Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 69 Issue 03 | 2026
DOI: 10.5281/zenodo.19252742

A QUASI–MONTE CARLO EXPECTATION–MAXIMIZATION FRAMEWORK FOR MULTIVARIATE JOINT MODELLING WITH PENALIZED COX REGRESSION

AZEEZ ADEBOYE*
GIT Research Unit, Faculty of Health Science, University of the Free State, Bloemfontein, Free State, South Africa. *Corresponding Author Email: azeez.an@ufs.ac.za

COLIN NOEL
GIT Research Unit, Department of Surgery, Faculty of Health Science, University of the Free State, Bloemfontein, Free State, South Africa.

Abstract

The joint modelling of longitudinal and time-to-event data provides an effective framework for analysing the association between biomarker trajectories and survival outcomes. However, when multiple longitudinal processes are modelled simultaneously, the resulting high-dimensional random effects structure leads to substantial computational challenges due to complex likelihood integrals. This paper proposes an efficient estimation framework for multivariate joint models using Quasi Monte Carlo (QMC) integration within a Monte Carlo Expectation Maximization (MCEM) algorithm. The approach employs deterministic low-discrepancy sequences, specifically Sobol and Halton sequences to approximate the intractable integrals in the E-step, thereby improving convergence speed and estimation accuracy compared to classical Monte Carlo (MC) methods. To enhance model flexibility and stability in the survival component, a ridge-penalized Cox regression is incorporated into the joint modelling framework. The penalty term regularizes the parameter estimates, mitigates overfitting, and improves generalization in the presence of correlated covariates or high-dimensional predictors. Extensive simulation studies demonstrate that QMC methods yield comparable or superior estimation accuracy relative to standard MC integration, with substantially reduced computation time. Among the QMC variants, the Halton sequence exhibits the most stable convergence and lowest estimation variance, particularly in small-sample and high-dimensional scenarios. The proposed approach is further illustrated using PBC dataset, where the multivariate joint model simultaneously links two or three longitudinal biomarkers with survival outcomes. The empirical results confirm that QMC integration significantly accelerates computation while maintaining robust and precise parameter estimates. This study highlights the potential of QMC-based numerical integration as a practical and efficient alternative for fitting high-dimensional joint models in biomedical applications..

Keywords: Joint modelling; Multivariate longitudinal data; Penalized Cox regression; Ridge penalty; Monte Carlo EM; Quasi–Monte Carlo integration.

1. INTRODUCTION

The joint modelling of longitudinal measurements and time-to-event outcomes has become a central topic in modern biostatistics and applied statistical methodology. Over the past two decades, substantial progress has been made in developing joint models that capture the association between repeated biomarker trajectories and survival processes [1–5]. Early work primarily focused on univariate joint models that link a single longitudinal covariate to a time-to-event outcome [4]. However, in many clinical studies, multiple biomarkers are repeatedly recorded, reflecting different aspects of disease progression. This motivates the development of multivariate joint models, which simultaneously account for multiple correlated longitudinal outcomes and survival events [6–9]. Modelling longitudinal and survival outcomes separately can be inefficient and may produce biased parameter estimates when the processes are correlated [3]. Joint models overcome this issue by incorporating shared random effects that link the longitudinal and survival submodels. Nonetheless, when several biomarkers are analysed jointly, the computational burden increases considerably due to the higher-dimensional random-effects structure. Each additional longitudinal outcome adds new random effects, complicating the integration over the random-effects distribution required to evaluate the joint likelihood [10–12]. Consequently, efficient numerical integration strategies are essential for the practical implementation of multivariate joint models.

A major challenge in this context lies in the evaluation of high-dimensional integrals. Several approaches have been proposed to approximate these integrals, including Bayesian smoothing spline methods [13]. Gaussian quadrature (GQ) and its adaptive [5] and pseudo-adaptive [14] variants; Laplace approximations [15], and Quasi-Monte Carlo (QMC) methods using quasi-random sequences [16]. Each of these approaches has distinct advantages and limitations. Smoothing spline and Bayesian approaches offer flexibility but can be computationally intensive and sensitive to model assumptions. Gaussian quadrature achieves high accuracy in low dimensions, but its performance deteriorates exponentially as dimensionality increases. Laplace approximations are efficient for high-dimensional data but can be inaccurate when subjects contribute few longitudinal measurements. Monte Carlo methods maintain a convergence rate independent of dimensionality, though convergence can be slow due to random sampling variability [17]. In contrast, Quasi-Monte Carlo (QMC) integration offers a promising alternative by using deterministic, low-discrepancy sequences (e.g., Sobol and Halton sequences) instead of pseudo-random samples. QMC methods produce sample points that cover the integration domain more uniformly, thereby improving convergence rates and reducing variance compared to standard MC methods [16]. Theoretically, QMC achieves convergence rates approaching $O(N^{-1}(\log N)^{d})$, outperformed the $O(N^{-1/2})$ rate typical of Monte Carlo. These advantages are particularly relevant for joint models, where the dimensionality of the random-effects space can be large. However, QMC methods lack built-in error estimation via random sampling, which can limit their practical implementation in some settings [18–21].

In this paper, we propose an efficient estimation framework for multivariate joint models of longitudinal and survival data using Quasi-Monte Carlo integration combined with a penalized Cox regression survival submodel. Specifically, we introduce QMC methods (Sobol and Halton sequences) into the Monte Carlo Expectation-Maximization (MCEM) algorithm for multivariate joint models, incorporate a ridge penalty into the Cox model to regularize estimation, handle high-dimensional or correlated covariates, and improve numerical stability, and systematically evaluate the computational efficiency and estimation accuracy of QMC methods relative to conventional Monte Carlo integration. Our framework assumes a linear mixed-effects structure for the longitudinal submodels and employs a penalized Cox regression for the event process, estimated via an iterative MCEM algorithm that alternates between expectation (E) and maximization (M) steps until convergence. To assess performance, we conduct extensive simulation studies in R, extended to accommodate the penalized Cox component and QMC integration. Finally, we demonstrate the applicability of the proposed methods through an analysis of the Mayo Clinic primary biliary cirrhosis (PBC) dataset. The remainder of this paper is organized as follows. Section 2 presents the methodological framework, including model formulation, penalized estimation, and QMC integration strategies, and the simulation design and performance evaluation. Section 4 applies the proposed model to the PBC dataset. Section 5 discusses the findings, limitations, and future research directions.

2. METHODS

2.1 Model Framework

We consider a joint modelling framework for multivariate longitudinal data and time-to-event outcomes, with a penalized Cox regression component to stabilize estimation and mitigate overfitting in the survival submodel. For subject $i = 1, .., n$, let

$$y_i = (y_{i1}^\top, \dots., y_{iK}^\top)^\top$$

represents a K-variate longitudinal outcome vector, where

$$y_{ik} = \left(y_{i1k}, \dots, y_{in_{ik}k}\right)^\top$$

represents the repeated measurements for the kth outcome observed at times $t_{ijk}, j = 1, .., n_{ik}$. Each subject also has a true event time T_i^* and a censoring time C_i. The observed event time is $T_i = \min(T_i^*, C_i)$, with the event indicator $\delta_i = I(T_i^* \le C_i)$. We assume noninformative and independent censoring.

This formulation extends standard joint models by Henderson et al. (2000) and Hickey et al. (2018) to the multivariate longitudinal setting, and introduces a regularization framework for the survival submodel. Although Philipson et al. (2020) introduced Quasi-Monte Carlo (QMC) integration within joint models, their work focused on univariate longitudinal outcomes and unpenalized Cox submodels. Our work generalizes this to multivariate longitudinal data and systematically evaluates the computational and inferential effects of QMC methods under penalized estimation.

2.2 Longitudinal Submodel

For each outcome $k = 1, .., K$, the longitudinal measurements are modelled as

$$y_{ik}(t) = \mu_{ik}(t) + \eta_{ik}(t) + \varepsilon_{ik}(t)$$

where, $\mu_{ik}(t) = \boldsymbol{x}_{ik}^\top(t)\boldsymbol{\beta}_k$ represents the fixed-effects component,

$$\eta_{ik}(t) = \boldsymbol{z}_{ik}^\top(t)\mathbf{b}_{ik}$$

is the subject-specific random-effects process, and $\varepsilon_{ik}(t) \sim \mathcal{N}(0, \sigma_k^2)$ are independent errors. To capture correlations among multiple outcomes within the same subject, we define

$$b_i = \left(\mathbf{b}_{i1}^{\top}, \dots \mathbf{b}_{iK}^{\top}\right)^{\top} \sim \mathcal{N}(0, \mathbf{D})$$

where the covariance matrix **D** includes both within- and between-outcome covariances, that is, $\mathrm{Cov}(\mathbf{b}_{ik}, \mathbf{b}_{ih}) = \mathbf{D}_{kh}$. This specification allows outcome specific latent processes to be correlated, accommodating shared unobserved influences across longitudinal trajectories.

2.3 Survival Submodel

The time-to-event process is modelled using a proportional hazards model of the form

$$h_i(t) = h_0(t)\exp\{\mathbf{w}_i^{\top}\gamma_{\varpi} + \eta_{2i}(t)\},$$

where $h_0(t)$ is an unspecified baseline hazard, $\mathbf{w}_i$ is a vector of baseline covariates, and γ_{ϖ} are the corresponding regression coefficients.

The term $\eta_{2i}(t)$ captures the dynamic association between the longitudinal and survival processes, defined as

$$\eta_{2i}(t) = \sum_{k=1}^{K} \gamma_{yk}\eta_{ik}(t),$$

with $\gamma_y = \left(\gamma_{y1}, \dots, \gamma_{yK}\right)^{\top}$ quantifying the contribution of each longitudinal outcome to the hazard. This flexible structure allows the event process to be influenced by multiple time-varying trajectories simultaneously.

2.4 Penalization of the Survival Submodel

To improve numerical stability and address possible multicollinearity or overparameterization in the survival model, we introduce an L_2 (ridge) penalty on the baseline covariate effects γ_{ϖ}. The penalized log-likelihood for the survival component is

$$\ell_{\text{surv}}^{\text{pen}} = \sum_{i=1}^{n}\left[\delta_i \log h_i(T_i) - \int_0^{T_i} h_i(\mu)\,d\mu\right] - \theta\|\gamma_{\varpi}\|_{2,}^2$$

where $\theta > 0$ controls the amount of shrinkage. Ridge regularization stabilizes estimation in high-dimensional or correlated settings and ensures smoother convergence in the EM algorithm.

Unlike Philipson et al. (2020), who used unpenalized estimation, our framework explicitly integrates penalization into the joint likelihood to evaluate its impact on bias and coverage

2.5 Joint Likelihood and Estimation

Let $\mathbf{X}_i$ and $\mathbf{Z}_i$ denote block-diagonal matrices of covariates for subject i, corresponding to the longitudinal design matrices. The conditional densities are given by

$$f\left(\mathbf{y}_i|\mathbf{b}_i;\boldsymbol{\theta}\right)=\prod_{k=1}^{K}\mathcal{N}\left(\mathbf{y}_{ik};\mathbf{X}_{ik}\boldsymbol{\beta}_k+\mathbf{Z}_{ik}\mathbf{b}_{ik},\sigma_k^2 I\right),$$

and

$$f(T_i,\delta_i|\mathbf{b}_i;\boldsymbol{\theta})=[h_i(T_i)]^{\delta_i}\text{exp}\left(-\int_0^{T_i}h_i(\mu)d\mu\right).$$

The marginal likelihood for all subjects is then

$$\mathcal{L}(\theta)=\prod_{i=1}^{n}\int f\left(\mathbf{y}_i|\mathbf{b}_i;\boldsymbol{\theta}\right)f(T_i,\delta_i|\mathbf{b}_i;\boldsymbol{\theta})f(\mathbf{b}_i;\boldsymbol{\theta})d\mathbf{b}_i$$

where $\boldsymbol{\theta}=\left(\beta,\text{vech}(\mathbf{D}),\gamma_{\varpi},\gamma_y,\sigma_1^2,\dots,\sigma_k^2,h_0(t)\right)$ denotes all model parameters.

2.6 Expectation–Maximization (EM) Estimation

Estimation proceeds via a penalized EM algorithm, iteratively maximizing the expected complete-data log-likelihood:

$$Q\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(\boldsymbol{m})}\right)=\mathbb{E}_{\mathbf{b}_i|\mathbf{y}_{i,}T_i,\delta_i;\boldsymbol{\theta}^{(m)}}\left[\text{log}f\left(\mathbf{y}_i,T_i,\delta_i,\mathbf{b}_i;\boldsymbol{\theta}\right)\right].$$

In the E-step, expectations with respect to the posterior distribution $f\left(\mathbf{b}_i|\mathbf{y}_i,T_i,\delta_i,;\boldsymbol{\theta}^{(\boldsymbol{m})}\right)$ are computed, while in the M-step, parameter updates are obtained by maximizing $Q\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(\boldsymbol{m})}\right)$ penalized by $\theta\|\gamma_{\varpi}\|_2^2$. To improve numerical integration over the random effects, the intractable integrals in the E-step are approximated using Monte Carlo and Quasi-Monte Carlo (QMC) methods.

2.7 Monte Carlo and Quasi-Monte Carlo Integration

Let

$$I[f]=\int\limits_{[0,1]^d}f(\chi)dx$$

denote an expectation over a *d*-dimensional unit cube, where f is a Lebesgue-integrable function. In the joint likelihood context, $f(\chi)$ corresponds to integrands involving random effects distributions.

(a) Vanilla Monte Carlo (MC):

A standard MC estimate draws $\{\chi_n\}_{n=1}^{N}\sim U(0,1)^d$ and computes

$$I_N^{\text{MC}}[f] = \frac{1}{N}\sum_{n=1}^{N} f(\chi_n),$$

with error $O\left(N^{-1/2}\right)$. However, convergence may be slow for high-dimensional or complex integrands.

(b) Quasi-Monte Carlo (QMC):

To improve efficiency, we replace random sampling by deterministic low-discrepancy sequences that cover the integration domain more uniformly:

- **Sobol Sequence**: A base-2 digital sequence ensuring uniform coverage of $[0,1]^d$. A scrambled variant [22] further reduces bias and improves convergence:

$$I_N^{\text{Sobol}}[f] = \frac{1}{N}\sum_{n=1}^{N} f(S(\chi_n)),$$

where $S(.)$ denotes the scrambling transformation.

- **Halton Sequence:** A generalization of the van der Corput sequence using coprime bases for each dimension:

$$I_N^{\text{Halton}}[f] = \frac{1}{N}\sum_{n=1}^{N} f(\chi_n)$$

which achieves faster convergence $(O(N^{-1}(\log N)^d))$ under smoothness conditions.

By embedding QMC integration within the EM algorithm, we improve numerical stability and reduce variance in likelihood approximation, particularly as the dimension of random effects increases. This contrasts with previous joint model implementations, which primarily relied on Gaussian quadrature or unsystematic MC integration.

2.8 Simulation Design

To evaluate the performance of the proposed penalized joint model under different integration schemes, we conducted simulation studies with $K = 2$ (bivariate) and $K = 3$ (trivariate) longitudinal outcomes and sample sizes $n \in \{250, 500, 800, 1000\}$.

Longitudinal outcomes were generated according to

$$y_{ijk} = \left(\beta_{0,k} + b_{i0k}\right) + \left(\beta_{1,k} + b_{i1k}\right)t_j + \beta_{2,k}x_{i1} + \beta_{3,k}x_{i2} + \varepsilon_{ijk}$$

where $x_{i1} \sim \mathcal{N}(0,1)$, $x_{i2} \sim \text{Bernoulli}(0.5)$, and $\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma_k^2)$.

The event times followed

$$h_i(t) = h_0(t)\exp\left(\gamma_v^{\top}\boldsymbol{x}_i + \sum_{k=1}^{K}\gamma_{yk}\left(b_{i0k} + b_{i1k}t\right)\right),$$

approximately 40% observed events. The ridge penalty parameter θ was selected adaptively based on model dimension ($\mathrm{df} = p/2$). Models were fitted using the joineRML R package [9], employing MC, Sobol QMC, and Halton QMC integration schemes. Performance was assessed in terms of bias, coverage, and computational efficiency. However, this methodological framework extends the current literature in several ways: extend the generalization of QMC-based joint models to multivariate longitudinal processes; integrate ridge penalization directly into the joint likelihood, allowing systematic control of overfitting in the survival submodel and regularised estimation; empirical and theoretical comparison of multiple QMC sequences (Sobol, Halton) against standard MC within the same estimation algorithm; and demonstration that QMC integration reduces variability in likelihood estimation and improves convergence behaviour in high-dimensional settings.

3. RESULTS

3.1 Simulation results

The simulation results for the convergence times and Interquartile range (IQR) both scenarios and across all three QMC approaches are given in Table 1. The results indicate that the time for the joint model to fit was much efficient and stable for the QMC Halton approach, followed by QMC Sobol, while the Vanilla MC method is the slowest and most viable as the sample size increases. As the sample size increases, all methods experience an increase in computation time, but their performance varies. This indicates that the Vanilla MC method becomes much slower as the sample size grows, with considerable variability in the computation times. Although Sobol starts off faster than Vanilla MC, it still experiences a substantial increase in computation time as the sample size increases, with increasing variability as well. The QMC Halton method is the fastest across all sample sizes and has the least variability in computation time.

Table 1: Convergence times with IQR (seconds) under Vanilla MC, QMC Sobol, and QMC Halton methods for both simulation studies and all sample sizes

Scenario	*n*	Vanilla MC	QMC Sobol	QMC Halton
1	250	96.80 (49.21-293.39)	9.36 (4.59-19.46)	9.21 (4.44-19.48)
	500	97.21 (19.61-295.01)	46.31 (9.06-118.34)	28.42 (8.23-98.56)
	800	142.12 (98.45-294.22)	99.51 (68.45-198.36)	84.62 (67.89-118.27)
	1000	280.01 (128.66-1088.29)	240.02 (121.47-978.79)	185 (87.34-282.16)
2	250	95.99 (47.61-295.44)	9.77 (4.11-28.99)	9.85 (4.32-29.21)
	500	290.21 (48.79-1008.22)	96.31 (28.95-498.22)	88.42 (28.01-487.98)
	800	292.12 (148.47-988.29)	99.51 (48.79-289.48)	174.62 (89.59-292.11)
	1000	295.01 (147.33-992.66)	196.02 (97.66-294.28)	89.86 (59.67-117.88)

Figure 1 showed boxplots comparing convergence times (in seconds) for the QMC methods across two scenarios. When $n = 250$, Sobol was slightly better compared to Halton in terms of lower median time and lower variability, making it the most efficient method, while Vanilla method is the worst choice, as it takes significantly longer and exhibits much greater variability. However, as the sample size increases, Halton was the overall best method for the lowest convergence times and Sobol significantly better than Vanilla MC method.

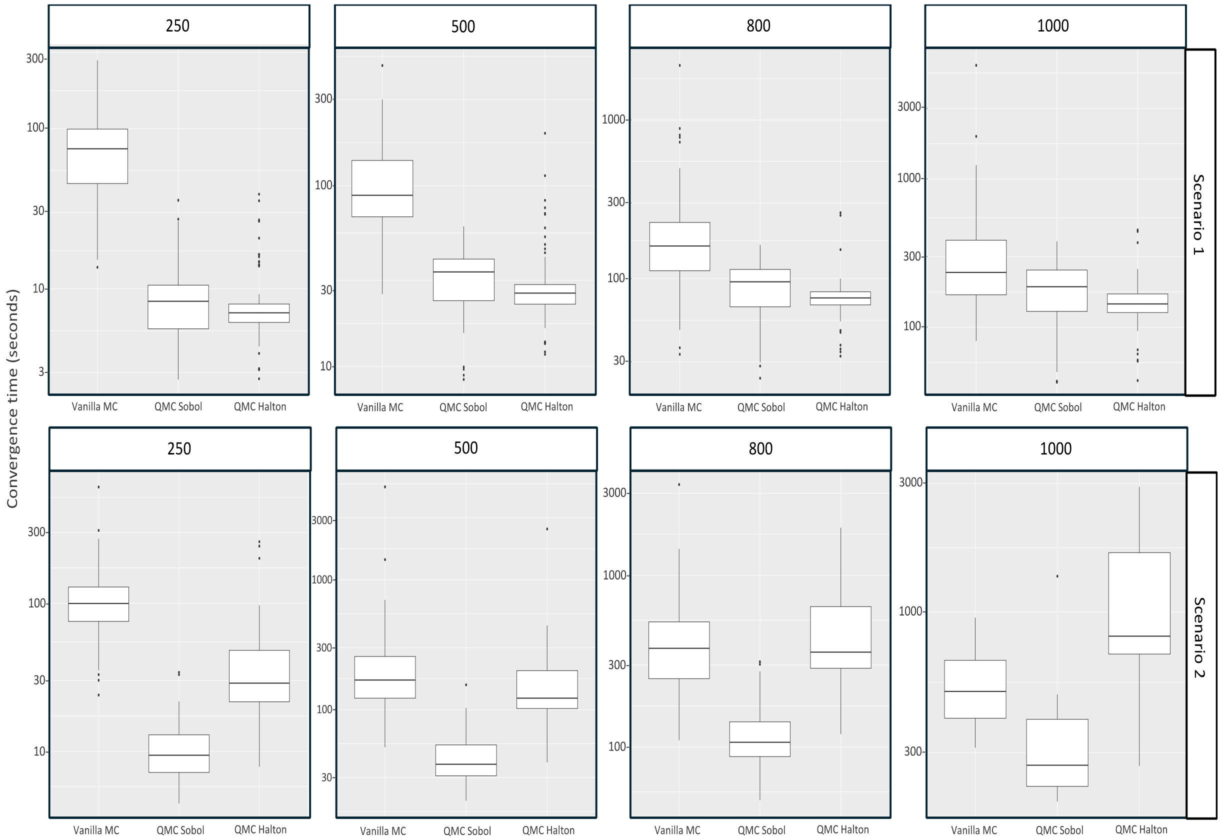

Figure 1: Comparison of the convergence times (seconds) over four different sample sizes (n=250, 500, 800, and 1000) simulated datasets using each MC method for scenario 1 (upper) and scenario 2 (lower) in the joint model

Moreover, we generated bias plots in Figure 2 to illustrate the bias in the estimates of model parameters for the three estimation methods. Typically, a bias plot shows the difference between the estimated values from each method and the true values.

In scenario 1, QMC Halton method performs best from all the sample sizes in the simulation with the median bias close to zero and the interquartile range smaller, indicating that the estimates are consistent.

Similarly, in scenario 2, QMC Halton method has the smallest bias and minimal variability and consider as the best performing method for estimating parameters in the joint model of multivariate longitudinal data and penalized Cox regression (Figure 2A).

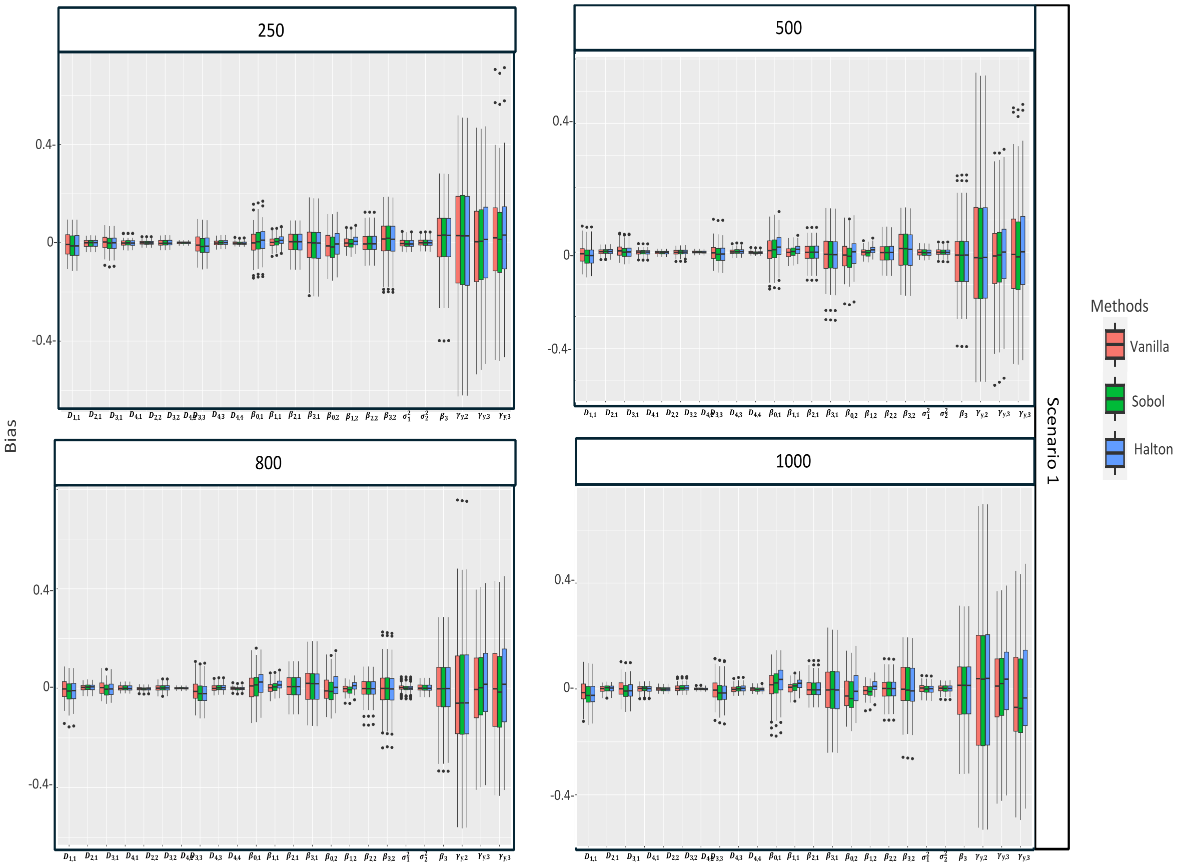

Figure 2: Comparison of the bias estimates of four different sample sizes (n=250, 500, 800, and 1000) simulated datasets using each MC method for scenario 1 in the joint model

The Table 2 and 3 present the coverage rates for both model and variance parameters of a joint model for longitudinal and survival data with ridge penalty using Vanilla MC, QMC Sobol, and QMC Halton methods across two simulation scenarios when n = 250, with equivalent results for $n = 500, n = 800, \text{and } n = 1000$ given in Table 1A to Table 3B. The coverage rates are generally high for most parameters, with values typically exceeding 0.95 for the three MC methods in scenario 1.

In contrast, Scenario 2 shows more variability in the coverage rates. While Sobol and Halton still perform relatively well, with coverage rates around 0.95 for most parameters, Vanilla MC yields significantly lower coverage for some parameters.

Across both scenarios, Halton and Sobol methods consistently provide reliable coverage, outperform the Vanilla method. However, Halton method frequently yields the highest coverage rates for parameters.

Table 2: coverage rates of model parameters for longitudinal and survival data across different simulation scenarios and methods (Vanilla, Sobol, and Halton): n=250 simulations

parameters	scenario 1			scenario 2		
	Vanilla	Sobol	Halton	Vanilla	Sobol	Halton
$\beta_{0,1}$	0.9869	0.9912	0.9974	0.9529	0.0471	0.9529
$\beta_{1,1}$	0.9942	0.9899	0.9842	0.9391	0.0609	0.9391
$\beta_{2,1}$	0.9953	0.9950	0.9953	0.9851	0.0149	0.9851
$\beta_{3,1}$	0.9883	0.9893	0.9912	0.9594	0.0406	0.9594
$\beta_{0,2}$	0.9615	0.9672	0.9550	0.9298	0.0702	0.9298
$\beta_{1,2}$	0.9958	0.9945	0.9953	0.9782	0.0218	0.9782
$\beta_{2,2}$	0.9527	0.9526	0.9525	0.9769	0.0231	0.9769
$\beta_{3,2}$	0.9031	0.9052	0.9023	0.9232	0.0768	0.9232
$\beta_{0,3}$	-	-	-	0.9362	0.0638	0.9362
$\beta_{1,3}$	-	-	-	0.9926	0.0074	0.9926
$\beta_{2,3}$	-	-	-	0.9575	0.0425	0.9575
$\beta_{3,3}$	-	-	-	0.8340	0.1660	0.8340
$\gamma_{\upsilon,1}$	0.9793	0.9799	0.9811	0.8846	0.1154	0.8846
$\gamma_{\upsilon,2}$	0.8849	0.8851	0.8856	0.7363	0.2637	0.7363
$\gamma_{y,1}$	0.9831	0.9764	0.9668	0.5305	0.4695	0.5305
$\gamma_{y,2}$	0.5551	0.5509	0.5684	0.7716	0.2284	0.7716
$\gamma_{y,3}$	-	-	-	0.9384	0.0616	0.9384

Scenario 1 (Table 3) indicate high and consistent coverage for all methods, with Halton performing slightly better. In Scenario 2, coverage rates for Vanilla drop for some parameters, suggesting that it may be more sensitive to changes in the model. Overall, Sobol and Halton methods provide more stable and reliable coverage across both scenarios, while Vanilla shows more variability.

Table 3: coverage rates of model variance parameters for longitudinal and survival data across different simulation scenarios and methods (Vanilla, Sobol, and Halton) when n=250 simulations

parameters	scenario 1			scenario 2		
	Vanilla	Sobol	Halton	Vanilla	Sobol	Halton
$D_{1,1}$	0.986	0.988	0.991	0.945	0.055	0.945
$D_{2,2}$	0.991	0.991	0.992	0.995	0.005	0.995
$D_{3,3}$	0.950	0.954	0.952	0.936	0.064	0.936
$D_{4,4}$	0.989	0.988	0.988	0.997	0.003	0.997
$D_{5,5}$	-	-	-	0.971	0.029	0.971
$D_{6,6}$	-	-	-	0.996	0.004	0.996
σ_1^2	0.999	0.999	0.999	0.976	0.024	0.976
σ_2^2	0.995	0.995	0.995	0.991	0.009	0.991
σ_3^2	-	-	-	0.992	0.008	0.992

3.2 Real Data Application

To illustrate the practical performance of the proposed penalized multivariate joint model and the comparative behaviour of the three Monte Carlo integration methods, we analysed data from the Mayo Clinic trial in primary biliary cirrhosis (PBC) of the liver (conducted between 1974-1984).

The analysis was conducted for bivariate and trivariate joint models using the mjoint() function in joineRML. The bivariate model jointly analysed bilirubin and albumin trajectories with a time-to-event outcome (time to death or censoring), while the trivariate extension included an additional longitudinal biomarker (prothrombin time). Each model incorporated fixed effects for year of observation, age, and treatment (drug), along with subject-specific random intercepts and slopes to account for within-subject heterogeneity over time. The survival submodel was specified as a penalized Cox regression, incorporating a ridge penalty on baseline covariates to regularize the estimation and prevent overfitting. Penalization modifies the partial likelihood by introducing a penalty term $\lambda_\tau \|\gamma_\varpi\|_2^2$, where λ_τ controls the degree of shrinkage. In this application, $\lambda_\tau = 0.1$ was selected based on cross-validation to balance bias and variance.

Model estimation was performed using the Vanilla Monte Carlo (MC), Sobol Quasi-Monte Carlo (QMC), and Halton QMC integration methods, each with 12,000 iterations. Convergence and burn-in parameters were tuned according to the control settings described in Section 2. The QMC methods, especially Halton sequences, demonstrated substantial computational advantages while maintaining stable estimation accuracy. For the bivariate model (K=2), computation times were 132.1s (MC), 52.9s (Sobol QMC), and 16.2s (Halton QMC). For the trivariate model (K = 3), fitting times increased to 390.7s, 78.4s, and 25.2s, respectively. These results align with the simulation findings, where Halton-based QMC achieved the most efficient convergence. Parameter estimates and standard errors for the bivariate model are presented in Table 4. Across all methods, coefficient estimates were highly consistent, though QMC methods, particularly Halton yielded smaller standard errors and more stable estimates.

Table 4: Coverage rates of model parameters (Estimate ± SE) using Vanilla MC, QMC Sobol, and QMC Halton methods for joint model with penalized Cox regression model application to PBC data

Variable	Bivariate			Trivariate		
	Vanilla	Sobol	Halton	Vanilla	Sobol	Halton
$\beta_{0,1}$	0.5085 ± 0.295	0.488 ±0.299	0.507 ±0.304	0.5308 ±0.304	0.4966 ±0.311	0.516 ±0.307
$\beta_{1,1}$	0.1883 ±0.01	0.1859 ±0.010	0.186 ±0.01	0.1945 ±0.012	0.1912 ±0.012	0.192 ±0.011
$\beta_{2,1}$	0.0004 ±0.006	0.001 ±0.006	0.0011 ±0.006	-0.0002 ±0.006	0.0005 ±0.006	0.0005 ±0.006
$\beta_{3,1}$	-0.1055 ±0.118	-0.1054 ±0.118	-0.1073 ±0.1203	-0.0898 ±0.119	-0.0932 ±0.122	-0.0944 ±0.121

$\beta_{0,2}$	3.9461 ±0.121	3.9431 ±0.120	3.9592 ±0.1208	3.9143 ±0.125	3.9161 ±0.126	3.9337 ±0.123
$\beta_{1,2}$	-0.1096 ±0.005	-0.1069 ±0.005	-0.1065 ±0.0052	-0.1114 ±0.006	-0.1091 ±0.006	-0.1088 ±0.006
$\beta_{2,2}$	-0.0082 ±0.002	-0.0083 ±0.002	-0.0083 ±0.002	-0.0076 ±0.002	-0.0077 ±0.002	-0.0077 ±0.002
$\beta_{3,2}$	0.0372 ±0.044	0.0369 ±0.043	0.0376 ±0.044	0.0355 ±0.044	0.0365 ±0.044	0.0364 ±0.043
$\beta_{0,3}$				1.0017 ±0.059	1.0058 ±0.059	1.0161 ±0.059
$\beta_{1,3}$				-0.0532 ±0.004	-0.0528 ±0.003	-0.0523 ±0.003
$\beta_{2,3}$				-0.0034 ±0.001	-0.0035 ±0.001	-0.0035 ±0.001
$\beta_{3,3}$				0.0249 ±0.024	0.0249 ±0.024	0.0254 ±0.024
$\gamma_{v,1}$	0.0651 ±0.0144	0.0657 ±0.0142	0.0659 ±0.0145	0.0641 ±0.015	0.0649 ±0.015	0.0652 ±0.015
$\gamma_{v,2}$	-0.2448 ±0.2908	-0.2451 ±0.2873	-0.2485 ±0.2912	-0.2138 ±0.298	-0.2193 ±0.296	-0.2217 ±0.297
$\gamma_{y,1}$	1.0187 ±0.1229	1.0174 ±0.1229	1.0239 ±0.1232	0.9284 ±0.145	0.9246 ±0.145	0.9322 ±0.146
$\gamma_{y,2}$	-2.4167 ±0.3306	-2.4149 ±0.3335	-2.3884 ±0.3311	-1.9907 ±0.452	-1.9648 ±0.456	-1.9817 ±0.452
$\gamma_{y,3}$				-1.7088 ±1.076	-1.7765 ±1.088	-1.7002 ±1.073

Table 5 summarizes the coverage rates for variance components. Again, both Sobol and Halton QMC methods provided similar or improved coverage relative to Vanilla MC, indicating higher numerical stability and reduced Monte Carlo variability. However, QMC Halton achieved the best balance between precision and computational efficiency, confirming its practical advantage in multivariate joint modelling contexts.

Table 5: Coverage rates of variance parameters using Vanilla MC, QMC Sobol, and QMC Halton methods for joint model with penalised Cox regression model application to PBC data

	Bivariate			**Trivariate**		
Variable	**Vanilla**	**Sobol**	**Halton**	**Vanilla**	**Sobol**	**Halton**
D_1,1	0.008	0.007	0.006	0.007	0.008	0.006
D_2,2	0.968	0.966	0.967	0.964	0.964	0.964
D_3,3	0.883	0.883	0.883	0.884	0.884	0.884
D_4,4	0.995	0.995	0.995	0.995	0.995	0.995
D_5,5	-	-	-	0.958	0.958	0.958
D_6,6	-	-	-	0.999	0.999	0.999
σ_1^2	0.881	0.880	0.880	0.881	0.881	0.881
σ_2^2	0.899	0.898	0.899	0.899	0.899	0.899
σ_3^2	-	-	-	0.975	0.975	0.975

4. DISCUSSION

This study presented an extended joint modelling framework that integrates ridge-penalized Cox regression with Monte Carlo (MC) and Quasi-Monte Carlo (QMC) integration for multivariate longitudinal and survival data. The proposed approach addresses key computational and inferential challenges in high-dimensional joint models by leveraging low-discrepancy sequences to improve numerical integration over random effects [23]. The findings showed that QMC integration significantly reduces computational time while maintaining or improving estimation accuracy compared to conventional MC. The deterministic nature of QMC nodes (Sobol and Halton sequences) ensures uniform coverage of the integration domain, leading to faster convergence, especially in higher-dimensional random-effects structures. The inclusion of a ridge penalty in the survival submodel enhances numerical stability and prevents overfitting, particularly in small samples or when baseline covariates are correlated. Penalization promotes smoother coefficient estimates and facilitates interpretation without inflating standard errors [24,25].

Both simulation and real data analyses demonstrated that QMC Halton performs comparably or better than standard MC and Sobol QMC in terms of precision, stability, and computational speed. The advantage was most pronounced in models with multiple longitudinal outcomes, confirming the scalability of the QMC approach. Quasi-Monte Carlo integration employs deterministic, low-discrepancy sequences that achieve improved convergence rates approaching $O(N^{-1}(\log N)^d)$ compared to the $O(N^{-1/2})$ rate of classical Monte Carlo [18,22,26–29]. The Koksma–Hlawka inequality provides the theoretical underpinning, linking integration error to the discrepancy of the sampling sequence [30]. Although widely applied in numerical analysis and finance, QMC methods remain underutilized in biostatistical models such as joint or frailty models, where they can substantially enhance EM algorithm performance.

In this context, QMC offers a deterministic and reproducible alternative for evaluating expectations in the E-step, improving convergence stability relative to pseudo-random sampling [31]. Furthermore, scrambling methods provide a means to estimate integration error through repeated randomization, supporting approximate variance estimation when bootstrap-based error quantification is infeasible [10,32–35]. The choice of the regularization parameter λ_τ remains crucial [36]. A large value can overshrink coefficients, while a small one risks overfitting. Cross-validation or information criteria (AIC/BIC) can guide selection, though at additional computational cost. In our analysis, manual tuning with cross-validation achieved satisfactory model performance.

Although QMC markedly improved computational efficiency, its advantages depend on model complexity, dimensionality, and the smoothness of the integrand. In very high-dimensional random-effects structures, the benefits may diminish due to increasing discrepancy in low-discrepancy sequences. Future work should explore adaptive QMC algorithms, scrambled sequence optimization, and Bayesian extensions where posterior distributions are approximated via QMC integration. In Bayesian joint models or nonlinear mixed-effects frameworks, QMC could improve high-dimensional integration accuracy and reduce sampling variability in posterior estimates. Moreover, combining QMC with modern acceleration strategies like stochastic approximation EM or quasi-Newton updates could further enhance scalability for large biostatistical datasets.

In conclusion, this work extends the existing joint modelling literature by introducing penalized survival components within a multivariate joint model, evaluating QMC integration methods within the estimation framework, and providing empirical evidence of their computational and inferential benefits. The findings suggest that QMC methods, particularly the Halton sequence, offer a reliable and efficient alternative to classical MC for high-dimensional joint models, supporting their broader adoption in complex biostatistical and clinical applications.

References

1) Alsefri M, Sudell M, García-Fiñana M, Kolamunnage-Dona R. Bayesian joint modelling of longitudinal and time to event data: a methodological review. BMC Med Res Methodol. 2020 Dec; 20(1):94.

2) Henderson R. Joint modelling of longitudinal measurements and event time data. Biostatistics. 2000 Dec 1; 1(4):465–80. A

3) Ibrahim JG, Chu H, Chen LM. Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data. J Clin Oncol. 2010 June 1;28(16):2796–801.

4) Rizopoulos D. JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. J Stat Softw. 2010;35(9).

5) Wulfsohn MS, Tsiatis AA. A Joint Model for Survival and Longitudinal Data Measured with Error. Biometrics. 1997 Mar;53(1):330.

6) Rustand D, Van Niekerk J, Krainski ET, Rue H, Proust-Lima C. Fast and flexible inference for joint models of multivariate longitudinal and survival data using integrated nested Laplace approximations. Biostatistics. 2024 Apr 15;25(2):429–48.

7) Medina-Olivares V, Lindgren F, Calabrese R, Crook J. Joint models of multivariate longitudinal outcomes and discrete survival data with INLA: An application to credit repayment behaviour. Eur J Oper Res. 2023 Oct;310(2):860–73.

8) Gupta R, Khoury JC, Altaye M, Jandarov R, Szczesniak RD. Flexible multivariate joint model of longitudinal intensity and binary process for medical monitoring of frequently collected data. Stat Med. 2021 Mar 30;40(7):1845–58.

9) Hickey GL, Philipson P, Jorgensen A, Kolamunnage-Dona R. joineRML: a joint model and software package for time-to-event and multivariate longitudinal outcomes. BMC Med Res Methodol. 2018 Dec;18(1):50.

10) Mauff K, Steyerberg E, Kardys I, Boersma E, Rizopoulos D. Joint models with multiple longitudinal outcomes and a time-to-event outcome: a corrected two-stage approach. Stat Comput. 2020 July;30(4):999–1014.

11) Crowther MJ. merlin—A unified modeling framework for data analysis and methods development in Stata. Stata J Promot Commun Stat Stata. 2020 Dec; 20(4):763–84.

12) Martin EC, Gasparini A, Crowther MJ. merlin: An R package for Mixed Effects Regression for Linear, Nonlinear and User-defined models [Internet]. arXiv; 2020 [cited 2025 Jan 9]. Available from: https://arxiv.org/abs/2007.14109

13) Ferede MM, Mwalili S, Dagne G, Karanja S, Hailu W, El-Morshedy M, Al-Bossly A. A Semiparametric Bayesian Joint Modelling of Skewed Longitudinal and Competing Risks Failure Time Data: With Application to Chronic Kidney Disease. Mathematics. 2022 Dec 18;10(24):4816.

14) Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data [Internet]. 0 ed. Chapman and Hall/CRC; 2012. Available from: https://www.taylorfrancis.com/books/9781439872871

15) Rizopoulos D, Verbeke G, Lesaffre E. Fully Exponential Laplace Approximations for the Joint Modelling of Survival and Longitudinal Data. J R Stat Soc Ser B Stat Methodol . 2009 June 1;71(3):637–54.

16) Philipson P, Hickey GL, Crowther MJ, Kolamunnage-Dona R. Faster Monte Carlo estimation of joint models for time-to-event and multivariate longitudinal data. Comput Stat Data Anal. 2020 Nov; 151:107010.

17) Wang X, Tan KS. Pricing and Hedging with Discontinuous Functions: Quasi–Monte Carlo Methods and Dimension Reduction. Manag Sci. 2013 Feb; 59(2):376–89.

18) Lamieux C. Monte Carlo and Quasi-Monte Carlo Sampling [Internet]. New York, NY: Springer New York; 2009. (Springer Series in Statistics). Available from: https://link.springer.com/10.1007/978-0-387-78165-5

19) Choi SCT, Ding Y, Hickernell FJ, Rathinavel J, Sorokin AG. Challenges in Developing Great Quasi-Monte Carlo Software. arXiv; 2023. Available from: https://arxiv.org/abs/2311.06162

20) Wang X, Sloan IH. Low discrepancy sequences in high dimensions: How well are their projections distributed? J Comput Appl Math. 2008 Apr; 213(2):366–86.

21) Zeng X, Geraci G, Eldred MS, Jakeman JD, Gorodetsky AA, Ghanem R. Multifidelity uncertainty quantification with models based on dissimilar parameters. Comput Methods Appl Mech Eng. 2023 Oct; 415:116205.

22) Owen AB. Monte Carlo, Quasi-Monte Carlo, and Randomized Quasi-Monte Carlo. In: Niederreiter H, Spanier J, editors. Monte-Carlo and Quasi-Monte Carlo Methods 1998 [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2000. p. 86–97.

23) Caillebotte A, Kuhn E, Lemler S. Estimation and variable selection in a joint model of survival times and longitudinal outcomes with random effects. arXiv; 2023. Available from: https://arxiv.org/abs/2306.16765

24) Horny G. Partial Likelihood Estimation of a Cox Model with Random Effects: an EM Algorithm based on Penalized Likelihood. Work Pap BETA. 2006; Available from: https://ideas.repec.org//p/ulp/sbbeta/2006-10.html

25) Ibrahim JG, Zhu H, Garcia RI, Guo R. Fixed and Random Effects Selection in Mixed Effects Models. Biometrics. 2011 June; 67(2):495–503. Available from: https://academic.oup.com/biometrics/article/67/2/495-503/7381000

26) You L, Qiu P. Joint modeling of multivariate nonparametric longitudinal data and survival data: A local smoothing approach. Stat Med. 2021 Dec 20; 40(29):6689–706.

27) Owen AB. Monte Carlo Variance of Scrambled Net Quadrature. SIAM J Numer Anal. 1997 Oct; 34(5):1884–910.

28) Gnewuch M, Dick J, Markhasin L, Sickel W. QMC integration based on arbitrary (t,m,s)-nets yields optimal convergence rates on several scales of function spaces. arXiv; 2024. Available from: https://arxiv.org/abs/2409.12879

29) Mascagni M, Yu H. Scrambled SoboĺÍ sequences via permutation. Monte Carlo Methods Appl. 2009 Jan; 15(4).

30) Bansal N, Jiang H. Quasi-Monte Carlo Beyond Hardy-Krause. arXiv; 2024. Available from: https://arxiv.org/abs/2408.06475

31) Bayousef M, Chi H. Computational Investigations of Optimal Scrambled Halton Sequences. In: Proceedings of the Practice and Experience on Advanced Research Computing. Pittsburgh PA USA: ACM; 2018. p. 1–4.

32) Cheng J, Druzdzel MJ. Computational Investigation of Low-Discrepancy Sequences in Simulation Algorithms for Bayesian Networks. arXiv; 2013. Available from: https://arxiv.org/abs/1301.3841

33) Lin H, McCulloch CE, Mayne ST. Maximum likelihood estimation in the joint analysis of time-to-event and multiple longitudinal variables. Stat Med. 2002 Aug 30; 21(16):2369–82.

34) Bayer C, Hammouda CB, Papapantoleon A, Samet M, Tempone R. Quasi-Monte Carlo for Efficient Fourier Pricing of Multi-Asset Options. arXiv; 2024. Available from: https://arxiv.org/abs/2403.02832

35) Kaarnioja V, Schillings C. Quasi-Monte Carlo for Bayesian design of experiment problems governed by parametric PDEs. arXiv; 2024. Available from: https://arxiv.org/abs/2405.03529

36) Ripatti S, Palmgren J. Estimation of Multivariate Frailty Models Using Penalized Partial Likelihood. Biometrics. 2000 Dec; 56(4):1016–22.

37) Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. Stat Sci. 1996 May 1; 11(2).

38) Hurvich CM, Simonoff JS, Tsai CL. Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. J R Stat Soc Ser B Stat Methodol. 1998 July 1; 60(2):271–93.

39) He Z, Wang X. An Integrated Quasi-Monte Carlo Method for Handling High Dimensional Problems with Discontinuities in Financial Engineering. Comput Econ. 2021 Feb; 57(2):693–718.

40) L'Ecuyer P. Randomized Quasi-Monte Carlo. In: Kenett RS, Longford NT, Piegorsch WW, Ruggeri F, editors. Wiley StatsRef: Statistics Reference Online. 1st ed. Wiley; 2020. p. 1–9.

41) Nevo D, Blacker D, Larson EB, Haneuse S. Modeling semi-competing risks data as a longitudinal bivariate process. Biometrics. 2022 Sept; 78(3):922–36.

42) Nevo D, Gorfine M. Causal inference for semi-competing risks data [Internet]. arXiv; 2020 [cited 2025 Jan 9]. Available from: https://arxiv.org/abs/2010.04485

43) Othus M, Li Y. A Gaussian Copula Model for Multivariate Survival Data. Stat Biosci [Internet]. 2010 Dec [cited 2025 Jan 9];2(2):154–79. Available from: http://link.springer.com/10.1007/s12561-010-9026-x

44) Ripatti S, Larsen K, Palmgren J. Maximum Likelihood Inference for Multivariate Frailty Models Using an Automated Monte Carlo EM Algorithm. Lifetime Data Anal [Internet]. 2002 [cited 2025 Jan 9];8(4):349–60. Available from: http://link.springer.com/10.1023/A:1020566821163

45) Owen AB. Scrambled net variance for integrals of smooth functions. Ann Stat [Internet]. 1997 Aug 1 [cited 2025 Jan 9];25(4). Available from: https://projecteuclid.org/journals/annals-of-statistics/volume-25/issue-4/Scrambled-net-variance-for-integrals-of-smooth-functions/10.1214/aos/1031594731.full