geqslant 0} (1 - e^{-\Delta c} - \Delta c) = 0.$$

Therefore,

$$\left| e^{-c_2} - e^{-c_1} \right| \leqslant \Delta c.$$

Last result holds also in the case of $c_2 > c_1 \geqslant 0$. □

Lemma 2 *Let $\Phi(c)$ be a Laplace exponent given by* (5), *and $\Lambda(t)$ is an increasing continuous function defined in $\mathbb{R}_+$. Then,*

$$\int_0^\tau \Phi'(\Lambda(\tau) - \Lambda(s)) ds \geqslant b\tau + \tau \int_0^\infty x e^{-\Lambda(\tau)x} \nu(dx). \tag{27}$$

Proof The statement of lemma follows directly from formula (5) and monotone increase of the function $\Lambda(.)$. □

The bivariate survival function for correlated frailty processes given by (11) is

$$\begin{aligned}
S(t_1, t_2) &= \mathbb{E} e^{-\int_0^{t_1} Z_1(\tau)\lambda_1(\tau)d\tau - \int_0^{t_2} Z_2(\tau)\lambda_2(\tau)d\tau} \\
&= e^{-\int_0^{t_1} \Phi_1(\Lambda_1(\tau, t_1))d\tau} e^{-\int_0^{t_2} \Phi_2(\Lambda_2(\tau, t_2))d\tau} \\
&\quad \times e^{-\int_0^{t_-} \Phi_0(\Lambda_1(\tau, t_1) + \alpha\Lambda_2(\tau, t_2))d\tau} e^{-\int_{t_-}^{t_+} \Phi_0(\Lambda_0(\tau, t_+))d\tau},
\end{aligned} \tag{28}$$

where $\Phi_i(.)$, $i = 0, 1, 2$, is the Laplace exponent for the frailty process $Y_i(t)$ and

$$\Lambda_0(\tau, t) = \begin{cases} \alpha\Lambda_2(\tau, t), & \text{if } t_2 > t_1, \\ \Lambda_1(\tau, t), & \text{otherwise.} \end{cases}$$

Using formula (28) and Assumptions 2-3 of Proposition 2, it can be shown that

$$\int_0^t \Phi_1(\Lambda_1(\tau, t))d\tau = -\ln\left(\lim_{t_2 \uparrow \infty} \frac{S(t, t_2)}{S(0, t_2)}\right) = g_1(t),$$

$$\int_0^t \Phi_2(\alpha\Lambda_2(\tau, t))d\tau = -\ln\left(\lim_{t_1 \uparrow \infty} \frac{S(t_1, t)}{S(t_1, 0)}\right) = g_2(t),$$

$$\int_0^{t_-} \Phi_0(\Lambda_1(\tau, t_1) + \alpha\Lambda_2(\tau, t_2))d\tau + \int_{t_-}^{t_+} \Phi_0(\Lambda_0(\tau, t_+))d\tau$$

$$= -\ln\left(\frac{S(t_1, t_2)}{\lim_{t_1 \uparrow \infty} \frac{S(t_1, t_2)}{S(t_1, 0)} \lim_{t_2 \uparrow \infty} \frac{S(t_1, t_2)}{S(0, t_2)}}\right) = g_0(t_1, t_2). \tag{29}$$

Note that expressions on the left hand of (29) are observable, $g_1(t) = g_0(t, 0)$ and $g_2(t) = g_0(0, t)$.

Let $\mathcal{L}_1$ be a set of all continuous increasing functions $\Lambda(t)$ on $\mathbb{R}_+$ satisfying $\Lambda(0) = 0$ and $\Lambda(t) \to \infty$ as $t \to \infty$. Given Laplace exponent Φ_0, consider the map A_1 defined by

$$(A_1\Lambda)(t) := \int_0^t \Phi_0(\Lambda(t) - \Lambda(\tau))d\tau. \tag{30}$$

Lemma 3 A_1 *is a bijective map from* $\mathcal{L}_1$ *to image* $A_1(\mathcal{L}_1)$ *of* $\mathcal{L}_1$ *under* A_1.

Proof It is easy to see that $(A_1\Lambda)(t)$ is a continuous monotone increasing function with $(A_1\Lambda)(0) = 0$ and $(A_1\Lambda)(t) \to \infty$ as $t \to \infty$ for $\Lambda \in \mathcal{L}_1$. That is, $A_1(\mathcal{L}_1) \subset \mathcal{L}_1$. For any $0 < T < \infty$, the set $\mathcal{L}_1^T$ of all continuous monotone increasing functions on $[0, T]$ satisfying $\Lambda(0) = 0$ is a complete metric space with usual supremum distance

$$d_T(\Lambda^a(.), \Lambda^b(.)) = \sup_{t \in [0.T]} |\Lambda^a(t) - \Lambda^b(t)| = ||f(.)||_{T,\infty}.$$

Let $g(t) = \int_0^t \Phi_0(\Lambda(t) - \Lambda(\tau))d\tau$ for $\Lambda \in \mathcal{L}_1$. Taking into account that $\Lambda(t)$ is differentiable in t a.e. and $\Phi_0(0) = 0$, we get that

$$g'(t) = \Lambda'(t) \int_0^t \Phi_0'(\Lambda(t) - \Lambda(\tau))d\tau \tag{31}$$

and, therefore,

$$\Lambda(t) = \int_0^t g'(\tau) \left(\int_0^\tau \Phi_0'(\Lambda(\tau) - \Lambda(s))ds \right)^{-1} d\tau. \tag{32}$$

In accordance with the mean value theorem applied to the integral on the right hand of (31), there exists $\tau' \in [0, \tau]$ such that

$$\tau^{-1}g'(\tau) = \Lambda'(\tau)\Phi_0'(\Lambda(\tau) - \Lambda(\tau')) \leqslant \Lambda'(\tau) \int_0^\infty x\nu(dx). \tag{33}$$

Since $\int_0^\infty x\nu(dx)$ is finite, the function on the left hand of (33) is integrable on $[0, t]$, $t < \infty$. Assume that two solutions Λ^a, Λ^b of (32) satisfy $\Lambda^a(t) = \Lambda^b(t)$ on the interval $[0, t_0]$, $t_0 \geqslant 0$. We shall show now that these solutions satisfy this property on the interval $[t_0, t_0 + t_\varepsilon]$ for some $t_\varepsilon > 0$ too. Let the map B_1 be given by

$$(B_1\Lambda)(t) = \Lambda(t_0) + \int_{t_0}^t g'(\tau) \left(\int_0^\tau \Phi_0'(\Lambda(\tau) - \Lambda(s))ds \right)^{-1} d\tau.$$

Taking into account $g'(\tau) \geqslant 0$ and inequalities (26) and (27), we get that

$$\left| \Lambda^a(t) - \Lambda^b(t) \right| = \left| (B_1\Lambda^a)(t) - (B_1\Lambda^b)(t) \right|$$

$$= \left| \int_{t_0}^t \frac{g'(\tau)d\tau}{\int_0^\tau \Phi_0'(\Lambda^a(\tau) - \Lambda^a(s))ds} - \int_{t_0}^t \frac{g'(\tau)d\tau}{\int_0^\tau \Phi_0'(\Lambda^b(\tau) - \Lambda^b(s))ds} \right|$$

$$= \left| \int_{t_0}^t \frac{g'(\tau)\left(\int_0^\tau \Phi_0'(\Lambda^b(\tau) - \Lambda^b(s))ds - \int_0^\tau \Phi_0'(\Lambda^a(\tau) - \Lambda^a(s))ds \right)d\tau}{\left(\int_0^\tau \Phi_0'(\Lambda^b(\tau) - \Lambda^b(s))ds \right)\left(\int_0^\tau \Phi_0'(\Lambda^a(\tau) - \Lambda^a(s))ds \right)} \right|$$

$$\leqslant \int_{t_0}^t \frac{g'(\tau)\left| \int_0^\tau \Phi_0'(\Lambda^b(\tau) - \Lambda^b(s))ds - \int_0^\tau \Phi_0'(\Lambda^a(\tau) - \Lambda^a(s))ds \right|d\tau}{\left| \int_0^\tau \Phi_0'(\Lambda^b(\tau) - \Lambda^b(s))ds \right| \left| \int_0^\tau \Phi_0'(\Lambda^a(\tau) - \Lambda^a(s))ds \right|}$$

$$\leqslant \int_{t_0}^t \frac{g'(\tau) \left(\int_0^\tau \int_0^\infty x \left|\exp(-(\Lambda^b(\tau) - \Lambda^b(s))x) - \exp(-(\Lambda^a(\tau) + \Lambda^a(s))x)\right| \nu(dx)ds\right) d\tau}{\left|\int_0^\tau \Phi_0'(\Lambda^b(\tau) - \Lambda^b(s))ds\right| \left|\int_0^\tau \Phi_0'(\Lambda^a(\tau) - \Lambda^a(s))ds\right|}$$

$$\leqslant \int_{t_0}^t \frac{g'(\tau) \left(\int_0^\tau \int_0^\infty x^2 ||\Lambda^a(.) - \Lambda^b(.)||_{\tau,\infty} \nu(dx)ds\right) d\tau}{\left(\tau \int_0^\infty x \exp(-x\Lambda^a(\tau))\nu(dx)\right) \left(\tau \int_0^\infty x \exp(-x\Lambda^b(\tau))\nu(dx)\right)}$$

$$\leqslant \frac{C_0 \left(\int_{t_0}^t \tau^{-1} g'(\tau)d\tau\right)}{C_a(t)C_b(t)} ||\Lambda^a(.) - \Lambda^b(.)||_{t,\infty},$$

where $0 < C_0 = \int_0^\infty x^2 \nu(dx) < \infty$, $0 < C_i(t) = \int_0^\infty x \exp(-x\Lambda^i(t))\nu(dx) < \infty$, $i = a, b$.

Let $C(t_0) = \Lambda^a(t_0) = \Lambda^b(t_0)$. Taking $t_{\varepsilon_1} > 0$ small enough, we can guarantee that $\Lambda^i(t) = (B_1\Lambda^i)(t) \leqslant 2C(t_0)$, $i = a, b$, and, therefore, $C_a(t)C_b(t) \geqslant C_3 > 0$ for all $t \in [t_0, t_0 + t_{\varepsilon_1}]$. Now we can find t_ε, $0 < t_\varepsilon \leqslant t_{\varepsilon_1}$, small enough so that

$$\left|(B_1\Lambda^a)(t) - (B_1\Lambda^b)(t)\right| \leqslant q||\Lambda^a(.) - \Lambda^b(.)||_{t_0+t_\varepsilon,\infty}$$

for all $t \in [0, t_0 + t_\varepsilon]$ and some positive $q < 1$. It means that

$$||\Lambda^a(t) - \Lambda^b(t)||_{t_0+t_\varepsilon,\infty} = ||(B_1\Lambda^a) - (B_1\Lambda^b)||_{t_0+t_\varepsilon,\infty}$$
$$\leqslant q||\Lambda^a(.) - \Lambda^b(.)||_{t_0+t_\varepsilon,\infty} < ||\Lambda^a(t) - \Lambda^b(t)||_{t_0+t_\varepsilon,\infty}.$$

From here, it follows that $||\Lambda^a(t) - \Lambda^b(t)||_{t_0+t_\varepsilon,\infty} = 0$. We continue this process until we get that $||\Lambda^a(t) - \Lambda^b(t)||_{\infty,\infty} = 0$. Hence, the map $A_1 : \mathcal{L}_1 \to A_1(\mathcal{L}_1)$ is injective. Since this map is also surjective, it is bijective $\qquad \square$

Define the set $\mathcal{L}_2$ of all functions $\Lambda_{12}(t_1, t_2, \alpha)$ such that

$$\Lambda_{12}(t_1, t_2) = \Lambda_1(t_1) + \alpha\Lambda_2(t_2)$$

for $\Lambda_1, \Lambda_2 \in \mathcal{L}_1$, and $\alpha > 0$. Given Laplace exponent Φ_0, consider the map A_2 defined by

$$(A_2\Lambda_{12})(t_1, t_2) := \int_0^{t_-} \Phi_0(\Lambda_1(\tau, t_1) + \alpha\Lambda_2(\tau, t_2))d\tau + \int_{t_-}^{t_+} \Phi_0(\Lambda_0(\tau, t_+))d\tau. \tag{34}$$

Lemma 4 *A_2 is a bijective map from $\mathcal{L}_2$ to image $A_2(\mathcal{L}_2)$ of $\mathcal{L}_2$ under A_2.*

Proof Since $\Lambda(t) = \Lambda_{12}(t, t) \in \mathcal{L}_1$, $\Lambda_1(t) = \Lambda_{12}(t, 0) \in \mathcal{L}_1$, and $\alpha\Lambda_2(t) = \Lambda_{12}(0, t) \in \mathcal{L}_1$ for $t \in [0, \infty)$, in accordance with Lemma 3, these functions can be uniquely defined by their images $g_0(t, t) = (A_2\Lambda_{1,2})(t, t) = (A_1\Lambda)(t)$, $g_0(t, 0) = (A_2\Lambda_{12})(t, 0) = (A_1\Lambda_1)(t)$, and $g_0(0, t) = (A_2(\alpha\Lambda_{12}))(0, t) = (A_1(\alpha\Lambda_2))(t)$, respectively. Assume now that $t_2 > t_1$ and the solution $\Lambda_{12}(t_1, t)$ to (34) is uniquely defined for $0 \leqslant t \leqslant t_0$ with $t_0 \geqslant t_1$. Define the map B_2 for $t > t_0$ by

$$(B_2\Lambda_{12})(t_1, t) = \Lambda_{12}(t_1, t_0)$$

$$+ \int_{t_0}^{t} \frac{dg_0(t_1, \tau)}{\int_0^{t_1} \Phi_0'(\Lambda_{12}(t_1, \tau) - \Lambda_{12}(s, s))ds + \int_{t_1}^{\tau} \Phi_0'(\Lambda_{12}(0, \tau) - \Lambda_{12}(0, s))ds}.$$

Similarly to the proof of Lemma 3, it can be shown that the solution to

$$\Lambda_{12}(t_1, t) = (B_2 \Lambda_{12})(t_1, t)$$

is also uniquely defined on the interval $[t_0, t_0 + t_\varepsilon]$ for some $t_\varepsilon > 0$ and, therefore, is uniquely defined for all $t \geqslant t_1$. The same result can be checked in the case of $t_1 > t_2$. □

For the sake of contradiction, assume now that there are functions $\Phi_0^a(.)$, $\Phi_0^b(.)$, $\Lambda_1^a(.)$, $\Lambda_1^b(.)$, $\Lambda_2^a(.)$, $\Lambda_2^b(.)$, and real positive numbers α^a and α^b such that

$$g_0(t_1, t_2) = (A_2^a \Lambda_{12}^a)(t_1, t_2) = (A_2^b \Lambda_{12}^b)(t_1, t_2) \tag{35}$$

holds for all $t_1, t_2 \geqslant 0$. In accordance with Lemmas 3–4, the maps A_i^a and A_i^b are invertible on their images of $\mathcal{L}_i$, $i = 1, 2$. Moreover, $g_0 \in A_2^a \mathcal{L}_2$, $g_0 \in A_2^b \mathcal{L}_2$, $g_i \in A_1^a \mathcal{L}_1$, and $g_i \in A_1^b \mathcal{L}_1$, $i = 1, 2$. Denote maps $(A_i^b)^{-1} A_i^a$ and $(A_i^a)^{-1} A_i^b$ by T_i^{ba} and T_i^{ab}, $i = 1, 2$, respectively. It holds that

$$\Lambda_{12}^b(t_1, t_2) = (T_2^{ba} \Lambda_{12}^a)(t_1, t_2) = ((A_2^b)^{-1} g_0)(t_1, t_2),$$
$$\Lambda_{12}^a(t_1, t_2) = (T_2^{ab} \Lambda_{12}^b)(t_1, t_2) = ((A_2^a)^{-1} g_0)(t_1, t_2),$$
$$\Lambda_i^a(t_i) = (T_1^{ab} \Lambda_i^b)(t_i) = ((A_1^a)^{-1} g_i)(t_i), i = 1, 2. \tag{36}$$

Note that $\Lambda_1^i(t_1) = \Lambda_{12}^i(t_1, 0)$ and $\alpha^i \Lambda_2^i(t_2) = \Lambda_{12}^i(0, t_2)$, $i \in \{a, b\}$. Taking into account that

$$\Lambda_1^b(t_1) = (T_1^{ba} \Lambda_1^a)(t_1), \qquad \alpha^b \Lambda_2^b(t_2) = (T_1^{ba}(\alpha^a \Lambda_2^a))(t_2),$$
$$\Lambda_1^a(t_1) = (T_1^{ab} \Lambda_1^b)(t_1), \qquad \alpha^a \Lambda_2^a(t_2) = (T_1^{ab}(\alpha^b \Lambda_2^b))(t_2) \tag{37}$$

we get from (36) that

$$\Lambda_1^b(t_1) + \alpha^b \Lambda_2^b(t_2) = (T_1^{ba} \Lambda_1^a)(t_1) + (T_1^{ba}(\alpha^a \Lambda_2^a))(t_2)$$
$$= (T_2^{ba}(\Lambda_1^a + \alpha^a \Lambda_2^a))(t_1, t_2),$$
$$\Lambda_1^a(t_1) + \alpha^a \Lambda_2^a(t_2) = (T_1^{ab} \Lambda_1^b)(t_1) + (T_1^{ab}(\alpha^b \Lambda_2^b))(t_2),$$
$$= (T_2^{ab}(\Lambda_1^b + \alpha^b \Lambda_2^b))(t_1, t_2). \tag{38}$$

Now we will prove that the action of the map T_2^{ba} (resp. T_2^{ab}) on the function Λ_{12}^a (resp. Λ_{12}^b) calculated in the point (t_1, t_2) depends only on the value of $\Lambda_{12}^b(t_1, t_2)$ (resp. $\Lambda_{12}^a(t_1, t_2)$).

Lemma 5 *Given the maps* $\Phi_0^a(.)$, $\Phi_0^b(.)$, $\Lambda_1^a(.)$, $\Lambda_1^b(.)$, $\Lambda_2^a(.)$, $\Lambda_2^b(.)$, *and real positive numbers* α^a *and* α^b, *there exist uniquely defined, monotone increasing, continuous functions* $f^{ab} : \mathbb{R}_+^1 \to \mathbb{R}_+^1$ *and* $f^{ba} : \mathbb{R}_+^1 \to \mathbb{R}_+^1$ *such that*

$$\Lambda_1^a(t_1) + \alpha^a \Lambda_2^a(t_2) = f^{ab}(\Lambda_1^b(t_1) + \alpha^b \Lambda_2^b(t_2)),$$
$$\Lambda_1^b(t_1) + \alpha^b \Lambda_2^a(t_2) = f^{ba}(\Lambda_1^a(t_1) + \alpha^a \Lambda_2^a(t_2)).$$

Proof Put $\alpha_1^a = \alpha_1^b = 1$. From (33) and Lemma 3, it follows that given the functions $g_i(t)$ the maps $(f_i^a)^{-1} : g_i(t) \to \alpha_i^a \Lambda_i^a(t)$ and $(f_i^b)^{-1} : g_i(t) \to \alpha_i^b \Lambda_i^b(t)$ are strictly monotone increasing, continuous, and uniquely defined. It holds that $g_i(t) = f_i^a(\alpha_i^a \Lambda_1^a(t)) = f_i^b(\alpha_i^b \Lambda_i^b(t))$. Therefore, the maps $f_i^{ba} = (f_i^b)^{-1} f_i^a : \mathbb{R}_+^1 \to \mathbb{R}_+^1$ and $f_i^{ab} = (f_i^a)^{-1} f_i^b : \mathbb{R}_+^1 \to \mathbb{R}_+^1$, $i = 1, 2$, are also strictly monotone increasing and continuous. Since $(T_i^{ba}(\alpha_i^a \Lambda_i^a))(t_i) = f_i^{ba}(\alpha_i^a \Lambda_i^a(t_i))$ and $(T_i^{ab}(\alpha_i^b \Lambda_i^b))(t_i) = f_i^{ab}(\alpha_i^b \Lambda_i^b(t_i))$ for $i = 1, 2$, the middle terms of Eq. (38) depend only on the values of $\alpha_i^b \Lambda_i^b(t_i)$ and $\alpha_i^a \Lambda_i^a(t_i)$ and do not depend on the behavior of the functions $\Lambda_i^a(t_i)$ and $\Lambda_i^b(t_i)$ on the intervals $[0, t_i)$. Therefore, we can rewrite equations in (38) as follows:

$$\begin{aligned}
\Lambda_1^b(t_1) + \alpha^b \Lambda_2^b(t_2) &= f_1^{ba}(\Lambda_1^a(t_1)) + f_2^{ba}(\alpha^a \Lambda_2^a(t_2)) \\
&= f^{ba}(\Lambda_1^a(t_1) + \alpha^a \Lambda_2^a(t_2)), \\
\Lambda_1^a(t_1) + \alpha^a \Lambda_2^a(t_2) &= f_1^{ab}(\Lambda_1^b(t_1)) + f_2^{ab}(\alpha^b \Lambda_2^b(t_2)), \\
&= f^{ab}(\Lambda_1^b(t_1) + \alpha^b \Lambda_2^b(t_2)).
\end{aligned} \tag{39}$$

for some continuous monotone increasing functions $f^{ba} : \mathbb{R}_+^1 \to \mathbb{R}_+^1$ and $f^{ab} : \mathbb{R}_+^1 \to \mathbb{R}_+^1$. Substituting in (39) $t_2 = 0$ or $t_1 = 0$, we get that $f_1^{ab} = f_2^{ab} = f^{ab}$ and $f_1^{ba} = f_2^{ba} = f^{ba}$. □

Proof of Proposition 2 Note firstly that in accordance with Assumption 4, the functions $\Phi_i(c)$, $i = 0, 1, 2$, are analytic ones on $\text{Re}(c) > -b$. Since $f^{ba}(.)$ and $f^{ab}(.)$ are defined on $\mathbb{R}_+^1$, additive and continuous, they are linear and it holds that $f^{ab}(x) = C_1^{ab} x + C_2^{ab}$ and $f^{ba}(x) = C_1^{ba} x + C_2^{ba}$ for some constant C_1^{ab}, C_2^{ab}, C_1^{ba}, and C_2^{ba}. As $f^{ab}(0) = f^{ba}(0) = 0$, we have that $C_2^{ab} = C_2^{ba} = 0$. Moreover, $C_1^{ab} = C_1^{ba} = 1$ and $\alpha^a = \alpha^b$ because $\Lambda_i^a(t^*) = \Lambda_i^b(t^*) = 1$ for some $t^* > 0$ and $i = 1, 2$. From here, it follows that $f^{ab}(.)$ and $f^{ba}(.)$ are identity transformations and $\Lambda_i^a(t) = \Lambda_i^b(t) = \Lambda_i(t)$ for all $t \geqslant 0$, $i = 1, 2$. It holds also that $(A_1^a \Lambda_i)(t) = (A_1^b \Lambda_i)(t)$ for all $t \geqslant 0$, $i = 1, 2$. To complete the proof of Proposition 2, we need to show that $\Phi_0^a(x) = \Phi_0^b(x)$ for any $x \in \mathbb{R}_+^1$. Indeed, the maps Φ_0^a

and Φ_0^b are real analytic functions on $\mathbb{R}_+^1$ and, therefore,

$$
\begin{aligned}
(A_1^a \Lambda_1)(t) &= \int_0^t \Phi_0^a(\Lambda_1(t) - \Lambda_1(\tau)) d\tau \\
&= \sum_{n=0}^{\infty} \frac{(\Phi_0^a)^{(n)}(0)}{n!} \int_0^t (\Lambda_1(t) - \Lambda_1(\tau))^n d\tau \\
(A_1^b \Lambda_1)(t) &= \int_0^t \Phi_0^b(\Lambda_1(t) - \Lambda_1(\tau)) d\tau \\
&= \sum_{n=0}^{\infty} \frac{(\Phi_0^b)^{(n)}(0)}{n!} \int_0^t (\Lambda_1(t) - \Lambda_1(\tau))^n d\tau
\end{aligned}
$$

for small values of $\Lambda_1(t)$. This holds iff $(\Phi_0^a)^{(n)}(0) = (\Phi_0^b)^{(n)}(0)$ for all nonnegative integer n and means that $\Phi_0^a(x) = \Phi_0^b(x) = \Phi_0(x)$ in some neighborhood of 0. Since $\Phi_0(x)$ is a real analytic function on $\mathbb{R}_+^1$, this function is uniquely defined in this area. Similarly, it can be proved that the functions $\Phi_i(x)$, $i = 1, 2$, are also uniquely defined in $\mathbb{R}_+^1$. $\qquad\square$

Appendix 4: Calculation of the conditional expectation of the log-density function

Case $t_2 \geqslant t_1$.
Let all failure and censoring times $t^{(k)}$, $0 \leqslant t^{(k)} \leqslant t_2$, $k = 1, \ldots, N_2$, be arranged in ascending order so that $0 = t^{(0)} \leqslant t^{(1)} \leqslant t^{(2)} \leqslant \cdots \leqslant t^{(N_2)} = t_2$ and $t_1 = t^{(N_1)}$, $N_1 \leqslant N_2$. The estimate of the cumulative hazard function $\hat{\Lambda}(t)$ is a step function with $\hat{\Lambda}(0_-) = 0$ and jumps (if non-censored) at times $t^{(k)}$, $k = 1, \ldots, N_2$. Since $Y_j(t)$, $j = 0, 1, 2$, are gamma processes, it holds that

$$
Y_j(t_l) = \sum_{k=1}^{N_{l_j}} \left(Y_j(t^{(k)}) - Y_j(t^{(k-1)}) \right), \quad j = 0, 1, 2, \ l_0 = l_2 = 2, \ l_1 = 1, \quad (40)
$$

where increments $\Delta Y_j^{(k-1)} = Y_j(t^{(k)}) - Y_j(t^{(k-1)})$ are independent gamma-distributed random variables with means $h_j \gamma^{-1} \Delta t^{(k-1)}$ and the variances $h_j \gamma^{-2} \Delta t^{(k-1)}$ for $\Delta t^{(k-1)} = t^{(k)} - t^{(k-1)}$, $h_1 = h_2 = (1 - \varrho)\gamma$ and $h_0 = \varrho\gamma$, $j = 0, 1, 2$, $k = 1, \ldots, N_{l_j}$.

Using decomposition (23) with independent gamma processes $Y_j(t)$, $j = 0, 1, 2$, we calculate the log-density function as follows:

$$
\begin{aligned}
&\ln(f(Z_1(.), Z_2(.)|t_1, t_2, \zeta)) \\
&= -\gamma \left(\sum_{k=0}^{N_1-1} \Delta Y_1^{(k)} + \sum_{k=0}^{N_2-1} \Delta Y_0^{(k)} + \sum_{k=0}^{N_2-1} \Delta Y_2^{(k)} \right) \\
&\quad - \sum_{k=0}^{N_1-1} \ln \left(\Gamma(\gamma(1-\varrho)\Delta t^{(k)}) \right) - \sum_{k=0}^{N_2-1} \ln \left(\Gamma(\gamma\varrho\Delta t^{(k)})\Gamma(\gamma(1-\varrho)\Delta t^{(k)}) \right) \\
&\quad + \gamma(1-\varrho) \left(\sum_{k=0}^{N_1-1} \Delta t^{(k)} \ln(\Delta Y_1^{(k)}) + \sum_{k=0}^{N_2-1} \Delta t^{(k)} \ln(\Delta Y_2^{(k)}) \right) \\
&\quad + \gamma\varrho \sum_{k=0}^{N_2-1} \Delta t^{(k)} \ln(\Delta Y_0^{(k)}) - \sum_{k=0}^{N_1-1} \ln(\Delta Y_1^{(k)}) - \sum_{k=0}^{N_2-1} \ln(\Delta Y_0^{(k)}) \\
&\quad - \sum_{k=0}^{N_2-1} \ln(\Delta Y_2^{(k)})
\end{aligned}
\tag{41}
$$

(index i denoting the twin pair is here omitted).

Given current estimates $\hat{\beta}$, $\hat{\Lambda}(t)$, $\hat{\zeta}$, and the observed data, the conditional expectation of the log-density function (41) (taking into account only summands depending on γ and ϱ) is equal to

$$
\begin{aligned}
&\mathbb{E} \ln(f(Z_1(.), Z_2(.)|t_1, t_2, \hat{\beta}, \hat{\Lambda}(t), \hat{\zeta})) \\
&= -\gamma \left(\sum_{k=0}^{N_1-1} \widehat{\Delta Y_1^{(k)}} + \sum_{k=0}^{N_2-1} \widehat{\Delta Y_0^{(k)}} + \sum_{k=0}^{N_2-1} \widehat{\Delta Y_2^{(k)}} \right) \\
&\quad - \sum_{k=0}^{N_1-1} \ln \left(\Gamma(\gamma(1-\varrho)\Delta t^{(k)}) \right) - \sum_{k=0}^{N_2-1} \ln \left(\Gamma(\gamma\varrho\Delta t^{(k)})\Gamma(\gamma(1-\varrho)\Delta t^{(k)}) \right) \\
&\quad + \gamma(1-\varrho) \left(\sum_{k=0}^{N_1-1} \Delta t^{(k)} \widehat{\ln(\Delta Y_1^{(k)})} + \sum_{k=0}^{N_2-1} \Delta t^{(k)} \widehat{\ln(\Delta Y_2^{(k)})} \right) \\
&\quad + \gamma\varrho \sum_{k=0}^{N_2-1} \Delta t^{(k)} \widehat{\ln(\Delta Y_0^{(k)})}.
\end{aligned}
\tag{42}
$$

Here, $\widehat{\Delta Y_i^k}$ and $\widehat{\ln(\Delta Y_i^k)}$ are conditional expectations of ΔY_i^k and $\ln(\Delta Y_i^k)$ given current estimates $\hat{\beta}$, $\hat{\Lambda}(t)$, $\hat{\zeta}$, and the observed data.

From formulas (21), (23), (24), and (40) after a number of transformations, we get for $\chi(u) = e^{\beta \mathbf{u}}$ that

$$
\mathbb{E} \exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^* \mathbf{u}}\lambda(\tau)d\tau\right) = e^{-E_j^0}
$$

$$
\mathbb{E}\Delta Y_j^{(k)} \exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^* \mathbf{u}}\lambda(\tau)d\tau\right) = \Delta t^{(k)} E_j^{1,k} e^{-E_j^0}
$$

$$
\mathbb{E}\left(\Delta Y_j^{(k)}\right)^2 \exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^* \mathbf{u}}\lambda(\tau)d\tau\right)
$$

$$
= \left(\left(\Delta t^{(k)} E_j^{1,k}\right)^2 - \Delta t^{(k)} E_j^{2,k}\right) e^{-E_j^0} \tag{43}
$$

$$
\mathbb{E}\left(\Delta Y_j^{(k)}\right)^3 \exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^* \mathbf{u}}\lambda(\tau)d\tau\right)
$$

$$
= \left(\left(\Delta t^{(k)} E_j^{1,k}\right)^3 - 3\left(\Delta t^{(k)}\right)^2 E_j^{1,k} E_j^{2,k} + \Delta t^{(k)} E_j^{3,k}\right) e^{-E_j^0}, \qquad j = 0, 1, 2,
$$

where

$$
E_j^0 = \sum_{k=0}^{N_{l_j}-1} \Delta t^{(k)} \Phi_G^{(0)}(0, t_{l_j}; \Lambda, \beta, \mathbf{u})
$$

$$
E_j^{m,k} = \Phi_G^{(m)}(t^{(k)}, t_{l_j}; \Lambda, \beta, \mathbf{u}), \qquad\qquad k = 0, \ldots, N_{l_j} - 1,
$$

$$
\Phi_G^{(m)}(\tau, t; \Lambda, \beta, \mathbf{u}) = \Phi_G^{(m)}\left(e^{\beta^* \mathbf{u}} \Lambda(\tau, t)\right), \qquad m = 1, 2, 3,
$$

and incomplete cumulative hazard function $\Lambda(\tau, t) = \int_\tau^t \lambda(s)ds$ is a step function, and $\Phi_G^{(m)}(.)$ stands for mth-order argument derivative.

For a gamma-distributed random variable Y with shape parameter h and the scale parameter γ, it holds that

$$
\mathbb{E} \ln(Y)e^{-cY} = \left[\psi(h) - \ln(\gamma + c)\right]e^{-\Phi_G^{(0)}(c)}
$$

$$
\mathbb{E} Y \ln(Y)e^{-cY} = \left[\psi(h + 1) - \ln(\gamma + c)\right]\Phi_G^{(1)}(c)e^{-\Phi_G^{(0)}(c)} \tag{44}
$$

$$
\mathbb{E} Y^2 \ln(Y)e^{-cY} = \left[\psi(h + 2) - \ln(\gamma + c)\right]\left[\Phi_G^{(1)}(c)^2 - \Phi_G^{(2)}(c)\right]e^{-\Phi_G^{(0)}(c)},
$$

where $c \geqslant 0$ and $\psi(.)$ is the digamma function defined by $\psi(x) = d \ln(\Gamma(x))/dx$. Last formulas can be obtained by differentiating the integral representation of the gamma function and not difficult transformations after that.

Using formulas (44), we get additionally to formulas (43)

$$
\mathbb{E}\ln(\Delta Y_j^{(k)})\exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^*\mathbf{u}}\lambda(\tau)d\tau\right) = F_j^{0,k}e^{-E_j^0}
$$

$$
\mathbb{E}\Delta Y_j^{(k)}\ln(\Delta Y_j^{(k)})\exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^*\mathbf{u}}\lambda(\tau)d\tau\right) = \Delta t^{(k)}E_j^{1,k}F_j^{1,k}e^{-E_j^0}
$$

$$
\mathbb{E}\left(\Delta Y_j^{(k)}\right)^2\ln(\Delta Y_j^{(k)})\exp\left(-\int_0^{t_{l_j}} Y_j(\tau)e^{\beta^*\mathbf{u}}\lambda(\tau)d\tau\right) \tag{45}
$$

$$
= \left(\left(\Delta t^{(k)}E_j^{1,k}\right)^2 - \Delta t^{(k)}E_j^{2,k}\right)F_j^{2,k}e^{-E_j^0},
$$

where

$$
F_j^{m,k} = \psi\left(h_j\Delta t^{(k)} + m\right) - \ln\left(\gamma + e^{\beta^*\mathbf{u}}\Lambda(t^{(k)}, t^{l_j})\right)
$$

for $m = 0, 1, 2$, $k = 0, \ldots, N_{l_j}$, and $j = 0, 1, 2$.

In accordance with Bayes' theorem, we get for the conditional density function that

$$
f\left(Z_1(.), Z_2(.)|\beta, \mathbf{u}_1, \mathbf{u}_2, \Lambda(.), \zeta, t_1, t_2, \delta_1, \delta_2\right)
$$

$$
= \frac{g\left(Z_1(.), Z_2(.), t_1, t_2|\beta, \mathbf{u}_1, \mathbf{u}_2, \Lambda(.), \zeta, \delta_1, \delta_2\right)}{\mathbb{E}g\left(Z_1(.), Z_2(.), t_1, t_2|\beta, \mathbf{u}_1, \mathbf{u}_2, \Lambda(.), \zeta, \delta_1, \delta_2\right)} \tag{46}
$$

with

$$
g\left(Z_1(.).Z_2(.)t_1, t_2|\beta, \mathbf{u}_1, \mathbf{u}_2, \Lambda(.), \zeta, \delta_1, \delta_2\right)
$$

$$
= \left(Z_1(t_1)e^{\beta^*\mathbf{u}_1}\lambda(t_1)\right)^{\delta_1}\left(Z_2(t_2)e^{\beta^*\mathbf{u}_2}\lambda(t_2)\right)^{\delta_2} f(Z_1(.), Z_2(.)|\zeta)
$$

$$
\times \exp\left(-\int_0^{t_1} Z_1(\tau)e^{\beta^*\mathbf{u}_1}\lambda(\tau)d\tau - \int_0^{t_2} Z_2(\tau)e^{\beta^*\mathbf{u}_2}\lambda(\tau)d\tau\right).
$$

Expression (46) can be calculated using Eq. (43) and the formulas

$$
Z_1(t_1) = \sum_{k=0}^{N_1-1}\left(Y_0^{(k)} + Y_1^{(k)}\right),
$$

$$
Z_2(t_1) = \sum_{k=0}^{N_2-1}\left(Y_0^{(k)} + Y_2^{(k)}\right),
$$

$$
\int_0^{t_1} Z_1(\tau)e^{\beta^*\mathbf{u}_1}\lambda(\tau)d\tau = \sum_{k=0}^{N_1-1} e^{\beta^*\mathbf{u}_1}\Lambda(t^{(k)}, t_1)\left(\Delta Y_0^{(k)} + \Delta Y_1^{(k)}\right),
$$

$$
\int_0^{t_2} Z_2(\tau)e^{\beta^*\mathbf{u}_2}\lambda(\tau)d\tau = \sum_{k=0}^{N_2-1} e^{\beta^*\mathbf{u}_2}\Lambda(t^{(k)}, t_2)\left(\Delta Y_0^{(k)} + \Delta Y_1^{(k)}\right),
$$

$$f(Z_1(.), Z_2(.)|\zeta) = \prod_{k=0}^{N_1-1} f(\Delta Y_1^{(k)}|\zeta) \prod_{k=0}^{N_2-1} f(\Delta Y_0^{(k)}|\zeta) \prod_{k=0}^{N_2-1} f(\Delta Y_2^{(k)}|\zeta).$$

Finally, we calculate the conditional log-density function (42) using formulas (43), (45), (46), and replacing β, ζ, and Λ with their current estimates $\hat{\beta}$, $\hat{\zeta}$, and $\hat{\Lambda}$.

References

Abbring, J.H., van den Berg, G.J.: The identifiability of the mixed proportional hazards competing risks model. J. R. Stat. Soc. B **65**(3), 701–710 (2003)

Ata, N., Özel, G.: Survival functions for the frailty models based on the discrete compound Poisson process. J. Stat. Comput. Simul. **83**(11), 2105–2116 (2012)

Breslow, N.E.: Discussion of the paper by D. R. Cox. J. R. Stat. Soc. B **34**, 216–217 (1972)

Brychkov, Y.A., Glaeske, H.-J., Prudnikov, A.P., Tuan, V.K.: Multidimensional Integral Transformations. Gordon and Breach Science Publishers, London (1992)

Elbers, C., Ridder, G.: True and spurious duration dependence: the identifiability of the proportional hazard model. Rev. Econ. Stud. **49**, 403–409 (1982)

Gjessing, H.K., Aalen, O.O., Hjort, N.L.: Frailty models based on Lévy processes. Adv. Appl. Probab. **35**, 532–550 (2003)

Gorfine, M., Hsu, L.: Frailty-based competing risks model for multivariate survival data. Biometrics **67**, 415–426 (2011)

Hauge, M.: The Danish twin register. In: Mednik, S.A., Baert, A.E., Backmann, B.P. (eds.) Prospective Longitudinal Research, an Empirical Basis for the Primary Prevention of Physiological Disorders, pp. 218–221. Oxford University Press, London (1981)

Honoré, B.E.: Identification results for duration models with multiple spells. Rev. Econ. Stud. **60**, 241–246 (1993)

Iachine, I. A.: Parameter estimation in the bivariate correlated frailty model with observed covariates via EM-Algorithm. Working Paper Series: Population Studies of Aging 16, CHS, Odense University (1995)

Wienke, A.: Frailty Models in Survival Analysis. Chapman & Hall, Boca Raton (2010)

Woodbury, M.A., Manton, K.G.: A random walk model of human mortality and aging. Theor. Popul. Biol. **11**, 37–48 (1997)

Yashin, A.I., Vaupel, J.W., Iachine, I.A.: Correlated individual frailty: an advantageous approach to survival analysis of bivariate data. Math. Popul. Stud. **5**, 145–159 (1995)

Yashin, A.I., Manton, K.G.: Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. Stat. Sci. **12**(1), 20–34 (1997)

Yashin, A.I., Iachine, I.A.: What difference does the dependence between durations make? Insights for population studies of aging. Lifetime Data Anal. **5**, 5–22 (1999)

Yashin, A.I., Arbeev, K.G., Akushevich, A., Kulminski, A., Ukraintseva, S.V., Stallard, E., Land, K.C.: The quadratic hazard model for analyzing longitudinal data on aging, health, and the life span. Phys. Life. Rev. **9**(2), 177–188 (2012)

Zeng, D., Lin, D.Y.: Maximum likelihood estimation in semiparametric regression models with censored data. J. R. Stat. Soc. B **69**, 507–564 (2007)