Accepted Manuscript

Estimating the Gerber–Shiu function in a Lévy risk model by Laguerre series expansion

Zhimin Zhang, Wen Su

PII: S0377-0427(18)30423-0
DOI: https://doi.org/10.1016/j.cam.2018.07.003
Reference: CAM 11789

To appear in: *Journal of Computational and Applied Mathematics*

Received date : 15 August 2017
Revised date : 27 April 2018

Please cite this article as: Z. Zhang, W. Su, Estimating the Gerber–Shiu function in a Lévy risk model by Laguerre series expansion, *Journal of Computational and Applied Mathematics* (2018), https://doi.org/10.1016/j.cam.2018.07.003

Estimating the Gerber-Shiu function in a Lévy risk model by Laguerre series expansion

Zhimin Zhang*, Wen Su

April 27, 2018

Abstract

In this paper, we provide a new method for estimating the Gerber-Shiu function in a pure jump Lévy risk model. First, we show that the Gerber-Shiu function can be expressed on the Laguerre basis and the Laguerre coefficients can be easily obtained by solving a linear system. Next, based on a high-frequency observation of the aggregate claims process, we estimate the Laguerre coefficients and this leads a new estimator of the Gerber-Shiu function. We derive the consistency property of this estimator when the sample size is large. Finally, we do some simulation studies to illustrate the finite sample size performance.

Keywords: Gerber-Shiu function, Estimate, Lévy risk model, Laguerre series.

1 Introduction

In this paper, we suppose that the surplus flow of an insurance company evolves as a Lévy process

$$U_t = u + ct - X_t, \quad t \geq 0, \tag{1.1}$$

where $u \geq 0$ is the initial surplus level, and $c > 0$ is the constant premium rate. The process $X = \{X_t\}_{t\geq 0}$ with $X_0 = 0$, representing the aggregate claims up to time t, is a pure-jump Lévy process. The characteristics of X are uniquely determined by the Laplace exponent which is defined by

$$\psi(s) := \frac{1}{t} \ln \mathbb{E}[e^{-sX_t}] = \int_0^\infty \left(e^{-sx} - 1\right) \nu(x)dx, \ s \geq 0, \tag{1.2}$$

where ν is a Lévy density supported on $(0, \infty)$ satisfying the usual condition $\int_0^\infty (1 \wedge x^2)\nu(x)dx < \infty$. For $k = 1, 2, \ldots$, define $\mu_k = \int_0^\infty x^k \nu(x)dx$, provided that these integrals are finite. In order to ensure that the surplus process has a positive drift with probability one, we suppose the following condition.

*College of Mathematics and Statistics, Chongqing University, Chongqing, 401331, P.R. China. Email: zmzhang@cqu.edu.cn.

Condition A.1 (net profit condition)

$$\text{The premium rate } c > \mu_1. \tag{1.3}$$

It is known that Condition **A.1** ensures that the surplus process $\{U_t\}$ has a positive drift w.r.t. t, however, it is also possible that the surplus level becomes negative sometimes due to some large claim sizes. We define the ruin time by

$$\tau = \inf\{t > 0 : \ U_t < 0\}$$

with the convention $\inf \emptyset = \infty$. The widely used risk measure to study ruin related quantities is the Gerber-Shiu discounted penalty function (Gerber and Shiu (1998)),

$$\phi(u) = \mathbb{E}[e^{-\delta\tau} w(U_{\tau-}, |U_\tau|)\mathbf{1}_{(\tau<\infty)}|U_0 = u], \quad u \geq 0, \tag{1.4}$$

where $\delta \geq 0$ is the interest force, $\mathbf{1}_{(A)}$ is the indicator function of an event A, and w is a nonnegative function defined on $[0, \infty) \times (0, \infty)$.

The Gerber-Shiu function is an important tool for studying ruin related quantities, such as the time to ruin, the surplus before and after ruin. Recently, a lot of contributions to the Gerber-Shiu function and its generalization have been made in various risk models. See e.g. Asmussen and Albrecher (2010). In risk theory, the Lévy process is a popular class of stochastic processes to model the surplus flow of the insurance company. See e.g. Yin and Wang (2009), Zhao and Yin (2010a), Yuen and Yin (2011), Shen et al. (2013), Yin and Wen (2013) and Li et al. (2017). In particular, for the study of the Gerber-Shiu function in Lévy risk model, we refer the interested readers to Zhao and Yin(2010b) and Kyprianou (2013). In the analysis of the Gerber-Shiu function, theoretical results are dependent on the information on the aggregate claims process, for example, the probability characteristics of the claim sizes and claim number process. However, such distributional information often includes some unknown parameters and quantities that have to be estimated from the historical data of the surplus process. Hence, statistical estimation of the Gerber-Shiu function is also very important. In particular, we remark that statistical estimation does not depend heavily on the specific models, and it can be applied as long as the historical sample is available.

Recently, statistical estimation of ruin probability has been studied by many authors. See, e.g. Politis (2003), Mnatsakanov et al. (2008), Shimizu (2009), Masiello (2014) and Zhang et al. (2014), Zhang (2016, 2017). In the Lévy risk model, Zhang and Yang (2013) proposed a nonparametric estimate of the ruin probability based on high-frequency observation of the surplus process; Zhang and Yang (2014) studied how to estimate the ruin probability via low-frequency sampling of the surplus process. Here, high frequency means that the sampling interval tends to zero as the sample size tends to infinity, while low frequency means that the sampling interval is a constant. For the Gerber-Shiu function, Shimizu (2011, 2012) considered its estimation in the Lévy risk model and perturbed compound Poisson model respectively, where the Laplace inversion method is used for constructing the estimate. However, the Laplace transform inversion method is not good, since it not only results in slow convergence rate, but also needs very long time for computation. To overcome these two drawbacks, Shimizu and Zhang (2017) estimate the Gerber-Shiu function by Fourier transform method, where the FFT algorithm is used to compute the estimator.

In this paper, we study how to estimate the Gerber-Shiu function by the Laguerre series expansion based on a random sample of the aggregate claims process. The motivation of this paper comes from

Comte et al. (2017), in which a Laplace deconvolution problem is solved by Laguerre series expansion method. Suppose that the premium rate c is known, but the Lévy density is unknown. Furthermore, suppose that the aggregate claims process X can be observed at a sequence of discrete time points so that the following sample is available

$$\boldsymbol{X}^n := \{X_{k\Delta}:\ k = 1, 2, \dots, n\},$$

where $\Delta := \Delta_n > 0$ is a sampling step. We estimate the Gerber-Shiu function under high-frequency observation assumption. More precisely, we shall consider the following condition.

Condition A.2 (High-frequency observation in a long term)

$$\lim_{n\to\infty} \Delta = 0, \quad \lim_{n\to\infty} n\Delta = \infty.$$

The above condition is also considered in Zhang and Yang (2013) and Shimizu and Zhang (2017). This condition is used to study the consistency property under the large sample size setting. In the simulation study, we find that the estimator behaves well even if Δ is not very small and $n\Delta$ is not very large.

The remainder of this paper is organized as follows. In Section 2, we introduce the Laguerre polynomials and express the Gerber-Shiu function on the Laguerre basis. In Section 3, we present a new Gerber-Shiu estimator based on Laguerre series expansion, and in Section 4 we study the convergence rate of our estimator. Finally, some simulation studies are given in Section 5 to illustrate the performance of our estimator when the sample size is finite.

2 Laguerre series expansion

Let $L^2(\mathbb{R}_+)$ denote the class of square integrable functions on the positive half-line, and denote the scalar product and L^2-norm on $L^2(\mathbb{R}_+)$ by

$$\langle p, q\rangle = \int_0^\infty p(x)q(x)dx, \ \|p\|^2 = \int_0^\infty p(x)^2 dx, \ \ \forall\, p, q \in \mathbb{L}^2(\mathbb{R}_+).$$

The Laguerre functions are defined by

$$\varphi_k(x) = \sqrt{2}L_k(2x)e^{-x},\ x \geq 0, \quad k = 0, 1, 2, \dots, \tag{2.1}$$

where $\{L_k(x)\}$ are the Laguerre polynomials given by

$$L_k(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{x^j}{j!}, \quad x \geq 0.$$

The Laguerre functions are uniformly bounded, i.e.

$$\sup_{x\in\mathbb{R}_+} |\varphi_k(x)| = \|\varphi_k\|_\infty \leq \sqrt{2}, \ \ \forall\, k \geq 0. \tag{2.2}$$

We also note that the following convolution relation holds true,

$$\int_0^x \varphi_k(x-y)\varphi_j(y)dy = \frac{1}{\sqrt{2}}[\varphi_{k+j}(x) - \varphi_{k+j+1}(x)]. \tag{2.3}$$

The above results can be found in Abramowitz and Stegun (1964).

It is known that the collection $\{\varphi_k\}_{k\geq 0}$ is a complete orthonormal basis of $L^2(\mathbb{R}_+)$ satisfying

$$\|\varphi_k\| = 1;\ \langle \varphi_k,\ \varphi_j\rangle = 0 \text{ for } k \neq j. \tag{2.4}$$

Hence, for every $f \in L^2(\mathbb{R}_+)$, we can develop it on the Laguerre basis, i.e.

$$f(x) = \sum_{k=0}^{\infty} a_{f,k}\varphi_k(x), \tag{2.5}$$

where $a_{f,k} = \langle f, \varphi_k\rangle = \int_0^\infty f(x)\varphi_k(x)dx$ for $k = 0, 1, 2, \ldots$. Furthermore, we introduce the space $\mathcal{S}_m = \text{Span}\{\varphi_0, \ldots, \varphi_{m-1}\}$, and for any function $f \in L^2(\mathbb{R}_+)$ we define its projection onto $\mathcal{S}_m$ by

$$f_m(x) = \sum_{k=0}^{m-1} a_{f,k}\varphi_k(x).$$

To evaluate the series truncation error in the above projection approximation, we introduce the Sobolev-Laguerre space that is defined by

$$W(\mathbb{R}_+, r, B) = \left\{ f : \mathbb{R}_+ \mapsto \mathbb{R},\ f \in L^2(\mathbb{R}_+),\ \sum_{k=0}^{\infty} k^r a_{f,k}^2 \leq B < \infty \right\},$$

where $0 < r, B < \infty$. Suppose that $f \in W(\mathbb{R}_+, r, B)$. Because of the orthonormal property of the Laguerre basis $\{\varphi_k\}$, we have the following approximation error,

$$\|f - f_m\|^2 = \sum_{k=m}^{\infty} a_{f,k}^2 \leq m^{-r} \sum_{k=0}^{\infty} k^r a_{f,k}^2 \leq Bm^{-r}. \tag{2.6}$$

Remark 2.1 *The Sobolev-Laguerre space was introduced by Bongioanni and Torrea (2009), and the link with the coefficients of a function on the Laguerre basis was studied by Comte and Genon-Catalet (2015). Suppose that $f \in L^2(\mathbb{R}_+)$ and r is a positive integer. Then the condition $\sum_{k=0}^{\infty} k^r a_{f,k}^2 < \infty$ is equivalent to the property that f admits derivatives up to order $r - 1$, with $f^{(r-1)}$ absolutely continuous and for $k = 0, 1, \ldots, r - 1$, the functions*

$$x^{\frac{k+1}{2}}(fe^x)^{(k+1)}e^{-x} = x^{\frac{k+1}{2}} \sum_{j=0}^{k+1} \binom{k+1}{j} f^{(j)}$$

belong to $L^2(\mathbb{R}_+)$, where $f^{(r)}$ is the Random-Nikodym derivative of $f^{(r-1)}$ w.r.t. the Lebesgue measure. We also know from Comte and Genon-Catalet (2015) and Zhang and Su (2018) that if f is a finite mixture of Erlang functions, then the bias in (2.6) has exponential decay rate.

Remark 2.2 *If $f \in W(\mathbb{R}_+, r, B)$ with $r > 1$, we can show that the infinite series in (2.5) is absolutely convergent uniformly for $x \geq 0$ as follows. Suppose that $f \in W(\mathbb{R}_+, r, B)$, then using Cauchy-Schwarz inequality we obtain*

$$\begin{aligned}\sum_{k=0}^{\infty} |a_{f,k}| \cdot |\varphi_k(x)| &\leq \sqrt{2} \sum_{k=0}^{\infty} |a_{f,k}| = \sqrt{2}|a_{f,0}| + +\sqrt{2} \sum_{k=1}^{\infty} k^{r/2} |a_{f,k}| k^{-r/2} \\ &\leq \sqrt{2}|a_{f,0}| + +\sqrt{2} \left(\sum_{k=1}^{\infty} k^r a_{f,k}^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} k^{-r} \right)^{1/2} \\ &\leq \sqrt{2}|a_{f,0}| + +\sqrt{2B} \left(\sum_{k=1}^{\infty} k^{-r} \right)^{1/2}.\end{aligned}$$

Since the infinite series $\sum_{k=1}^{\infty} k^{-r}$ is convergent for $r > 1$, then the infinite series $\sum_{k=0}^{\infty} a_{f,k}\varphi_k(x)$ is absolutely convergent uniformly for $x \geq 0$.

We shall present a Laguerre series expansion of the Gerber-Shiu function, which plays an important role in estimating the Gerber-Shiu function. It follows from Garrido and Morales (2006) and Shimizu and Zhang (2017) that the Gerber-Shiu function satisfies the following renewal equation

$$\phi(u) = \int_0^u \phi(u-x)g(x)dx + h(u), \ u \geq 0, \tag{2.7}$$

where the functions g and h in (2.7) are defined by

$$\begin{aligned} g(x) &= \frac{1}{c} \int_x^{\infty} e^{-\rho(y-x)} \nu(y)dy, \\ h(u) &= \frac{1}{c} \int_u^{\infty} e^{-\rho(x-u)} \int_x^{\infty} w(x, y-x)\nu(y)dydx, \end{aligned}$$

where ρ is the unique positive root of equation (in s)

$$cs + \psi(s) = \delta. \tag{2.8}$$

Note that $\rho = 0$ as $\delta = 0$. It is known that equation (2.7) is defective since $\int_0^{\infty} g(x)dx < 1$ under condition $\delta > 0$ or Condition **A.1**.

In order to use the Laguerre series expansion method, we need to check the square integrability of the functions ϕ, g and h. To this end, we need the following conditions.

Condition A.3 (Integrability for w) For the penalty function w,

$$\int_0^{\infty} \int_0^{\infty} (1+x)w(x,y)\nu(x+y)dydx < \infty. \tag{2.9}$$

Lemma 1 *Under conditions **A.1** and **A.3**, we have $\phi, g, h \in L^2(\mathbb{R}_+)$.*

Proof. This follows from the same arguments used in Section 2 in Zhang and Su (2018). ■

In the sequel, suppose that $\phi, g, h \in L^2(\mathbb{R}_+)$. Furthermore, we suppose that the following condition on ϕ and g holds, which is easily satisfied due to Remark 2.1.

Condition A.4 $\phi, g \in W(\mathbb{R}_+, r, B)$ for some $r > 1$ and $0 < B < \infty$.

Now we develop ϕ, g and h on the Laguerre basis, i.e.

$$\phi(u) = \sum_{k=0}^{\infty} a_{\phi,k}\varphi_k(u), \quad g(x) = \sum_{k=0}^{\infty} a_{g,k}\varphi_k(x), \quad h(u) = \sum_{k=0}^{\infty} a_{h,k}\varphi_k(u), \quad u, x \geq 0. \tag{2.10}$$

We can use the projection ϕ_m to approximate ϕ. Furthermore, for $\phi \in W(\mathbb{R}_+, r, B)$, we have

$$\|\phi - \phi_m\|^2 \leq Bm^{-r} \tag{2.11}$$

due to (2.6).

Now we study how to determine the constants $a_{\phi,k}$ in the projection function ϕ_m. Under Condition **A.4** and using Remark 2.2, we know that the following two infinite series are absolutely convergent uniformly for $x \geq 0$,

$$\sum_{k=0}^{\infty} a_{\phi,k}\varphi_k(x), \quad \sum_{k=0}^{\infty} a_{g,k}\varphi_k(x).$$

Hence, using the convolution formula (2.3) and changing the order of summation and integration we have

$$\begin{aligned}
\int_0^u \phi(u-x)g(x)dx &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} a_{\phi,k}a_{g,j}\int_0^u \varphi_k(u-x)\varphi_j(x)dx \\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} 2^{-1/2}a_{\phi,k}a_{g,j}[\varphi_{k+j}(u) - \varphi_{k+j+1}(u)] \\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{k} 2^{-1/2}[a_{g,k-j} - a_{g,k-j-1}]a_{\phi,j}\varphi_k(u),
\end{aligned} \tag{2.12}$$

where we have used the convention $a_{g,-1} = 0$. Plugging the formulas in (2.10) and (2.12) into the integral equation (2.7) gives

$$\sum_{k=0}^{\infty} a_{\phi,k}\varphi_k(u) = \sum_{k=0}^{\infty}\sum_{j=0}^{k} 2^{-1/2}[a_{g,k-j} - a_{g,k-j-1}]a_{j,\phi}\varphi_k(u) + \sum_{k=0}^{\infty} a_{h,k}\varphi_k(u). \tag{2.13}$$

Comparing the coefficients of φ_k in (2.13) gives for $k = 0, 1, 2, \ldots$

$$a_{\phi,k} = \sum_{j=0}^{k} 2^{-1/2}[a_{g,k-j} - a_{g,k-j-1}]a_{\phi,j} + a_{h,k}.$$

This yields an infinite linear triangular system

$$\boldsymbol{A}_\infty \vec{\boldsymbol{a}}_{\phi,\infty} = \vec{\boldsymbol{a}}_{h,\infty}, \tag{2.14}$$

where $\vec{\boldsymbol{a}}_{\phi,\infty} = (a_{\phi,0}, a_{\phi,1}, a_{\phi,2}, \ldots)^T$, $\vec{\boldsymbol{a}}_{h,\infty} = (a_{h,0}, a_{h,1}, a_{h,2}, \ldots)^T$, and the elements in $\boldsymbol{A}_\infty$ are given by

$$[\boldsymbol{A}_\infty]_{k,j} = \begin{cases} 1 - 2^{-1/2} a_{g,0}, & \text{if } k = j, \\ 2^{-1/2}[a_{g,k-j-1} - a_{g,k-j}], & \text{if } k > j, \\ 0, & \text{otherwise.} \end{cases}$$

We consider the truncate version of (2.14),

$$\boldsymbol{A}_m \vec{\boldsymbol{a}}_{\phi,m} = \vec{\boldsymbol{a}}_{h,m}, \tag{2.15}$$

where $\vec{\boldsymbol{a}}_{\phi,m} = (a_{\phi,0}, a_{\phi,1}, \ldots, a_{\phi,m-1})^T$, $\vec{\boldsymbol{a}}_{h,m} = (a_{h,0}, a_{h,1}, \ldots, a_{h,m-1})^T$, and $\boldsymbol{A}_m$ is the sub-block of $\boldsymbol{A}_\infty$. Note that $\boldsymbol{A}_m$ is a lower triangular Toeplitz matrix, and all the diagonal elements in it equal to $1 - 2^{-1/2} a_{g,0}$ and

$$1 - 2^{-1/2} a_{g,0} = 1 - 2^{-1/2} \int_0^\infty g(x) \varphi_0(x) dx = 1 - \int_0^\infty e^{-x} g(x) dx > 1 - \int_0^\infty g(x) dx > 0,$$

since g is a defective density function. As a result, the matrix $\boldsymbol{A}_m$ is invertible, and we can solve equation (2.15) to obtain $\vec{\boldsymbol{a}}_{\phi,m} = \boldsymbol{A}_m^{-1} \vec{\boldsymbol{a}}_{h,m}$.

3 The estimator

By the definition of ϕ_m, it suffices to estimate the coefficients $a_{\phi,k}$ for $k = 0, 1, \ldots, m-1$, or the vector $\vec{\boldsymbol{a}}_{\phi,m}$. Furthermore, due to $\vec{\boldsymbol{a}}_{\phi,m} = \boldsymbol{A}_m^{-1} \vec{\boldsymbol{a}}_{h,m}$ and the definitions of $\boldsymbol{A}_m$ and $\vec{\boldsymbol{a}}_{h,m}$, we have to estimate the following quantities

$$a_{g,k}, \quad a_{h,k}, \quad k = 0, 1, \ldots, m-1.$$

For $a_{g,k}$, we have

$$\begin{aligned} a_{g,k} &= \int_0^\infty g(x) \varphi_k(x) dx = \frac{1}{c} \int_0^\infty \int_x^\infty e^{-\rho(y-x)} \nu(y) dy \varphi_k(x) dx \\ &= \int_0^\infty Q_k(y, \rho) \nu(y) dy, \end{aligned} \tag{3.1}$$

where by changing the order of integrals,

$$Q_k(y, \rho) = \frac{1}{c} \int_0^y e^{-\rho(y-x)} \varphi_k(x) dx, \quad y > 0.$$

Moreover, by some careful calculations we have for $\rho \neq 1$

$$\begin{aligned} Q_k(y, \rho) &= \frac{\sqrt{2}}{c} e^{\rho y} \sum_{j=0}^k (-2)^j \binom{k}{j} \int_0^y \frac{x^j}{j!} e^{-(1-\rho)x} dx \\ &= \frac{1}{c} \frac{\sqrt{2}}{1-\rho} \sum_{j=0}^k \left(\frac{2}{\rho - 1}\right)^j \binom{k}{j} \left(e^{-\rho y} - \sum_{l=0}^j \frac{[(1-\rho)y]^l}{l!} e^{-y} \right). \end{aligned} \tag{3.2}$$

For $a_{h,k}$ we have

$$\begin{aligned} a_{h,k} &= \int_0^\infty h(u)\varphi_k(u)du = \frac{1}{c}\int_0^\infty \int_u^\infty e^{-\rho(x-u)} \int_x^\infty w(x,y-x)\nu(y)dydx\varphi_k(u)du \\ &= \int_0^\infty R_k(y,\rho)\nu(y)dy, \end{aligned} \tag{3.3}$$

where by changing the order of integrals,

$$R_k(y,\rho) = \frac{1}{c}\int_0^y \int_u^y e^{-\rho(x-u)} w(x,y-x)\varphi_k(u)dxdu, \quad y > 0. \tag{3.4}$$

We remark that for most of the interesting penalty functions used in ruin theory, we can compute the double integrals in (3.4) to obtain closed-form expressions for $R_k(y,\rho)$.

Now we use formulas (3.1) and (3.3) to estimate $a_{g,k}$ and $a_{h,k}$. First, we estimate the Lundberg exponent ρ. Recall that we have the sample $\{X_{k\Delta}:\ k = 0,1,\ldots,n\}$. For convenience, we put

$$Z_k = X_{k\Delta} - X_{(k-1)\Delta}, \quad k = 1,\ldots,n.$$

As in Shimizu and Zhang (2017), the estimate of ρ, denoted by $\widehat{\rho}$, is defined to be the positive root of the following equation (in s)

$$cs + \frac{1}{n\Delta}\sum_{j=1}^n [e^{-sZ_j} - 1] = \delta.$$

Further, we put $\widehat{\rho} = 0$ as $\delta = 0$, since $\rho = 0$ in this case. The following lemma gives the consistency property of $\widehat{\rho}$. See Shimizu and Zhang (2017).

Lemma 2 *Suppose that Conditions* **A.1**, **A.2** *and* $\mu_2 < \infty$ *hold. Then for* $\delta > 0$ *we have*

$$\widehat{\rho} - \rho = O_p((n\Delta)^{-\frac{1}{2}} + \Delta).$$

It is known that under the high-frequency condition **A.2** that $\frac{1}{n\Delta}\sum_{j=1}^n \delta_{Z_j}(dx)$ converges weakly to the measure $\nu(x)dx$, where δ_x denotes the Dirac measure at x. Then by formulae (3.1) and (3.3) we can estimate $a_{g,k}$ and $a_{h,k}$ by

$$\widehat{a}_{g,k} = \frac{1}{n\Delta}\sum_{j=1}^n Q_k(Z_j,\widehat{\rho}), \quad \widehat{a}_{h,k} = \frac{1}{n\Delta}\sum_{j=1}^n R_k(Z_j,\widehat{\rho}).$$

We define the estimates of $\boldsymbol{A}_m$ and $\vec{\boldsymbol{a}}_{h,m}$ by replacing $a_{g,k}$ and $a_{h,k}$ by $\widehat{a}_{g,k}$ and $\widehat{a}_{h,k}$, and denote them by $\widehat{\boldsymbol{A}}_m$ and $\widehat{\vec{\boldsymbol{a}}}_{h,m}$, i.e.

$$[\widehat{\boldsymbol{A}}_m]_{k,j} = \begin{cases} 1 - 2^{-1/2}\widehat{a}_{g,0}, & \text{if } k = j, \\ 2^{-1/2}[\widehat{a}_{g,k-j-1} - \widehat{a}_{g,k-j}], & \text{if } k > j, \\ 0, & \text{otherwise,} \end{cases}$$

and $\widehat{\vec{\boldsymbol{a}}}_{m,h} = (\widehat{a}_{h,0},\widehat{a}_{h,1},\ldots,\widehat{a}_{h,m-1})^T$. Finally, we estimate ϕ_m by

$$\widehat{\phi}_m(u) = \sum_{k=0}^{m-1} \widehat{a}_{\phi,k}\varphi_k(u), \quad u \geq 0, \tag{3.5}$$

where

$$\widehat{\vec{\boldsymbol{a}}}_{\phi,m} = (\widehat{a}_{\phi,0},\widehat{a}_{\phi,1},\ldots,\widehat{a}_{\phi,m-1})^T = \widehat{\boldsymbol{A}}_m^{-1}\widehat{\vec{\boldsymbol{a}}}_{h,m}.$$

4 Consistency property

In the sequel, we use C to denote a generic positive constant that may take different values at different steps. For two positive sequences (a_n) and (b_n), we use $a_n \lesssim b_n$ to mean $a_n \leqslant C \cdot b_n$ uniformly in $n \in \mathbb{N}^+$ for some $C > 0$. Similarly, for two positive functions f_1, f_2 on $\mathbb{R}_+$, we use $f_1 \lesssim f_2$ to denote $f_1(x) \leq C f_2(x)$ uniformly in $x \in \mathbb{R}_+$. For any vector $\vec{\boldsymbol{x}}$, its Euchidean norm $\|\vec{\boldsymbol{x}}\|_2$ is defined by $\|\vec{\boldsymbol{x}}\|_2 = \vec{\boldsymbol{x}}^T \vec{\boldsymbol{x}}$. For a matrix $\boldsymbol{B}$, its operator norm is defined by

$$\|\boldsymbol{B}\|_{op} = \max_{\|\vec{\boldsymbol{x}}\|_2=1} \|\boldsymbol{B}\vec{\boldsymbol{x}}\|_2 = \sqrt{\lambda_{max}(\boldsymbol{B}^T\boldsymbol{B})},$$

where $\lambda_{max}(\boldsymbol{B}^T\boldsymbol{B})$ is the largest eigenvalue of $\boldsymbol{B}^T\boldsymbol{B}$. The Frobenius norm of matrix $\boldsymbol{B}$ is defined by $\|\boldsymbol{B}\|_F = \sqrt{\sum_{ij} b_{ij}^2}$, where b_{ij} are elements of $\boldsymbol{B}$. For two matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, we have

$$\|\boldsymbol{B}_1\boldsymbol{B}_2\|_F \leq \|\boldsymbol{B}_1\|_F\|\boldsymbol{B}_2\|_{op} \quad \text{and} \quad \|\boldsymbol{B}_1\boldsymbol{B}_2\|_F \leq \|\boldsymbol{B}_1\|_{op}\|\boldsymbol{B}_2\|_F. \tag{4.1}$$

If $\boldsymbol{B}$ is a square matrix of dimension m, we have

$$\frac{1}{\sqrt{m}}\|\boldsymbol{B}\|_F \leq \|\boldsymbol{B}\|_{op} \leq \|\boldsymbol{B}\|_F. \tag{4.2}$$

Now we derive the asymptotic properties of our estimate $\widehat{\phi}_m(u)$. First, by Pythagoras principle we have

$$\|\phi - \widehat{\phi}_m\|^2 = \|\phi - \phi_m\|^2 + \|\phi_m - \widehat{\phi}_m\|^2. \tag{4.3}$$

The first term on the right hand side of (4.3) can be bounded by (2.11) if $\phi \in W(\mathbb{R}_+, r, B)$. Let us consider the second term on the right hand size of (4.3). Using triangle inequality and Jensen's inequality we have

$$\begin{aligned}
\|\phi_m - \widehat{\phi}_m\|^2 &= \|\boldsymbol{A}_m^{-1}\vec{\boldsymbol{a}}_{h,m} - \widehat{\boldsymbol{A}}_m^{-1}\vec{\widehat{\boldsymbol{a}}}_{h,m}\|_2^2 \\
&= \|\boldsymbol{A}_m^{-1}(\vec{\boldsymbol{a}}_{h,m} - \vec{\widehat{\boldsymbol{a}}}_{h,m}) + (\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1})(\vec{\widehat{\boldsymbol{a}}}_{h,m} - \vec{\boldsymbol{a}}_{h,m}) + (\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1})\vec{\boldsymbol{a}}_{h,m}\|_2^2 \\
&\leq 3\|\boldsymbol{A}_m^{-1}(\vec{\boldsymbol{a}}_{h,m} - \vec{\widehat{\boldsymbol{a}}}_{h,m})\|_2^2 + 3\|(\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1})(\vec{\widehat{\boldsymbol{a}}}_{h,m} - \vec{\boldsymbol{a}}_{h,m})\|_2^2 + 3\|(\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1})\vec{\boldsymbol{a}}_{h,m}\|_2^2 \\
&\leq 3\|\boldsymbol{A}_m^{-1}\|_{op}^2 \cdot \|\vec{\boldsymbol{a}}_{h,m} - \vec{\widehat{\boldsymbol{a}}}_{h,m}\|_2^2 + 3\|\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1}\|_{op}^2 \cdot \|\vec{\widehat{\boldsymbol{a}}}_{h,m} - \vec{\boldsymbol{a}}_{h,m}\|_2^2 \\
&\quad + 3\|\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1}\|_{op}^2 \cdot \|\vec{\boldsymbol{a}}_{h,m}\|_2^2.
\end{aligned} \tag{4.4}$$

To continue with, we need some conditions and lemmas to show the consistency property.

Condition B.1 For some $\alpha_1, \alpha_2 > 0$,

$$w(x_1, x_2) \leq C(1 + x_1)^{\alpha_1}(1 + x_2)^{\alpha_2}, \qquad \left|\frac{d}{dx_2}w(x_1, x_2)\right| \leq C(1 + x_1)^{\alpha_1}(1 + x_2)^{\alpha_2}.$$

Condition B.2(k) For some integer $k \geq 1$, $\mu_k < \infty$.

Condition B.3 The Lévy density ν is continuous on $(0, \infty)$, and for some $\alpha \in (0, 1)$,

$$\lim_{\Delta \to 0} \Delta^{2-\alpha}\nu(\Delta) = 0.$$

Remark 4.1 *Condition **B.1** is not very restrictive. For example, it is satisfied by the penalty functions used in ruin probability, the expected claim size causing ruin and the joint moments (or Laplace transform) of the deficit at ruin and the deficit at ruin. Condition **B.3** is not very restrictive too, and it is satisfied by many widely used subordintors in ruin theory, e.g. compound Poisson, Lévy-Gamma and NIG subordinators.*

Lemma 3 *Under Condition **A.1**, we have for all* $m \geq 1$

$$\|\boldsymbol{A}_m^{-1}\|_{op} \leq \frac{2c}{c-\mu_1}.$$

Proof. This can be easily obtained by using the same arguments as in Zhang and Su (2017). ∎

Lemma 4 *Under Condition **A.3**, we have* $\|\vec{\boldsymbol{a}}_{h,m}\|_2^2 \leq \|h\|^2 < \infty$.

Proof. This holds since

$$\|\vec{\boldsymbol{a}}_{h,m}\|_2^2 = \sum_{k=0}^{m-1} a_{h,k}^2 \leq \sum_{k=0}^{\infty} a_{h,k}^2 = \|h\|^2 < \infty.$$

This completes the proof. ∎

Lemma 5 *Under Conditions* **A.2, B.1, B.2**$(2(\alpha_1+\alpha_2+2))$ *and* **B.3**, *we have*

$$\|\vec{\boldsymbol{a}}_{h,m} - \widehat{\vec{\boldsymbol{a}}}_{h,m}\|_2^2 = O_p(m((n\Delta)^{-1} + \Delta^2)) + o(m\Delta^{2\alpha}). \tag{4.5}$$

Proof. For each k we have

$$a_{h,k} - \widehat{a}_{h,k} = \mathrm{I}_{k,1} + \mathrm{I}_{k,2} + \mathrm{I}_{k,3},$$

where

$$\begin{aligned}
\mathrm{I}_{k,1} &= \frac{1}{n\Delta}\sum_{j=1}^{n}(R_k(Z_j,\rho) - R_k(Z_j,\widehat{\rho})),\\
\mathrm{I}_{k,2} &= \frac{1}{n\Delta}\sum_{j=1}^{n}\left(\mathbb{E}[R_k(Z_j,\rho)] - R_k(Z_j,\rho)\right),\\
\mathrm{I}_{k,3} &= \int_0^{\infty} R_k(y,\rho)\nu(y)dy - \frac{1}{\Delta}\mathbb{E}[R_k(Z_1,\rho)].
\end{aligned}$$

Then using Jensen's inequality we have

$$\|\vec{\boldsymbol{a}}_{h,m} - \widehat{\vec{\boldsymbol{a}}}_{h,m}\|_2^2 = \sum_{k=0}^{m-1}(a_{h,k} - \widehat{a}_{h,k})^2 = \sum_{k=0}^{m-1}(\mathrm{I}_{k,1} + \mathrm{I}_{k,2} + \mathrm{I}_{k,3})^2$$

$$\leq \quad 3\sum_{k=0}^{m-1}(\mathrm{I}_{k,1}^2+\mathrm{I}_{k,2}^2+\mathrm{I}_{k,3}^2). \tag{4.6}$$

First, we consider the sum $\sum_{k=0}^{m-1}\mathrm{I}_{k,1}^2$. Recall that $\rho=\widehat{\rho}=0$ as $\delta=0$. Hence, we have $\mathrm{I}_{k,1}=0$ as $\delta=0$, and this leads to $\sum_{k=0}^{m-1}\mathrm{I}_{k,1}^2=0$. When $\delta>0$, by the mean value theory we know that there exists a random number $\rho^*\geq\delta/c$ (due to $\widehat{\rho},\rho\geq\delta/c$. See e.g. Shimizu and Zhang (2017)) such that for $x\geq u$

$$|e^{-\rho(x-u)}-e^{-\widehat{\rho}(x-u)}|=|e^{-\rho^*(x-u)}(x-u)(\rho-\widehat{\rho})|\leq e^{-\frac{\delta}{c}(x-u)}(x-u)|\rho-\widehat{\rho}|\leq\frac{c}{e\delta}|\widehat{\rho}-\rho|. \tag{4.7}$$

As a result, by the uniform upper bound (2.2) and Condition **B.1** we have

$$\begin{aligned}
\sup_{k\geq0}|\mathrm{I}_{k,1}| &\leq \frac{1}{c}\frac{1}{n\Delta}\sum_{j=1}^{n}\int_0^{Z_j}\int_u^{Z_j}|e^{-\rho(x-u)}-e^{-\widehat{\rho}(x-u)}|w(x,Z_j-x)\sup_{k\geq0}|\varphi_k(u)|dxdu\\
&\lesssim \frac{1}{n\Delta}\sum_{j=1}^{n}\int_0^{Z_j}\int_u^{Z_j}w(x,Z_j-x)dxdu\cdot|\rho-\widehat{\rho}|\\
&\lesssim \frac{1}{n\Delta}\sum_{j=1}^{n}\int_0^{Z_j}\int_u^{Z_j}(1+x)^{\alpha_1}(1+Z_j-x)^{\alpha_2}dxdu\cdot|\rho-\widehat{\rho}|\\
&\lesssim \frac{1}{n\Delta}\sum_{j=1}^{n}Z_j^2(1+Z_j)^{\alpha_1+\alpha_2}\cdot|\rho-\widehat{\rho}|. \qquad (4.8)
\end{aligned}$$

Under Condition **B.2(**$\alpha_1+\alpha_2+2$**)**, we have

$$\frac{1}{n\Delta}\sum_{j=1}^{n}\mathbb{E}[Z_j^2(1+Z_j)^{\alpha_1+\alpha_2}]\lesssim\frac{1}{n\Delta}\sum_{j=1}^{n}\mathbb{E}[Z_j^2+Z_j^{\alpha_1+\alpha_2+2}]<\infty$$

due to Lemma 8, then Markov's inequality gives $\frac{1}{n\Delta}\sum_{j=1}^{n}[Z_j^2(1+Z_j)^{\alpha_1+\alpha_2}]=O_p(1)$. This together with Lemma 2 and (4.8) gives

$$\sup_{k\geq0}|\mathrm{I}_{k,1}|=O_p(1)\cdot O_p((n\Delta)^{-\frac{1}{2}}+\Delta)=O_p((n\Delta)^{-\frac{1}{2}}+\Delta),$$

which yields

$$\sum_{k=0}^{m-1}\mathrm{I}_{k,1}^2=O_p\left(m((n\Delta)^{-1}+\Delta^2)\right). \tag{4.9}$$

Next, by the uniform upper bound (2.2) and Condition **B.1** we have

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=0}^{m-1}\mathrm{I}_{k,2}^2\right) &= \sum_{k=0}^{m-1}Var(\mathrm{I}_{k,2})=\sum_{k=0}^{m-1}\frac{1}{c^2}\frac{1}{n\Delta^2}Var\left(\int_0^{Z_1}\int_u^{Z_1}e^{-\rho(x-u)}w(x,Z_1-x)dx\varphi_k(u)du\right)\\
&\leq \sum_{k=0}^{m-1}\frac{1}{c^2}\frac{1}{n\Delta^2}\mathbb{E}\left(\int_0^{Z_1}\int_u^{Z_1}e^{-\rho(x-u)}w(x,Z_1-x)dx\varphi_k(u)du\right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{k=0}^{m-1} \frac{1}{c^2}\frac{2}{n\Delta^2}\mathbb{E}\left(\int_0^{Z_1}\int_u^{Z_1} w(x, Z_1-x)dxdu\right)^2 \\
&\lesssim \sum_{k=0}^{m-1} \frac{1}{n\Delta^2}\mathbb{E}\left(\int_0^{Z_1}\int_u^{Z_1} (1+x)^{\alpha_1}(1+Z_1-x)^{\alpha_2}dxdu\right)^2 \\
&\lesssim \sum_{k=0}^{m-1} \frac{1}{n\Delta^2}\mathbb{E}[Z_1^4(1+Z_1)^{2(\alpha_1+\alpha_2)}] \lesssim \frac{m}{n\Delta},
\end{aligned}
$$

where the last step holds under Condition **B.2**$(2(\alpha_1+\alpha_2+2))$ due to Lemma 8. Then we have

$$
\sum_{k=0}^{m-1} \mathrm{I}_{k,2}^2 = O_p\left(m(n\Delta)^{-1}\right) \tag{4.10}
$$

thanks to Markov's inequality.

Third, we consider $\mathrm{I}_{k,3}$. By the uniform upper bound (2.2) and Condition **B.1** we have

$$
\begin{aligned}
\sup_{k\geq 0}|R_k(y,\rho)| &\leq \frac{1}{c}\int_0^y \left|\int_u^y e^{-\rho(x-u)}w(x,y-x)dx\right| \cdot \sup_{k\geq 0}|\varphi_k(u)|du \\
&\leq \frac{\sqrt{2}}{c}\int_0^y \left|\int_u^y e^{-\rho(x-u)}w(x,y-x)dx\right| du \\
&\lesssim \int_0^y\int_u^y (1+x)^{\alpha_1}(1+y-x)^{\alpha_2}dxdu \\
&\lesssim \int_0^y (y-u)(1+y)^{\alpha_1+\alpha_2}du \\
&\lesssim y^2(1+y)^{\alpha_1+\alpha_2}
\end{aligned}
$$

and

$$
\begin{aligned}
\sup_{k\geq 0}\left|\frac{d}{dy}R_k(y,\rho)\right| &\leq \frac{1}{c}\int_0^y \left|\frac{d}{dy}\int_u^y e^{-\rho(x-u)}w(x,y-x)dx\right| \cdot \sup_{k\geq 0}|\varphi_k(u)|du \\
&\leq \frac{\sqrt{2}}{c}\int_0^y \left|e^{-\rho(y-u)}w(y,0) + \int_u^y e^{-\rho(x-u)}\frac{d}{dy}w(x,y-x)dx\right| du \\
&\lesssim \int_0^y \left((1+y)^{\alpha_1} + \int_u^y (1+x)^{\alpha_1}(1+y-x)^{\alpha_2}dx\right)du \\
&\leq y(1+y)^{\alpha_1} + y^2(1+y)^{\alpha_1+\alpha_2}.
\end{aligned}
$$

Then using Lemma 9 we obtain $\sup_{k\geq 0}|\mathrm{I}_{k,3}| = o(\Delta^\alpha)$ due to $\mu_{\alpha_1+\alpha_2+2} < \infty$. As a result, we have

$$
\sum_{k=0}^{m} \mathrm{I}_{k,3}^2 = o(m\Delta^{2\alpha}). \tag{4.11}
$$

Finally, plugging (4.9)-(4.11) into (4.6) we obtain (4.5). This completes the proof. ■

Lemma 6 *Under Conditions* **A.2, B.2**(2) *and* **B.3***, we have*

$$\|\boldsymbol{A}_m - \widehat{\boldsymbol{A}}_m\|_F^2 = O_p(m^2((n\Delta)^{-1} + \Delta^2)) + o(m^2\Delta^{2\alpha}). \tag{4.12}$$

Proof. Note that

$$[\boldsymbol{A}_m - \widehat{\boldsymbol{A}}_m]_{k,j} = \begin{cases} 2^{-1/2}[\widehat{a}_{g,0} - a_{g,0}], & \text{if } k = j, \\ 2^{-1/2}[\widehat{a}_{g,k-j} - a_{g,k-j}] - 2^{-1/2}[\widehat{a}_{g,k-j-1} - a_{g,k-j-1}], & \text{if } k > j, \\ 0, & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{aligned} \|\boldsymbol{A}_m - \widehat{\boldsymbol{A}}_m\|_F^2 &= \sum_{k=1}^m \sum_{j=1}^k [\boldsymbol{A}_m - \widehat{\boldsymbol{A}}_m]_{k,j}^2 \\ &= \sum_{k=1}^m \sum_{j=1}^k 2^{-1} \left([\widehat{a}_{g,k-j} - a_{g,k-1}] + [\widehat{a}_{g,k-j-1} - a_{g,k-j-1}]\right)^2 \\ &\leq \sum_{k=1}^m \sum_{j=1}^k \left([\widehat{a}_{g,k-j} - a_{g,k-j}]^2 + [\widehat{a}_{g,k-j-1} - a_{g,k-j-1}]^2\right), \end{aligned} \tag{4.13}$$

with the understanding that $a_{g,-1} = \widehat{a}_{g,-1} = 0$. Furthermore, for each k we have

$$\widehat{a}_{g,k} - a_{g,k} = \mathrm{II}_{k,1} + \mathrm{II}_{k,2} + \mathrm{II}_{k,3},$$

where

$$\begin{aligned} \mathrm{II}_{k,1} &= \frac{1}{n\Delta} \sum_{j=1}^n [Q_k(Z_j, \widehat{\rho}) - Q_k(Z_j, \rho)], \\ \mathrm{II}_{k,2} &= \frac{1}{n\Delta} \sum_{j=1}^n \left(Q_k(Z_j, \rho) - \mathbb{E}[Q_k(Z_j, \rho)]\right), \\ \mathrm{II}_{k,3} &= \frac{1}{n\Delta} \sum_{j=1}^n \mathbb{E}[Q_k(Z_j, \rho)] - \int_0^\infty Q_k(y, \rho)\nu(y)dy. \end{aligned}$$

Then (4.13) together with Jensen's inequality yields

$$\begin{aligned} \|\boldsymbol{A}_m - \widehat{\boldsymbol{A}}_m\|_F^2 &\leq \sum_{k=1}^m \sum_{j=1}^k \left([\mathrm{II}_{k-j,1} + \mathrm{II}_{k-j,2} + \mathrm{II}_{k-j,3}]^2 + [\mathrm{II}_{k-j-1,1} + \mathrm{II}_{k-j-1,2} + \mathrm{II}_{k-j-1,3}]^2\right) \\ &\leq 6 \sum_{k=1}^m \sum_{j=1}^k \left(\mathrm{II}_{k-j,1}^2 + \mathrm{II}_{k-j,2}^2 + \mathrm{II}_{k-j,3}^2\right) \\ &\leq 6m \sum_{k=0}^{m-1} (\mathrm{II}_{k,1}^2 + \mathrm{II}_{k,2}^2 + \mathrm{II}_{k,3}^2). \end{aligned} \tag{4.14}$$

First, we consider the sum $\sum_{k=0}^{m-1} \mathrm{II}_{k,1}^2$. Note that we have $\sum_{k=0}^{m-1} \mathrm{II}_{k,1}^2 = 0$ as $\delta = 0$ since $\rho = \widehat{\rho} = 0$. Let us consider the case $\delta > 0$. By inequalities (2.2) and (4.7) we have

$$\sup_{k\geq 0} |\mathrm{II}_{k,1}| \leq \frac{1}{c} \frac{1}{n\Delta} \sum_{j=1}^n \int_0^{Z_j} |e^{-\widehat{\rho}(Z_j - x)} - e^{-\rho(Z_j - x)}| \cdot \sup_{k\geq 0} |\varphi_k(x)| dx$$

$$
\begin{aligned}
&\leq \frac{\sqrt{2}}{c}\frac{1}{n\Delta}\sum_{j=1}^{n}\int_{0}^{Z_j}(Z_j-x)dx\cdot|\widehat{\rho}-\rho| \\
&= \frac{\sqrt{2}}{2c}\frac{1}{n\Delta}\sum_{j=1}^{n}Z_j^2\cdot|\widehat{\rho}-\rho|,
\end{aligned}
$$

which together with Markov's inequality and Lemmas 2 and 8 gives

$$\sup_{k\geq 0}|\mathrm{II}_{k,1}| = O_p(1)\cdot O_p((n\Delta)^{-\frac{1}{2}}+\Delta) = O_p((n\Delta)^{-\frac{1}{2}}+\Delta).$$

Then we have

$$\sum_{k=0}^{m-1}\mathrm{II}_{k,1}^2 = O_p(m((n\Delta)^{-1}+\Delta^2)). \tag{4.15}$$

Next, we consider the sum $\sum_{k=0}^{m-1}\mathrm{II}_{k,2}^2$. We have

$$
\begin{aligned}
\mathbb{E}\left(\sum_{k=0}^{m-1}\mathrm{II}_{k,2}^2\right) &= \frac{1}{n\Delta^2}\sum_{k=0}^{m-1}Var(Q_k(Z_1,\rho)) \leq \frac{1}{n\Delta^2}\sum_{k=0}^{m-1}\mathbb{E}|Q_k(Z_1,\rho)|^2 \\
&\leq \frac{1}{c^2n\Delta^2}\sum_{k=0}^{m-1}\mathbb{E}\left(\int_0^{Z_1}e^{-\rho(Z_1-x)}|\varphi_k(x)|dx\right)^2 \\
&\leq \frac{2m}{c^2n\Delta^2}\mathbb{E}Z_1^2 \\
&= \frac{2m}{c^2n\Delta}(\mu_2+\Delta\mu_1^2),
\end{aligned}
$$

where the last equality follows from Lemma 8. Then by Markov's inequality we obtain

$$\sum_{k=0}^{m-1}\mathrm{II}_{k,2}^2 = O_p(m(n\Delta)^{-1}). \tag{4.16}$$

Third, for $\mathrm{II}_{k,3}$ we can prove (see Appendix B)

$$\sup_{k\geq 0}|\mathrm{II}_{k,3}| = o(\Delta^{\alpha}), \tag{4.17}$$

which yields

$$\sum_{k=0}^{m-1}\mathrm{II}_{k,3}^2 = o(m\Delta^{2\alpha}). \tag{4.18}$$

Finally, substituting (4.15), (4.16) and (4.18) into (4.14) we complete the proof. ■

Lemma 7 *Suppose that $m^2(n\Delta)^{-1} = o(1)$ and $m^2\Delta^{2\alpha} = o(1)$. Then under Conditions* **A.1, A.2, B.2(**2**)** *and* **B.3**, *we have*

$$\|\boldsymbol{A}_m^{-1}-\widehat{\boldsymbol{A}}_m^{-1}\|_{op}^2 = O_p(m^2((n\Delta)^{-1}+\Delta^2))+o(m^2\Delta^{2\alpha}).$$

Proof. By Lemma 10 in Appendix A we have

$$\begin{aligned}
\|\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1}\|_{op} &= \|\boldsymbol{A}_m^{-1} - (\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m + \boldsymbol{A}_m)^{-1}\|_{op} \\
&\le \frac{\|\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m\|_{op} \cdot \|\boldsymbol{A}_m^{-1}\|_{op}^2}{1 - \|\boldsymbol{A}_m^{-1}(\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m)\|_{op}} \\
&\le \frac{4c^2}{(c-\mu_1)^2} \frac{\|\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m\|_{op}}{1 - \|\boldsymbol{A}_m^{-1}(\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m)\|_{op}} \\
&\le \frac{4c^2}{(c-\mu_1)^2} \frac{\|\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m\|_{F}}{1 - \|\boldsymbol{A}_m^{-1}(\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m)\|_{op}}.
\end{aligned} \tag{4.19}$$

Furthermore, under conditions $m^2(n\Delta)^{-1} = o(1)$ and $m^2\Delta^{2\alpha} = o(1)$, we have

$$\|\boldsymbol{A}_m^{-1}(\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m)\|_{op} \le \|\boldsymbol{A}_m^{-1}\|_{op} \cdot \|\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m\|_{op} \le \frac{2c}{c-\mu_1}\|\widehat{\boldsymbol{A}}_m - \boldsymbol{A}_m\|_{op} = o_p(1).$$

This together with (4.19) and Lemma 6 completes the proof. ■

Theorem 4.1 *Suppose that $m^2(n\Delta)^{-1} = o(1)$ and $m^2\Delta^{2\alpha} = o(1)$. Then under Conditions* **A.1-A.4, B.1, B.2**$(2(\alpha_1+\alpha_2+2))$ *and* **B.3**, *we have*

$$\|\phi - \widehat{\phi}_m\|^2 = \|\phi - \phi_m\|^2 + O_p(m^2((n\Delta)^{-1} + \Delta^2)) + o(m^2\Delta^{2\alpha}). \tag{4.20}$$

Proof. By Lemmas 3-6 we have

$$\begin{aligned}
\|\boldsymbol{A}_m^{-1}\|_{op}^2 \cdot \|\vec{\boldsymbol{a}}_{h,m} - \vec{\widehat{\boldsymbol{a}}}_{h,m}\|_2^2 &= O_p(m((n\Delta)^{-1} + \Delta^2)) + o(m\Delta^{2\alpha}), \\
\|\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1}\|_{op}^2 \cdot \|\vec{\boldsymbol{a}}_{h,m}\|_2^2 &= O_p(m^2((n\Delta)^{-1} + \Delta^2)) + o(m^2\Delta^{2\alpha}),
\end{aligned}$$

and

$$\|\boldsymbol{A}_m^{-1} - \widehat{\boldsymbol{A}}_m^{-1}\|_{op}^2 \cdot \|\vec{\widehat{\boldsymbol{a}}}_{h,m} - \vec{\boldsymbol{a}}_{h,m}\|_2^2 = o_p\left(\|\boldsymbol{A}_m^{-1}\|_{op}^2 \cdot \|\vec{\boldsymbol{a}}_{h,m} - \vec{\widehat{\boldsymbol{a}}}_{h,m}\|_2^2\right).$$

Substituting the above results back into (4.4) and comparing the convergence rates we can complete the proof. ■

Remark 4.2 *Suppose the conditions in Theorem 4.1. Then by* (2.11) *and* (4.20) *we have*

$$\|\phi - \widehat{\phi}_m\|^2 = O(m^{-r}) + O_p(m^2((n\Delta)^{-1} + \Delta^2)) + o(m^2\Delta^{2\alpha}). \tag{4.21}$$

Omitting the term $o(m^2\Delta^{2\alpha})$ and minimizing the order $O(m^{-r}) + O_p(m^2((n\Delta)^{-1} + \Delta^2))$ w.r.t. m we find that the optimal truncate parameter, say m_{op}, has order $O(((n\Delta)^{-1} + \Delta^2)^{-\frac{1}{r+2}})$.

5 Simulation studies

In this section, let us present some simulation examples to illustrate the performance of our estimator under finite sample setting. We consider the following two Lévy risk models.

(1) The compound Poisson risk model with premium rate $c = 30$, Poisson intensity $\lambda = 20$ and exponential jumps with mean $\mu = 2/3$. In this case, the Levy density $\nu(x) = \lambda/\mu e^{-x/\mu}$, $x > 0$.

(2) The Lévy-Gamma risk model with premium rate $c = 1.5$ and Gamma-type density $\nu(x) = x^{-1}e^{-x}$, $x > 0$.

As in Shimizu and Zhang (2017), we estimate the following three classes of Gerber-Shiu functions:

- ruin probability (RP): $\phi(u) = \mathbb{P}(\tau < \infty | U_0 = u)$ with $\delta = 0$ and $w(x, y) \equiv 1$;
- expected claim size causing ruin (ECS): $\phi(u) = \mathbb{E}\left[(U_{\tau-} + |U_\tau|)\mathbf{1}_{(\tau<\infty)} | U_0 = u\right]$ with $\delta = 0$ and $w(x, y) = x + y$;
- Laplace transform of ruin time (LT): $\phi(u) = \mathbb{E}\left[e^{-\delta\tau}\mathbf{1}_{(\tau<\infty)} | U_0 = u\right]$ with $\delta = 0.1$ and $w(x, y) \equiv 1$.

For the compound Poisson model with exponential claims, true value of the Gerber-Shiu functions can be computed based on the following formulae that are provided in Shimizu and Zhang (2017),

- $\phi(u) = \frac{\lambda\mu}{c} e^{-(1/\mu - \lambda/c)u}$;
- $\phi(u) = \mu(1 + 2\frac{\lambda\mu}{c})e^{-(1/\mu-\lambda/c)u} - \mu e^{-u/\mu}$;
- $\phi(u) = \frac{\lambda\mu}{c(1+\rho\mu)} e^{-(\rho+1/\mu-(\lambda+\delta)/c)u}$.

For the Lévy-Gamma risk model, explicit formulae for these Gerber-Shiu functions do not exist. However, we can use ϕ_m to approximate the true value. On the other hand, we can also use the FFT algorithm proposed in Shimizu and Zhang (2017) to approximate the true value. In Figure 1, we illustrate these two methods for comparison. In the Laguerre series expansion method, we set the truncate parameter $m = 20$, while for the FFT method, we use formula (4.1) in Shimizu and Zhang (2017) with $m = 50$ and $K = 2^{13}$. It follows from Figure 1 that the approximate curves nearly coincide, but when u is small the Laguerre series expansion can result in smoother curves than the FFT algorithm. In the sequel of this section we shall use the Laguerre series expansion with truncate parameter $m = 20$ to compute the reference value for the Lévy-Gamma risk model.

We shall consider $(n, \Delta) = (400, 0.05), (1000, 0.02), (2500, 0.01), (5000, 0.01)$, where $n\Delta = 20, 20, 25, 50$, respectively. For the truncate parameter m it is hard to find a data driven method for selection. By Remark 4.2 we know that when $\phi \in W(\mathbb{R}_+, r, B)$, the optimal truncate parameter m_{op} has order $O(((n\Delta)^{-1} + \Delta^2)^{-\frac{1}{r+2}})$. In the sequel we set $m = \lfloor 5((n\Delta)^{-1} + \Delta^2)^{-\frac{1}{10}} \rfloor$, where $\lfloor\cdot\rfloor$ denotes the integer part.

To show the performance of the estimate, we perform 1000 experiments and analyze mean value, the integrated mean square error (IMSE) and mean relative error, which are computed by

$$\frac{1}{1000}\sum_{j=1}^{1000} \widehat{\phi}_{m,j}(u), \quad \frac{1}{1000}\sum_{j=1}^{1000}\int_0^{30} |\widehat{\phi}_{m,j}(u) - \phi(u)|^2 du, \quad \frac{1}{1000}\sum_{j=1}^{1000}\left|\frac{\widehat{\phi}_{m,j}(u)}{\phi(u)} - 1\right|,$$

where $\widehat{\phi}_{m,j}(u)$ denotes the estimated Gerber-Shiu function in the j-th experiment. Note that we compute the integral on a finite domain $[0, 30]$, because both the reference value and the estimator are very close to zero as u increase.

First, we display the mean value curves of the estimated Gerber-Shiu functions based on 1000 sample paths and compare them with the reference value curves in Figures 2 and 3. We observe that the mean value curves converge to the reference value curves as $n\Delta$ increases. Furthermore, when the initial surplus u is large, it is hard to distinguish the mean value curves and the reference value curves. Next, we provide some values of IMSEs in Table 1, where we easily observe that the IMSEs decrease as $n\Delta$ increases for each Gerber-Shiu function and each Lévy risk model. Finally, we present some results on mean relative error curves in Figures 4 and 5. Again, we find that the mean relative errors decrease as $n\Delta$ increases, but we also find that the mean relative error curves increase as the initial surplus value becomes large, which is due to that the denominator $\phi(u)$ in the definition of mean relative error converges to zero when u becomes large.

Now we compare Laguerre series expansion method with FFT method used in Shimizu and Zhang (2017). For the FFT method, we use formula (4.2) in Shimizu and Zhang (2017) to compute the estimate and we use the same parameter setting as in Shimizu and Zhang (2017). First, in Table 1 we provide the IMSE values for both methods, and by comparison we find that the Laguerre series expansion method can lead to smaller IMSEs. Next, in Figures 6 and 7 we compare these two methods in terms of mean relative errors. where we find the Laguerre series expansion method can yield smaller mean relative errors compared with the FFT method.

Table 1: IMSEs for the estimation of Gerber-Shiu function.

	(n, Δ)	Laguerre			FFT		
		RP	ECS	LT	RP	ECS	LT
Compound Poisson	$(400, 0.05)$	0.016810	0.34640	0.015880	0.017340	0.35390	0.016730
	$(1000, 0.02)$	0.005424	0.07218	0.004641	0.006139	0.07571	0.005827
	$(2500, 0.01)$	0.002317	0.02792	0.002303	0.003528	0.02795	0.003597
	$(5000, 0.01)$	0.000604	0.01147	0.000863	0.002282	0.02150	0.002311
Lévy-Gamma	$(400, 0.05)$	0.231650	2.71708	0.052180	0.386713	3.39299	0.076717
	$(1000, 0.02)$	0.190217	2.31552	0.050884	0.314151	2.96768	0.061428
	$(2500, 0.01)$	0.139269	1.43479	0.038353	0.300377	2.13981	0.053417
	$(5000, 0.01)$	0.066753	0.36252	0.019565	0.131593	1.12507	0.025959

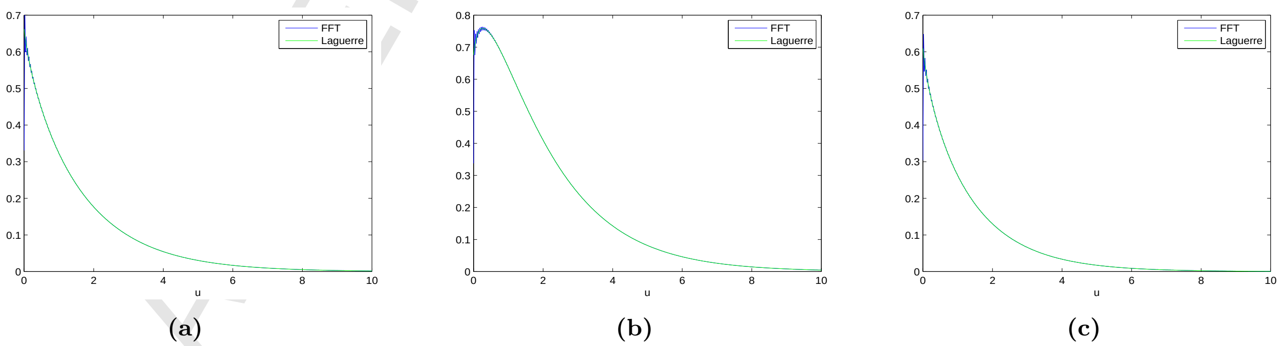
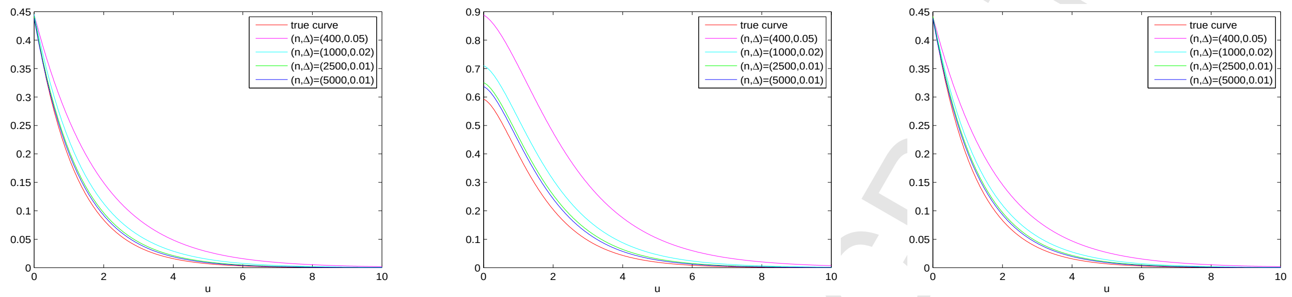
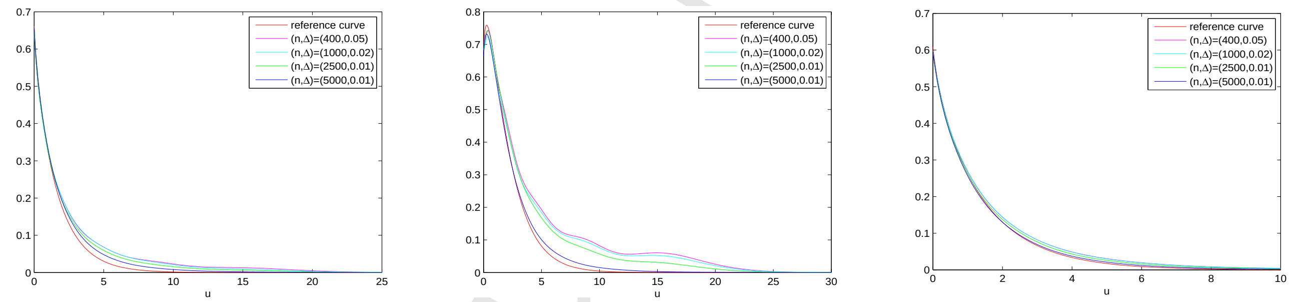


Figure 1: Comparing with FFT method for Lévy-Gamma risk model: reference value curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.



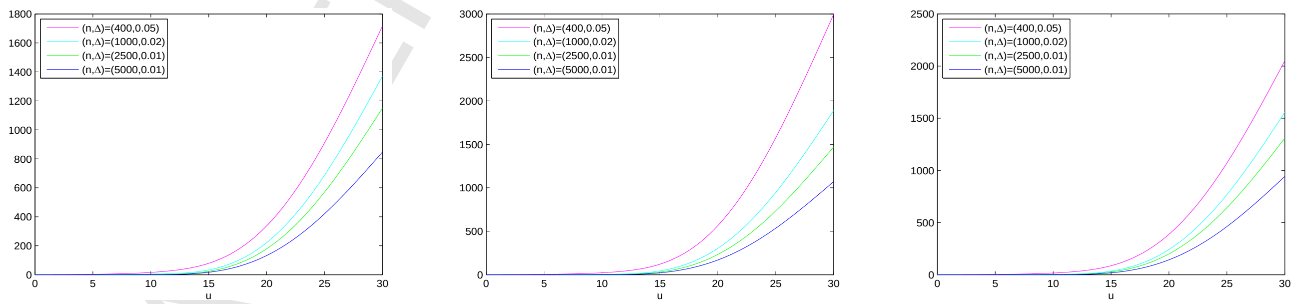
(a) (b) (c)

Figure 2: Estimating the Gerber-Shiu function for compound Poisson risk model: mean value curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.



(a) (b) (c)

Figure 3: Estimating the Gerber-Shiu function for Lévy-Gamma risk model: mean value curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.



(a) (b) (c)

Figure 4: Estimating the Gerber-Shiu function for compound Poisson risk model: mean relative error curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.

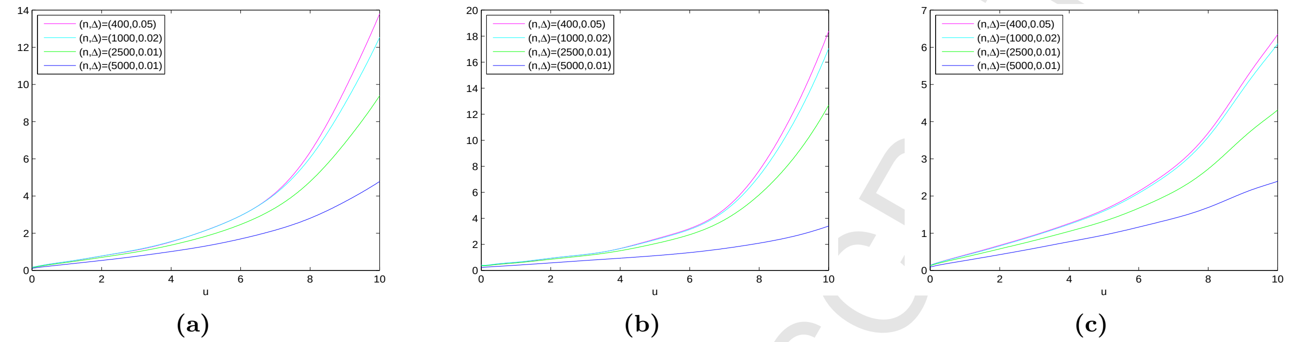


Figure 5: Estimating the Gerber-Shiu function for Lévy-Gamma risk model: mean relative error curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.

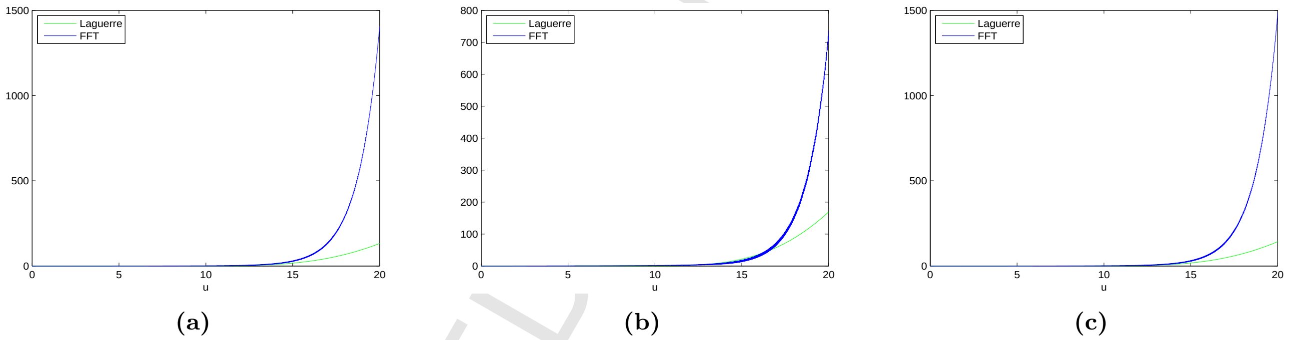


Figure 6: Comparing with FFT method for compound Poisson risk model: mean relative error curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.

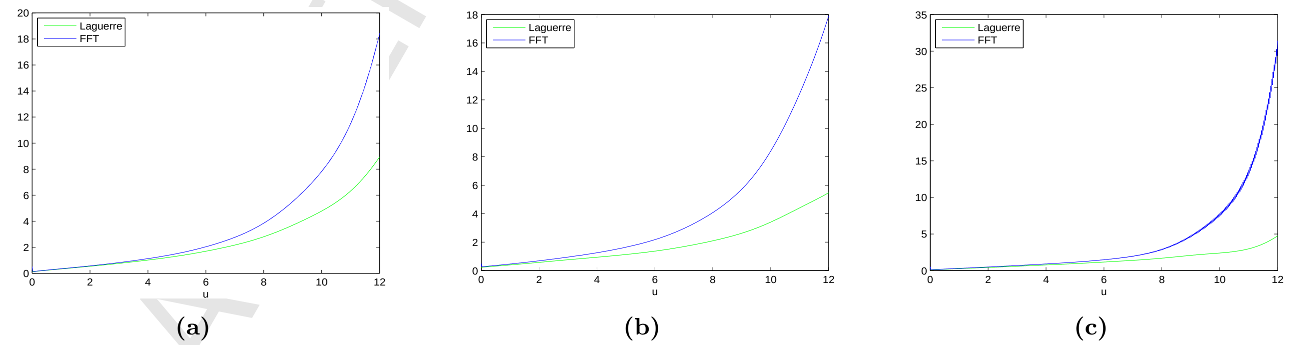


Figure 7: Comparing with FFT method for Lévy-Gamma risk model: mean relative error curves. (a) Ruin probability; (b) Expected claim size causing ruin; (c) Laplace transform of ruin time.

Acknowledgements

The authors would like to thank the two anonymous reviewers for their helpful comments and suggestions, which have significantly improved this paper. This research was supported by the National Natural Science Foundation of China (11471058, 11101451), MOE(Ministry of Education in China) Project of Humanities and Social Sciences (16YJC910005) and Fundamental Research Funds for the Central Universities (106112017CDJZRPY0201).

References

[1] Abramowitz, M., Stegun, I.A., 1964. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* National Bureau of Standards Applied Mathematics Series.

[2] Asmussen, S., Albrecher, H., 2010. *Ruin Probabilities*, Second Edition. World Scientific, Singapore.

[3] Bongioanni, B., Torrea, J.L., 2009. What is a Sobolev space for the Laguerre function system? *Studia Mathematica*, 192(2), 147–172.

[4] Comte F., Genon-Catalot, V., 2009. Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Processes and Their Applications*, 119, 4088–4123.

[5] Comte, F., Genon-Catalot, V., 2015. Adaptive Laguerre density estimation for mixed Poisson models. *Electronic Journal of Statistics*, 2015(9), 1112–1148.

[6] Comte, F., Cuenod, C.A., Pensky, M., Rozenholc, Y., 2017. Laplace deconvolution on the basis of time domain data and its application to Dynamic Contrast Enhanced imaging. *Journal of the Royal Statistical Society: Series B*, 79(1), 69–94.

[7] Gerber, H.U., Shiu, E.S.W., 1998. On the time value of ruin. *North American Actuarial Journal*, 2(1), 48-78.

[8] Mabon, G., 2016. Adaptive deconvolution of linear functionals on the nonnegative real line. *Journal of Statistical Planning and Inference*, 178, 1-23.

[9] Masiello, E., 2014. On semiparametric estimation of ruin probabilities in the classical risk model. *Scandinavian Actuarial Journal*, 2014(4), 283–308.

[10] Mnatsakanov, R., Ruymgaart, L. L., Ruymgaart, F. H., 2008. Nonparametric estimation of ruin probabilities given a random sample of claims. *Mathematical Methods of Statististics*, 17(1), 35–43.

[11] Politis, K., 2003. Semiparametric estimation for non-ruin probabilities. *Scandinavian Actuarial Journal*, 2003(1), 75–96.

[12] Kyprianou, A., 2013. *Gerber-Shiu Risk Theory.* Springer.

[13] Li, Y., Yin, C. and Zhou, X., 2017. On the last exit times for spectrally negative Lévy processes. *Journal of Applied Probability*, 54, 474-489.

[14] Shen, Y., Yin, C. and Yuen, K.C., 2013. Alternative approach to the optimality of the threshold strategy for spectrally negative Lévy processes. *Acta Mathematicae Applicatae Sinica, English Series*, 29(4), 705-716.

[15] Shimizu, Y., 2009. A new aspect of a risk process and its statistical inference. *Insurance: Mathematics and Economics*, 44(1), 70–77.

[16] Shimizu, Y., 2011. Estimation of the expected discounted penalty function for Lévy insurance risks. *Mathematican Methods of Statistics*, 20(2), 125–149.

[17] Shimizu, Y., 2012. Nonparametric estimation of the Gerber-Shiu function for the Winer-Poisson risk model. *Scandinavian Actuarial Journal*, 2012(1), 56–69.

[18] Shimizu, Y., Zhang, Z., 2017. Estimating Gerber–Shiu functions from discretely observed Lévy driven surplus. *Insurance: Mathematics and Economics*, 74, 84–98,

[19] Stewart, G.W. and Sun, J.G., 1990. *Matrix perturbation theory*. Academic Press, Beijing.

[20] Yin, C. and Wang C., 2009. Optimality of the barrier strategy in de Finettis dividend problem for spectrally negative Lévy processes: An alternative approach. *Journal of Computational and Applied Mathematics*, 233, 482-491.

[21] Yin, C. and Wen, Y., 2013. Optimal dividend problem with a terminal value for spectrally positive Lévy processes. *Insurance: Mathematics and Economics*, 53, 769-773.

[22] Yuen, K.C. and Yin, C., 2011. On optimality of the barrier strategy for a general Lévy risk process. *Mathematical and Computer Modelling*, 53, 1700-1707.

[23] Zhang, Z., 2017. Estimating the Gerber-Shiu function by Fourier-Sinc series expansion. *Sandinavian Actuarial Journal*, 2017(10), 898-919

[24] Zhang, Z., 2017. Nonparametric estimation of the finite time ruin probability in the classical risk model. *Sandinavian Actuarial Journal*, 2017(5), 452–469.

[25] Zhang, Z. and Su, W., 2018. A new efficient method for estimating the Gerber-Shiu function in the classical risk model. *Sandinavian Actuarial Journal*, 2018(5), 426–449.

[26] Zhang, Z., Yang, H. and Yang, H., 2014. On a nonparametric estimator for ruin probability in the classical risk model. *Scandinavian Actuarial Journal*, 2014(4), 309–338.

[27] Zhang, Z. and Yang, H., 2013. Nonparametric estimate of the ruin probability in a pure-jump Lévy risk model. *Insurance: Mathematics and Economics*, 53, 24–35.

[28] Zhang, Z. and Yang, H., 2014. Nonparametric estimation for the ruin probability in a Lévy risk model under low-frequency observation. *Insurance: Mathematics and Economics*, 59, 168–177.

[29] Zhao, Y. and Yin C., 2010a. Expected discounted Penalty function for Lévy insurancerisk processes, *Acta Mathematicae Sinica, English Series*, 26(4), 575-586.

[30] Zhao, Y. and Yin C., 2010b. The Gerber-Shiu expected discounted penalty function for Lévy insurance risk processes. *Acta Mathematicae Applicatae Sinica, English Series*, 26(4), 575-586.

Appendix

A Some useful lemmas

Lemma 8 *(Proposition 2.2 in Comte and Genon-Catalot (2009)) Let $k \geq 1$ be an integer. If $\mu_k < \infty$, then $\mathbb{E}Z_1^k < \infty$, and for $1 \leq l \leq k$,*

$$\mathbb{E}Z_1^l = \Delta\mu_l + o(\Delta).$$

In particular, if $\mu_2 < \infty$, then

$$\mathbb{E}Z_1 = \Delta\mu_1, \quad \mathbb{E}Z_1^2 = \Delta\mu_2 + \Delta^2\mu_1^2.$$

Lemma 9 *(Shimizu and Zhang (2017)) Let $\{f_k(x)\}_{k\geq 0}$ be a sequence of differentiable functions on $(0,\infty)$ such that for some positive integers κ_0, κ_1*

$$\sup_{k\geq 0} |f_k(x)| \lesssim x^{\kappa_0}, \quad \sup_{k\geq 0} |f_k'(x)| \lesssim x^{\kappa_1}.$$

Assume that $\mu_{\kappa_0}, \mu_{\kappa_1} < \infty$, and $\lim_{\Delta\to 0} \Delta^{2-\alpha}\nu(\Delta) = 0$ for some $0 < \alpha < 1$. Then we have

$$\lim_{\Delta\to 0} \sup_{k\geq 0} \left| \frac{1}{\Delta^\alpha} \left(\frac{1}{\Delta}\mathbb{E}[f_k(X_\Delta)] - \int_0^\infty f_k(x)\nu(x)dx \right) \right| = 0.$$

Proof. This can be proved after a minor revision of Lemma B.1 in Shimizu and Zhang (2017). ■

Lemma 10 *(Stewart and Sun (1990)) Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrices with dimension m. If $\boldsymbol{A}$ is invertible and $\|\boldsymbol{A}^{-1}\boldsymbol{B}\|_{op} < 1$, then $\tilde{\boldsymbol{A}} := \boldsymbol{A} + \boldsymbol{B}$ is invertible and it holds*

$$\|\tilde{\boldsymbol{A}}^{-1} - \boldsymbol{A}^{-1}\|_{op} \leq \frac{\|\boldsymbol{B}\|_{op}\|\boldsymbol{A}^{-1}\|_{op}^2}{1 - \|\boldsymbol{A}^{-1}\boldsymbol{B}\|_{op}}.$$

B Proof of (4.17)

We cannot use Lemma 9 to prove (4.17) since we cannot find a positive integer κ such that

$$\sup_{k\geq 0} \left| \frac{d}{dy} Q_k(y,\rho) \right| \lesssim y^\kappa.$$

We shall follow the same approach adopted in Shimizu and Zhang (2017) to prove (4.17) under Conditions **B.2**(1) and **B.3**.

For every fixed $\Delta > 0$, we split the subordinator X into one part (X_t^Δ) with jumps smaller than Δ, and another part $(\tilde{X}_t^\Delta)$ which is a compound Poisson process with intensity of jumps $\lambda_\Delta = \int_{\{x\geq\Delta\}} \nu(x)dx$

and jump density function $\frac{1}{\lambda_\Delta}\nu(x)\mathbf{1}_{(x\geq\Delta)}$. For the compound Poisson process, we have $\tilde{X}_t^\Delta = \sum_{j=1}^{N_t^\Delta}\xi_j^\Delta$, where $(N_t^\Delta)_{t\geq 0}$ is a homogeneous Poisson process with intensity λ_Δ and $\{\xi_j^\Delta\}$ are i.i.d. with density function $\frac{1}{\lambda_\Delta}\nu(x)\mathbf{1}_{(x\geq\Delta)}$. Hence, we have

$$\begin{aligned}\frac{1}{\Delta^\alpha}\mathrm{II}_{k,3} &= \frac{1}{\Delta^\alpha}\left\{\frac{1}{\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta+\tilde{X}_\Delta^\Delta,\rho)] - \int_0^\infty Q_k(y,\rho)\nu(y)dy\right\}\\ &= \frac{1}{\Delta^{\alpha+1}}e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta,\rho)] - \frac{1}{\Delta^\alpha}\int_0^\Delta Q_k(y,\rho)\nu(y)dy\\ &\quad+\frac{1}{\Delta^\alpha}\left\{\lambda_\Delta e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta+\xi_1^\Delta,\rho)] - \int_\Delta^\infty Q_k(y,\rho)\nu(y)dy\right\}\\ &\quad+\frac{1}{\Delta^{\alpha+1}}\sum_{n=2}^\infty\frac{\lambda_\Delta^n\Delta^n}{n!}e^{-\lambda_\Delta\Delta}\mathbb{E}\left[Q_k\left(X_\Delta^\Delta+\sum_{j=1}^n\xi_j^\Delta,\rho\right)\right].\end{aligned}$$

The proof is completed if we can prove the following limits,

$$\lim_{\Delta\to0}\sup_{k\geq0}\left|\frac{1}{\Delta^{\alpha+1}}e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta,\rho)]\right| = 0, \tag{B.1}$$

$$\lim_{\Delta\to0}\sup_{k\geq0}\left|\frac{1}{\Delta^\alpha}\int_0^\Delta Q_k(y,\rho)\nu(y)dy\right| = 0, \tag{B.2}$$

$$\lim_{\Delta\to0}\sup_{k\geq0}\left|\frac{1}{\Delta^\alpha}\left\{\lambda_\Delta e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta+\xi_1^\Delta,\rho)] - \int_\Delta^\infty Q_k(y,\rho)\nu(y)dy\right\}\right| = 0, \tag{B.3}$$

$$\lim_{\Delta\to0}\sup_{k\geq0}\left|\frac{1}{\Delta^{\alpha+1}}\sum_{n=2}^\infty\frac{\lambda_\Delta^n\Delta^n}{n!}e^{-\lambda_\Delta\Delta}\mathbb{E}\left[Q_k\left(X_\Delta^\Delta+\sum_{j=1}^n\xi_j^\Delta,\rho\right)\right]\right| = 0. \tag{B.4}$$

First, we consider (B.1). Note that

$$\sup_{k\geq0}|Q_k(y,\rho)| \leq \frac{1}{c}\int_0^y e^{-\rho(y-x)}\cdot\sup_{k\geq0}|\varphi_k(x)|dx \leq \frac{\sqrt{2}}{c}y \tag{B.5}$$

due to the uniform upper bound (2.2). Then we have

$$\begin{aligned}\sup_{k\geq0}\left|\frac{1}{\Delta^{\alpha+1}}e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta,\rho)]\right| &\leq \frac{1}{\Delta^{\alpha+1}}\mathbb{E}[\sup_{k\geq0}|Q_k(X_\Delta^\Delta,\rho)|] \leq \frac{\sqrt{2}}{c}\frac{1}{\Delta^{\alpha+1}}\mathbb{E}[X_\Delta^\Delta]\\ &= \frac{\sqrt{2}}{c}\frac{\Delta\int_0^\Delta x\nu(x)dx}{\Delta^{\alpha+1}} = \frac{\sqrt{2}}{c}\frac{\int_0^\Delta x\nu(x)dx}{\Delta^\alpha},\end{aligned} \tag{B.6}$$

where the equality follows from Lemma 8. Hence, by L'Hôpital's rule and Conditions **B.2(1)** and **B.3** we have

$$\lim_{\Delta\to0}\sup_{k\geq0}\left|\frac{1}{\Delta^{\alpha+1}}e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta,\rho)]\right| \leq \lim_{\Delta\to0}\frac{\sqrt{2}}{c}\frac{\int_0^\Delta x\nu(x)dx}{\Delta^\alpha} = \frac{\sqrt{2}}{\alpha c}\lim_{\Delta\to0}\Delta^{2-\alpha}\nu(\Delta) = 0,$$

which proves (B.1).

Second, we consider (B.2). Using (B.5) we obtain

$$\sup_{k\geq 0}\left|\frac{1}{\Delta^\alpha}\int_0^\Delta Q_k(y,\rho)\nu(y)dy\right| \leq \frac{\sqrt{2}}{c}\frac{\int_0^\Delta y\nu(y)dy}{\Delta^\alpha} \to 0 \quad \text{as } \Delta\to 0,$$

where the last step is due to L'Hôpital's rule and Conditions **B.2(1)** and **B.3**.

Third, we consider (B.3). Note that

$$\begin{aligned}
& \sup_{k\geq 0}\left|\frac{1}{\Delta^\alpha}\left\{\lambda_\Delta e^{-\lambda_\Delta\Delta}\mathbb{E}[Q_k(X_\Delta^\Delta+\xi_1^\Delta,\rho)] - \int_\Delta^\infty Q_k(y,\rho)\nu(y)dy\right\}\right| \\
= & \sup_{k\geq 0}\left|\frac{e^{-\lambda_\Delta\Delta}-1}{\Delta^\alpha}\int_\Delta^\infty Q_k(y,\rho)\nu(y)dy + \frac{e^{-\lambda_\Delta\Delta}}{\Delta^\alpha}\int_\Delta^\infty(\mathbb{E}[Q_k(X_\Delta^\Delta+y,\rho)]-Q_k(y,\rho))\nu(y)dy\right| \\
= & \frac{1-e^{-\lambda_\Delta\Delta}}{\Delta^\alpha}\int_\Delta^\infty \sup_{k\geq 0}|Q_k(y,\rho)|\nu(y)dy + \frac{e^{-\lambda_\Delta\Delta}}{\Delta^\alpha}\int_\Delta^\infty \sup_{k\geq 0}|\mathbb{E}[Q_k(X_\Delta^\Delta+y,\rho)-Q_k(y,\rho)]|\nu(y)dy.
\end{aligned}$$

By (B.5) we have

$$\frac{1-e^{-\lambda_\Delta\Delta}}{\Delta^\alpha}\int_\Delta^\infty \sup_{k\geq 0}|Q_k(y,\rho)|\nu(y)dy \leq \frac{\sqrt{2}\mu_1}{c}\frac{1-e^{-\lambda_\Delta\Delta}}{\Delta^\alpha} \to 0 \ \text{ as } \Delta\to 0,$$

where the limit is due to L'Hôpital's rule and Conditions **B.2(1)** and **B.3**. See Appendix B of Shimizu and Zhang (2017). By the uniform upper bound (2.2) we have

$$\begin{aligned}
& \sup_{k\geq 0}|Q_k(X_\Delta^\Delta+y,\rho)]-Q_k(y,\rho)| \\
= & \sup_{k\geq 0}\left|\frac{1}{c}\int_y^{X_\Delta^\Delta+y} e^{-\rho(X_\Delta^\Delta+y-x)}\varphi_k(x)dx + \frac{1}{c}\int_0^y[e^{-\rho(X_\Delta^\Delta+y-x)}-e^{-\rho(y-x)}]\varphi_k(x)dx\right| \\
\leq & \frac{1}{c}\int_y^{X_\Delta^\Delta+y} e^{-\rho(X_\Delta^\Delta+y-x)}\cdot\sup_{k\geq 0}|\varphi_k(x)|dx + \frac{1}{c}\int_0^y |e^{-\rho(X_\Delta^\Delta+y-x)}-e^{-\rho(y-x)}|\cdot\sup_{k\geq 0}|\varphi_k(x)|dx \\
\leq & \frac{\sqrt{2}}{c}X_\Delta^\Delta + \frac{\sqrt{2}}{c}\int_0^y |e^{-\rho(X_\Delta^\Delta+y-x)}-e^{-\rho(y-x)}|dx.
\end{aligned}$$

Furthermore, using the mean value theory we have

$$|e^{-\rho(X_\Delta^\Delta+y-x)}-e^{-\rho(y-x)}| = |-X_\Delta^\Delta\rho e^{-\rho x^*}| \leq \rho X_\Delta^\Delta,$$

where $x^*>0$ is a random number between $y-x$ and $X_\Delta^\Delta+y-x$. Then we have

$$\sup_{k\geq 0}|Q_k(X_\Delta^\Delta+y,\rho)]-Q_k(y,\rho)| \leq \frac{\sqrt{2}}{c}X_\Delta^\Delta + \frac{\sqrt{2}}{c}\rho X_\Delta^\Delta y.$$

Using this upper bound we have

$$\begin{aligned}
& \frac{e^{-\lambda_\Delta\Delta}}{\Delta^\alpha}\int_\Delta^\infty \sup_{k\geq 0}|\mathbb{E}[Q_k(X_\Delta^\Delta+y,\rho)-Q_k(y,\rho)]|\nu(y)dy \\
\leq & \frac{\sqrt{2}}{c}\frac{\mathbb{E}X_\Delta^\Delta}{\Delta^\alpha}\int_\Delta^\infty \nu(y)dy + \frac{\sqrt{2}}{c}\rho\frac{\mathbb{E}X_\Delta^\Delta}{\Delta^\alpha}\int_\Delta^\infty y\nu(y)dy
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sqrt{2}}{c}\frac{\Delta\int_0^\Delta y\nu(y)dy}{\Delta^\alpha}\int_\Delta^\infty \nu(y)dy + \frac{\sqrt{2}}{c}\rho\frac{\Delta\int_0^\Delta y\nu(y)dy}{\Delta^\alpha}\int_\Delta^\infty y\nu(y)dy \\
&\leq \frac{\sqrt{2}}{c}\mu_1\frac{\int_0^\Delta y\nu(y)dy}{\Delta^\alpha} + \frac{\sqrt{2}}{c}\rho\mu_1\Delta^{1-\alpha}\int_0^\Delta y\nu(y)dy \\
&\to 0 \quad \text{as } \Delta \to 0,
\end{aligned}
$$

thanks to L'Hôpital's rule and Conditions **B.2(1)** and **B.3**.

Finally, we consider (B.4). By (B.5) we have

$$
\begin{aligned}
&\sup_{k\geq 0}\left|\frac{1}{\Delta^{\alpha+1}}\sum_{n=2}^\infty \frac{\lambda_\Delta^n\Delta^n}{n!}e^{-\lambda_\Delta\Delta}\mathbb{E}\left[Q_k\left(X_\Delta^\Delta + \sum_{j=1}^n \xi_j^\Delta, \rho\right)\right]\right| \\
&\leq \frac{1}{\Delta^{\alpha+1}}\sum_{n=2}^\infty \frac{\lambda_\Delta^n\Delta^n}{n!}e^{-\lambda_\Delta\Delta}\mathbb{E}\left[\sup_{k\geq 0}\left|Q_k\left(X_\Delta^\Delta + \sum_{j=1}^n \xi_j^\Delta, \rho\right)\right|\right] \\
&\leq \frac{\sqrt{2}}{c}\frac{1}{\Delta^{\alpha+1}}\sum_{n=2}^\infty \frac{\lambda_\Delta^n\Delta^n}{n!}e^{-\lambda_\Delta\Delta}\mathbb{E}\left(X_\Delta^\Delta + \sum_{j=1}^n \xi_j^\Delta\right).
\end{aligned}
$$

Then using the same arguments as in the proof of Lemma B.1 in Shimizu and Zhang (2017) can prove (B.4). This completes the proof.