

PREDICTION IN A POISSON CLUSTER MODEL

MUNEYA MATSUI,* *Keio University*

THOMAS MIKOSCH,** *University of Copenhagen*

Abstract

We consider a Poisson cluster model, motivated by insurance applications. At each claim arrival time, modeled by the point of a homogeneous Poisson process, we start a cluster process which represents the number or amount of payments triggered by the arrival of a claim in a portfolio. The cluster process is a Lévy or truncated compound Poisson process. Given the observations of the process over a finite interval, we consider the expected value of the number and amount of payments in a future time interval. We also give bounds for the error encountered in this prediction procedure.

Keywords: Poisson cluster model; prediction; conditional expectation; claims reserving; insurance; Lévy process; shot noise

2010 Mathematics Subject Classification: Primary 60K30

Secondary 60G25; 60G55

1. Introduction

In this paper we consider the model

$$M(t) = \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq t\}} L_k(t - T_k), \quad t \geq 1, \quad (1.1)$$

where $0 < T_1 < T_2 < \dots$ are the points of a homogeneous Poisson process with intensity $\lambda > 0$ and (L_k) is a sequence of independent and identically distributed (i.i.d.) stochastic processes independent of (T_k) and such that $L_k(t) = 0$, $t \leq 0$, almost surely (a.s.). Writing N for the counting process generated by the points T_k , $k = 1, 2, \dots$, M takes the form

$$M(t) = \sum_{k=1}^{N(t)} L_k(t - T_k), \quad t \geq 1. \quad (1.2)$$

Processes of this type are related to Poisson shot noise processes, which have many applications in rather different areas. For example, in an insurance context, T_k may describe the arrival of a claim in an insurance portfolio and $(L_k(t - T_k))_{t \geq T_k}$ is the corresponding payment process from the insurer to the insured starting at time T_k . Alternatively, $(L_k(t - T_k))_{t \geq T_k}$ can be the counting process for these payments. The main focus of this paper is on applications in an insurance context.

Received 26 October 2009; revision received 27 January 2010.

* Current address: Mathematical Engineering No. 4 Laboratory, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. Email address: mmuneya@gmail.com
Research supported by a JSPS Research Fellowship for Young Scientists.

** Postal address: Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. Email address: mikosch@math.ku.dk

Research partly supported by the Danish Research Council (FNU) grants 272-06-0442 and 09-072331.

Early on, shot noise processes were used for modeling purposes in many fields of applied probability such as bunching in traffic (see [1]), computer failure times (see [17]), and earthquake aftershocks (see [28]). More recently, shot noise processes have been used for modeling large computer networks such as the Internet; see, for example, [12] and [13] for some early work. In the context of the workload of large computer networks, shot noise processes arise as aggregated versions of the ON/OFF or infinite-source Poisson models, also known as $M/G/\infty$ models; see, for example, [16], [23], and [24], and, for further extensions, [2] and [22]. Other applications include finance (see [8] and [25]) and physics (see [3]). Most of the papers mentioned above aim at an asymptotic theory for the shot noise process, deriving Gaussian or Lévy process limits, or at the asymptotics of the extremes of such processes; see also [4], [5], [9], [10], [11], [14], [15], [20], and [27].

The focus of this paper is not on the asymptotic properties of the process M defined in (1.1), but on precise results about the prediction of the increments

$$M(t, t + s) = M(t + s) - M(t), \quad t \geq 1, s > 0,$$

i.e. we will calculate $E[M(t, t + s) | \mathcal{F}_t]$ for some suitable σ -fields \mathcal{F}_t , where we do not necessarily assume that $M(t, t + s)$ has finite variance. Results of this kind are surprisingly explicit due to the Poisson structure underlying the process M .

The particular structure of the process M and the prediction problem are motivated by reserving problems in insurance. Here the points $T_k \leq 1$ describe the arrival of claims in a portfolio in a given period, one year say, and $M(t)$ is the number or amount of payments for the claims arriving in $[0, 1]$ and being paid off in $[0, t]$, $t \geq 1$. Correspondingly, $M(t, t + s)$ is the number or amount of payments executed in the interval $(t, t + s]$, $s > 0$.

Problems of this kind are related to *claims reserving*. Traditionally, claims reserving uses a statistical model in which conditions of the type $E[M(t + s) | M(t)] = f_{t,s}M(t)$ for suitable constants $f_{t,s}$ are assumed to hold. These constants are then estimated based on several years of data available (the so-called *chain ladder*). We refer the reader to Chapter 11 of [21] for an introduction to the topic. Jessen *et al.* [6] considered a stochastic process based model with Poisson components which allows one to derive explicit expressions for the prediction of payment numbers and amounts, and the corresponding prediction errors. This model requires that time series of annual observations on payment numbers and amounts are available; the structure of this model is significantly simpler than (1.1).

In the present paper we follow a different path that has already been mentioned in Section 11.3.3 of [21]. There model (1.1) was considered for i.i.d. homogeneous Poisson processes L_k , $k = 1, 2, \dots$, and explicit expressions for the prediction $E[M(t, t + s) | M(t)]$ were obtained. First, in Section 2, we show that this approach can be extended to general Lévy processes L_k . The homogeneous Poisson case is again a benchmark (Section 2.3). In this case we can also give a recursive algorithm for determining the prediction. In Section 3 we consider modifications to model (1.2). Instead of a Lévy process L_k we allow for Lévy processes which are truncated at a random level or which take into account a delay in the reporting of a claim. In Section 3 we also consider different σ -fields \mathcal{F}_t , and calculate the corresponding predictions $E[M(t, t + s) | \mathcal{F}_t]$ and their conditional or unconditional prediction errors. Depending on \mathcal{F}_t , the prediction of $M(t, t + s)$ and the prediction error often have a structure which is simpler than those for $E[M(t, t + s) | M(t)]$.

2. Prediction in a Poisson cluster process with Lévy clusters

In this section we consider model (1.1) and we assume that an activity process L_k starts at the Poisson point $T_k \in [0, 1]$. In this context, (L_k) is a sequence of i.i.d. Lévy processes with generic element L ; we refer the reader to [26] for the definition and properties of Lévy processes. Then $M(t, t + s)$, $t \geq 1, s > 0$, can be interpreted as the measurement of activities initiated in $[0, 1]$ which are still alive in $(t, t + s]$. In the insurance context mentioned in the introduction, $M(t, t + s)$ is interpreted as the *number or amount of payments* for claims that occurred in one year say. Other interpretations are possible as well. For example, $M(t, t + s]$ may describe the workload to be managed by a large computer network for sources that started an activity (such as sending packets to other sources) in the interval $[0, 1]$. The Lévy process condition on L_k is assumed for convenience; in this case we obtain explicit expressions for the prediction of $M(t, t + s)$ given $M(t)$, $t \geq 1$ (see Section 2.2).

We start with an analysis of the moment structure of M .

2.1. The first and second moments of M

The following result is elementary.

Lemma 2.1. *Consider model (1.2) with i.i.d. Lévy processes L_k , $k = 1, 2, \dots$, and a homogeneous Poisson process N with intensity $\lambda > 0$.*

- (i) *Assume that $\mu = E[L(1)]$ exists and is finite. Then $E[M(t)] = \lambda\mu(t - 0.5)$, $t \geq 1$.*
- (ii) *Assume that $\sigma^2 = \text{var}(L(1))$ is finite. Then, for $1 \leq t_1 \leq t_2$,*

$$\text{cov}(M(t_1), M(t_2)) = \lambda\sigma^2(t_1 - 0.5) + \lambda\mu^2 \int_0^1 (t_1 - s)(t_2 - s) ds.$$

In particular, $\text{var}(M(t)) = \lambda\sigma^2(t - 0.5) + \lambda\mu^2(t^2 - t + \frac{1}{3})$, $t \geq 1$.

We see that process M is overdispersed in the sense that $\lim_{t \rightarrow \infty} \text{var}(M(t))/E[M(t)] = \infty$.

2.2. Prediction of future increments

In this section we derive explicit expressions for the quantities $E[M(t, t + s) | M(t)]$ for $t \geq 1$ and $s > 0$. We assume that the general conditions of Lemma 2.1 hold and also that $\mu = E[L(1)]$ is finite.

We start with a simple observation. By the definition of $M(t)$, $t \geq 1$, in (1.2), the σ -field generated by $M(t)$ is contained in the σ -field generated by $(L_i(t - T_i))$ and (T_i) . Therefore, writing $L_i(t, t + s) = L_i(t + s) - L_i(t)$ for the increments of L_i , for $t \geq 1$ and $s > 0$,

$$\begin{aligned} E[M(t, t + s) | M(t)] &= \sum_{k=1}^{\infty} E[E[\mathbf{1}_{\{T_k \leq 1\}} L_k(t - T_k, t + s - T_k) | (L_i(t - T_i), (T_i)) | M(t)]] \\ &= \sum_{k=1}^{\infty} E[\mathbf{1}_{\{T_k \leq 1\}} E[L_k(t - T_k, t + s - T_k) | L_k(t - T_k), T_k] | M(t)] \\ &= \mu s \sum_{k=1}^{\infty} E[\mathbf{1}_{\{T_k \leq 1\}} | M(t)] \\ &= \mu s E[N(1) | M(t)]. \end{aligned} \tag{2.1}$$

In the first step we used dominated convergence and in the last step we used the stationary independent increments of the Lévy process L_k . Thus, we are left to calculate $E[N(1) \mid M(t)]$ for $t \geq 1$. We mention at this point that in an insurance context the number $N(1)$ of claims arriving in the interval $[0, 1]$ is in general not observable at time $t \geq 1$. It usually takes a much longer time period than just one year before the claim number for the first year $[0, 1]$ is known (the so-called incurred but not reported effect). Therefore, the quantity $E[N(1) \mid M(t)]$, $t \geq 1$, has the intuitive meaning that one gathers information about the payments in the period $[0, t]$, represented by the quantity $M(t)$, in order to gather information about the (generally unknown) claim number $N(1)$.

It is our aim to express the prediction of $M(t, t + s)$ given $M(t)$ as explicitly as possible. Motivated by the previous calculations, we will determine

$$\hat{M}_A(t, t + s) = E[M(t, t + s) \mid M(t) \in A], \quad t \geq 1, s > 0,$$

for any Borel set A . Later, in Section 2.3, we will specify the set A . The quantity $\hat{M}_A(t, t + s)$ is the conditional first moment of $M(t, t + s)$. In order to get an idea of the conditional prediction error, we are also interested in the conditional second moment of $M(t, t + s)$. Both moments can easily be derived from the characteristic function of $M(t, t + s)$ given $M(t)$.

Lemma 2.2. *Consider model (1.2) with i.i.d. Lévy processes L_k , $k = 1, 2, \dots$, and a homogeneous Poisson process N with intensity $\lambda > 0$. For any Borel set A , the characteristic function of $M(t, t + s)$ given $\{M(t) \in A\}$ has the form*

$$\begin{aligned} \hat{f}_A(x) &= E[e^{ixM(t,t+s)} \mid M(t) \in A] \\ &= \frac{E[(E[e^{ixL(s)}])^{N(1)} P(L(R_{N(1)}(t)) \in A \mid N(1))]}{P(L(R_{N(1)}(t)) \in A)}, \quad x \in \mathbb{R}, t \geq 1, s > 0. \end{aligned}$$

Here we assume that the denominator does not vanish and that

$$R_r(t) = \sum_{i=1}^r (t - U_i), \quad t \geq 1, r = 0, 1, \dots, \tag{2.2}$$

for an i.i.d. $U(0, 1)$ sequence (U_i) such that L , N , and (U_i) are independent.

Remark 2.1. The condition $P(L(R_{N(1)}(t)) \in A) > 0$ is needed for a proper definition of the conditional characteristic function $\hat{f}_A(x)$. From the proof below, it follows that this condition is equivalent to $P(M(t) \in A) > 0$.

Proof of Lemma 2.2. The same argument leading to (2.1) yields, for $x \in \mathbb{R}$, $t \geq 1$, and $s > 0$,

$$\begin{aligned} &E[e^{ixM(t,t+s)} \mid M(t)] \\ &= E[E[e^{ixM(t,t+s)} \mid (L_k(t - T_k)), (T_k)] \mid M(t)] \\ &= E\left[E\left[\prod_{k=1}^{\infty} \exp\{ix \mathbf{1}_{\{T_k \leq 1\}} L_k(t - T_k, t + s - T_k)\} \mid (L_k(t - T_k)), (T_k)\right] \mid M(t)\right] \\ &= E\left[E\left[\prod_{k=1}^{\infty} (\mathbf{1}_{\{T_k > 1\}} + \mathbf{1}_{\{T_k \leq 1\}} E[e^{ixL(s)}]) \mid (T_k)\right] \mid M(t)\right] \\ &= E[(E[e^{ixL(s)}])^{N(1)} \mid M(t)]. \end{aligned} \tag{2.3}$$

Therefore, we will calculate the following quantities for any Borel set A (assuming that the denominator does not vanish):

$$\begin{aligned} \mathbb{P}(N(1) = r \mid M(t) \in A) &= \frac{\mathbb{P}(N(1) = r, \sum_{k=1}^{N(1)} L_k(t - T_k) \in A)}{\mathbb{P}(\sum_{k=1}^{N(1)} L_k(t - T_k) \in A)} \\ &= \frac{\mathbb{P}(N(1) = r, L(\sum_{k=1}^{N(1)}(t - T_k)) \in A)}{\mathbb{P}(L(\sum_{k=1}^{N(1)}(t - T_k)) \in A)}, \quad r = 0, 1, \dots \end{aligned}$$

In the last step we used the independent stationary increments of the i.i.d. Lévy processes $L_i, i = 1, 2, \dots$, conditional on (T_k) . Conditioning on $N(1)$ and using the order statistics property of a homogeneous Poisson process, we obtain

$$\mathbb{P}\left(\sum_{k=1}^{N(1)}(t - T_k) \in \cdot \mid N(1) = r\right) = \mathbb{P}(R_r(t) \in \cdot), \quad r = 0, 1, \dots,$$

where R_r is defined in (2.2). We conclude that

$$\mathbb{P}(N(1) = r \mid M(t) \in A) = \frac{\mathbb{P}(N(1) = r) \mathbb{P}(L(R_r(t)) \in A)}{\mathbb{P}(L(R_{N(1)}(t)) \in A)}. \tag{2.4}$$

Now substitute (2.4) into (2.3) to obtain the desired expression for $\hat{f}_A(x)$.

Since we know the characteristic function $\hat{f}_A(x), x \in \mathbb{R}$, of $M(t, t + s]$ given $\{M(t) \in A\}$, we can derive the moments of the prediction $\hat{M}_A(t, t + s]$ by differentiating $\hat{f}_A(x)$ at $x = 0$ sufficiently often. The following result summarizes the analysis of the first and second conditional moments.

Theorem 2.1. Consider model (1.2) with i.i.d. Lévy processes $L_k, k = 1, 2, \dots$, and a homogeneous Poisson process N with intensity $\lambda > 0$.

- (i) Assume that $\mu = E[L(1)]$ exists and is finite. Then the prediction $\hat{M}_A(t, t + s]$ of $M(t, t + s]$ given $\{M(t) \in A\}$ has the following form for any Borel set A :

$$\hat{M}_A(t, t + s] = \mu s \frac{E[N(1) \mathbb{P}(L(R_{N(1)}(t)) \in A \mid N(1))]}{\mathbb{P}(L(R_{N(1)}(t)) \in A)}, \quad t \geq 1, s > 0. \tag{2.5}$$

- (ii) Assume that $\sigma^2 = \text{var}(L(1))$ is finite. Then the conditional variance of $M(t, t + s]$ given $\{M(t) \in A\}$ has the following form for any Borel set A :

$$\begin{aligned} \text{var}(M(t, t + s] \mid M(t) \in A) &= \frac{\sigma^2}{\mu} \hat{M}_A(t, t + s] + (\mu s)^2 \frac{E[(N(1))^2 \mathbb{P}(L(R_{N(1)}(t)) \in A \mid N(1))]}{\mathbb{P}(L(R_{N(1)}(t)) \in A)} \\ &\quad - (\hat{M}_A(t, t + s])^2, \quad t \geq 1, s > 0. \end{aligned} \tag{2.6}$$

Here we have assumed that the probability $\mathbb{P}(L(R_{N(1)}(t)) \in A)$ does not vanish.

Remark 2.2. The variance of $M(t, t + s]$ conditional on $M(t)$ gives us a certain measure for the uncertainty of the prediction $\hat{M}_A(t, t + s]$. In general, this conditional variance is difficult to obtain. However, if L is a homogeneous Poisson process, we can derive recursive algorithms

for determining this quantity; see Section 2.3. It would be desirable to obtain an expression for the unconditional mean square error

$$\begin{aligned} & E[(M(t, t + s) - E[M(t, t + s) | M(t)])^2] \\ &= E[\text{var}(M(t, t + s) | M(t))] \\ &= E[(E[M(t, t + s)] - E[E[M(t, t + s) | M(t)])^2]. \end{aligned}$$

The second moment $E[(M(t, t + s)]^2$ provides an upper bound for the mean square error. It can be derived from Lemma 2.1. The quantity $E[(E[M(t, t + s) | M(t)])^2]$ seems intractable.

2.3. Poisson clusters

In this section we assume that L is a homogeneous Poisson process with intensity $\gamma > 0$. In this case we can give explicit expressions for the prediction

$$\hat{M}_m(t, t + s) = E[M(t, t + s) | M(t) = m], \quad t \geq 1, s > 0, m = 0, 1, \dots,$$

and the conditional prediction error. In what follows, it will be convenient to use the Laplace–Stieltjes transform of any nonnegative random variable Y :

$$\phi_Y(z) = E[e^{-zY}], \quad z \geq 0,$$

and its m th derivatives $\phi_Y^{(m)}(z) = (-1)^m E[Y^m e^{-zY}]$, $m = 0, 1, 2, \dots$

Theorem 2.2. Assume that L is a homogeneous Poisson process with intensity $\gamma > 0$. Then the prediction of $M(t, t + s]$ given $\{M(t) = m\}$ has the form

$$\hat{M}_m(t, t + s) = \lambda \gamma s \frac{\phi_{R_{N(1)+1}(t)}^{(m)}(\gamma)}{\phi_{R_{N(1)}(t)}^{(m)}(\gamma)}, \quad t \geq 1, s > 0, m = 0, 1, \dots,$$

and the conditional mean square error is given by

$$\begin{aligned} \text{var}(M(t, t + s) | M(t) = m) &= (\lambda \gamma s)^2 \frac{\phi_{R_{N(1)+2}(t)}^{(m)}(\gamma)}{\phi_{R_{N(1)}(t)}^{(m)}(\gamma)} + (1 + \gamma s) \hat{M}_m(t, t + s) \\ &\quad - (\hat{M}_m(t, t + s))^2, \quad t \geq 1, s > 0, m = 0, 1, \dots, \end{aligned}$$

where R_r is defined in (2.2).

Proof. According to (2.5) and (2.6), we need to evaluate

$$E[(N(1))^i P(L(R_{N(1)}(t)) = m | N(1))], \quad i = 0, 1, 2.$$

We have

$$\begin{aligned} E[N(1) P(L(R_{N(1)}(t)) = m | N(1))] &= \sum_{r=1}^{\infty} r P(N(1) = r) P(L(R_r(t)) = m) \\ &= \lambda \sum_{r=0}^{\infty} e^{-\lambda} \frac{\lambda^r}{r!} \frac{E[(\gamma R_{r+1}(t))^m e^{-\gamma R_{r+1}(t)}]}{m!} \\ &= \lambda E \left[\frac{(\gamma R_{N(1)+1}(t))^m e^{-\gamma R_{N(1)+1}(t)}}{m!} \right] \\ &= \frac{\lambda(-\gamma)^m}{m!} \phi_{R_{N(1)+1}(t)}^{(m)}. \end{aligned}$$

In a similar way, we calculate

$$\begin{aligned}
 P(L(R_{N(1)}(t)) = m) &= \frac{(-\gamma)^m}{m!} \phi_{R_{N(1)}(t)}^{(m)}(\gamma), \\
 E[(N(1))^2 P(L(R_{N(1)}(t)) = m \mid N(1))] &= \frac{\lambda(-\gamma)^m}{m!} \phi_{R_{N(1)+1}(t)}^{(m)}(\gamma) \\
 &\quad + \frac{\lambda^2(-\gamma)^m}{m!} \phi_{R_{N(1)+2}(t)}^{(m)}(\gamma).
 \end{aligned}$$

This concludes the proof.

In view of Theorem 2.2 it is crucial to be able to evaluate the derivatives of $\phi_{R_{N(1)}(t)}$. Although we have the representations

$$\phi_{R_{N(1)}}(\gamma) = e^{-\lambda(1-\phi_{t-U}(\gamma))} \quad \text{and} \quad \phi_{t-U}(\gamma) = \gamma^{-1} e^{-\gamma t} (e^\gamma - 1),$$

their derivatives are complicated and not necessarily useful. Fortunately, the following proposition yields a recursive scheme for determining these derivatives.

Proposition 2.1. *Let $\ell = 1, 2, \dots$, and let $\phi_{t-U}^{(\ell)}(\gamma)$ and $\phi_{R_{N(1)+r}(t)}^{(\ell)}(\gamma)$, $r = 0, 1, 2$, be the ℓ th derivatives of ϕ_{t-U} and $\phi_{R_{N(1)+r}(t)}$, $r = 0, 1, 2$. Then the following recursive relations are valid:*

$$\begin{aligned}
 \phi_{t-U}^{(\ell)}(\gamma) &= (-1)^\ell \gamma^{-\ell-1} [\Gamma(\ell + 1, \gamma(t - 1)) - \Gamma(\ell + 1, \gamma t)] \\
 \phi_{R_{N(1)}(t)}^{(\ell)}(\gamma) &= \lambda \sum_{k=1}^{\ell} \binom{\ell - 1}{k - 1} \phi_{t-U}^{(k)}(\gamma) \phi_{R_{N(1)}(t)}^{(\ell-k)}(\gamma), \\
 \phi_{R_{N(1)+1}(t)}^{(\ell)}(\gamma) &= \lambda \sum_{k=0}^{\ell} \binom{\ell}{k} \phi_{t-U}^{(k)}(\gamma) \phi_{R_{N(1)}(t)}^{(\ell-k)}(\gamma), \\
 \phi_{R_{N(1)+2}(t)}^{(\ell)}(\gamma) &= \lambda \sum_{k=0}^{\ell} \binom{\ell}{k} \phi_{t-U}^{(k)}(\gamma) \phi_{R_{N(1)+1}(t)}^{(\ell-k)}(\gamma),
 \end{aligned}$$

where $\Gamma(\alpha, x) = \int_x^\infty e^{-y} y^{\alpha-1} dy$, $x > 0$, is the incomplete gamma function.

Proof. Observe that, for $t \geq 1$,

$$\phi_{t-U}(\gamma) = \int_t^\infty e^{-\gamma u} (e^\gamma - 1) du. \tag{2.7}$$

Then Leibniz’s rule yields, for $u > 0$ and $\ell = 0, 1, \dots$,

$$\begin{aligned}
 [e^{-\gamma u} (e^\gamma - 1)]^{(\ell)} &= \sum_{k=0}^{\ell} \binom{\ell}{k} (e^{-\gamma u})^{(k)} (e^\gamma - 1)^{(\ell-k)} \\
 &= e^{-\gamma(u-1)} (1 - u)^\ell - (-u)^\ell e^{-\gamma u}.
 \end{aligned}$$

Now, interchanging the integral and the derivative in (2.7), we have, for $t \geq 1$,

$$\begin{aligned}
 \phi_{t-U}^{(\ell)}(\gamma) &= \int_t^\infty [e^{-\gamma(u-1)} (1 - u)^\ell - (-u)^\ell e^{-\gamma u}] du \\
 &= (-1)^\ell \int_{t-1}^t e^{-\gamma u} u^\ell du \\
 &= (-1)^\ell \gamma^{-\ell-1} [\Gamma(\ell + 1, \gamma(t - 1)) - \Gamma(\ell + 1, \gamma t)].
 \end{aligned}$$

This is the desired formula for the derivatives of ϕ_{t-U} . By the definition of $\phi_{R_{N(1)}(t)}$, observe that

$$\phi'_{R_{N(1)}(t)} = \lambda \phi'_{t-U} \phi_{R_{N(1)}(t)},$$

where the prime notation denotes the first derivative. Another application of Leibniz's rule yields the desired formula for the derivatives of $\phi_{R_{N(1)}(t)}$. We also observe that

$$\phi_{R_{N(1)}(t)+1} = \phi_{R_{N(1)}(t)} \phi_{t-U} \quad \text{and} \quad \phi_{R_{N(1)}(t)+2} = \phi_{R_{N(1)}(t)+1} \phi_{t-U}.$$

Applications of the Leibniz rule yield the desired expressions for the derivatives.

Remark 2.3. Assume that the homogeneous Poisson process L_k has the arrivals $0 < \Gamma_{k1} < \Gamma_{k2} < \dots$, $k = 1, 2, \dots$. Let $(X_{ki})_{k,i=1,2,\dots}$ be an i.i.d. sequence of random variables such that $\nu = E X_{11}$ exists and is finite. Moreover, we assume that (T_k) , (X_{ki}) , and $(\Gamma_{ki})_{k,i=1,2,\dots}$ are mutually independent. Then it is possible to define the reward process

$$S(t) = \sum_{k=1}^{N(1)} \sum_{i=1}^{\infty} X_{ki} \mathbf{1}_{\{T_k + \Gamma_{ki} \leq t\}}, \quad t \geq 1.$$

The reward $S(t)$ can be interpreted as the total amount of payments executed at the times $T_k + \Gamma_{ki} \leq t$, $t \geq 1$, $i = 1, 2, \dots$, for claims arriving in an insurance portfolio at times $T_k \leq 1$. A conditioning argument shows that

$$E[S(t, t + s) \mid M(t)] = \nu E[M(t, t + s) \mid M(t)], \quad t \geq 1, s > 0.$$

The i.i.d.-ness condition of the X_{ki} s can be further relaxed. For example, if we look at the conditional expectations $E[S(\ell, \ell + 1) \mid M(\ell)]$, $\ell = 1, 2, \dots$, then we may allow the distribution of the i.i.d. payments X_{ki} executed in the interval $(\ell, \ell + 1]$ to depend on ℓ . The conditional variance $\text{var}(S(t, t + s) \mid M(t))$ can be calculated as well. Some of these calculations are provided in Section 11.3.3 of [21]. We omit details.

3. Prediction with different information sets and non-Lévy clusters

3.1. Prediction with compound Poisson clusters

In this section we consider model (1.2) with a sequence (L_k) of i.i.d. compound Poisson processes. We assume the representation

$$L_k(t) = \sum_{i=1}^{\infty} X_{ki} \mathbf{1}_{\{\Gamma_{ki} \leq t\}}, \quad t \geq 1, \tag{3.1}$$

where (X_{ki}) is a double array of i.i.d. random variables, independent of the double array (Γ_{ki}) and the Poisson points (T_k) . The points $0 < \Gamma_{k1} < \Gamma_{k2} < \dots$ constitute the homogeneous Poisson process N_k with intensity $\gamma > 0$ underlying the compound Poisson process L_k . Throughout, we assume that $\mu = E[L(1)] = E X_{11} \gamma$ exists and is finite.

In Section 2.3 we mentioned that in an insurance context it is natural to assume that the number of claims $N(1)$ is unobservable at time $t \geq 1$. It is common for claims to be reported long after they were incurred. In this section we replace the condition $M(t)$ in the prediction of $M(t, t + s)$ for $t \geq 1$ and $s > 0$ by the number of claims that were incurred in $[0, 1]$ and reported

by time $t \geq 1$. We say that the k th claim is reported if $T_k + \Gamma_{k1} \leq t$, i.e. if the first payment has been executed by time t . For the corresponding counting process of reported claims, we write

$$\bar{N}_0(t) = \sum_{k=1}^{N(1)} \mathbf{1}_{\{T_k + \Gamma_{k1} \leq t\}} = \sum_{k=1}^{N(1)} \mathbf{1}_{\{N_k(t - T_k) \geq 1\}}, \quad t \geq 1. \tag{3.2}$$

We focus on the calculations of the conditional expectation

$$\hat{M}_\ell(t, t + s) = E[M(t, t + s) \mid \bar{N}_0(t) = \ell], \quad \ell = 0, 1, \dots,$$

and the corresponding prediction error.

In view of (3.2) the σ -field generated by $\bar{N}_0(t)$ is contained in the σ -field \mathcal{F}_t generated by $(\Gamma_{ki})_{k,i \geq 1, T_k + \Gamma_{ki} \leq t}$ and (T_k) . Therefore, for $t \geq 1$ and $s > 0$,

$$\begin{aligned} \hat{M}_\ell(t, t + s) &= E \left[\sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq 1\}} E[L_k(t - T_k, t + s - T_k) \mid \mathcal{F}_t] \mid \bar{N}_0(t) = \ell \right] \\ &= \mu s E[N(1) \mid \bar{N}_0(t) = \ell]. \end{aligned}$$

The latter conditional expectation will be evaluated below. Since we are also interested in other conditional moments of $M(t, t + s)$, we calculate the conditional characteristic function $E[e^{ixM(t,t+s)} \mid \bar{N}_0(t) = \ell]$, $x \in \mathbb{R}$. We use some properties of the Poisson random measure and refer the reader to [7].

Lemma 3.1. *The conditional characteristic function of $M(t, t + s)$ given $\{\bar{N}_0(t) = \ell\}$ has the form*

$$E[e^{ixM(t,t+s)} \mid \bar{N}_0(t) = \ell] = (E[e^{ixL(s)}])^\ell \exp\{-\lambda_0(t)(1 - E[e^{ixL(s)}])\}, \quad x \in \mathbb{R}, \tag{3.3}$$

where

$$\lambda_0(t) = \lambda \frac{e^{-\gamma t}}{\gamma} (e^\gamma - 1). \tag{3.4}$$

Proof. We proceed similarly to the proof of Lemma 2.2:

$$\begin{aligned} E[e^{ixM(t,t+s)} \mid \bar{N}_0(t)] &= E[E[e^{ixM(t,t+s)} \mid \mathcal{F}_t] \mid \bar{N}_0(t)] \\ &= E[(E[e^{ixL(s)}])^{N(1)} \mid \bar{N}_0(t)] \\ &= (E[e^{ixL(s)}])^{\bar{N}_0(t)} E[(E[e^{ixL(s)}])^{N(1) - \bar{N}_0(t)} \mid \bar{N}_0(t)]. \end{aligned} \tag{3.5}$$

Let Q be a Poisson random measure on the state space $E = [0, 1] \times [0, \infty)$ with mean measure $\nu = \lambda \text{Leb} \times F$, where F denotes the distribution of Γ_{k1} . Then $N(1)$ and $\bar{N}_0(t)$ have the Poisson integral representations

$$\begin{aligned} N(1) &= \int_E Q(ds, dy) \\ &= \int_{s=0}^1 \int_{y=0}^{t-s} Q(ds, dy) + \int_{s=0}^1 \int_{y=t-s}^{\infty} Q(ds, dy) \\ &= \bar{N}_0(t) + [N(1) - \bar{N}_0(t)]. \end{aligned}$$

Due to the splitting property of the Poisson process and since the integrals above are defined on disjoint subsets of the state space, the random variables $\bar{N}_0(t)$ and $N(1) - \bar{N}_0(t)$ are independent and Poisson distributed with parameters

$$E[N(1) - \bar{N}_0(t)] = \lambda \int_0^1 \int_{t-s}^\infty F(dy) ds = \lambda \frac{e^{-\gamma t}}{\gamma} (e^\gamma - 1) = \lambda_0(t)$$

and $E[\bar{N}_0(t)] = \lambda - \lambda_0(t)$.

Therefore, we conclude from (3.5) that

$$E[e^{ixM(t,t+s)} | \bar{N}_0(t)] = (E[e^{ixL(s)}])^{\bar{N}_0(t)} E[(E[e^{ixL(s)}])^{N(1)-\bar{N}_0(t)}].$$

The latter relation yields the desired conditional characteristic function.

The following result can now be obtained by differentiating the characteristic function (3.3) sufficiently often and then considering the derivatives at 0.

Theorem 3.1. *Assume that L is a compound Poisson process with underlying Poisson intensity γ .*

- (i) *If $\mu = E X_{11}\gamma$ exists and is finite, then the prediction of $M(t, t + s]$ given $\{\bar{N}_0(t) = \ell\}$ has the form*

$$\hat{M}_\ell(t, t + s) = \mu s[\lambda_0(t) + \ell], \quad t \geq 1, s > 0, \ell = 0, 1, \dots, \tag{3.6}$$

where $\lambda_0(t)$ is given in (3.4).

- (ii) *If, in addition, $\sigma^2 = \text{var}(L(1)) = E[X_{11}^2]\gamma < \infty$ then*

$$\begin{aligned} \text{var}(M(t, t + s) | \bar{N}_0(t) = \ell) \\ = \sigma^2 s(\ell + \lambda_0(t)) + (\mu s)^2 \lambda_0(t), \quad t \geq 1, s > 0, \ell = 0, 1, \dots \end{aligned} \tag{3.7}$$

Remark 3.1. Note that the unconditional prediction error of $M(t, t + s]$ given $\bar{N}_0(t)$ has the form

$$\begin{aligned} E[\text{var}(M(t, t + s) | \bar{N}_0(t))] &= (\sigma^2 s + (\mu s)^2) \lambda_0(t) + \sigma^2 s E[\bar{N}_0(t)] \\ &= \sigma^2 s \lambda + (\mu s)^2 \lambda_0(t). \end{aligned}$$

Remark 3.2. In a next step we include more information on the condition of the expected value of $M(t, t + s]$, i.e. we focus on $E[M(t, t + s) | M(t), \bar{N}_0(t)]$. Both processes $M(t)$ and $\bar{N}_0(t)$ are assumed to be observable at time t . Then the σ -field \mathcal{G}_t generated by (T_k) and $((\Gamma_{ki}, X_{ki}))_{k,i \geq 1, T_k + \Gamma_{ki} \leq t}$ is larger than the σ -field generated by $M(t)$ and $\bar{N}_0(t)$. Therefore,

$$\begin{aligned} E[M(t, t + s) | M(t), \bar{N}_0(t)] &= \mu s E[E[N(1) | \mathcal{G}_t] | M(t), \bar{N}_0(t)] \\ &= \mu s E[N(1) | M(t), \bar{N}_0(t)]. \end{aligned} \tag{3.8}$$

Consider the Poisson random measure \tilde{Q} with points $(T_k, (\Gamma_{ki})_{i=1,2,\dots}, (X_{ki})_{i=1,2,\dots})$ on the state space $[0, \infty) \times [0, \infty)^\infty \times \mathbb{R}^\infty$. Then $N(1)$, $\bar{N}_0(t)$, and $M(t)$ have Poisson integral

representations with respect to \tilde{Q} :

$$\begin{aligned}
 N(1) &= \int_{[0,1] \times [0,\infty)^\infty \times \mathbb{R}^\infty} \tilde{Q}(ds, d(\gamma_i), d(x_i)), \\
 \bar{N}_0(t) &= \int_{[0,1] \times [0,\infty)^\infty \times \mathbb{R}^\infty} \mathbf{1}_{\{s+\gamma_1 \leq t\}}((s, (\gamma_i), (x_i))) \tilde{Q}(ds, d(\gamma_i), d(x_i)), \\
 M(t) &= \int_{[0,1] \times [0,\infty)^\infty \times \mathbb{R}^\infty} \sum_{i=1}^\infty x_i \mathbf{1}_{\{s+(\gamma_1+\dots+\gamma_i) \leq t\}}((s, (\gamma_i), (x_i))) \tilde{Q}(ds, d(\gamma_i), d(x_i)).
 \end{aligned}$$

Note that

$$N(1) - \bar{N}_0(t) = \int_{[0,1] \times [0,\infty)^\infty \times \mathbb{R}^\infty} \mathbf{1}_{\{s+\gamma_1 > t\}}((s, (\gamma_i), (x_i))) \tilde{Q}(ds, d(\gamma_i), d(x_i)).$$

The support of the integrand in $N(1) - \bar{N}_0(t)$ is disjoint from the supports of the integrands in $\bar{N}_0(t)$ and $M(t)$, and, therefore, $N(1) - \bar{N}_0(t)$ is independent of $\bar{N}_0(t)$ and $M(t)$. We conclude from (3.8) that

$$\begin{aligned}
 E[M(t, t + s) \mid M(t), \bar{N}_0(t)] &= \mu s [E[N(1) - \bar{N}_0(t) \mid M(t), \bar{N}_0(t)] \\
 &\quad + E[\bar{N}_0(t) \mid M(t), \bar{N}_0(t)]] \\
 &= \mu s [E[N(1) - \bar{N}_0(t)] + \bar{N}_0(t)] \\
 &= \mu s [\lambda_0(t) + \bar{N}_0(t)].
 \end{aligned} \tag{3.9}$$

Surprisingly, this is the same as (3.6), the formula for $E[M(t, t + s) \mid \bar{N}_0(t)]$. Hence, taking into account information additional to $\bar{N}_0(t)$ does not change the prediction of $M(t, t + s)$. A similar calculation shows that the prediction error remains the same. Also, note that we may conclude from (3.9) that

$$E[M(t, t + s) \mid M(t)] = \mu s [\lambda_0(t) + E[\bar{N}_0(t) \mid M(t)]].$$

The latter relation sheds some light on the prediction formula (2.5).

3.2. Prediction with delay in reporting

In this section we assume that the cluster process starting at T_k has the form $L_k(t - T_k) = R_k(t - T_k - D_k)$, where R_k is a Lévy process on $[0, \infty)$ with the convention that $R_k(s) = 0$ a.s. for $s \leq 0$ and D_k is a positive random variable with distribution F_D . We write $R = R_0$ for a generic element of the sequence (R_k) . We also assume that the i.i.d. sequence (D_k) , the sequence of the claim arrivals (T_k) , and the i.i.d. sequence $(R_k)_{k=0,1,\dots}$ of Lévy processes are independent. We interpret D_k as the time that elapses between the arrival time T_k of the k th claim and its reporting time $T_k + D_k$. Thus, the inhomogeneous Poisson process

$$\bar{N}_1(t) = \#\{k \geq 1 : T_k + D_k \leq t, T_k \in [0, 1]\}, \quad t \geq 1,$$

is observable at time $t \geq 1$, whereas the claim number $N(1)$ is not necessarily observable. In what follows, we will give expressions for the prediction of $M(t, t + s)$ given $\bar{N}_1(t)$:

$$\tilde{M}_\ell(t, t + s) = E[M(t, t + s) \mid \bar{N}_1(t) = \ell], \quad \ell = 0, 1, \dots, t \geq 1, s > 0.$$

The key to the derivation of the prediction and the prediction error is again an expression for the characteristic function of $M(t, t + s)$ given $\bar{N}_1(t)$.

Lemma 3.2. *The conditional characteristic function of $M(t, t + s)$ given $\{\bar{N}_1(t) = \ell\}$ has the form*

$$\begin{aligned} & \mathbb{E}[e^{ixM(t,t+s)} \mid \bar{N}_1(t) = \ell] \\ &= (\mathbb{E}[e^{ixR(s)}])^\ell \exp\left\{-\lambda \int_{v=0}^1 \int_{r=t-v}^{t+s-v} (1 - \mathbb{E}[e^{ixR(t+s-v-r)}]) \, dv F_D(dr)\right\} \end{aligned} \tag{3.10}$$

for $\ell = 0, 1, \dots, t \geq 1$, and $s > 0$.

Proof. We start by calculating the characteristic function of $M(t, t + s)$ conditional on (T_j) and (D_j) :

$$\begin{aligned} & \mathbb{E}[e^{ixM(t,t+s)} \mid (T_j), (D_j)] \\ &= \mathbb{E}\left[\prod_{k=1}^\infty \exp\{ixR_k(t - T_k - D_k, t + s - T_k - D_k)\mathbf{1}_{\{T_k \leq 1\}}\} \mid (T_j), (D_j)\right] \\ &= \prod_{k=1}^{N(1)} \mathbb{E}[e^{ixR(t-T_k-D_k, t+s-T_k-D_k)} \mid T_k, D_k] \\ &= (\mathbb{E}[e^{ixR(s)}])^{\bar{N}_1(t)} \prod_{\{k: T_k \leq 1, T_k + D_k > t\}} \mathbb{E}[e^{ixR(t+s-T_k-D_k)} \mid T_k, D_k]. \end{aligned} \tag{3.11}$$

In the last step we used the fact that $R(t - T_k - D_k) = 0$ a.s. for k such that $T_k + D_k \geq t$ a.s. Write \tilde{Q} for the Poisson process of the points (T_k, D_k) on the state space $[0, 1] \times (0, \infty)$ with mean measure $\lambda \text{Leb} \times F_D$. Then the Poisson process \bar{N}_1 has representation

$$\bar{N}_1(t) = \int_{v=0}^1 \int_{r=0}^{t-v} \tilde{Q}(dv, dr), \quad t \geq 1. \tag{3.12}$$

Since the σ -field generated by $\bar{N}_1(t)$ is contained in the σ -field generated by (T_k) and (D_k) , we have, in view of (3.11),

$$\begin{aligned} & \mathbb{E}[e^{ixM(t,t+s)} \mid \bar{N}_1(t)] \\ &= (\mathbb{E}[e^{ixR(s)}])^{\bar{N}_1(t)} \mathbb{E}\left[\prod_{\{k: T_k \leq 1, T_k + D_k > t\}} \mathbb{E}[e^{ixR(t+s-T_k-D_k)} \mid T_k, D_k] \mid \bar{N}_1(t)\right] \\ &= (\mathbb{E}[e^{ixR(s)}])^{\bar{N}_1(t)} \\ &\quad \times \mathbb{E}\left[\exp\left\{\sum_{\{k: t < T_k + D_k \leq t+s\}} \log \mathbb{E}[e^{ixR(t+s-T_k-D_k)} \mid T_k, D_k]\right\} \mid \bar{N}_1(t)\right] \\ &= (\mathbb{E}[e^{ixR(s)}])^{\bar{N}_1(t)} \\ &\quad \times \mathbb{E}\left[\exp\left\{\int_{v=0}^1 \int_{r=t-v}^{t+s-v} \log \mathbb{E}[e^{ixR(t+s-v-r)}] \tilde{Q}(dv, dr)\right\} \mid \bar{N}_1(t)\right]. \end{aligned}$$

We observe that the Poisson integrals in the last expression and in (3.12) have disjoint supports; hence, they are independent and

$$\begin{aligned} & \mathbb{E}[e^{ixM(t,t+s)} \mid \bar{N}_1(t)] \\ &= (\mathbb{E}[e^{ixR(s)}])^{\bar{N}_1(t)} \mathbb{E}\left[\exp\left\{\int_{v=0}^1 \int_{r=t-v}^{t+s-v} \log \mathbb{E}[e^{ixR(t+s-v-r)}] \tilde{Q}(dv, dr)\right\}\right]. \end{aligned}$$

Direct calculation for $t \geq 1$ and $s > 0$ yields (3.10).

Remark 3.3. Note that, for $t \geq 1$ and $s > 0$,

$$\int_{v=0}^1 \int_{r=t-v}^{t+s-v} (1 - E[e^{ixR(t+s-v-r)}]) \, dv F_D(dr) = \int_t^{t+s} (1 - E[e^{ixR(t+s-z)}]) G(dz),$$

where $G = F_U * F_D$. Therefore,

$$\begin{aligned} & \exp \left\{ -\lambda \int_{v=0}^1 \int_{r=t-v}^{t+s-v} (1 - E[e^{ixR(t+s-v-r)}]) \, dv F_D(dr) \right\} \\ &= \exp \left\{ -E[\bar{N}_1(t, t+s)] \int_t^{t+s} \frac{1 - E[e^{ixR(t+s-z)}]}{E[\bar{N}_1(t, t+s)]} G(dz) \right\} \\ &= E[(E[e^{ixR(t+s-Z)}])^{\bar{N}_1(t, t+s)}], \end{aligned}$$

where Z is independent of R and has distribution $G(dz)/E[\bar{N}_1(t, t+s)]$ on $(t, t+s]$. This expression is an alternative formula for the second term in (3.10).

Differentiation of the conditional characteristic function at $x = 0$ sufficiently often yields the following result.

Theorem 3.2. Consider model (1.2) with the delayed Lévy processes $L_k(t) = R_k(t - D_k)$, $k = 1, 2, \dots$, as cluster processes, where, by convention, $R_k(s) = 0$ a.s. for $s \leq 0$, (D_k) constitutes an i.i.d. sequence of positive random variables with distribution F_D , and the sequences (T_k) , (D_k) , and (R_k) are independent.

- (i) Assume that $\mu = E[R(1)]$ exists and is finite. Then the prediction $\tilde{M}_\ell(t, t+s]$ of $M(t, t+s]$ given $\{\bar{N}_1(t) = \ell\}$ has the following form for $\ell = 0, 1, \dots$:

$$\tilde{M}_\ell(t, t+s] = \mu s \ell + \mu J_1, \quad t \geq 1, s > 0, \tag{3.13}$$

where

$$J_i = J_i(t, s) = \lambda \int_{v=0}^1 \int_{r=t-v}^{t+s-v} (t+s-r-v)^i F_D(dr) \, dv, \quad i = 1, 2.$$

- (ii) If, in addition, $\sigma^2 = \text{var}(R(1)) < \infty$ then

$$\text{var}(M(t, t+s] \mid \bar{N}_1(t) = \ell) = \ell s \sigma^2 + \sigma^2 J_1 + \mu^2 J_2. \tag{3.14}$$

Remark 3.4. The prediction formulae (3.13) and (3.6) are rather similar. Both are linear functions of $\bar{N}_1(t) = \ell$ or $\bar{N}_0(t) = \ell$, respectively. This is in agreement with the assumptions of the chain ladder, which is a standard technique for claims reserving; see [18] and [19]. A particularly interesting case occurs when the delay in reporting variable D is $U(0, a)$ distributed for some $a > 0$. Then

$$\tilde{M}_\ell(t, t+s] = \mu s \ell + \mu \lambda a^{-1} \int_0^1 \int_{r=(t-v) \wedge a}^{(t+s-v) \wedge a} (t+s-r-v) \, dr \, dv.$$

A comparison of the conditional prediction errors (3.7) and (3.14) shows that both are linear functions of ℓ as well. Since $\bar{N}_1(t)$ is Poisson distributed with parameter

$$E[\bar{N}_1(t)] = \lambda \int_{v=0}^1 \int_{r=0}^{t-v} \, dv F_D(dr) = \lambda \int_{t-1}^t F_D(v) \, dv, \quad t \geq 1,$$

a straightforward calculation yields the unconditional prediction error

$$E[\text{var}(M(t, t + s) \mid \bar{N}_1(t))] = \lambda\sigma^2s \int_{t-1}^t F_D(v) \, dv + \sigma^2 J_1 + \mu^2 J_2.$$

3.3. Truncated compound Poisson clusters

In this section we consider another modification of the Poisson cluster process. We again consider model (1.2), but here the cluster processes L_k are not Lévy processes. We assume that the i.i.d. processes L_i , $i = 1, 2, \dots$, have the following structure:

$$L_i(t) = K_i \wedge R_i(t), \quad t \geq 1, \tag{3.15}$$

where K_i is a nonnegative random variable and R_i is a nonnegative compound Poisson process with representation given on the right-hand side of (3.1). Moreover, we assume that $(K_i)_{i=0,1,\dots}$ is an i.i.d. sequence independent of the i.i.d. sequence $(R_i)_{i=0,1,\dots}$. We also write N_i for the homogeneous Poisson process with intensity $\gamma > 0$ which underlies the compound Poisson process R_i . If $N_i = R_i$ and K_i is a nonnegative integer-valued random variable, K_i is interpreted as the total number of payments for the i th claim. In particular, if $K_i = k_0$ is a constant integer, there are exactly k_0 payments for each claim.

As explained before, in practice we are often not informed at time $t = 1$ whether a claim has happened, i.e. the arrival times T_k are often unknown until some future instant of time. Therefore, in Section 3.1 we only consider those claims for which $T_k + \Gamma_{k1} \leq t$, and in Section 3.2 we only consider those claims for which $T_k + D_k \leq t$. Here we will consider the prediction of $M(t, t + s)$ given $\bar{N}_0(t)$ defined in (3.2) and $(R_k(t - T_k))_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}}$. We write \mathcal{H}_t for the σ -field generated by these quantities which are observable at time t . Here and in what follows, we use the notation of Section 3.1.

Theorem 3.3. *Assume that the i.i.d. sequence (L_i) has the structure described in (3.15).*

(i) *If $E[K_1] < \infty$ then the prediction of $M(t, t + s)$, $t \geq 1$, $s > 0$, given \mathcal{H}_t has the form*

$$\begin{aligned} E[M(t, t + s) \mid \mathcal{H}_t] &= \lambda_0(t) E[K_0 \wedge R_0(s)] \\ &+ \sum_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} E[K_0 \wedge (R_0(s) + R_k(t - T_k)) \\ &\quad - K_0 \wedge R_k(t - T_k) \mid R_k(t - T_k)], \end{aligned}$$

where $\lambda_0(t)$ is defined in (3.4). Here $(R_k)_{k=0,1,\dots}$, (T_k) , and $(K_k)_{k=0,1,\dots}$ are independent.

(ii) *If, in addition, $\text{var}(K_1) < \infty$ then, for $t \geq 1$ and $s > 0$,*

$$\begin{aligned} \text{var}(M(t, t + s) \mid \mathcal{H}_t) &= \lambda_0(t) E[(K_0 \wedge R_0(s))^2] \\ &+ \sum_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} \text{var}(K_0 \wedge (R_0(s) + R_k(t - T_k)) \\ &\quad - K_0 \wedge R_k(t - T_k) \mid R_k(t - T_k)). \end{aligned}$$

This result again follows by differentiation at 0 of the characteristic function of $M(t, t + s)$ given \mathcal{H}_t .

Lemma 3.3. *The conditional characteristic function of $M(t, t + s]$ given \mathcal{H}_t has the form*

$$\mathbb{E}[e^{ixM(t,t+s)} \mid \mathcal{H}_t] \tag{3.16}$$

$$= \prod_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} \mathbb{E}[e^{ix[K_0 \wedge (R_0(s) + R_k(t - T_k)) - K_0 \wedge R_k(t - T_k)]} \mid R_k(t - T_k)] \\ \times \exp\{-\lambda_0(t)(1 - \mathbb{E} e^{ixK_0 \wedge R_0(s)})\}. \tag{3.17}$$

Proof. We start by observing that

$$M(t, t + s) = \sum_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} [K_k \wedge R_k(t + s - T_k) - K_k \wedge R_k(t - T_k)] \\ + \sum_{\{k: T_k \leq 1, T_k + \Gamma_{k1} > t\}} K_k \wedge R_k(t + s - T_k) \\ = \sum_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} [K_k \wedge R_k(t + s - T_k) - K_k \wedge R_k(t - T_k)] \\ + \sum_{\{k: T_k \leq 1, T_k + \Gamma_{k1} > t\}} K_k \wedge R_k(t - T_k, t + s - T_k).$$

The σ -field \mathcal{H}_t is contained in the σ -field \mathcal{G}_t generated by (T_k) and $((\Gamma_{ki}, X_{ki}))_{k,i \geq 1, T_k + \Gamma_{ki} \leq t}$. Therefore,

$$\mathbb{E}[e^{ixM(t,t+s)} \mid \mathcal{G}_t] \\ = \prod_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} \mathbb{E}[e^{ix[K_0 \wedge (R_0(s) + R_k(t - T_k)) - K_0 \wedge R_k(t - T_k)]} \mid R_k(t - T_k)] \\ \times \prod_{\{k: T_k \leq 1, T_k + \Gamma_{k1} > t\}} \mathbb{E}[e^{ixK_0 \wedge R_k(t - T_k, t + s - T_k)} \mid T_k, R_k(t - T_k)] \\ = \prod_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} \mathbb{E}[e^{ix[K_0 \wedge (R_0(s) + R_k(t - T_k)) - K_0 \wedge R_k(t - T_k)]} \mid R_k(t - T_k)] \\ \times (\mathbb{E}[e^{ixK_0 \wedge R_0(s)}])^{N(1) - \bar{N}_0(t)}.$$

The first factor is measurable with respect to \mathcal{H}_t , and \mathcal{H}_t is independent of $N(1) - \bar{N}_0(t)$. Therefore, we have

$$\mathbb{E}[e^{ixM(t,t+s)} \mid \mathcal{H}_t] \\ = \prod_{\{k: T_k \leq 1, T_k + \Gamma_{k1} \leq t\}} \mathbb{E}[e^{ix[K_0 \wedge (R_0(s) + R_k(t - T_k)) - K_0 \wedge R_k(t - T_k)]} \mid R_k(t - T_k)] \\ \times \mathbb{E}[(\mathbb{E}[e^{ixK_0 \wedge R_0(s)}])^{N(1) - \bar{N}_0(t)}].$$

The latter relation implies the desired result (3.16).

Remark 3.5. The unconditional prediction error is given by

$$\begin{aligned} & E[\text{var}(M(t, t + s) \mid \mathcal{H}_t)] \\ &= \lambda_0(t) E[(K_0 \wedge R_0(s))^2] \\ &+ E \left[\sum_{k=1}^{N(1)} \mathbf{1}_{\{N_k(t-T_k) \geq 1\}} \text{var}(K_0 \wedge (R_0(s) + R_k(t - T_k)) \right. \\ &\quad \left. - K_0 \wedge R_k(t - T_k) \mid R_k(t - T_k)) \right]. \end{aligned}$$

Using the order statistics property of the homogeneous Poisson process N , we obtain

$$\begin{aligned} & E[\text{var}(M(t, t + s) \mid \mathcal{H}_t)] - \lambda_0(t) E[(K_0 \wedge R_0(s))^2] \\ &= \lambda \int_0^1 E[\mathbf{1}_{\{N_1(t-y) \geq 1\}} \text{var}(K_0 \wedge (R_0(s) + R_1(t - y)) \\ &\quad - K_0 \wedge R_1(t - y) \mid R_1(t - y))] dy. \end{aligned}$$

Acknowledgements

A major part of this paper was written when Muneya Matsui visited the Department of Mathematics at the University of Copenhagen. Its hospitality and the financial support from the FNU grant 272-06-0442 are gratefully acknowledged. We would also like to thank the anonymous referee for helpful comments.

References

- [1] BARTLETT, M. S. (1963). The spectral analysis of point processes. *J. R. Statist. Soc. B* **25**, 264–296.
- [2] FAÏ, G., GONZÁLEZ-ARÉVALO, B., MIKOSCH, T. AND SAMORODNITSKY, G. (2006). Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Systems* **54**, 121–140.
- [3] GIRAITIS, L., MOLCHANOV, S. A. AND SURGAILIS, D. (1993). Long memory shot noises and limit theorems with application to Burgers’ equation. In *New Directions in Time Series Analysis, Part II* (IMA Vol. Math. Appl. **46**), Springer, New York, pp. 153–176.
- [4] HEINRICH, L. AND SCHMIDT, V. (1985). Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. Appl. Prob.* **17**, 709–730.
- [5] HSING, T. AND TEUGELS, J. L. (1989). Extremal properties of shot noise processes. *Adv. Appl. Prob.* **21**, 513–525.
- [6] JESSEN, A. H., MIKOSCH, T. AND SAMORODNITSKY, G. (2009). Prediction of outstanding payments in a Poisson cluster model. Preprint. Available at www.math.ku.dk/~mikosch/Preprint/Prediction/prediction260309.pdf.
- [7] KALLENBERG, O. (1983). *Random Measures*, 3rd edn. Akademie, Berlin.
- [8] KLÜPPELBERG, C. AND KÜHN, C. (2004). Fractional Brownian motion as a weak limit of Poisson shot noise processes—with applications to finance. *Stoch. Process. Appl.* **113**, 333–351.
- [9] KLÜPPELBERG, C. AND MIKOSCH, T. (1995). Delay in claim settlement and ruin probability approximations. *Scand. Actuarial J.* **1995**, 154–168.
- [10] KLÜPPELBERG, C. AND MIKOSCH, T. (1995). Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli* **1**, 125–147.
- [11] KLÜPPELBERG, C., MIKOSCH, T. AND SCHÄRF, A. (2003). Regular variation in the mean and stable limits for Poisson shot noise. *Bernoulli* **9**, 467–496.
- [12] KONSTANTOPOULOS, T. AND LIN, S.-J. (1998). Macroscopic models for long-range dependent network traffic. *Queueing Systems* **28**, 215–243.
- [13] KURTZ, T. G. (1997). Limit theorems for workload input models. In *Stochastic Networks. Theory and Applications*, eds F. P. Kelly, S. Zachary and I. Ziedins, Oxford University Press, pp. 119–139.
- [14] LANE, J. A. (1984). The central limit theorem for the Poisson shot-noise process. *J. Appl. Prob.* **21**, 287–301.
- [15] LANE, J. A. (1987). The Berry–Esseen bound for the Poisson shot-noise. *Adv. Appl. Prob.* **19**, 512–514.
- [16] LEVY, J. B. AND TAQUU, M. S. (2000). Renewal reward processes with heavy-tailed interrenewal times and heavy-tailed rewards. *Bernoulli* **6**, 23–44.

- [17] LEWIS, P. A. W. (1964). A branching Poisson process model for the analysis of computer failure patterns. *J. R. Statist. Soc. B* **26**, 398–456.
- [18] MACK, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bull.* **23**, 213–225.
- [19] MACK, T. (1994). Which stochastic model is underlying the chain ladder method? *Insurance Math. Econom.* **15**, 133–138.
- [20] MCCORMICK, W. P. (1997). Extremes for shot noise processes with heavy tailed amplitudes. *J. Appl. Prob.* **34**, 643–656.
- [21] MIKOSCH, T. (2009). *Non-Life Insurance Mathematics*, 2nd edn. Springer, Berlin.
- [22] MIKOSCH, T. AND SAMORODNITSKY, G. (2007). Scaling limits for cumulative input processes. *Math. Operat. Res.* **32**, 890–918.
- [23] MIKOSCH, T., RESNICK, S., ROOTZÉN, H. AND STEGEMAN, A. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Prob.* **12**, 23–68.
- [24] PIRAS, V. AND TAQQU, M. S. (2000). The limit of a renewal-reward process with heavy-tailed rewards is not a linear fractional stable motion. *Bernoulli* **6**, 607–614.
- [25] SAMORODNITSKY, G. (1996). A class of shot noise models for financial applications. In *Athens Conference on Applied Probability and Time Series Analysis* (Lecture Notes Statist. **114**), Vol. I, Springer, New York, pp. 332–353.
- [26] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- [27] STEGEMAN, A. (2002). Extremal behavior of heavy-tailed on-periods in a superposition of on/off processes. *Adv. Appl. Prob.* **34**, 179–204.
- [28] VERE-JONES, D. (1970). Stochastic models for earthquake occurrence. *J. R. Statist. Soc. B* **32**, 1–62.