Supplement to "Statistical Modeling of the Effectiveness of Preventive Maintenance for Repairable Systems"

Xin YE[†], Jiaxiang CAI[†], Loon Ching TANG[†], and Zhi-Sheng YE[†]

[†]Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore, 117576

S1 The Bayesian Inference

In this section, we present the Bayesian inference procedure. In the Bayesian setting, one very first tough question concerns the choice of the prior distribution for parameters. Without sufficient prior information, we may simply use a uniform distribution over the whole parameter space as a non-informative prior (Syversveen, 1998), or leverage the Jeffreys prior that is proportional to the square root of the determinant of the Fisher information matrix (Gelman, 2009), or construct the reference prior by maximizing the missing information of model parameters (Berger et al., 2009). For our service vehicle study, there is no prior information on the reliability of service vehicles and on the preventive maintenance effectiveness. Hence, non-informative priors are considered in our Bayesian inference procedure, which is introduced in greater detail in the sequel.

S1.1 Bayesian formulation

Consider a generic system that follows the failure process in Section 2.1 of the paper. Recall that the AIC values in our real data example favor the model with a power-law baseline ROCOF $\lambda_0(t) = (\beta/\eta)\,(t/\eta)^{\beta-1}$ and the Lognormal(μ, σ^2) multiplicative PM random effects, i.e., the PLP-Lognormal model. For comparison, we next conduct the Bayesian inference procedure under this PLP-Lognormal model setting and examine its performance. The observed data from this system is $\mathbf{D} = \{L, R, N(t) : t \in [L, R]\}$. The set of PM random effects is $\mathbf{A} = \{A_k : k\Delta \leq R\}$. When n independent systems are under observation, let $\mathbb{D} = \{\mathbf{D}_i : i = 1, 2, \cdots, n\}$ and $\mathbb{A} = \{\mathbf{A}_i : i = 1, 2, \cdots, n\}$ denote the observed data and the PM random effects from all the n systems, respectively. Conditioning on the PM random effects, the likelihood function can be written as

$$L(\mathbb{D} \mid \mathbb{A}, \beta, \eta) = \prod_{i=1}^{n} \exp(\Lambda_i(L_i) - \Lambda_i(R_i)) \prod_{t \in [L_i, R_i]} \lambda_i(t)^{dN_i(t)}. \tag{1}$$

In the Bayesian framework, the parameters β, η, μ, σ^2 and $\mathbb{A}$ are all assumed random variables. The posterior distribution is

$$\begin{aligned}\pi(\mathbb{A}, \beta, \eta, \mu, \sigma^2 \mid \mathbb{D}) &\propto \pi(\mathbb{D} \mid \mathbb{A}, \beta, \eta)\pi(\mathbb{A}, \beta, \eta, \mu, \sigma^2) \\ &= L(\mathbb{D} \mid \mathbb{A}, \beta, \eta)\pi(\mathbb{A}, \beta, \eta, \mu, \sigma^2).\end{aligned}$$

We assume that β, η, $\mu \mid \sigma^2$, σ^2, and $\mathbb{A} \mid \mu, \sigma^2$ are independent, and the prior is given by

$$\pi(\mathbb{A}, \beta, \eta, \mu, \sigma^2) = \pi(\mathbb{A} \mid \mu, \sigma^2)\pi(\mu \mid \sigma^2)\pi(\sigma^2)\pi(\beta)\pi(\eta).$$

The priors are chosen as follows.

- $\pi(A_{ik} \mid \mu, \sigma^2) \sim \text{Lognormal}(\mu, \sigma^2)$, where μ is the mean and σ^2 is the variance.
- $\pi(\mu \mid \sigma^2) \sim N(a, \sigma^2/b)$, where a is the mean and σ^2/b is the variance.
- $\pi(\sigma^2) \sim \text{Inv-Gamma}(\alpha, \nu)$, i.e. inverse gamma distribution, where α is the shape parameter and ν is the scale parameter (see also Gelman's Prior distributions for variance parameters for a complete exposition in Gelman (2006)).
- $\pi(\beta) \sim \text{Gamma}(c, d)$, where c is the shape parameter and d is the rate parameter.
- $\pi(\eta) \sim \text{Gamma}(e, f)$, where e is the shape parameter and f is the rate parameter.

The values of hyperparameters a, b, α, ν, c, d, e and f are carefully set to ensure non-informative priors. The specific values of hyperparameters are introduced in Section S1.4. The dependence of the variables are shown in Figure 1. With the above settings and the independent and identically

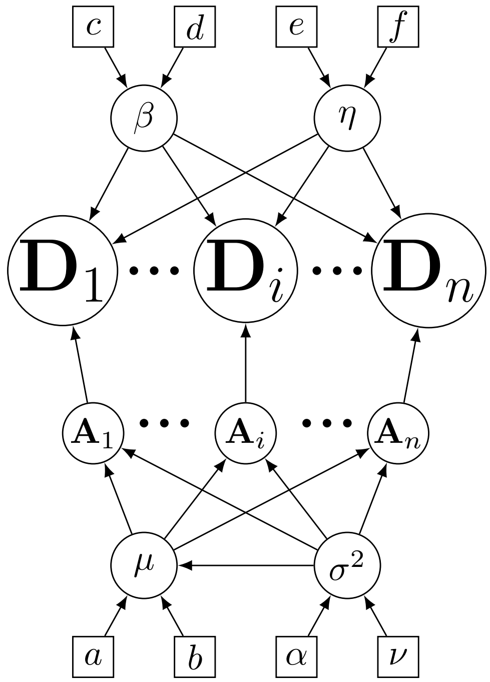

Figure 1: The dependence of the variables. In the directed graph, arrows run into nodes from their direct influences. Rectangle nodes are fixed constants (i.e. hyperparameters), and circular nodes are variables or observed data.

distributed (i.i.d.) unobserved PM random effects A_{ik} assumption, we have

$$\pi(\mathbb{A} \mid \mu, \sigma^2) = \prod_{i=1}^{n} \prod_{k=1}^{\lfloor R_i/\Delta \rfloor} \pi(A_{ik} \mid \mu, \sigma^2) = \prod_{i=1}^{n} \prod_{k=1}^{\lfloor R_i/\Delta \rfloor} \frac{a_{ik}^{-1}}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(\ln(a_{ik}) - \mu)^2}{\sigma^2}\right), \tag{2}$$

As can be seen, combining $\pi(\mu \mid \sigma^2)$ and $\pi(\sigma^2)$, we have $\pi(\mu, \sigma^2) \sim \text{NIG}(a, b, \alpha, \nu)$, i.e., normal-inverse gamma distribution as below.

$$\pi(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{b}}{\sigma} \exp\left(-\frac{1}{2}\frac{b\,(\mu - a)^2}{\sigma^2}\right) \times \frac{\nu^\alpha \left(\sigma^2\right)^{-\alpha-1}}{\Gamma\left(\alpha\right)} \exp\left(-\frac{\nu}{\sigma^2}\right). \tag{3}$$

We carefully choose the normal-inverse gamma distribution for (μ, σ^2) to make our formulas analytically tractable. The detailed derivations are introduced in the sequel.

There is no analytical expression for the posterior distribution $\pi(\mathbb{A}, \beta, \eta, \mu, \sigma^2 \mid \mathbb{D})$. As such, we use the Markov Chain Monte Carlo (MCMC) algorithm to generate samples from this posterior distribution. These posterior samples are further used to compute the point and interval estimations of the model parameters, which is introduced in the next section.

S1.2 Gibbs sampling algorithm

To be specific, considering the difficulty in directly sampling from the posterior, we use the Gibbs sampling algorithm (George and McCulloch, 1993). The pseudocode of the Gibbs sampling algorithm is presented in Algorithm 1.

Algorithm 1 Gibbs sampling algorithm

Input: Initial values $(\mathbb{A}^{(0)}, \beta^{(0)}, \eta^{(0)}, \mu^{(0)}, (\sigma^2)^{(0)})$, number of burn-ins N_0 and number of total iterations N.

while $1 \leq k \leq N$ **do**

$$\mathbb{A}^{(k)} \sim \pi(\mathbb{A} \mid \beta^{(k-1)}, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{D}) \tag{4}$$

$$\beta^{(k)} \sim \pi(\beta \mid \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)}, \mathbb{D}) \tag{5}$$

$$\eta^{(k)} \sim \pi(\eta \mid \beta^{(k)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)}, \mathbb{D}) \tag{6}$$

$$\mu^{(k)} \sim \pi(\mu \mid \beta^{(k)}, \eta^{(k)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)}, \mathbb{D}) \tag{7}$$

$$(\sigma^2)^{(k)} \sim \pi(\sigma^2 \mid \beta^{(k)}, \eta^{(k)}, \mu^{(k)}, \mathbb{A}^{(k)}, \mathbb{D}) \tag{8}$$

Output: Sample values $(\mathbb{A}^{(k)}, \beta^{(k)}, \eta^{(k)}, \mu^{(k)}, (\sigma^2)^{(k)})$, $k = \{N_0 + 1, \cdots, N\}$.

Next, we investigate the above full conditional distributions (4)-(8). By combining Equations (1) and (2), we have

$$\pi(\mathbb{A} \mid \beta, \eta, \mu, \sigma^2, \mathbb{D}) \propto \prod_{i=1}^{n} \Bigg(\exp(\Lambda_i(L_i) - \Lambda_i(R_i)) \prod_{t \in [L_i, R_i]} \lambda_i(t)^{dN_i(t)} \prod_{k=1}^{\lfloor R_i/\Delta \rfloor} \frac{a_{ik}^{-1}}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(\ln(a_{ik}) - \mu)^2}{\sigma^2} \right) \Bigg). \tag{9}$$

By combining Equation (1) and $\pi(\beta) \sim \text{Gamma}(c, d)$, we have

$$\pi(\beta \mid \eta, \mu, \sigma^2, \mathbb{A}, \mathbb{D}) \propto \prod_{i=1}^{n} \left(\exp(\Lambda_i(L_i) - \Lambda_i(R_i)) \prod_{t \in [L_{i0}, R_i]} \lambda_i(t)^{dN_i(t)} \right) \frac{d^c}{\Gamma(c)} \beta^{c-1} \exp(-d\beta), \tag{10}$$

and by combining Equation (1) and $\pi(\eta) \sim \text{Gamma}(e, f)$, we have

$$\pi(\eta \mid \beta, \mu, \sigma^2, \mathbb{A}, \mathbb{D}) \propto \prod_{i=1}^{n} \left(\exp(\Lambda_i(L_i) - \Lambda_i(R_i)) \prod_{t \in [L_i, R_i]} \lambda_i(t)^{dN_i(t)} \right) \frac{f^e}{\Gamma(e)} \eta^{e-1} \exp(-f\eta). \tag{11}$$

By combining Equations (2) and (3), we have

$$\pi(\sigma^2 \mid \beta, \eta, \mu, \mathbb{A}, \mathbb{D}) \propto \left(\sigma^2\right)^{-\left(1+\alpha+\frac{1}{2}+\frac{1}{2}\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor\right)} \times \\ \exp\left(-\frac{\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\left(\ln(a_{ik})-\mu\right)^2+\frac{1}{2}b\left(\mu-a\right)^2+\nu}{\sigma^2}\right).$$

As can be seen, $\pi(\sigma^2 \mid \beta, \eta, \mu, \mathbb{A}, \mathbb{D})$ follows an inverse gamma distribution:

$$\text{Inv-Gamma}\left(\alpha+\frac{1}{2}+\frac{1}{2}\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor, \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\left(\ln(a_{ik})-\mu\right)^2+\frac{1}{2}b\left(\mu-a\right)^2+\nu\right).$$

Similarly, by combining Equations (2) and (3), we can obtain

$$\begin{aligned}
\pi(\mu \mid \beta, \eta, \sigma^2, \mathbb{A}, \mathbb{D}) &\propto \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\frac{\left(\ln(a_{ik})-\mu\right)^2}{\sigma^2}-\frac{1}{2}\frac{b\left(\mu-a\right)^2}{\sigma^2}\right) \\
&\propto \exp\left(-\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\frac{\mu^2-2\mu\ln(a_{ik})+\ln^2(a_{ik})}{2\sigma^2}-\frac{b\mu^2-2ab\mu+a^2b}{2\sigma^2}\right) \\
&\propto \exp\left(-\frac{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}{2\sigma^2}\left[\mu^2-2\left(\frac{\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\ln(a_{ik})+ab}{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}\right)\mu\right.\right. \\
&\quad \left.\left.+\frac{\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\ln^2(a_{ik})+a^2b}{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}\right]\right) \\
&\propto \exp\left(-\frac{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}{2\sigma^2}\left[\mu-\frac{\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\ln(a_{ik})+ab}{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}\right]^2\right).
\end{aligned}$$

Hence, $\pi(\mu \mid \beta, \eta, \sigma^2, \mathbb{A}, \mathbb{D})$ follows a normal distribution:

$$N\left(\frac{\sum_{i=1}^{n}\sum_{k=1}^{\lfloor R_i/\Delta\rfloor}\ln(a_{ik})+ab}{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}, \frac{\sigma^2}{b+\sum_{i=1}^{n}\lfloor R_i/\Delta\rfloor}\right).$$

As above, (9), (10) and (11) do not have the analytical expressions and the Metropolis-Hastings (M-H) algorithm is thus used to sample from them at every iteration $1 \le k \le N$ of the above Gibbs sampling algorithm. Consider the kth iteration of the Gibbs sampling algorithm, and thus we have $\mathbb{A}^{(k)} \sim \pi(\mathbb{A} \mid \beta^{(k-1)}, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{D})$. To generate the kth sample from $\pi(\mathbb{A} \mid \beta^{(k-1)}, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{D})$, the M-H algorithm first generates a candidate $\mathbb{A}^*$ from the pre-specified proposal distribution $h(\mathbb{A}) = \pi\left(\mathbb{A} \mid \mu^{(k-1)}, (\sigma^2)^{(k-1)}\right)$, and then calculates the probability of accepting this candidate using the acceptance function

$$f\left(\mathbb{A}^{(k-1)}, \mathbb{A}^*\right) = \min\left\{1, \frac{\pi\left(\mathbb{D} \mid \mathbb{A}^*, \beta^{(k-1)}, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}\right)}{\pi\left(\mathbb{D} \mid \mathbb{A}^{(k-1)}, \beta^{(k-1)}, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}\right)}\right\}.$$

A random number u is then generated from the standard uniform distribution to determine whether to accept this generated candidate. For instance, if $u < f\left(\mathbb{A}^{(k-1)}, \mathbb{A}^*\right)$, then the candidate value $\mathbb{A}^*$ is accepted and we set $\mathbb{A}^{(k)} = \mathbb{A}^*$. Otherwise, the candidate value is discarded and we set $\mathbb{A}^{(k)} = \mathbb{A}^{(k-1)}$.

Similarly, to generate the kth samples $\beta^{(k)}$ and $\eta^{(k)}$, we run the random walk M-H algorithm to first generate candidates β^* and η^* from the proposal functions $h(\beta^* \mid \beta^{(k-1)})$ and $h(\eta^* \mid \eta^{(k-1)})$, respectively. The common choice of $h(\cdot)$ is a normal distribution such that $h(\beta^* \mid \beta^{(k-1)}) = N(\beta^{(k-1)}, \sigma_\beta^2)$ and $h(\eta^* \mid \eta^{(k-1)}) = N(\eta^{(k-1)}, \sigma_\eta^2)$, where σ_β^2 and σ_η^2 are constants that are carefully set in Section S1.4. Thus, the acceptance functions are respectively given by

$$f\left(\beta^{(k-1)}, \beta^*\right) = \min\left\{1, \frac{\pi(\mathbb{D} \mid \beta^*, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)})\pi(\beta^*)}{\pi(\mathbb{D} \mid \beta^{(k-1)}, \eta^{(k-1)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)})\pi(\beta^{(k-1)})}\right\},$$

and

$$f\left(\eta^{(k-1)}, \eta^*\right) = \min\left\{1, \frac{\pi(\mathbb{D} \mid \eta^*, \beta^{(k)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)})\pi(\eta^*)}{\pi(\mathbb{D} \mid \eta^{(k-1)}, \beta^{(k)}, \mu^{(k-1)}, (\sigma^2)^{(k-1)}, \mathbb{A}^{(k)})\pi(\eta^{(k-1)})}\right\}.$$

We accept the candidate if the value of the acceptance function exceeds a random number generated from the standard uniform distribution. Otherwise, the candidate value is discarded.

S1.3 Bayesian parameter estimation

In this section, simulation is conducted to check the performance of the Bayesian parameter estimation under the PLP-Lognormal setting. To be specific, we consider the power-law baseline ROCOF $\lambda_0(t) = (\beta/\eta)\,(t/\eta)^{\beta-1}$ with $(\beta, \eta) = (1.25, 100)$ and the Lognormal(μ, σ^2) PM random effect with $(\mu, \sigma) \in \{(-1/5, 1/2), (-1/8, 1/2), (-1/10, 1/2)\}$. As with Section 5.1 of the paper, three levels of the sample size are considered: $n \in \{30, 45, 60\}$, four levels of the PM time interval are examined: $\Delta \in \{365, 450, 540, 630\}$, and we let the observation window of $n/3$ systems be $[400, 2000]$, $n/3$ systems be $[400, 1200]$, and the remaining $n/3$ systems be $[200, 1800]$.

For each setting above, we run $1,000$ Monte Carlo replications. For each replication, we implement the Gibbs sampling algorithm to sample posterior samples. The detailed discussions on the convergence of the chain induced by the Gibbs sampling algorithm are deferred to Section S1.4. The mean values of posterior samples are used as the point estimations of parameters, and the 95% highest posterior density intervals are computed as the credible interval estimations of parameters. Based on the $1,000$ replications, we calculate the bias, root mean squared error (RMSE), and the coverage probability of the 95% highest posterior density interval. All computations were conducted on an Intel(R) Xeon(R) CPU E5-2698 v4 (2.20 GHz).

The estimation results and computation time results are presented in Tables 1-4. These results reveal that the biases and RMSEs are generally small. The coverage probabilities of the 95% highest posterior density intervals for μ and σ are closer to the nominal value than β and η. This might be attributed to the inadequate mixing property of the chain for parameters β and η. To support our speculation, we conduct the convergence analysis in Section S1.4, and the convergence diagnostic plots indeed show the inadequate mixing property of the chain and the high level of autocorrelation among samples. In addition, by checking the computation time results, we note in passing that the Bayesian estimation method is generally less computationally efficient than the proposed EM estimation method. Overall, tuning the MCMC algorithm for such a complex posterior in our study is hard and time-consuming.

Table 1: Biases, RMSEs, the coverage probability (CP) of the 95% highest density interval and average computation time (in hours), based on $1,000$ Monte Carlo replications, under the PLP-Lognormal model ($\Delta = 365$).

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/5, 1/2)$	3.025	bias ($\times 10^{-2}$)	0.523	82.527	-0.252	0.207
			RMSE	0.093	23.473	0.069	0.065
			CP	0.935	0.928	0.947	0.949
	$(1.25, 100, -1/8, 1/2)$	3.321	bias ($\times 10^{-2}$)	0.293	89.521	0.392	-0.472
			RMSE	0.108	22.194	0.075	0.052
			CP	0.935	0.937	0.944	0.959
	$(1.25, 100, -1/10, 1/2)$	3.240	bias ($\times 10^{-2}$)	0.674	95.361	-0.341	-0.452
			RMSE	0.094	22.093	0.044	0.046
			CP	0.937	0.931	0.943	0.958
45	$(1.25, 100, -1/5, 1/2)$	4.574	bias ($\times 10^{-2}$)	0.426	63.215	-0.083	0.035
			RMSE	0.088	20.405	0.052	0.056
			CP	0.945	0.941	0.942	0.958
	$(1.25, 100, -1/8, 1/2)$	4.329	bias ($\times 10^{-2}$)	0.212	111.532	0.231	0.178
			RMSE	0.079	18.184	0.042	0.037
			CP	0.937	0.940	0.953	0.955
	$(1.25, 100, -1/10, 1/2)$	4.925	bias ($\times 10^{-2}$)	0.805	122.623	0.124	0.401
			RMSE	0.115	21.162	0.048	0.036
			CP	0.935	0.937	0.956	0.953
60	$(1.25, 100, -1/5, 1/2)$	6.715	bias ($\times 10^{-2}$)	0.079	124.921	0.073	0.044
			RMSE	0.073	18.491	0.039	0.038
			CP	0.941	0.945	0.956	0.957
	$(1.25, 100, -1/8, 1/2)$	6.532	bias ($\times 10^{-2}$)	0.125	87.315	0.088	0.213
			RMSE	0.079	18.262	0.047	0.032
			CP	0.944	0.943	0.949	0.953
	$(1.25, 100, -1/10, 1/2)$	6.357	bias ($\times 10^{-2}$)	0.942	107.459	0.205	0.378
			RMSE	0.071	16.145	0.032	0.033
			CP	0.947	0.949	0.952	0.953

Table 2: Biases, RMSEs, the coverage probability (CP) of the 95% highest density interval and average computation time (in hours), based on $1,000$ Monte Carlo replications, under the PLP-Lognormal model ($\Delta = 450$).

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/5, 1/2)$	2.934	bias ($\times 10^{-2}$)	0.113	48.431	0.096	0.062
			RMSE	0.101	21.455	0.048	0.038
			CP	0.935	0.934	0.946	0.961
	$(1.25, 100, -1/8, 1/2)$	3.046	bias ($\times 10^{-2}$)	0.105	35.329	0.077	0.068
			RMSE	0.095	20.968	0.042	0.060
			CP	0.937	0.936	0.940	0.964
	$(1.25, 100, -1/10, 1/2)$	3.032	bias ($\times 10^{-2}$)	0.436	85.591	-0.124	0.325
			RMSE	0.096	25.630	0.075	0.057
			CP	0.935	0.945	0.944	0.953
45	$(1.25, 100, -1/5, 1/2)$	4.645	bias ($\times 10^{-2}$)	0.095	37.481	0.102	0.055
			RMSE	0.068	18.453	0.051	0.053
			CP	0.941	0.938	0.947	0.957
	$(1.25, 100, -1/8, 1/2)$	4.527	bias ($\times 10^{-2}$)	0.084	30.531	0.062	0.043
			RMSE	0.083	15.451	0.041	0.036
			CP	0.945	0.943	0.948	0.960
	$(1.25, 100, -1/10, 1/2)$	4.801	bias ($\times 10^{-2}$)	0.341	46.682	0.117	-0.245
			RMSE	0.118	22.456	0.046	0.052
			CP	0.938	0.941	0.948	0.954
60	$(1.25, 100, -1/5, 1/2)$	5.371	bias ($\times 10^{-2}$)	0.076	36.257	0.121	0.027
			RMSE	0.061	17.627	0.034	0.040
			CP	0.940	0.947	0.951	0.955
	$(1.25, 100, -1/8, 1/2)$	5.554	bias ($\times 10^{-2}$)	0.069	25.412	0.071	0.038
			RMSE	0.082	16.641	0.042	0.033
			CP	0.942	0.937	0.954	0.948
	$(1.25, 100, -1/10, 1/2)$	5.386	bias ($\times 10^{-2}$)	0.143	32.674	0.142	-0.227
			RMSE	0.085	20.943	0.032	0.051
			CP	0.943	0.950	0.944	0.948

Table 3: Biases, RMSEs, the coverage probability (CP) of the 95% highest density interval and average computation time (in hours), based on $1,000$ Monte Carlo replications, under the PLP-Lognormal model ($\Delta = 540$).

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/5, 1/2)$	2.813	bias ($\times 10^{-2}$)	0.065	87.551	0.142	-0.721
			RMSE	0.084	19.571	0.053	0.048
			CP	0.939	0.943	0.952	0.943
	$(1.25, 100, -1/8, 1/2)$	2.653	bias ($\times 10^{-2}$)	0.083	47.682	0.053	0.851
			RMSE	0.073	20.372	0.053	0.049
			CP	0.932	0.937	0.946	0.945
	$(1.25, 100, -1/10, 1/2)$	3.204	bias ($\times 10^{-2}$)	0.135	41.463	-0.121	-0.401
			RMSE	0.086	21.571	0.053	0.032
			CP	0.935	0.933	0.939	0.941
45	$(1.25, 100, -1/5, 1/2)$	4.671	bias ($\times 10^{-2}$)	0.057	77.581	0.121	0.510
			RMSE	0.067	16.025	0.045	0.035
			CP	0.938	0.942	0.959	0.945
	$(1.25, 100, -1/8, 1/2)$	4.324	bias ($\times 10^{-2}$)	0.076	55.251	0.044	-0.521
			RMSE	0.067	17.856	0.048	0.036
			CP	0.936	0.941	0.952	0.949
	$(1.25, 100, -1/10, 1/2)$	4.463	bias ($\times 10^{-2}$)	0.094	38.593	0.117	-0.301
			RMSE	0.052	19.659	0.049	0.035
			CP	0.938	0.939	0.942	0.947
60	$(1.25, 100, -1/5, 1/2)$	5.397	bias ($\times 10^{-2}$)	0.047	56.138	0.091	0.305
			RMSE	0.041	14.648	0.042	0.039
			CP	0.945	0.941	0.960	0.944
	$(1.25, 100, -1/8, 1/2)$	5.523	bias ($\times 10^{-2}$)	0.083	41.543	0.035	-0.355
			RMSE	0.056	13.372	0.038	0.035
			CP	0.943	0.944	0.949	0.952
	$(1.25, 100, -1/10, 1/2)$	5.301	bias ($\times 10^{-2}$)	0.044	45.102	0.098	-0.283
			RMSE	0.057	17.150	0.040	0.039
			CP	0.945	0.946	0.947	0.952

Table 4: Biases, RMSEs, the coverage probability (CP) of the 95% highest density interval and average computation time (in hours), based on 1,000 Monte Carlo replications, under the PLP-Lognormal model ($\Delta = 630$).

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/8, 1/2)$	2.610	bias ($\times 10^{-2}$)	0.243	41.572	0.075	0.347
			RMSE	0.082	25.421	0.049	0.052
			CP	0.936	0.932	0.942	0.955
	$(1.25, 100, -1/10, 1/2)$	2.849	bias ($\times 10^{-2}$)	0.323	42.549	-0.461	0.668
			RMSE	0.070	26.674	0.038	0.053
			CP	0.933	0.944	0.948	0.957
45	$(1.25, 100, -1/5, 1/2)$	4.124	bias ($\times 10^{-2}$)	0.302	24.594	0.141	-0.119
			RMSE	0.032	15.148	0.059	0.045
			CP	0.939	0.928	0.944	0.951
	$(1.25, 100, -1/8, 1/2)$	4.276	bias ($\times 10^{-2}$)	0.145	27.921	0.044	-0.218
			RMSE	0.065	21.343	0.045	0.036
			CP	0.935	0.947	0.958	0.943
	$(1.25, 100, -1/10, 1/2)$	4.352	bias ($\times 10^{-2}$)	0.156	32.814	0.315	-0.406
			RMSE	0.051	23.245	0.029	0.048
			CP	0.929	0.945	0.947	0.958
60	$(1.25, 100, -1/5, 1/2)$	5.548	bias ($\times 10^{-2}$)	0.073	26.447	0.041	-0.107
			RMSE	0.036	16.847	0.041	0.026
			CP	0.942	0.938	0.945	0.953
	$(1.25, 100, -1/8, 1/2)$	5.218	bias ($\times 10^{-2}$)	0.075	17.234	0.030	0.163
			RMSE	0.026	14.548	0.037	0.032
			CP	0.939	0.951	0.949	0.954
	$(1.25, 100, -1/10, 1/2)$	5.155	bias ($\times 10^{-2}$)	0.123	23.217	0.157	-0.314
			RMSE	0.051	22.534	0.027	0.042
			CP	0.943	0.946	0.950	0.949

S1.4 Convergence analysis

Next, we summarize the chain induced by the Gibbs sampling algorithm and assess its convergence. The Bayesian parameter estimation is implemented with priors introduced in Section S1.1. To set up non-informative priors, the hyperparameters are set as follows (Spiegelhalter et al., 1996; Gelman, 2006),

$$a = 10^{-3}, b = 10^{-3}, \alpha = 10^{-3}, \nu = 10^{-3}, c = 10^{-4}, d = 10^{-4}, e = 10^{-4}, f = 10^{-4}.$$

The parameters of proposal functions are set as $\sigma_\beta = 0.22$ and $\sigma_\eta = 7.20$. To define the stopping criteria for the algorithm, we run the Geweke diagnostic (Geweke, 1992) to examine the convergence of the Markov chain induced by the proposed Gibbs sampling algorithm. Specifically, we conduct the Geweke test by taking the first 10% and the last 50% part of the chain into account. Besides, the Gelman and Rubin's statistic (Gelman and Rubin, 1992; Brooks and Gelman, 1998) is also computed to check the convergence of multiple chains from the overdispersed starting points. As suggested by Brooks and Gelman (1998), the convergence is reached if the Gelman-Rubin convergence diagnostic value is less than 1.2 for all model parameters. According to our comprehensive numerical experience, we need to run the Gibbs sampling algorithm for a total number of $N = 500,000$ iterations to ensure the Gelman-Rubin convergence diagnostic values are less than 1.2 and p-values of the Geweke test statistic for all model parameters β, η, μ and σ are above the pre-specified threshold $\delta = 0.05$ under all the settings. We then discard the first $N_0 = 100,000$ iterations as burn-in. After burn-in, we thin the chain by setting the value of thinning as 100 to reduce the autocorrelation of posterior samples. As such, a chain of $4,000$ iterations is obtained after burn-in and thinning.

However, we caution that the mixing property of the induced chain might be still poor, even though they have passed the Geweke test and their Gelman-Rubin convergence diagnostic values are less than 1.2 (Gong and Flegal, 2016). To illustrate that, we present a realization of the convergence diagnostic plots of the chain for model parameters β, η, μ and σ in Figure 2. As can be seen, the trace plots and autocorrelation plots for parameters β and η show the inadequate mixing property of the chain and the high level of autocorrelation among samples. To overcome this, more iterations are required, which, however, renders the Bayesian estimation procedure even more computationally cumbersome. Alternative strategies such as other carefully selected proposal functions for the M-H algorithm, the adaptive Metropolis sampler (Haario et al., 1999, 2001) and the delayed rejection mechanism (Tierney and Mira, 1999; Green and Mira, 2001) can be considered in the future to improve the mixing efficiency of the chain, and thus the precision of the Bayesian parameter estimation can be imrpoved.

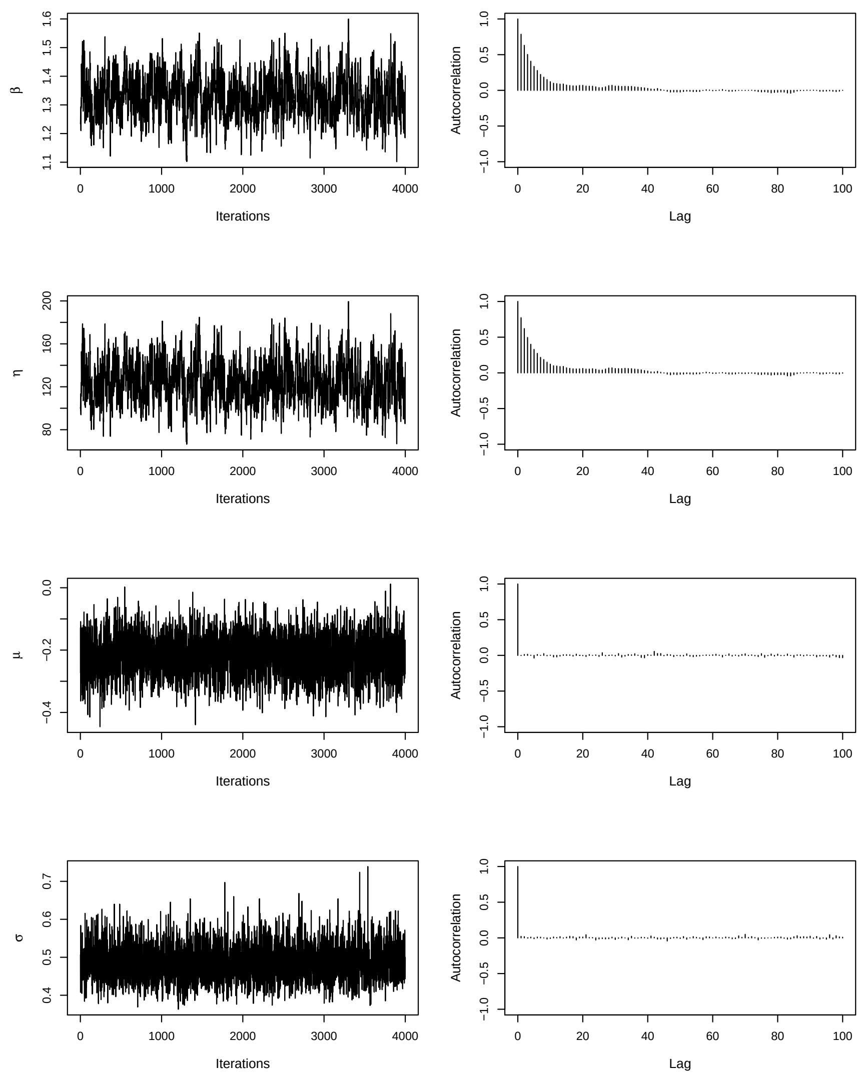

Figure 2: Trace plots (left) and corresponding autocorrelation plots (right) of the chain after burn-in and thinning, under the PLP-Lognormal model setting with $(\beta, \eta, \mu, \sigma) = (1.25, 100, -0.2, 0.5)$, $n = 30$ and $\Delta = 630$.

S2 Technical Notes

S2.1 Intensity of the failure process

In this section, we add some more discussions on the system ROCOF. Let $\mathcal{F}_t$ be the filtration (i.e. the history just up to time t) of the process up to and including time t. For the filtration, we have $\mathcal{F}_{t^-} = \bigcup_{s<t} \mathcal{F}_s$. The assumption that two failure events cannot occur at exactly the same time is plausible in most settings and is also retained in our study. The intensity for the failure process is defined formally as

$$\lambda(t \mid \mathcal{F}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\{\Delta N(t) = 1 \mid \mathcal{F}_{t^-}\}}{\Delta t}. \tag{12}$$

The system ROCOF $\lambda(t)$ in Equation (1) of the paper is accordingly defined by $\lambda(t \mid \mathcal{F}_{t^-})$.

S2.2 Product integral

In this section, we clarify the product integral, i.e., the term $\prod_{u \in [L,R]} \{\lambda(u)\}^{dN(u)}$ in Equation (2) of the paper. Consider a generic system that follows the defined failure model in the paper. Suppose m failures occur in the observation window $[L, R]$ denoted by an ordered collection $\{t_l : 1 \le l \le m\}$, $m = N(R)$. Consider partitions $L = u_0 < u_1 < \cdots < u_S = R$ of the observation window $[L, R]$ and define $\Delta u_s = u_{s+1} - u_s$, $s = 0, 1, \cdots, S$, where $u_{S+1} = u_S^+$. As introduced in Section 2.6 in Andersen et al. (2012), the term $\prod_{u \in [L,R]} \{\lambda(u)\}^{dN(u)}$ in Equation (2) of the paper is actually a limit of approximating the finite product $\lim_{S \to \infty} \prod_{s=0}^{S} \{\lambda(u_s)\}^{\Delta N(u_s)}$. As $S \to \infty$ and $\Delta u_s \to 0$, the m intervals that contain the event times $t_1, \ldots, t_m$ have $\Delta N(u_s) = 1$; for all others $\Delta N(u_s) = 0$. Hence, we have

$$\begin{aligned}
&\prod_{u \in [L,R]} \{\lambda(u)\}^{dN(u)} \\
=& \lim_{S \to \infty} \prod_{s=0}^{S} \{\lambda(u_s)\}^{\Delta N(u_s)} \\
=& \prod_{l=1}^{m} \lambda(t_l).
\end{aligned}$$

According to Theorem 2.1 in Cook et al. (2007), the probability density of the observed failure events $\{t_l : 1 \le l \le m\}$ for the failure process with intensity (12) over the observation window

$[L, R]$ is

$$\prod_{l=1}^{m} \lambda\left(t_l\right) \cdot \exp\left\{-\int_L^R \lambda(u) du\right\}. \tag{13}$$

With the above settings, the likelihood function in Equation (2) of the paper is obtained by integrating $\mathbf{A}$ in (13) out:

$$L(\boldsymbol{\theta}; \mathbf{D}) = \underbrace{\int_0^\infty \cdots \int_0^\infty}_{\lfloor R/\Delta \rfloor} \exp(\Lambda(L) - \Lambda(R)) \prod_{l=1}^{m} \lambda\left(t_l\right) dF_A\big(a_1\big) \cdots dF_A\big(a_{\lfloor R/\Delta \rfloor}\big).$$

S2.3 Complete-data log-likelihood function

This section presents detailed derivations to obtain complete-data log-likelihood function. Consider a generic system. Denote $\mathbf{D} = \{L, R, N(t) : \ t \in [L, R]\}$ as the observed data and $\mathbf{A} = \{A_k : k\Delta \leq R\}$ as the set of unobserved PM random effects. When n independent systems are available, the recurrent failure data from n systems are thus $\mathbb{D} = \{\mathbf{D}_i : i = 1, 2, \cdots, n\}$, and $\mathbb{A} = \{\mathbf{A}_i : i = 1, 2, \cdots, n\}$ is the collection of all the PM random effects. Based on the complete data $\mathbb{C} = \mathbb{D} \cup \mathbb{A}$, the complete-data likelihood function is given by

$$L(\boldsymbol{\theta}; \mathbb{C}) = \prod_{i=1}^{n} L(\boldsymbol{\theta}; \mathbf{D}_i, \mathbf{A}_i) = \prod_{i=1}^{n} \left[\exp(\Lambda(L_i) - \Lambda(R_i)) \prod_{t \in [L_i, R_i]} \lambda_i(t)^{dN_i(t)} \prod_{k=1}^{\lfloor R_i/\Delta \rfloor} f_A(a_{ik})\right].$$

Therefore, the complete-data log-likelihood function $\ell\left(\boldsymbol{\theta}; \mathbb{C}\right)$ in Equation (3) of the paper is obtained by simply taking the log of the above equation.

S2.4 Q-function

Detailed derivations to obtain the Q-function (i.e. Equation (4) of the paper) are given by

$$\begin{aligned} \mathrm{E}\left[\ell\left(\boldsymbol{\theta}; \mathbb{C}\right) \mid \mathbb{D}, \boldsymbol{\theta}^{(j)}\right] &= \sum_{i=1}^{n} \mathrm{E}\left[\ell(\boldsymbol{\theta}; \mathbf{D}_i, \mathbf{A}_i) \mid \mathbf{D}_i, \boldsymbol{\theta}^{(j)}\right] \\ &= \sum_{i=1}^{n} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{\lfloor R_i/\Delta \rfloor} \ell(\boldsymbol{\theta}; \mathbf{D}_i, \mathbf{A}_i) f_{\mathbf{A}_i}\left(\mathbf{a}_i \mid \mathbf{D}_i, \boldsymbol{\theta}^{(j)}\right) da_{i1} \cdots da_{i\lfloor R_i/\Delta \rfloor}, \end{aligned}$$

where the first equality follows from the independence among n systems and $f_{\mathbf{A}_i}(\mathbf{a}_i \mid \mathbf{D}_i, \boldsymbol{\theta}^{(j)})$ is the conditional PDF of random effects (i.e. Equation (5) of the paper) , which can be expressed as

$$f_{\mathbf{A}_i}\left(\mathbf{a}_i \mid \mathbf{D}_i, \boldsymbol{\theta}^{(j)}\right) = \frac{L(\boldsymbol{\theta}^{(j)}; \mathbf{D}_i, \mathbf{A}_i)}{L(\boldsymbol{\theta}^{(j)}; \mathbf{D}_i)}.$$

S3 The Non-periodic or Condition-based PM Setting

In this section, we discuss the applicability of the proposed method to the non-periodic PM setting or the condition-based PM setting. In keeping with the notation in Doyen and Gaudoin (2006) and Doyen and Gaudoin (2011), we define the failure process $N(t)$ that counts failures, the PM counting process $M(t)$ that counts PM actions and $K(t)$ that counts both failures and PM actions over $[0, t]$. Denote $\{C_i\}_{i\geq 0}$ the failure and PM time sequence ($C_0 = 0$). Denote $\mathcal{F}_t, t \geq 0$, as the natural filtration generated by the history of the processes $N(t)$ and $M(t)$ and the hypothetical observation of PM random effects $\{A_j\}_{j\geq 1}$ up to and including time t: $\mathcal{F}_t = \sigma(\{N(s), M(s), A_{M(s)}\}_{0\leq s\leq t})$. For the filtration, let $\mathcal{F}_{t^-} = \bigcup_{s<t} \mathcal{F}_s$.

- *Non-periodic PM policy.*

  Consider a repairable system. The intensity for the failure process is defined by

$$\lambda(t \mid \mathcal{F}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\{\Delta N(t) = 1 \mid \mathcal{F}_{t^-}\}}{\Delta t}.$$

  Next, when we consider a non-periodic time-based PM program to mitigate the risk of failures, the PM time sequence is also predictable with respect to the filtration, i.e., the next PM time is a deterministic function of the history of the PM and failure processes (Doyen and Gaudoin, 2011). Following the model setting in the paper, we consider the minimal repair assumption and explicitly model the effect of PM as a multiplicative random effect on ROCOFs. Thus, for the non-periodic PM policy, the system ROCOF is formulated as

$$\lambda(t) \triangleq \lambda(t \mid \mathcal{F}_{t^-}) = \prod_{j=1}^{M(t^-)} A_j \lambda_0(t),$$

where random effects A_j, $j = 1, \cdots, M(t^-)$, are assumed to be i.i.d. nonnegative random variables representing the effects of each PM action.

- *Condition-based PM policy.*

  Next, we consider a repairable system under a condition-based PM program. In this case, the PM time sequence is not predictable with respect to the filtration, i.e., PM times are not deterministic. The intensity function for the failure process is defined by

$$\lambda^N(t \mid \mathcal{F}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\{\Delta N(t) = 1 \mid \mathcal{F}_{t^-}\}}{\Delta t}.$$

  The intensity function for the condition-based PM process is defined by

$$\lambda^M(t \mid \mathcal{F}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\{\Delta M(t) = 1 \mid \mathcal{F}_{t^-}\}}{\Delta t}.$$

  The failure and PM intensities completely characterize the failure and PM processes. The competing risks approach developed in the context of maintenance is commonly used to derive the failure and PM intensities (Doyen and Gaudoin, 2006; Dijoux and Gaudoin, 2009). To use the competing risks approach, we introduce the concept of risk variables. After the kth PM or repair action, the time to the next failure (i.e., the next repair action) is a random variable Z_{k+1}. However, the failure can be avoided by a condition-based PM action at a random time, Y_{k+1}. The time until the next failure or PM is $W_{k+1} = \min(Z_{k+1}, Y_{k+1})$. The random variables Y_{k+1} and Z_{k+1} are called the risk variables.

  The subhazard rates of the first failure and PM risk variables Z_1 and Y_1 are accordingly defined as below (Doyen and Gaudoin, 2006):

$$\begin{aligned}
\lambda_c(w) &= \lim_{\Delta w \to 0} \frac{1}{\Delta w} \Pr\left(w < Z_1 \le w + \Delta w, Z_1 < Y_1 \mid W_1 > w\right), \\
\lambda_p(w) &= \lim_{\Delta w \to 0} \frac{1}{\Delta w} \Pr\left(w < Y_1 \le w + \Delta w, Y_1 \le Z_1 \mid W_1 > w\right).
\end{aligned} \tag{14}$$

  Following the generalized virtual age model in Doyen and Gaudoin (2006), we define a sequence of random variables $\{V_k\}_{k \ge 1}$, with $V_0 = 0$, called effective ages. We assume that after the kth failure or PM the risk variables Y_{k+1} and Z_{k+1} behave like the risk variables of a new system, never maintained before V_k. Under the assumption, the connections between failure and PM intensities and subhazard rates are established by the following equations:

$$\begin{aligned} \lambda^N(t \mid \mathcal{F}_{t^-}) &= \lambda_c\left(V_{K(t^-)} + t - C_{K(t^-)}\right), \\ \lambda^M(t \mid \mathcal{F}_{t^-}) &= \lambda_p\left(V_{K(t^-)} + t - C_{K(t^-)}\right), \end{aligned} \tag{15}$$

where $\lambda_c(\cdot)$ and $\lambda_p(\cdot)$ are the subhazard rates of the first latent repair and PM times, respectively, defined in (14). As with our proposed model for the periodic PM setting, we can still consider the minimal repair assumption and explicitly model the PM random effects, which accounts for the potential PM adverse effects as well. By blending the idea of ARA$_\infty$ model from Doyen and Gaudoin (2004), we have

$$V_{K(t)} = \begin{cases} A_{M(t)}(V_{K(t)-1} + C_{K(t)} - C_{K(t)-1}), & dM(t) = 1, \\ V_{K(t)-1} + C_{K(t)} - C_{K(t)-1}, & dM(t) = 0, \end{cases}$$

where random effects $A_{M(t)}$ are assumed to be i.i.d. nonnegative random variables representing the effects of each PM action and $dM(t) = \lim_{\Delta t \to 0} M((t + \Delta t)^-) - M(t^-)$.

We next discuss the modeling method on the dependence between PM and repair actions. It is sufficient to express the dependence between the risk variables Y_1 and Z_1 since the failure and PM intensities are determined by them, as shown in (15). Though the independent risks assumption is common in competing risks approach literature (Cox, 1959; Gail, 1975; Crowder, 2001), it is not realistic since PM and repair actions are linked through the degradation process. To characterize the dependence between PM and repair actions, the dependent competing risks model, e.g., the alert-delay model, is introduced in Dijoux and Gaudoin (2009). To be specific, the alert-delay model assumes that the link between the PM and repair action risk variables is as follows:

$$Y_1 = pZ_1 + \mathcal{E},$$

where $p \in [0, 1]$ and Z_1 and $\mathcal{E}$ are two independent positive random variables. Commonly used distributions for Z_1 and $\mathcal{E}$ are the exponential distribution. The identifiability of the alert-delay model is also proved for $p \neq 1$ in Dijoux and Gaudoin (2009). Under this dependence assumption, the subhazard rates are readily derived, and thus the failure and PM intensities can be computed by (15). The detailed derivations can be found in Dijoux and Gaudoin (2009).

S4 Gain of the quasi-Monte Carlo method

To illustrate the gain of the quasi-Monte Carlo method, we carefully investigate the Monte Carlo method and quasi-Monte Carlo method from both theoretical and empirical perspectives, as detailed below.

- *Theoretical perspective.* When using N samples, the Monte Carlo integration method yields a probabilistic error bound of $O(N^{-1/2})$ independent of dimensions (Niederreiter, 1992; Caflisch, 1998). The quasi-Monte Carlo integration method can yield a deterministic error bound of $O(N^{-1}(\log N)^d)$, where d is the dimension of the integral (Caflisch, 1998). To be exact, this is the upper bound of the error. The convergence rate of the quasi-Monte Carlo integration method is generally faster than that theoretical bound (Asmussen and Glynn, 2007).

  The accuracy of the quasi-Monte Carlo method generally increases faster than Monte Carlo method as N increases. However, this advantage should be checked judiciously since it is only guaranteed if N is large enough and the dimension of the integral d is not large. It is widely believed that quasi-Monte Carlo method is applicable to problems in a dimension of moderate size, say, for $d < 15$ (Wang and Fang, 2003). To check that, the dimension of each integral in Equation (4) of the paper (i.e. Q-function) is less than 10 in our studies. As such, with the carefully chosen sample size N, the advantage of the faster convergence rate of the quasi-Monte Carlo method can be well retained in our studies.

- *Empirical perspective.* To illustrate the gain of quasi-Monte Carlo method, we conduct simulation studies to check the average calculation time of Monte Carlo EM (MCEM) and quasi-Monte Carlo EM (QMCEM) algorithms. We consider the power-law baseline ROCOF $\lambda_0(t) = (\beta/\eta)\,(t/\eta)^{\beta-1}$ with $(\beta, \eta) = (1.25, 100)$ and Lognormal(μ, σ^2) PM random effect with $(\mu, \sigma) \in \{(-1/5, 1/2), (-1/8, 1/2), (-1/10, 1/2)\}$ in the simulation. Two levels of the

Table 5: The average calculation time (in seconds) of Monte Carlo EM (MCEM) and quasi-Monte Carlo EM (QMCEM) algorithms under the power-law baseline ROCOF and lognormal PM random effect setting.

n	$(\beta,\ \eta,\ \mu,\ \sigma)$	Method	$\Delta = 365$	$\Delta = 540$
30	$(1.25, 100, -1/5, 1/2)$	MCEM algorithm	1554.39	1063.51
		QMCEM algorithm	1482.60	1215.37
	$(1.25, 100, -1/8, 1/2)$	MCEM algorithm	3197.65	1404.68
		QMCEM algorithm	2354.67	1270.73
	$(1.25, 100, -1/10, 1/2)$	MCEM algorithm	4719.59	1220.60
		QMCEM algorithm	3706.83	1324.08
45	$(1.25, 100, -1/5, 1/2)$	MCEM algorithm	1925.21	1155.26
		QMCEM algorithm	1561.40	1354.13
	$(1.25, 100, -1/8, 1/2)$	MCEM algorithm	3205.59	1388.43
		QMCEM algorithm	2627.82	1279.86
	$(1.25, 100, -1/10, 1/2)$	MCEM algorithm	5361.79	1366.29
		QMCEM algorithm	3992.40	1537.02
60	$(1.25, 100, -1/5, 1/2)$	MCEM algorithm	2211.40	1227.00
		QMCEM algorithm	1657.05	1471.24
	$(1.25, 100, -1/8, 1/2)$	MCEM algorithm	3171.34	1363.68
		QMCEM algorithm	3066.79	1306.54
	$(1.25, 100, -1/10, 1/2)$	MCEM algorithm	6075.84	1481.73
		QMCEM algorithm	4276.59	1622.46

PM time interval are considered: $\Delta \in \{365, 540\}$. The sample size design and observation window design are the same as in the Section 5.1 of the paper. All computations were conducted on an Intel(R) Xeon(R) CPU E5-2698 v4 (2.20 GHz). The results are summarized in Table 5. As can be seen, the average calculation time of MCEM algorithm is generally longer than QMCEM algorithm under different settings.

In addition, to further illustrate the convergence rate in Monte Carlo method and quasi-Monte Carlo method, we next compute the relative error of the Monte Carlo and quasi-Monte Carlo approximations of the Q-function, $|Q^\star(\boldsymbol{\theta} \mid \tilde{\boldsymbol{\theta}}) - Q(\boldsymbol{\theta} \mid \tilde{\boldsymbol{\theta}})|/Q(\boldsymbol{\theta} \mid \tilde{\boldsymbol{\theta}})$, where $Q^\star(\boldsymbol{\theta} \mid \tilde{\boldsymbol{\theta}})$ denotes the approximations of Q-function. For example, we plot the relative approximation error as a function of the sample size N from 10^1 to 10^4 by considering the power-law baseline ROCOF and lognormal PM random effect with $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}} = (\beta, \eta, \mu, \sigma) = (1.25, 100, -1/5, 1/2)$, PM time interval $\Delta = 365$, and sample size $n = 30$. The results are presented in Figure 3. As can be seen, the relative approximation error of the Monte Carlo method decreases as the increase of sample size N, roughly at a rate of $1/\sqrt{N}$, and the quasi-Monte Carlo method is generally better at all sample sizes and appears to show a faster convergence rate.

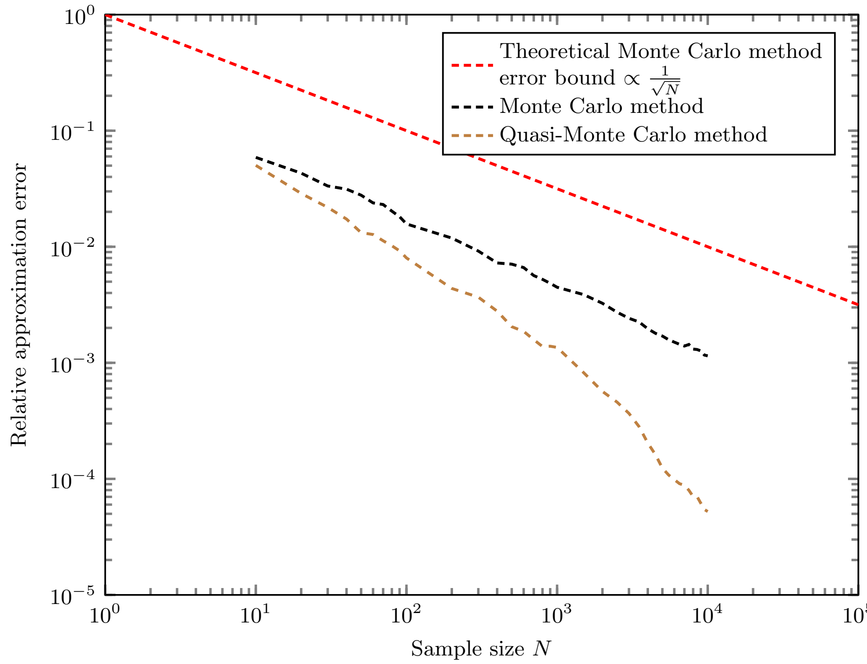

Figure 3: Relative approximation error as a function of the sample size N from 10^1 to 10^4.

S5 Martingale residuals

In this section, we introduce the concept of martingale residuals. Denote $\Lambda = (\Lambda(t))_{t\geq 0}$ as the cumulative intensity process of the counting process $N = (N(t))_{t\geq 0}$, such that $\Lambda(t) = \int_0^t \lambda(s)ds$ and $\lambda(s)$ is the intensity. The counting process martingale is defined by $M(t) = N(t) - \Lambda(t)$.

Consider a generic repairable system that follows the defined failure model in the paper. We denote $\{t_l : 1 \leq l \leq m\}$ as the ordered collection of its failure time epochs over the observation window $[L, R]$, where $m = N(R)$ (i.e., the number of failures observed from the system) and $t_0 \triangleq L$. With the assumed model, the failure process follows a Poisson process after conditioning on the frailties (i.e., PM random effects) and the observation windows. By plugging in parameter estimates, the corresponding martingale residuals are obtained as $\widehat{M}(t_l) = N(t_l) - \Lambda(t_l; \hat{\boldsymbol{\theta}}, \hat{\mathbf{A}})$, $1 \leq l \leq m$. Based on the martingale residuals, we can simply adopt the Kolmogorov-Smirnov (KS) type test statistic (see also Chauvel et al. (2016)) to evaluate the proximities between the observed data and the estimated models, which is defined as

$$\begin{aligned}\mathrm{KS}(\hat{\boldsymbol{\theta}}, \hat{\mathbb{A}}) &= \sup_{1\le l\le m} \left|\widehat{M}(t_l)\right| \\ &= \sup_{1\le l\le m} \left|N(t_l) - \Lambda(t_l; \hat{\boldsymbol{\theta}}, \hat{\mathbf{A}})\right|.\end{aligned}$$

As with Section 4 of the paper, a parametric bootstrap method (Efron and Tibshirani, 1994) can be next employed to compute the quantiles of the above test statistic distribution.

S6 Addtional Figures and Tables

S6.1 Additional parameter estimation results

As indicated in Section 5.1 of the paper, when using `optim()` in R (R Core Team, 2020), the constraints on the range of parameters need to be first specified. The range of parameters in the baseline ROCOF is as below:

- For *power-law process* $\lambda_0(t) = (\beta/\eta)\,(t/\eta)^{\beta-1}$, we have parameters $\eta, \beta > 0$.
- For *log-linear law process* $\lambda_0(t) = \exp(\eta + \beta t)$, we have parameters $\eta, \beta \in \mathbb{R}$.

The range of parameters in the distribution of PM random effects is as below:

- For the Gamma(α, ν) random effect, we have the shape parameter $\alpha > 0$ and the rate parameter $\nu > 0$.
- For the Lognormal(μ, σ^2) random effect, we have parameters $\mu \in \mathbb{R}$ and $\sigma > 0$.

The additional parameter estimation results under PLP-Lognormal, PLP-Gamma, LLP-Lognormal and LLP-Gamma model settings are presented in Tables 6-21. Besides, we compute the relative estimation error $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})/\boldsymbol{\theta}$ in each replication, and then present box plots of the relative estimation error under the PLP-Lognormal model to check how the estimated parameters spread out. The box plots of the relative estimation error based on the 1,000 replications are plotted in Figures 4-7. As demonstrated by the figures, the relative estimation error generally decreases as n increases under the PLP-Lognormal model setting. Overall, the performance of the EM algorithm in parameter estimations is satisfactory under different settings.

S6.2 Numerical issues

In this section, we discuss the numerical instability problems, e.g., arithmetic overflows, which arise when considering the log-linear baseline model. As shown in Tables 14-21, when examining the log-linear law baseline ROCOF model, the true magnitudes of the parameters β and η are set to be 0.0005 and -4.5, respectively. The numerical problems may arise due to huge differences in scales of the parameters. For example, when using quasi-Newton type method in `optim()`, the convergence can be compromised due to the ill-conditioned problem on the Hessian matrix of the Q-function. Besides, similar problems can also arise when computing the inverse of the Hessian matrix of the Q-function to obtain confidence intervals. To overcome these hurdles, we resort to the built-in control parameter "parscale" in `optim()` to rescale the parameters so that unit change of rescaled parameters have nearly the same impact on the objective function.

In addition, according to our numerical experience, the numerical stability issues seem to arise more frequently in the quasi-Newton BFGS method than the Nelder-Mead method. The is because BFGS method performs a line search in the direction of the gradient, which can lead to extreme parameter values in the line search step and cause numerical issues. Luckily, we note that the `optim()` can generally handle it without generating errors even though the objective function returns an "Inf" or "NaN" value. This built-in feature of the `optim()` in R (R Core Team, 2020) helps to maintain the numerical stability.

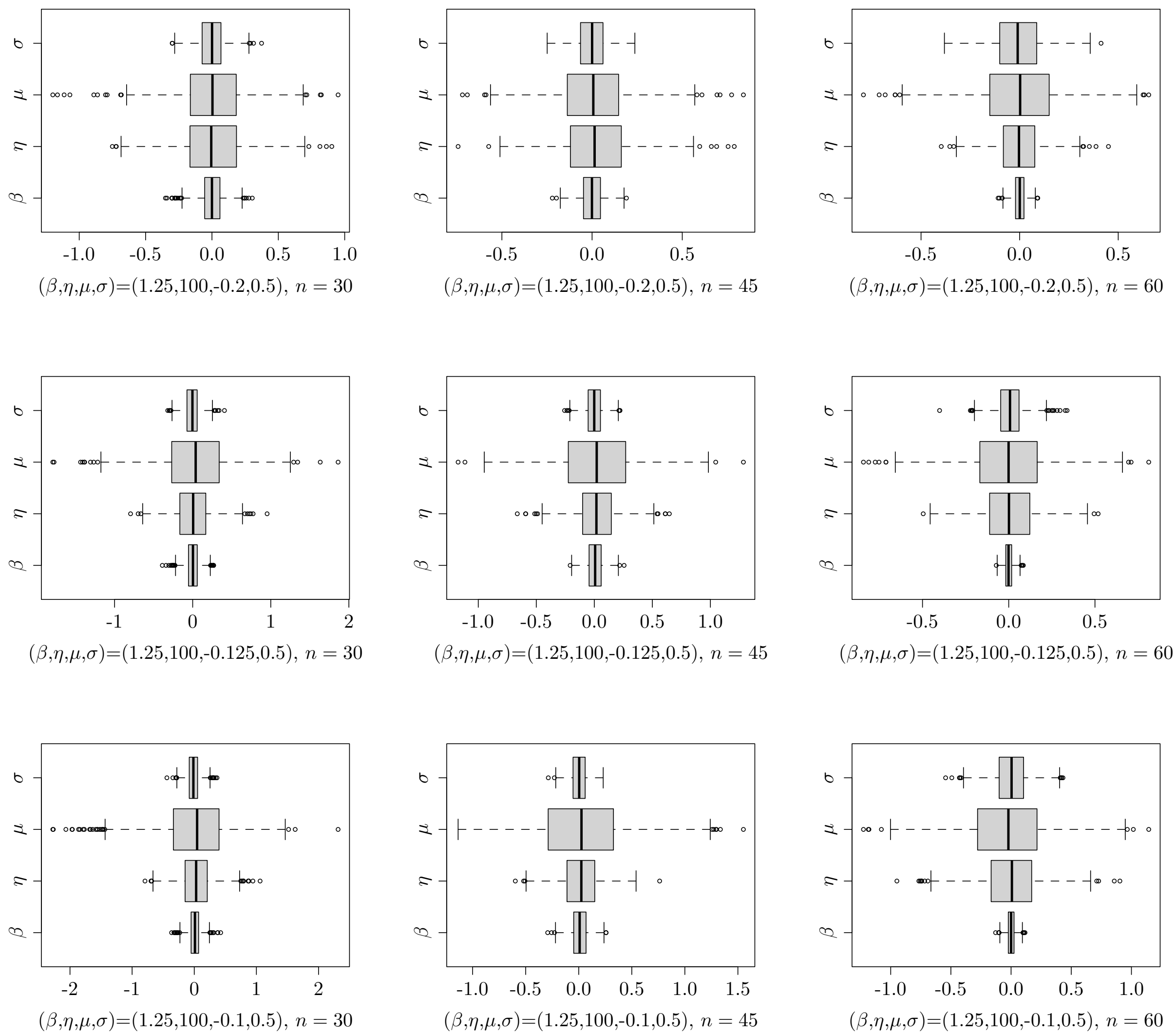

Figure 4: Box plots of the relative estimation error of $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\eta}, \hat{\mu}, \hat{\sigma})$, under the PLP-Lognormal model with PM time interval $\Delta = 365$.

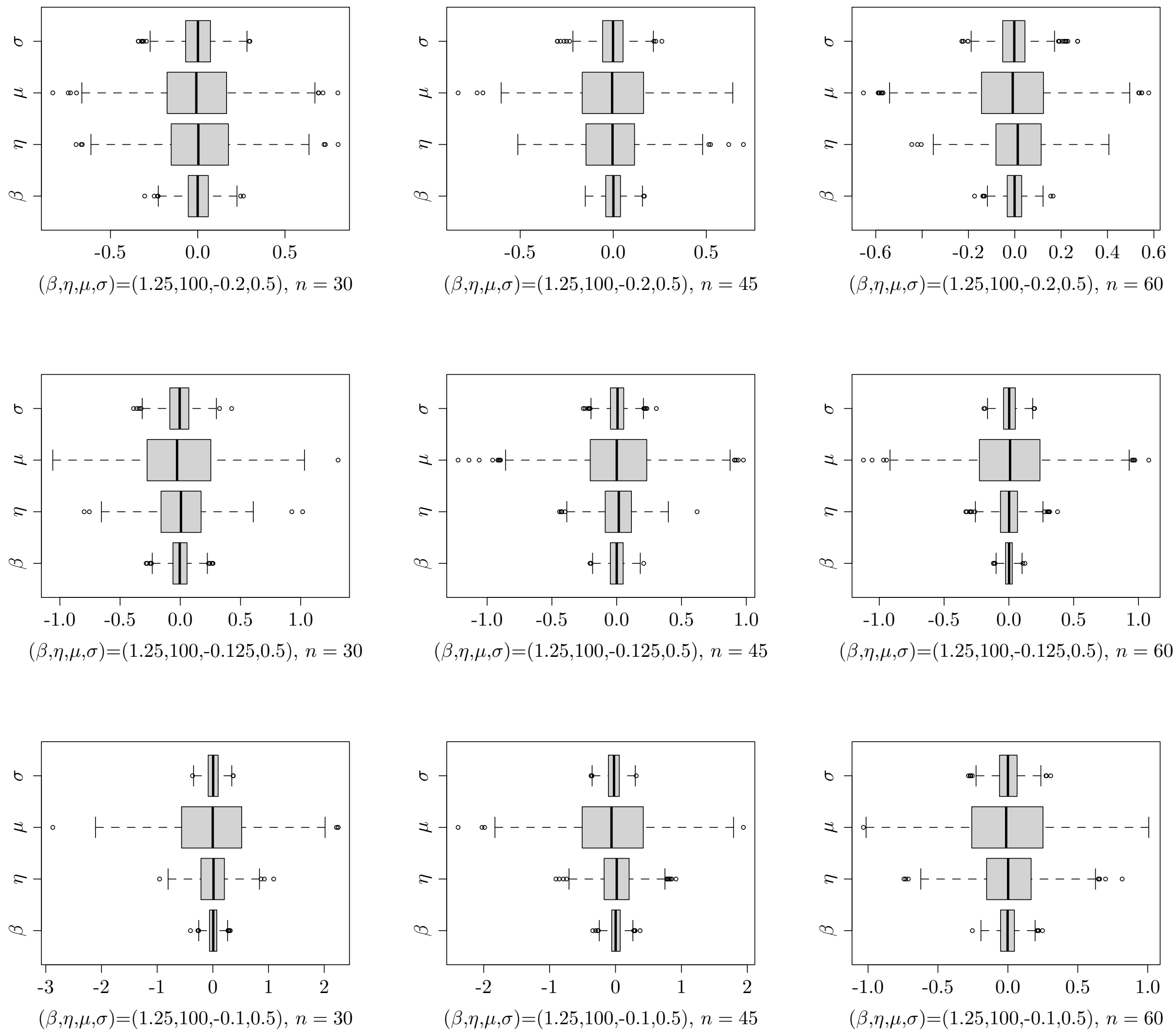

Figure 5: Box plots of the relative estimation error of $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\eta}, \hat{\mu}, \hat{\sigma})$, under the PLP-Lognormal model with PM time interval $\Delta = 450$.

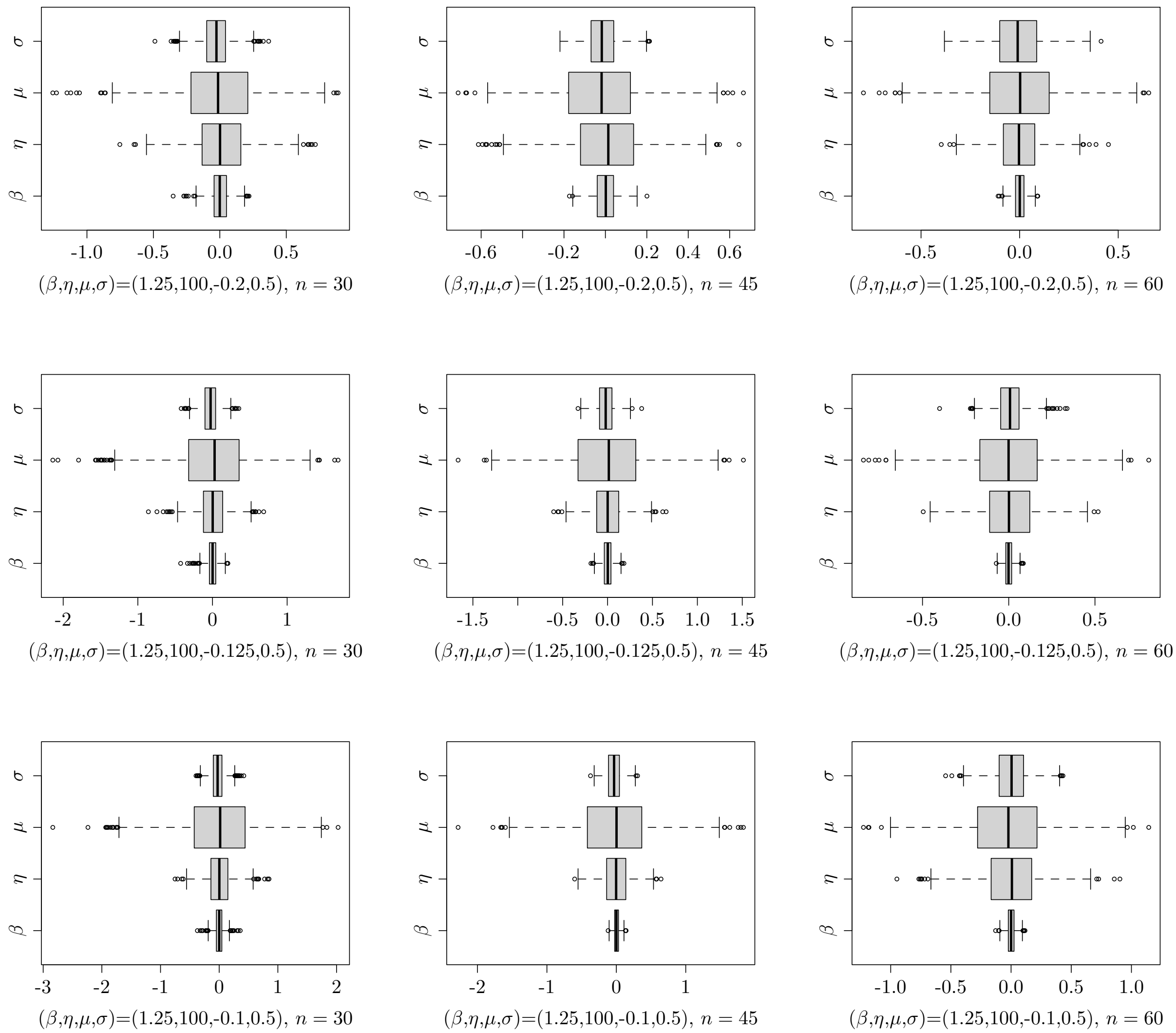

Figure 6: Box plots of the relative estimation error of $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\eta}, \hat{\mu}, \hat{\sigma})$, under the PLP-Lognormal model with PM time interval $\Delta = 540$.

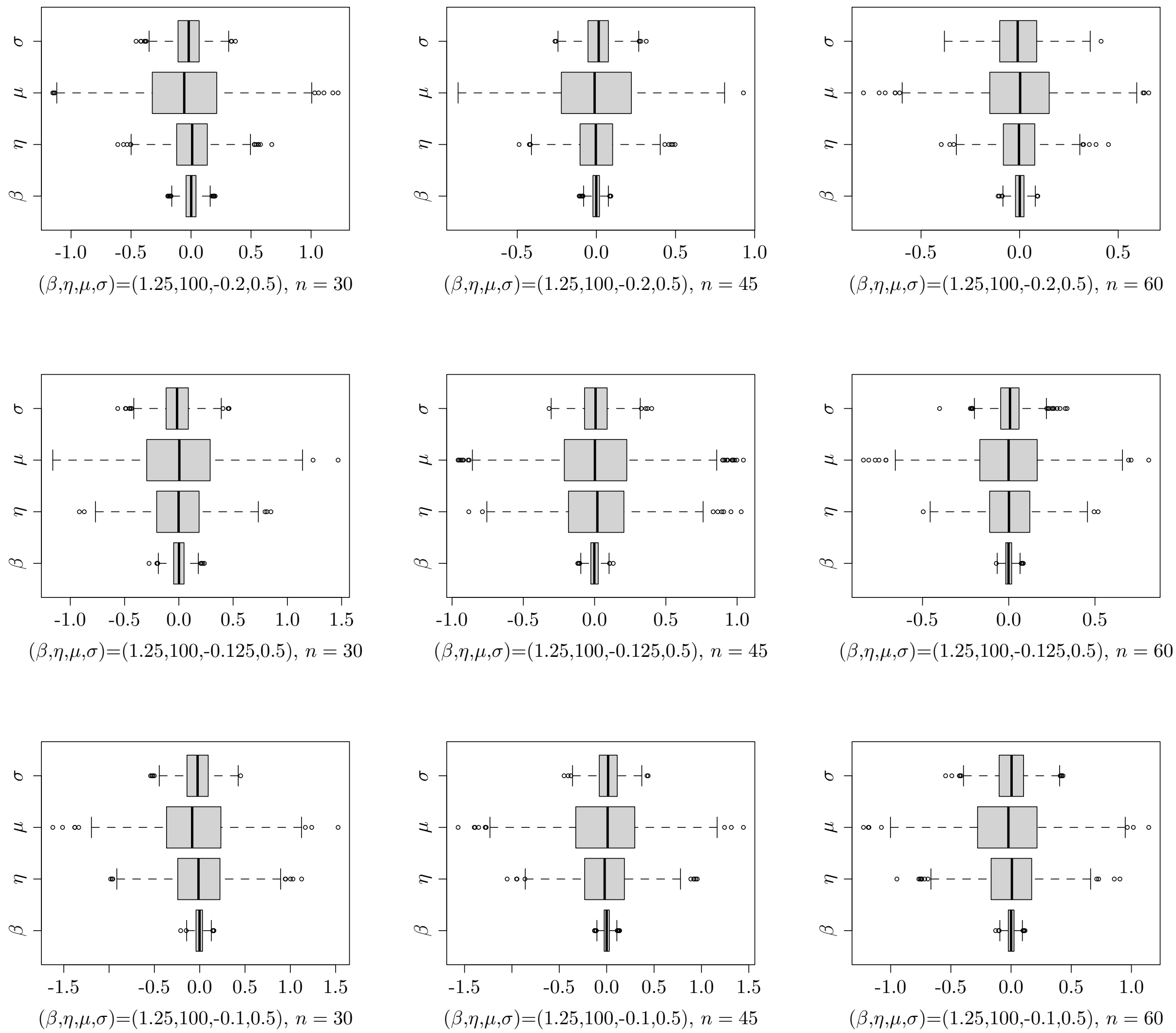

Figure 7: Box plots of the relative estimation error of $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\eta}, \hat{\mu}, \hat{\sigma})$, under the PLP-Lognormal model with PM time interval $\Delta = 630$.

Table 6: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the PLP-Lognormal model with the PM time interval: $\Delta = 365$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/5, 1/2)$	1482.60	bias ($\times 10^{-2}$)	0.145	156.898	-0.086	-0.019
			RMSE	0.115	26.084	0.053	0.051
			CP (asymptotic CI)	0.930	0.941	0.955	0.977
			CP (bootstrap CI)	0.938	0.949	0.957	0.973
	$(1.25, 100, -1/8, 1/2)$	2354.67	bias ($\times 10^{-2}$)	0.286	113.190	-0.405	-0.334
			RMSE	0.112	24.902	0.057	0.048
			CP (asymptotic CI)	0.915	0.937	0.939	0.972
			CP (bootstrap CI)	0.936	0.946	0.951	0.971
	$(1.25, 100, -1/10, 1/2)$	3706.83	bias ($\times 10^{-2}$)	1.265	334.40	-0.200	-0.704
			RMSE	0.126	26.776	0.061	0.050
			CP (asymptotic CI)	0.906	0.914	0.911	0.946
			CP (bootstrap CI)	0.931	0.937	0.936	0.961
45	$(1.25, 100, -1/5, 1/2)$	1561.40	bias ($\times 10^{-2}$)	0.131	167.256	-0.052	0.016
			RMSE	0.087	20.458	0.044	0.044
			CP (asymptotic CI)	0.955	0.959	0.951	0.970
			CP (bootstrap CI)	0.952	0.963	0.955	0.972
	$(1.25, 100, -1/8, 1/2)$	2627.82	bias ($\times 10^{-2}$)	0.195	166.732	-0.262	-0.278
			RMSE	0.095	19.142	0.045	0.039
			CP (asymptotic CI)	0.925	0.950	0.949	0.965
			CP (bootstrap CI)	0.933	0.948	0.957	0.977
	$(1.25, 100, -1/10, 1/2)$	3992.40	bias ($\times 10^{-2}$)	1.203	322.932	-0.202	0.661
			RMSE	0.112	19.973	0.047	0.042
			CP (asymptotic CI)	0.920	0.934	0.941	0.953
			CP (bootstrap CI)	0.936	0.944	0.955	0.962
60	$(1.25, 100, -1/5, 1/2)$	1657.05	bias ($\times 10^{-2}$)	0.119	260.642	-0.031	0.055
			RMSE	0.074	17.664	0.036	0.032
			CP (asymptotic CI)	0.959	0.957	0.956	0.973
			CP (bootstrap CI)	0.954	0.958	0.965	0.970
	$(1.25, 100, -1/8, 1/2)$	3066.79	bias ($\times 10^{-2}$)	0.173	246.082	-0.146	0.453
			RMSE	0.079	18.062	0.040	0.035
			CP (asymptotic CI)	0.924	0.940	0.941	0.969
			CP (bootstrap CI)	0.938	0.952	0.949	0.971
	$(1.25, 100, -1/10, 1/2)$	4276.59	bias ($\times 10^{-2}$)	1.471	352.213	-0.462	0.737
			RMSE	0.083	18.606	0.037	0.035
			CP (asymptotic CI)	0.918	0.931	0.942	0.963
			CP (bootstrap CI)	0.932	0.941	0.959	0.960

Table 7: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the PLP-Lognormal model with the PM time interval: $\Delta = 450$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/5, 1/2)$	1346.54	bias ($\times 10^{-2}$)	0.142	41.868	−0.171	0.086
			RMSE	0.106	22.942	0.054	0.053
			CP (asymptotic CI)	0.954	0.952	0.938	0.959
			CP (bootstrap CI)	0.957	0.949	0.952	0.963
	$(1.25, 100, -1/8, 1/2)$	1853.72	bias ($\times 10^{-2}$)	0.093	55.931	0.087	−0.077
			RMSE	0.115	24.651	0.047	0.058
			CP (asymptotic CI)	0.952	0.961	0.963	0.955
			CP (bootstrap CI)	0.956	0.958	0.967	0.958
	$(1.25, 100, -1/10, 1/2)$	1452.41	bias ($\times 10^{-2}$)	0.893	113.262	0.323	0.431
			RMSE	0.125	31.256	0.078	0.062
			CP (asymptotic CI)	0.931	0.940	0.923	0.948
			CP (bootstrap CI)	0.944	0.947	0.952	0.965
45	$(1.25, 100, -1/5, 1/2)$	1622.76	bias ($\times 10^{-2}$)	0.135	52.633	0.166	−0.075
			RMSE	0.072	18.453	0.047	0.042
			CP (asymptotic CI)	0.947	0.948	0.941	0.953
			CP (bootstrap CI)	0.949	0.948	0.955	0.947
	$(1.25, 100, -1/8, 1/2)$	2341.74	bias ($\times 10^{-2}$)	0.090	34.931	−0.062	0.059
			RMSE	0.086	15.144	0.044	0.039
			CP (asymptotic CI)	0.947	0.955	0.959	0.961
			CP (bootstrap CI)	0.953	0.959	0.951	0.949
	$(1.25, 100, -1/10, 1/2)$	1789.68	bias ($\times 10^{-2}$)	0.504	73.262	0.120	−0.309
			RMSE	0.123	28.251	0.066	0.058
			CP (asymptotic CI)	0.944	0.958	0.941	0.960
			CP (bootstrap CI)	0.953	0.966	0.953	0.970
60	$(1.25, 100, -1/5, 1/2)$	2044.81	bias ($\times 10^{-2}$)	0.096	60.649	0.180	−0.073
			RMSE	0.058	13.894	0.041	0.037
			CP (asymptotic CI)	0.955	0.950	0.947	0.961
			CP (bootstrap CI)	0.947	0.948	0.956	0.968
	$(1.25, 100, -1/8, 1/2)$	2637.54	bias ($\times 10^{-2}$)	0.086	13.258	0.027	0.047
			RMSE	0.051	10.152	0.045	0.033
			CP (asymptotic CI)	0.944	0.947	0.952	0.965
			CP (bootstrap CI)	0.957	0.953	0.946	0.954
	$(1.25, 100, -1/10, 1/2)$	2683.84	bias ($\times 10^{-2}$)	0.123	34.243	−0.105	0.289
			RMSE	0.092	23.561	0.037	0.046
			CP (asymptotic CI)	0.952	0.959	0.953	0.967
			CP (bootstrap CI)	0.955	0.957	0.962	0.973

Table 8: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the PLP-Lognormal model with the PM time interval: $\Delta = 540$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/5, 1/2)$	1215.37	bias ($\times 10^{-2}$)	0.081	145.603	0.237	−1.332
			RMSE	0.091	21.965	0.064	0.058
			CP (asymptotic CI)	0.961	0.953	0.955	0.953
			CP (bootstrap CI)	0.968	0.957	0.963	0.965
	$(1.25, 100, -1/8, 1/2)$	1270.73	bias ($\times 10^{-2}$)	−0.071	83.930	0.043	−1.414
			RMSE	0.086	20.233	0.068	0.058
			CP (asymptotic CI)	0.947	0.954	0.932	0.949
			CP (bootstrap CI)	0.954	0.961	0.951	0.959
	$(1.25, 100, -1/10, 1/2)$	1324.08	bias ($\times 10^{-2}$)	−0.210	63.181	0.223	−1.349
			RMSE	0.098	22.293	0.070	0.060
			CP (asymptotic CI)	0.927	0.937	0.912	0.941
			CP (bootstrap CI)	0.945	0.951	0.937	0.956
45	$(1.25, 100, -1/5, 1/2)$	1354.13	bias ($\times 10^{-2}$)	0.121	84.521	0.235	−0.920
			RMSE	0.073	19.529	0.055	0.045
			CP (asymptotic CI)	0.961	0.963	0.959	0.971
			CP (bootstrap CI)	0.958	0.966	0.953	0.967
	$(1.25, 100, -1/8, 1/2)$	1279.86	bias ($\times 10^{-2}$)	0.068	73.451	0.035	−0.946
			RMSE	0.071	18.791	0.059	0.048
			CP (asymptotic CI)	0.953	0.952	0.966	0.952
			CP (bootstrap CI)	0.960	0.947	0.968	0.971
	$(1.25, 100, -1/10, 1/2)$	1537.02	bias ($\times 10^{-2}$)	0.133	52.881	0.302	−1.431
			RMSE	0.052	20.274	0.059	0.057
			CP (asymptotic CI)	0.951	0.958	0.949	0.967
			CP (bootstrap CI)	0.963	0.961	0.955	0.948
60	$(1.25, 100, -1/5, 1/2)$	1471.24	bias ($\times 10^{-2}$)	−0.108	64.945	0.161	−0.496
			RMSE	0.057	14.245	0.043	0.039
			CP (asymptotic CI)	0.968	0.962	0.964	0.970
			CP (bootstrap CI)	0.972	0.965	0.968	0.970
	$(1.25, 100, -1/8, 1/2)$	1306.54	bias ($\times 10^{-2}$)	0.091	56.144	0.117	−0.492
			RMSE	0.059	14.567	0.041	0.040
			CP (asymptotic CI)	0.966	0.965	0.961	0.959
			CP (bootstrap CI)	0.960	0.968	0.959	0.963
	$(1.25, 100, -1/10, 1/2)$	1622.46	bias ($\times 10^{-2}$)	0.028	73.015	0.333	−0.668
			RMSE	0.063	15.207	0.046	0.037
			CP (asymptotic CI)	0.949	0.956	0.943	0.959
			CP (bootstrap CI)	0.954	0.961	0.949	0.964

Table 9: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the PLP-Lognormal model with the PM time interval: $\Delta = 630$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(1.25, 100, -1/8, 1/2)$	1361.95	bias ($\times 10^{-2}$)	0.092	38.547	0.082	-0.630
			RMSE	0.088	28.724	0.053	0.079
			CP (asymptotic CI)	0.954	0.960	0.947	0.933
			CP (bootstrap CI)	0.955	0.965	0.954	0.952
	$(1.25, 100, -1/10, 1/2)$	2193.20	bias ($\times 10^{-2}$)	-0.426	56.236	0.543	-0.978
			RMSE	0.063	35.721	0.045	0.082
			CP (asymptotic CI)	0.935	0.941	0.937	0.946
			CP (bootstrap CI)	0.947	0.958	0.9432	0.950
45	$(1.25, 100, -1/5, 1/2)$	1241.92	bias ($\times 10^{-2}$)	0.073	26.251	0.131	0.653
			RMSE	0.042	16.025	0.063	0.044
			CP (asymptotic CI)	0.955	0.957	0.952	0.946
			CP (bootstrap CI)	0.956	0.949	0.963	0.958
	$(1.25, 100, -1/8, 1/2)$	1567.05	bias ($\times 10^{-2}$)	0.037	29.043	-0.067	0.572
			RMSE	0.046	27.420	0.041	0.055
			CP (asymptotic CI)	0.957	0.953	0.957	0.939
			CP (bootstrap CI)	0.955	0.961	0.955	0.952
	$(1.25, 100, -1/10, 1/2)$	2499.63	bias ($\times 10^{-2}$)	0.194	41.853	0.327	0.635
			RMSE	0.053	30.542	0.047	0.069
			CP (asymptotic CI)	0.946	0.953	0.948	0.955
			CP (bootstrap CI)	0.952	0.966	0.953	0.964
60	$(1.25, 100, -1/5, 1/2)$	1572.35	bias ($\times 10^{-2}$)	0.034	19.962	0.056	-0.342
			RMSE	0.038	12.056	0.045	0.032
			CP (asymptotic CI)	0.951	0.956	0.947	0.945
			CP (bootstrap CI)	0.957	0.958	0.946	0.951
	$(1.25, 100, -1/8, 1/2)$	1736.82	bias ($\times 10^{-2}$)	0.044	22.451	0.032	0.296
			RMSE	0.031	16.583	0.031	0.045
			CP (asymptotic CI)	0.948	0.955	0.944	0.946
			CP (bootstrap CI)	0.942	0.963	0.948	0.969
	$(1.25, 100, -1/10, 1/2)$	2731.36	bias ($\times 10^{-2}$)	0.083	35.765	0.184	0.428
			RMSE	0.045	26.209	0.036	0.077
			CP (asymptotic CI)	0.948	0.949	0.957	0.952
			CP (bootstrap CI)	0.954	0.963	0.956	0.955

Table 10: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the PLP-Gamma model with the PM time interval: $\Delta = 365$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(1.25, 100, 4, 5)$	1988.53	bias ($\times 10^{-2}$)	0.468	253.675	3.774	5.238
			RMSE	0.084	25.401	1.042	1.217
			CP (asymptotic CI)	0.937	0.936	0.945	0.957
			CP (bootstrap CI)	0.942	0.943	0.958	0.953
	$(1.25, 100, 5, 5)$	2762.04	bias ($\times 10^{-2}$)	0.249	136.057	3.509	4.264
			RMSE	0.073	17.488	1.480	1.611
			CP (asymptotic CI)	0.945	0.932	0.939	0.951
			CP (bootstrap CI)	0.953	0.949	0.941	0.955
45	$(1.25, 100, 4, 5)$	2604.52	bias ($\times 10^{-2}$)	-0.252	171.432	3.063	4.738
			RMSE	0.055	19.932	1.307	0.843
			CP (asymptotic CI)	0.936	0.945	0.961	0.972
			CP (bootstrap CI)	0.949	0.952	0.965	0.956
	$(1.25, 100, 5, 5)$	3542.67	bias ($\times 10^{-2}$)	0.175	95.323	2.891	4.032
			RMSE	0.053	16.360	1.030	1.361
			CP (asymptotic CI)	0.943	0.946	0.941	0.959
			CP (bootstrap CI)	0.955	0.954	0.953	0.948
60	$(1.25, 100, 4, 5)$	3259.05	bias ($\times 10^{-2}$)	0.219	123.643	1.854	2.438
			RMSE	0.031	18.950	0.712	1.258
			CP (asymptotic CI)	0.944	0.951	0.957	0.968
			CP (bootstrap CI)	0.958	0.957	0.961	0.952
	$(1.25, 100, 5, 5)$	4746.24	bias ($\times 10^{-2}$)	-0.187	72.523	1.671	3.467
			RMSE	0.041	15.432	0.798	1.654
			CP (asymptotic CI)	0.939	0.947	0.952	0.953
			CP (bootstrap CI)	0.947	0.950	0.959	0.962

Table 11: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the PLP-Gamma model with the PM time interval: $\Delta = 450$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(1.25, 100, 4, 5)$	1451.76	bias ($\times 10^{-2}$)	0.435	202.532	12.128	7.432
			RMSE	0.093	22.816	2.341	1.933
			CP (asymptotic CI)	0.941	0.939	0.941	0.958
			CP (bootstrap CI)	0.952	0.946	0.953	0.969
	$(1.25, 100, 5, 5)$	1236.64	bias ($\times 10^{-2}$)	0.334	145.251	4.537	3.447
			RMSE	0.060	15.753	0.746	0.771
			CP (asymptotic CI)	0.965	0.946	0.947	0.953
			CP (bootstrap CI)	0.957	0.953	0.951	0.956
45	$(1.25, 100, 4, 5)$	1156.04	bias ($\times 10^{-2}$)	-0.372	121.432	3.063	4.738
			RMSE	0.052	15.932	1.307	0.843
			CP (asymptotic CI)	0.936	0.945	0.961	0.972
			CP (bootstrap CI)	0.949	0.952	0.965	0.956
	$(1.25, 100, 5, 5)$	1266.91	bias ($\times 10^{-2}$)	0.275	95.323	2.891	4.032
			RMSE	0.053	16.360	1.030	1.361
			CP (asymptotic CI)	0.943	0.946	0.941	0.959
			CP (bootstrap CI)	0.955	0.954	0.953	0.948
60	$(1.25, 100, 4, 5)$	2634.67	bias ($\times 10^{-2}$)	0.129	93.643	1.854	2.438
			RMSE	0.031	18.950	0.712	1.258
			CP (asymptotic CI)	0.944	0.951	0.957	0.968
			CP (bootstrap CI)	0.958	0.957	0.961	0.952
	$(1.25, 100, 5, 5)$	2978.40	bias ($\times 10^{-2}$)	-0.261	62.523	1.671	3.467
			RMSE	0.041	15.432	0.798	1.423
			CP (asymptotic CI)	0.939	0.947	0.952	0.953
			CP (bootstrap CI)	0.947	0.950	0.959	0.962

Table 12: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on $1,000$ Monte Carlo replications, under the PLP-Gamma model with the PM time interval: $\Delta = 540$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(1.25, 100, 4, 5)$	1067.32	bias ($\times 10^{-2}$)	-0.336	-156.334	4.501	3.176
			RMSE	0.076	18.958	1.232	0.838
			CP (asymptotic CI)	0.965	0.937	0.943	0.946
			CP (bootstrap CI)	0.954	0.946	0.957	0.955
	$(1.25, 100, 5, 5)$	1205.46	bias ($\times 10^{-2}$)	0.223	163.675	17.832	16.310
			RMSE	0.074	18.401	0.989	1.127
			CP (asymptotic CI)	0.949	0.957	0.960	0.938
			CP (bootstrap CI)	0.954	0.963	0.958	0.951
45	$(1.25, 100, 4, 5)$	1176.01	bias ($\times 10^{-2}$)	0.203	132.201	3.036	2.621
			RMSE	0.058	18.830	0.754	0.652
			CP (asymptotic CI)	0.953	0.948	0.961	0.949
			CP (bootstrap CI)	0.961	0.959	0.968	0.952
	$(1.25, 100, 5, 5)$	1365.87	bias ($\times 10^{-2}$)	0.132	107.753	14.740	13.954
			RMSE	0.073	15.134	0.800	0.862
			CP (asymptotic CI)	0.947	0.943	0.964	0.957
			CP (bootstrap CI)	0.958	0.945	0.966	0.960
60	$(1.25, 100, 4, 5)$	1684.63	bias ($\times 10^{-2}$)	0.137	91.854	7.154	8.607
			RMSE	0.060	15.753	0.625	0.771
			CP (asymptotic CI)	0.964	0.955	0.962	0.957
			CP (bootstrap CI)	0.958	0.958	0.963	0.964
	$(1.25, 100, 5, 5)$	1479.65	bias ($\times 10^{-2}$)	0.085	79.923	10.219	9.018
			RMSE	0.056	13.926	0.743	0.776
			CP (asymptotic CI)	0.958	0.947	0.950	0.953
			CP (bootstrap CI)	0.961	0.950	0.953	0.949

Table 13: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on $1,000$ Monte Carlo replications, under the PLP-Gamma model with the PM time interval: $\Delta = 630$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(1.25, 100, 4, 5)$	1025.51	bias ($\times 10^{-2}$)	0.366	139.856	12.293	14.475
			RMSE	0.085	19.669	1.633	1.346
			CP (asymptotic CI)	0.947	0.935	0.965	0.959
			CP (bootstrap CI)	0.958	0.943	0.968	0.963
	$(1.25, 100, 5, 5)$	988.80	bias ($\times 10^{-2}$)	0.476	98.562	10.385	8.051
			RMSE	0.091	20.859	1.102	0.738
			CP (asymptotic CI)	0.955	0.947	0.971	0.948
			CP (bootstrap CI)	0.951	0.959	0.968	0.954
45	$(1.25, 100, 4, 5)$	1123.57	bias ($\times 10^{-2}$)	0.223	82.392	6.934	8.203
			RMSE	0.076	18.473	0.894	1.032
			CP (asymptotic CI)	0.936	0.955	0.957	0.948
			CP (bootstrap CI)	0.944	0.959	0.966	0.959
	$(1.25, 100, 5, 5)$	1089.83	bias ($\times 10^{-2}$)	0.249	75.391	9.305	8.264
			RMSE	0.076	18.536	0.869	1.024
			CP (asymptotic CI)	0.953	0.943	0.957	0.939
			CP (bootstrap CI)	0.959	0.952	0.964	0.944
60	$(1.25, 100, 4, 5)$	1278.64	bias ($\times 10^{-2}$)	0.161	76.205	7.885	10.417
			RMSE	0.067	14.646	0.794	0.634
			CP (asymptotic CI)	0.957	0.953	0.950	0.948
			CP (bootstrap CI)	0.959	0.962	0.953	0.945
	$(1.25, 100, 5, 5)$	1548.97	bias ($\times 10^{-2}$)	0.134	67.264	5.937	6.205
			RMSE	0.057	14.759	0.842	0.835
			CP (asymptotic CI)	0.963	0.947	0.956	0.946
			CP (bootstrap CI)	0.967	0.958	0.960	0.948

Table 14: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on $1,000$ Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 365$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(0.0005, -4.5, -0.2, 0.5)$	1520.53	bias ($\times 10^{-2}$)	6.234×10^{-3}	1.832	0.179	0.145
			RMSE	3.689×10^{-4}	0.082	0.066	0.079
			CP (asymptotic CI)	0.931	0.939	0.943	0.958
			CP (bootstrap CI)	0.951	0.945	0.950	0.963
	$(0.0005, -4.5, -0.125, 0.5)$	2261.48	bias ($\times 10^{-2}$)	7.346×10^{-3}	2.674	-0.158	0.234
			RMSE	8.458×10^{-5}	0.094	0.049	0.068
			CP (asymptotic CI)	0.921	0.937	0.958	0.962
			CP (bootstrap CI)	0.938	0.946	0.961	0.957
	$(0.0005, -4.5, -0.1, 0.5)$	3215.62	bias ($\times 10^{-2}$)	5.850×10^{-3}	2.159	-0.433	0.524
			RMSE	3.405×10^{-4}	0.125	0.059	0.061
			CP (asymptotic CI)	0.925	0.933	0.948	0.961
			CP (bootstrap CI)	0.940	0.945	0.952	0.975
45	$(0.0005, -4.5, -0.2, 0.5)$	2294.30	bias ($\times 10^{-2}$)	3.696×10^{-3}	1.341	0.088	0.133
			RMSE	2.548×10^{-4}	0.055	0.079	0.063
			CP (asymptotic CI)	0.935	0.944	0.940	0.951
			CP (bootstrap CI)	0.945	0.952	0.954	0.956
	$(0.0005, -4.5, -0.125, 0.5)$	3231.43	bias ($\times 10^{-2}$)	6.457×10^{-3}	3.749	0.122	-0.146
			RMSE	6.759×10^{-4}	0.123	0.074	0.058
			CP (asymptotic CI)	0.932	0.945	0.955	0.958
			CP (bootstrap CI)	0.943	0.953	0.952	0.969
	$(0.0005, -4.5, -0.1, 0.5)$	4534.56	bias ($\times 10^{-2}$)	8.254×10^{-4}	-0.353	-0.310	0.358
			RMSE	1.116×10^{-4}	0.113	0.042	0.046
			CP (asymptotic CI)	0.920	0.934	0.941	0.953
			CP (bootstrap CI)	0.936	0.944	0.955	0.962
60	$(0.0005, -4.5, -0.2, 0.5)$	2453.10	bias ($\times 10^{-2}$)	8.439×10^{-6}	0.647	0.012	-0.023
			RMSE	7.259×10^{-5}	0.092	0.060	0.034
			CP (asymptotic CI)	0.941	0.943	0.956	0.962
			CP (bootstrap CI)	0.954	0.951	0.950	0.970
	$(0.0005, -4.5, -0.125, 0.5)$	3671.09	bias ($\times 10^{-2}$)	3.604×10^{-4}	1.516	-0.042	0.039
			RMSE	8.046×10^{-5}	0.070	0.054	0.038
			CP (asymptotic CI)	0.956	0.943	0.953	0.948
			CP (bootstrap CI)	0.944	0.945	0.957	0.960
	$(0.0005, -4.5, -0.1, 0.5)$	5269.36	bias ($\times 10^{-2}$)	5.631×10^{-4}	0.387	0.087	-0.138
			RMSE	2.193×10^{-4}	0.132	0.025	0.036
			CP (asymptotic CI)	0.927	0.943	0.947	0.963
			CP (bootstrap CI)	0.933	0.952	0.951	0.959

Table 15: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 450$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(0.0005, -4.5, -0.2, 0.5)$	1184.03	bias ($\times 10^{-2}$)	-4.125×10^{-3}	2.836	0.131	0.193
			RMSE	3.069×10^{-4}	0.176	0.102	0.121
			CP (asymptotic CI)	0.927	0.942	0.956	0.939
			CP (bootstrap CI)	0.948	0.945	0.971	0.960
	$(0.0005, -4.5, -0.125, 0.5)$	1548.29	bias ($\times 10^{-2}$)	7.127×10^{-3}	2.142	0.135	0.224
			RMSE	2.481×10^{-4}	0.243	0.102	0.119
			CP (asymptotic CI)	0.921	0.935	0.942	0.943
			CP (bootstrap CI)	0.945	0.958	0.946	0.949
	$(0.0005, -4.5, -0.1, 0.5)$	1865.18	bias ($\times 10^{-2}$)	6.813×10^{-3}	1.598	0.275	-0.312
			RMSE	2.941×10^{-4}	0.153	0.089	0.057
			CP (asymptotic CI)	0.916	0.941	0.952	0.960
			CP (bootstrap CI)	0.946	0.949	0.958	0.969
45	$(0.0005, -4.5, -0.2, 0.5)$	1994.23	bias ($\times 10^{-2}$)	2.129×10^{-3}	1.322	0.119	0.173
			RMSE	4.691×10^{-4}	0.123	0.081	0.054
			CP (asymptotic CI)	0.930	0.949	0.955	0.942
			CP (bootstrap CI)	0.946	0.953	0.962	0.954
	$(0.0005, -4.5, -0.125, 0.5)$	2418.62	bias ($\times 10^{-2}$)	5.385×10^{-3}	0.932	0.122	0.185
			RMSE	9.210×10^{-5}	0.152	0.063	0.091
			CP (asymptotic CI)	0.929	0.938	0.946	0.947
			CP (bootstrap CI)	0.951	0.945	0.957	0.969
	$(0.0005, -4.5, -0.1, 0.5)$	2711.10	bias ($\times 10^{-2}$)	-1.168×10^{-3}	0.149	0.157	-0.193
			RMSE	1.007×10^{-4}	0.102	0.051	0.049
			CP (asymptotic CI)	0.922	0.939	0.945	0.953
			CP (bootstrap CI)	0.937	0.944	0.953	0.967
60	$(0.0005, -4.5, -0.2, 0.5)$	2305.27	bias ($\times 10^{-2}$)	8.754×10^{-4}	1.823	0.089	0.121
			RMSE	9.814×10^{-5}	0.095	0.054	0.072
			CP (asymptotic CI)	0.934	0.953	0.953	0.949
			CP (bootstrap CI)	0.939	0.949	0.957	0.963
	$(0.0005, -4.5, -0.125, 0.5)$	2892.37	bias ($\times 10^{-2}$)	-1.594×10^{-3}	0.502	0.052	-0.081
			RMSE	6.127×10^{-5}	0.077	0.061	0.083
			CP (asymptotic CI)	0.941	0.943	0.955	0.949
			CP (bootstrap CI)	0.942	0.955	0.963	0.970
	$(0.0005, -4.5, -0.1, 0.5)$	3021.69	bias ($\times 10^{-2}$)	7.183×10^{-4}	0.374	0.077	0.109
			RMSE	5.127×10^{-5}	0.083	0.042	0.055
			CP (asymptotic CI)	0.936	0.944	0.947	0.951
			CP (bootstrap CI)	0.953	0.949	0.956	0.943

Table 16: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on $1,000$ Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 540$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(0.0005, -4.5, -0.2, 0.5)$	1252.73	bias ($\times 10^{-2}$)	6.323×10^{-3}	2.317	0.127	0.218
			RMSE	5.846×10^{-4}	0.104	0.083	0.064
			CP (asymptotic CI)	0.923	0.951	0.948	0.937
			CP (bootstrap CI)	0.941	0.960	0.953	0.951
	$(0.0005, -4.5, -0.125, 0.5)$	1410.22	bias ($\times 10^{-2}$)	8.852×10^{-3}	1.768	0.071	0.323
			RMSE	3.549×10^{-4}	0.152	0.115	0.087
			CP (asymptotic CI)	0.925	0.946	0.939	0.940
			CP (bootstrap CI)	0.941	0.953	0.944	0.946
	$(0.0005, -4.5, -0.1, 0.5)$	1926.51	bias ($\times 10^{-2}$)	-3.573×10^{-3}	4.561	0.115	0.241
			RMSE	9.605×10^{-5}	0.136	0.092	0.083
			CP (asymptotic CI)	0.932	0.947	0.955	0.960
			CP (bootstrap CI)	0.949	0.952	0.959	0.953
45	$(0.0005, -4.5, -0.2, 0.5)$	1358.64	bias ($\times 10^{-2}$)	-2.564×10^{-3}	3.519	0.089	0.114
			RMSE	3.792×10^{-4}	0.079	0.099	0.038
			CP (asymptotic CI)	0.929	0.958	0.951	0.943
			CP (bootstrap CI)	0.953	0.961	0.948	0.946
	$(0.0005, -4.5, -0.125, 0.5)$	1512.64	bias ($\times 10^{-2}$)	5.742×10^{-3}	2.833	0.117	0.185
			RMSE	3.182×10^{-4}	0.137	0.093	0.103
			CP (asymptotic CI)	0.923	0.952	0.951	0.944
			CP (bootstrap CI)	0.955	0.961	0.957	0.949
	$(0.0005, -4.5, -0.1, 0.5)$	1493.20	bias ($\times 10^{-2}$)	-1.659×10^{-3}	3.176	0.218	-0.258
			RMSE	7.718×10^{-5}	0.104	0.056	0.066
			CP (asymptotic CI)	0.933	0.949	0.956	0.961
			CP (bootstrap CI)	0.946	0.948	0.960	0.974
60	$(0.0005, -4.5, -0.2, 0.5)$	1523.65	bias ($\times 10^{-2}$)	9.174×10^{-4}	-2.312	0.213	-0.091
			RMSE	7.184×10^{-5}	0.041	0.074	0.052
			CP (asymptotic CI)	0.937	0.954	0.946	0.949
			CP (bootstrap CI)	0.947	0.958	0.952	0.955
	$(0.0005, -4.5, -0.125, 0.5)$	1956.72	bias ($\times 10^{-2}$)	-1.035×10^{-3}	1.605	0.088	0.122
			RMSE	9.349×10^{-5}	0.082	0.061	0.071
			CP (asymptotic CI)	0.929	0.948	0.953	0.946
			CP (bootstrap CI)	0.948	0.957	0.959	0.965
	$(0.0005, -4.5, -0.1, 0.5)$	2130.28	bias ($\times 10^{-2}$)	7.126×10^{-4}	1.857	0.163	0.115
			RMSE	6.523×10^{-5}	0.077	0.065	0.032
			CP (asymptotic CI)	0.938	0.963	0.951	0.954
			CP (bootstrap CI)	0.942	0.967	0.955	0.962

Table 17: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 630$.

n	$(\beta, \eta, \mu, \sigma)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\mu}$	$\hat{\sigma}$
30	$(0.0005, -4.5, -0.2, 0.5)$	1352.49	bias ($\times 10^{-2}$)	3.751×10^{-3}	0.731	0.316	0.438
			RMSE	4.285×10^{-4}	0.073	0.065	0.089
			CP (asymptotic CI)	0.928	0.954	0.941	0.950
			CP (bootstrap CI)	0.954	0.958	0.951	0.966
	$(0.0005, -4.5, -0.125, 0.5)$	1410.22	bias ($\times 10^{-2}$)	-1.648×10^{-3}	1.452	0.235	-0.522
			RMSE	1.208×10^{-4}	0.098	0.082	0.066
			CP (asymptotic CI)	0.915	0.941	0.944	0.963
			CP (bootstrap CI)	0.943	0.955	0.953	0.972
	$(0.0005, -4.5, -0.1, 0.5)$	1926.51	bias ($\times 10^{-2}$)	-2.503×10^{-3}	2.389	-0.233	0.534
			RMSE	9.481×10^{-5}	0.104	0.075	0.062
			CP (asymptotic)	0.920	0.947	0.936	0.941
			CP (bootstrap)	0.942	0.949	0.953	0.948
45	$(0.0005, -4.5, -0.2, 0.5)$	1358.64	bias ($\times 10^{-2}$)	1.471×10^{-3}	1.151	0.210	0.331
			RMSE	2.587×10^{-4}	0.085	0.055	0.062
			CP (asymptotic CI)	0.932	0.951	0.949	0.957
			CP (bootstrap CI)	0.958	0.963	0.947	0.962
	$(0.0005, -4.5, -0.125, 0.5)$	1512.57	bias ($\times 10^{-2}$)	8.517×10^{-4}	2.185	0.188	0.324
			RMSE	7.362×10^{-5}	0.074	0.074	0.053
			CP (asymptotic CI)	0.927	0.968	0.940	0.951
			CP (bootstrap CI)	0.940	0.951	0.958	0.965
	$(0.0005, -4.5, -0.1, 0.5)$	1493.20	bias ($\times 10^{-2}$)	8.546×10^{-4}	1.439	0.224	0.283
			RMSE	4.574×10^{-5}	0.161	0.065	0.045
			CP (asymptotic)	0.943	0.937	0.948	0.955
			CP (bootstrap)	0.953	0.949	0.954	0.960
60	$(0.0005, -4.5, -0.2, 0.5)$	1528.65	bias ($\times 10^{-2}$)	1.158×10^{-3}	0.452	0.113	0.154
			RMSE	2.576×10^{-4}	0.035	0.031	0.045
			CP (asymptotic CI)	0.937	0.944	0.956	0.940
			CP (bootstrap CI)	0.949	0.957	0.963	0.954
	$(0.0005, -4.5, -0.125, 0.5)$	1956.72	bias ($\times 10^{-2}$)	5.034×10^{-4}	1.309	0.074	-0.262
			RMSE	6.385×10^{-5}	0.042	0.069	0.068
			CP (asymptotic CI)	0.935	0.955	0.946	0.956
			CP (bootstrap CI)	0.948	0.947	0.954	0.966
	$(0.0005, -4.5, -0.1, 0.5)$	2130.28	bias ($\times 10^{-2}$)	5.289×10^{-4}	0.578	0.120	-0.187
			RMSE	2.193×10^{-5}	0.132	0.025	0.036
			CP (asymptotic CI)	0.927	0.949	0.952	0.954
			CP (bootstrap CI)	0.945	0.946	0.953	0.965

Table 18: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 365$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(0.0005, -4.5, 4, 5)$	1482.60	bias ($\times 10^{-2}$)	3.578×10^{-4}	1.132	7.631	11.272
			RMSE	1.543×10^{-4}	0.181	1.287	1.352
			CP (asymptotic CI)	0.918	0.944	0.953	0.962
			CP (bootstrap CI)	0.947	0.952	0.947	0.973
	$(0.0005, -4.5, 5, 5)$	2319.67	bias ($\times 10^{-2}$)	6.451×10^{-3}	1.745	3.451	-6.923
			RMSE	2.947×10^{-4}	0.153	0.891	1.035
			CP (asymptotic CI)	0.925	0.942	0.956	0.961
			CP (bootstrap CI)	0.931	0.947	0.959	0.965
45	$(0.0005, -4.5, 4, 5)$	1766.25	bias ($\times 10^{-2}$)	6.984×10^{-3}	-0.461	4.175	-8.238
			RMSE	9.737×10^{-5}	0.110	0.715	0.851
			CP (asymptotic CI)	0.926	0.950	0.945	0.953
			CP (bootstrap CI)	0.937	0.953	0.949	0.948
	$(0.0005, -4.5, 5, 5)$	2832.71	bias ($\times 10^{-2}$)	5.275×10^{-3}	-1.278	2.923	-3.040
			RMSE	1.104×10^{-4}	0.116	0.767	0.691
			CP (asymptotic CI)	0.928	0.937	0.949	0.952
			CP (bootstrap CI)	0.935	0.942	0.956	0.955
60	$(0.0005, -4.5, 4, 5)$	2489.21	bias ($\times 10^{-2}$)	8.437×10^{-4}	0.559	2.135	-2.467
			RMSE	6.746×10^{-5}	0.094	0.632	0.644
			CP (asymptotic CI)	0.940	0.953	0.949	0.964
			CP (bootstrap CI)	0.945	0.951	0.953	0.958
	$(0.0005, -4.5, 5, 5)$	3251.04	bias ($\times 10^{-2}$)	8.042×10^{-4}	0.541	1.692	-1.816
			RMSE	7.763×10^{-5}	0.093	0.558	0.720
			CP (asymptotic CI)	0.938	0.942	0.956	0.948
			CP (bootstrap CI)	0.946	0.943	0.967	0.959

Table 19: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 450$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(0.0005, -4.5, 4, 5)$	1346.54	bias ($\times 10^{-2}$)	5.759×10^{-3}	3.712	3.144	5.175
			RMSE	4.109×10^{-4}	0.153	1.104	1.503
			CP (asymptotic CI)	0.922	0.941	0.948	0.956
			CP (bootstrap CI)	0.928	0.956	0.955	0.943
	$(0.0005, -4.5, 5, 5)$	1745.30	bias ($\times 10^{-2}$)	-7.351×10^{-3}	4.386	3.181	3.465
			RMSE	4.166×10^{-4}	0.152	1.256	1.761
			CP (asymptotic CI)	0.927	0.937	0.944	0.938
			CP (bootstrap CI)	0.934	0.943	0.953	0.946
45	$(0.0005, -4.5, 4, 5)$	1724.51	bias ($\times 10^{-2}$)	2.085×10^{-3}	2.871	-2.109	-3.156
			RMSE	3.671×10^{-4}	0.115	0.886	1.342
			CP (asymptotic CI)	0.927	0.953	0.939	0.958
			CP (bootstrap CI)	0.939	0.957	0.948	0.967
	$(0.0005, -4.5, 5, 5)$	2306.41	bias ($\times 10^{-2}$)	6.573×10^{-3}	3.006	2.367	2.142
			RMSE	3.341×10^{-4}	0.130	1.438	0.942
			CP (asymptotic CI)	0.934	0.931	0.953	0.936
			CP (bootstrap CI)	0.931	0.955	0.964	0.949
60	$(0.0005, -4.5, 4, 5)$	1974.93	bias ($\times 10^{-2}$)	8.410×10^{-4}	-1.632	1.123	2.110
			RMSE	1.042×10^{-4}	0.094	0.926	1.192
			CP (asymptotic CI)	0.923	0.967	0.942	0.950
			CP (bootstrap CI)	0.945	0.963	0.945	0.956
	$(0.0005, -4.5, 5, 5)$	2792.48	bias ($\times 10^{-2}$)	9.208×10^{-4}	-1.150	-1.110	-1.104
			RMSE	7.796×10^{-5}	0.084	0.894	0.859
			CP (asymptotic CI)	0.930	0.963	0.941	0.958
			CP (bootstrap CI)	0.937	0.967	0.953	0.964

Table 20: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 540$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(0.0005, -4.5, 4, 5)$	915.11	bias ($\times 10^{-2}$)	5.853×10^{-3}	1.457	3.783	5.325
			RMSE	3.653×10^{-4}	0.142	1.244	1.571
			CP (asymptotic CI)	0.922	0.944	0.951	0.942
			CP (bootstrap CI)	0.947	0.954	0.963	0.970
	$(0.0005, -4.5, 5, 5)$	1109.93	bias ($\times 10^{-2}$)	-1.392×10^{-5}	0.794	4.681	6.750
			RMSE	1.162×10^{-4}	0.115	1.698	1.847
			CP (asymptotic CI)	0.925	0.947	0.944	0.958
			CP (bootstrap CI)	0.953	0.961	0.954	0.960
45	$(0.0005, -4.5, 4, 5)$	997.70	bias ($\times 10^{-2}$)	-2.320×10^{-3}	0.813	3.512	4.708
			RMSE	1.079×10^{-4}	0.105	1.035	1.189
			CP (asymptotic CI)	0.923	0.963	0.959	0.961
			CP (bootstrap CI)	0.947	0.954	0.963	0.970
	$(0.0005, -4.5, 5, 5)$	1436.84	bias ($\times 10^{-2}$)	8.466×10^{-4}	1.156	2.464	3.512
			RMSE	7.641×10^{-5}	0.103	1.450	1.953
			CP (asymptotic CI)	0.932	0.958	0.941	0.942
			CP (bootstrap CI)	0.940	0.961	0.939	0.947
60	$(0.0005, -4.5, 4, 5)$	1657.03	bias ($\times 10^{-2}$)	-5.492×10^{-4}	0.572	1.401	2.538
			RMSE	7.875×10^{-5}	0.077	0.782	1.163
			CP (asymptotic CI)	0.943	0.946	0.955	0.966
			CP (bootstrap CI)	0.944	0.955	0.961	0.972
	$(0.0005, -4.5, 5, 5)$	1810.24	bias ($\times 10^{-2}$)	3.092×10^{-4}	-0.714	1.246	1.289
			RMSE	5.134×10^{-5}	0.069	1.019	1.064
			CP (asymptotic CI)	0.944	0.960	0.952	0.957
			CP (bootstrap CI)	0.939	0.965	0.947	0.946

Table 21: Biases and RMSEs of the MLE, the coverage probability (CP) of the 95% asymptotic confidence interval (CI) and 95% bootstrap confidence interval and average computation time (in seconds), computed based on 1,000 Monte Carlo replications, under the LLP-Lognormal model with the PM time interval: $\Delta = 630$.

n	$(\beta, \eta, \alpha, \nu)$	Run time		$\hat{\beta}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\nu}$
30	$(0.0005, -4.5, 4, 5)$	779.96	bias ($\times 10^{-2}$)	-2.178×10^{-3}	1.687	3.246	3.322
			RMSE	8.652×10^{-4}	0.135	1.783	1.862
			CP (asymptotic CI)	0.912	0.938	0.967	0.959
			CP (bootstrap CI)	0.944	0.943	0.958	0.963
	$(0.0005, -4.5, 5, 5)$	932.16	bias ($\times 10^{-2}$)	-4.841×10^{-4}	1.528	2.785	2.891
			RMSE	2.193×10^{-4}	0.124	1.260	1.135
			CP (asymptotic CI)	0.927	0.953	0.956	0.943
			CP (bootstrap CI)	0.934	0.957	0.965	0.957
45	$(0.0005, -4.5, 4, 5)$	974.70	bias ($\times 10^{-2}$)	7.945×10^{-4}	1.470	0.983	2.618
			RMSE	3.731×10^{-4}	0.103	1.044	1.236
			CP (asymptotic CI)	0.926	0.945	0.954	0.948
			CP (bootstrap CI)	0.949	0.952	0.953	0.966
	$(0.0005, -4.5, 5, 5)$	1035.32	bias ($\times 10^{-2}$)	-8.396×10^{-5}	-1.008	1.294	1.312
			RMSE	7.232×10^{-5}	0.088	1.300	1.394
			CP (asymptotic CI)	0.930	0.965	0.957	0.952
			CP (bootstrap CI)	0.942	0.956	0.951	0.968
60	$(0.0005, -4.5, 4, 5)$	1338.29	bias ($\times 10^{-2}$)	-1.302×10^{-5}	0.849	1.145	1.176
			RMSE	7.839×10^{-3}	0.078	0.916	1.174
			CP (asymptotic CI)	0.923	0.958	0.933	0.950
			CP (bootstrap CI)	0.940	0.954	0.941	0.968
	$(0.0005, -4.5, 5, 5)$	1624.81	bias ($\times 10^{-2}$)	6.544×10^{-5}	0.756	0.889	1.075
			RMSE	4.113×10^{-5}	0.054	0.752	0.408
			CP (asymptotic CI)	0.939	0.947	0.952	0.961
			CP (bootstrap CI)	0.945	0.957	0.958	0.963

References

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (2012), *Statistical Models based on Counting Processes*, Springer Science & Business Media.

Asmussen, S. and Glynn, P. W. (2007), *Stochastic Simulation: Algorithms and Analysis*, vol. 57, Springer.

Berger, J. O., Bernardo, J. M., and Sun, D. (2009), "The formal definition of reference priors," *The Annals of Statistics*, 37(2), 905–938.

Brooks, S. P. and Gelman, A. (1998), "General methods for monitoring convergence of iterative simulations," *Journal of Computational and Graphical Statistics*, 7(4), 434–455.

Caflisch, R. E. (1998), "Monte Carlo and quasi-Monte Carlo methods," *Acta Numerica*, 7, 1–49.

Chauvel, C., Dauxois, J.-Y., Doyen, L., and Gaudoin, O. (2016), "Parametric bootstrap goodness-of-fit tests for imperfect maintenance models," *IEEE Transactions on Reliability*, 65(3), 1343–1359.

Cook, R. J., Lawless, J. F., et al. (2007), *The Statistical Analysis of Recurrent Events*, Springer.

Cox, D. R. (1959), "The analysis of exponentially distributed life-times with two types of failure," *Journal of the Royal Statistical Society: Series B (Methodological)*, 21(2), 411–421.

Crowder, M. J. (2001), *Classical Competing Risks*, CRC Press.

Dijoux, Y. and Gaudoin, O. (2009), "The alert-delay competing risks model for maintenance analysis," *Journal of Statistical Planning and Inference*, 139(5), 1587–1603.

Doyen, L. and Gaudoin, O. (2004), "Classes of imperfect repair models based on reduction of failure intensity or virtual age," *Reliability Engineering & System Safety*, 84(1), 45–56.

— (2006), "Imperfect maintenance in a generalized competing risks framework," *Journal of Applied Probability*, 43(3), 825–839.

— (2011), "Modeling and assessment of aging and efficiency of corrective and planned preventive maintenance," *IEEE Transactions on Reliability*, 60(4), 759–769.

Efron, B. and Tibshirani, R. J. (1994), *An Introduction to the Bootstrap*, CRC press.

Gail, M. (1975), "A review and critique of some models used in competing risk analysis," *Biometrics*, 209–222.

Gelman, A. (2006), "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)," *Bayesian Analysis*, 1(3), 515–534.

— (2009), "Bayes, Jeffreys, prior distributions and the philosophy of statistics," *Statistical Science*, 24(2), 176–178.

Gelman, A. and Rubin, D. B. (1992), "Inference from iterative simulation using multiple sequences," *Statistical Science*, 457–472.

George, E. I. and McCulloch, R. E. (1993), "Variable selection via Gibbs sampling," *Journal of the American Statistical Association*, 88(423), 881–889.

Geweke, J. (1992), "Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments," *Bayesian Statistics*, 4, 641–649.

Gong, L. and Flegal, J. M. (2016), "A practical sequential stopping rule for high-dimensional Markov chain Monte Carlo," *Journal of Computational and Graphical Statistics*, 25(3), 684–700.

Green, P. J. and Mira, A. (2001), "Delayed rejection in reversible jump Metropolis–Hastings," *Biometrika*, 88(4), 1035–1053.

Haario, H., Saksman, E., and Tamminen, J. (1999), "Adaptive proposal distribution for random walk Metropolis algorithm," *Computational Statistics*, 14(3), 375–395.

— (2001), "An adaptive Metropolis algorithm," *Bernoulli*, 223–242.

Niederreiter, H. (1992), *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM.

R Core Team (2020), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Spiegelhalter, D., Thomas, A., Best, N., and Gilks, W. (1996), "BUGS 0.5: Bayesian inference using Gibbs sampling manual (version ii)," *MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK*, 1–59.

Syversveen, A. R. (1998), "Noninformative bayesian priors. interpretation and problems with construction and applications," *Preprint Statistics*, 3(3), 1–11.

Tierney, L. and Mira, A. (1999), "Some adaptive Monte Carlo methods for Bayesian inference," *Statistics in Medicine*, 18(17-18), 2507–2515.

Wang, X. and Fang, K.-T. (2003), "The effective dimension and quasi-Monte Carlo integration," *Journal of Complexity*, 19(2), 101–124.