Modeling and Estimation of Stochastic Transition Rates in Life Insurance with Regime Switching Based on Generalized Cox Processes

David Baños[1] Erik Bølviken[2], Sindre Duedahl[3] and Frank Proske[2]

Abstract

In this paper we aim at modeling stochastic transition rates of state processes in life insurance by using generalized Cox processes. A feature of our non-Gaussian model is that it can be used to capture "regime switching" effects of data which may be due to regulatory changes in insurance markets or external "shocks" caused e.g. by an economical crisis, natural disasters or epidemics. We propose a method how to estimate the unknown parameters of our model for stochastic transition rates from insurance data by using non-linear filtering techniques for Lévy processes. As a result we also obtain an explicit formula for the unnormalized density of a filtering problem with singular coefficients.

Key words and phrases: Life insurance, Stochastic Transition Rates, Lévy processes, non-linear filtering

AMS 2000 classification: 60G51; 60G35; 60H15; 60H40; 60H15; 91B70

1 Introduction

An important challenge in the risk analysis and risk management of life insurance companies worldwide has been the accurate modeling of transition

[1]Inland Norway University of Applied Sciences. PO Box 400, 2418 Elverum, Norway
E-mail address: david.banos@inn.no

[2]Centre of Mathematics for Applications (CMA), Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N-0316 Oslo, Norway.
E-mail address: davidru@math.uio.no, erikb@math.uio.no, proske@math.uio.no

[3]Institute of Computer Science II, Friedrich-Ebert-Alle 144, University of Bonn, 53113 Bonn, Germany.
E-mail address: duedahl@cs.uni-bonn.de

rates as e.g. mortality rates or disability transition rates in the calculation of insurance premiums. Compared to financial risk of technical interest rates longevity risk e.g. , which is due to increasing life expectancy of policy holders and pensioners, is a source of insurance risk, which has been systematically underestimated for many years. A reason for the negligence of this type of risk in the insurance business has also been due to the use of deterministic models for mortality rates as e.g. the classical Gompertz-Makeham model. The latter models however, which cannot capture the uncertainty of the future dynamics of mortality rates, have led to a miscalculation of insurance premiums with respect to defined-benefit pension plans and annuities, from which many insurance companies have suffered substantial losses.

In order to overcome the deficiencies of deterministic models for transition rates, there have been various attempts in the literature in recent years to describe the dynamics of future transition rates rates by using stochastic models. See e.g.the models of Lee, Carter [**?**] or Cairns, Blake, Dowd [**?**] in the case of mortality rates.

In this paper we want to study a non-Gaussian stochastic model for stochastic transition rates, which allows for the modeling of "regime switching" effects of data or more precisely "regime switching" effects of the jump behaviour or the tails of the distribution of data which may be due to different types of influence factors as e.g. regulatory changes in insurance markets or external "shocks" caused by a financial or political crisis, natural disasters or epidemics.

To be more specific, we consider in the following a cádlág stochastic process $Z_t, 0 \leq t \leq T$ with a finite state space S on some probability space $(\Omega, \mathcal{F}, P)$, which is used as a model for the state of the insured dynamically in time. Further, we denote by $N_{ik}(t)$ the process which counts the number of transitions from state i to k of the state process $Z_t, 0 \leq t \leq T$ in the time interval $(0, t]$. In a regular insurance model with a Markovian state process it is well known that

$$N_{ik}(t) - \int_0^t \mu_{ik}(s)ds, 0 \leq t \leq T$$

is a P-martingale with respect to the natural filtration $\left\{\mathcal{F}_t^Z\right\}_{0 \leq t \leq T}$, where $\mu_{ik}(s)$ is the transition rate at time s with respect to a transition from i to j. See e.g. [**?**].

One of the deficiencies of such a model as mentioned is that the deterministic transition rates may not capture the actual future transition rates.

Therefore it is reasonable to assume a stochastic model for the transition rates $\mu_{ik}(t), 0 \leq t \leq T$:

In the sequel, let $\mu_{ik}(t,x), 0 \leq t \leq T$ be the transition rate at time t of an insured aged x years with respect to a transition from state i to state j, $i,j \in S$. In particular, the state space S of the insured in the case of a permanent disability insurance consists of the states $*$ ("alive"), $\diamond$ ("permanently disabled") and $\dagger$ ("dead").

In order to estimate stochastic transition rates from insurance data one may think of $\mu_{ik}(t,x), 0 \leq t \leq T$ as a result of a "parametrization" of the deterministic transition rates by means of an unknown "parametrization process" $X_t, 0 \leq t \leq T$.

More precisely, if $S = \{*, \diamond, \dagger\}$ one could assume that

$$\begin{aligned} \mu_{*\diamond}(t,x) &= Y_t^{(1)} + 10^{Y_t^{(2)} + Y_t^{(3)} x}, \\ \mu_{*\dagger}(t,x) &= \mu_{\diamond\dagger}(t,x) = Y_t^{(4)} + 10^{Y_t^{(5)} + Y_t^{(6)} x}, \end{aligned}$$

where $Y_t = (Y_t^{(1)}, ..., Y_t^{(6)}), 0 \leq t \leq T$ is a generalized Cox process given by

$$dY_t = h(t, X_t)dt + dB_t^Y + \int_{\mathbb{R}^6} \varsigma N_\lambda(dt, d\varsigma)$$

and $X_t, 0 \leq t \leq T$ the unknown "parametrization process" modeled by the stochastic differential equation (SDE)

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X \tag{1}$$

for Borel functions h, b and σ, where $B_t^Y \in \mathbb{R}^6$, $B_t^X \in \mathbb{R}^d$ are independent Brownian motions and where N_λ is the jump measure of a "generalized Cox process" with a predictable compensator $\widehat{\mu}$ given by

$$\widehat{\mu}(dt, d\varsigma, \omega) = \lambda(t, X_t, \varsigma)dt\nu(d\varsigma) \tag{2}$$

for a Lévy measure ν and a Borel function λ.

More generally, we may assume in this paper that stochastic transition rates $\mu_{ik}(t,x), 0 \leq t \leq T, i,k \in S$ are described by a stochastic Gompertz-Makeham model $GM(r,s)$ given by

$$\mu_{ik}(t,x) = h_{ik}^{1,r}(t,x) + \exp(h_{ik}^{2,s}(t,x)), \tag{3}$$

where $h_{ik}^{1,r}(t,x)$, $h_{ik}^{2,s}(t,x)$ are time-dependent stochastic polynomials of degree r and s, respectively, that is

$$h_{ik}^{1,r}(t,x) = \sum_{l=0}^{r} Y_t^{(l)} x^l$$

and

$$h_{ik}^{2,s}(t,x) = \sum_{l=0}^{s} Y_t^{(r+1+l)} x^l$$

for all $i, k \in S$.

In order to estimate the unknown "parametrization" process $X_t, 0 \leq t \leq T$ from (indirectly) observed insurance data

$$Y_t = (Y_t^{(0)}, ..., Y_t^{(r)}, Y_t^{(r+1)}, ..., Y_t^{(r+s)})^*, 0 \leq t \leq T, \tag{4}$$

where $*$ denotes transposition, one can apply non-linear filtering techniques for Lévy processes as proposed in [**?**] to the *signal process* $X_t \in \mathbb{R}^n, 0 \leq t \leq T$ and the *observation process* $Y_t \in \mathbb{R}^m, 0 \leq t \leq T$:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X, \tag{5}$$

$$dY_t = h(t, X_t)dt + dB_t^Y + \int_{\mathbb{R}^m} \varsigma N_\lambda(dt, d\varsigma), \tag{6}$$

where $m = r + s + 2$.

Using the latter non-Gaussian filtering framework, we want to model stochastic transition rates, which are subject to regime switching effects of insurance data. In modeling this phenomenon one could e.g. assume that the "parametrization" process $X_t, 0 \leq t \leq T$ is described by

$$dX_t = b(X_t)dt + dB_t^X, \tag{7}$$

where the drift coefficient $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a *discontinuous* vector field. An example of such a discontinuous vector field is

$$b(t,x) = \begin{cases} a_1 & \text{, if } \|x\| \geq \tau \\ a_2 & \text{else} \end{cases} .$$

Here the vectors $a_1, a_2 \in \mathbb{R}^n$ stand for the different regime switching states the parametrization process X_t will assume, if it exceeds a certain threshold τ at time t, that is $\|X_t\| \geq \tau$, or not.

Another example of such a drift coefficient in the case $n = 1$, which exhibits the feature of *mean-reversion* in connection with regime switching effects is

$$b(x) = \begin{cases} a(b_1 - x) & \text{, if } x \geq \tau \\ a(b_2 - x) & \text{else} \end{cases} .$$

for $a, b_1, b_2 \geq 0$. In this case the parametrization process X_t may be interpreted as a mean-reverting process with a mean reversion coefficients a and different long-run average levels b_1, b_2 depending on the threshold τ.

The parameters a, b_1, b_2 and the threshold τ in the above examples are a priori unknown and will be estimated from insurance data by using non-linear filtering techniques.

The non-linear filtering problem for our model is to find the least square estimate to the (possibly transformed) signal process X_t at time t, given the history of the observation process up to time t, that is to determine the conditional expectation

$$E[f(X_t) \,|\mathcal{F}_t^Y] \,,$$

where f is a given Borel function and where $\mathcal{F}_t^Y$ is the $\sigma-$algebra, generated by $\{Y_s, 0 \leq s \leq t\}$.

One of the objectives of this paper is the derivation of an explicit representation of the unnormalized conditional density with respect to the optimal filter of the filter problem (**??**) and (**??**), when the drift coefficient b in (**??**) is merely (bounded and) Borel measurable. In solving this problem, we explicitly construct a (weak) solution to a stochastic partial differential equation given by the Duncan-Mortensen-Zakai or shortly Zakai equation for the conditional unnormalized density, which can be regarded as a weak solution to a stochastic Fokker-Planck equation with singular coefficients. See [**?**] in the deterministic case. Our method relies on a representation formula of the unnormalized conditional density found in [**?**] in the case of regular coefficients and finite Lévy measures, which we want to invoke in connection with an approximation argument and local time techniques. As a result we give an explicit representation of the unnormalized conditional density associated with the least square estimate of the unknown parametrization process X_t of the generalized Cox process (**??**) in our model for the dynamics of stochastic transition rates. In contrast to [**?**] we do not require in this paper that b is regular in the sense of Lipschitz continuity or that the Lévy measure ν in (**??**) is finite.

We remark that non-linear filtering has been intensively studied in the literature since the 1960's. See e.g. Lipster and Shiryaev [**?**], Kallianpur [**?**], Fleming and Rishel [**?**], Xiong [**?**] and the references therein. See also the innovation approach for the conditional density of the filter process by Fujisaki, Kallianpur and Kunita (see e.g.[**?**]). As for solutions of the Zakai equation in the Gaussian case we refer the reader to Zakai [**?**], Gyöngy, Krylov [**?**], [**?**], Pardoux [**?**], [**?**], Kunita [**?**]. See also [**?**], [**?**] or the works [**?**], [**?**], which give a generalization of results in [**?**], [**?**] to the non-Gaussian case.

The main objective of this paper is to introduce a model for the dynamics

of stochastic transition rates which is able to describe "regime switching" effects of the jump or tail distribution behaviour of e.g. observed mortality rates or transition rates in disability insurance by using the generalized Cox process (**??**) in the framework of non-linear filtering for Lévy processes, where the signal process, that is the parametrization process X_t of (**??**) is modeled by a SDE with singular coefficients.

A popular model for stochastic transition rates in the case of mortality rates was proposed by Lee, Carter [**?**]. In this discrete-time model the error terms are Gaussian distributed. A generalization of the Lee-Carter model is the Gaussian two-factor stochastic mortality model by Cairns, Blake and Dowd [**?**], which is used to describe the different behaviours of mortality rates at lower and higher ages. Reasons for the success of these models in life insurance is the simplicity of their implementation and their prediction reliability in forecasting mortality rates under "usual" circumstances. However, a disadvantage of these models is that they cannot capture e.g. the observed skewness and (semi-) heavy tailed innovation distributions of data coming from cohort effects or short term catastrophic events as e.g. the Tsunami in 2004. In recent years there have been therefore several attempts to tackle this problem in the literature. In order to model heavy-tailed distributions of mortality data Giacometti et al. [**?**] generalized the Lee-Carter model by modeling the distributional behaviour of the error terms by infinitely divisible distributions in the case of Normal Inverse Gaussian laws. Another model in this direction, which is based on non-Gaussian distributions for error terms in the framework of [**?**], is the paper of Wang et al. [**?**]. See also the approach in [**?**] based on Markov regime switching models or [**?**], where the authors employ jump diffusions to describe age-adjusted mortality rates.

Contrary to our model (**??**) and (**??**), however, the above mentioned models cannot be used to model the rather complex phenomenon of the occurrence of changing types of jumps or types of heavy-tailedness of distributions of real data as a result of different types of "external" shocks. The reason for this is that these models are finite-dimensional models (in discrete time). Our model can be regarded as an infinite dimensional model for stochastic transition rates, since one of the unknown parameters is given by the parametrization process $X_t, 0 \leq t \leq T$. In this paper we use the powerful tool of non-linear filtering for Lévy processes to efficiently estimate this process from constantly updated observations. Therefore we may expect that our approach is more flexible than those mentioned and also suitable for the modeling of other types of stochastic transition rates beyond mortality rates.

Our paper is organized as follows:

In Section 2 we introduce the framework of our paper and derive an explicit representation of the unnormalized conditional density associated with the least square estimate of the parametrization process $X_t, 0 \leq t \leq T$ by constructing an explicit (weak) solution of a Zakai equation with singular coefficients. Further, we study the regularity of the obtained solution. Using the results of Section 2, we finally want to discuss in Section 3 various specifications of our model and its implementation in life insurance based on Monte-Carlo simulation.

2 Framework and Main Results

In this Section we want to introduce the mathematical framework of our general model for stochastic transition rates and to discuss the estimation of the unknown parameters or parameter processes of the the model from constantly updated observations in connection with a non-linear filtering problem for Lévy processes. In solving this problem we derive an explicit representation of the optimal filter of the filtering problem by constructing a (weak) L^p-solution of the Zakai equation for the unnormalized conditional density of the filter process with initial Lévy noise and singular coefficients.

In what follows we consider a Lévy process $L_t \in \mathbb{R}^m, 0 \leq t \leq T$, that is a stochastically continuous process with stationary independent increments starting in zero defined on a filtered complete probability space

$$(\Omega^*, \mathcal{F}^*, \pi^*), \{\mathcal{F}_t^*\}_{0 \leq t \leq T},$$

where $\{\mathcal{F}_t^*\}_{0 \leq t \leq T}$ is a π^*-augmented filtration generated by $L_.$.

We may here assume from now on that $L_t, 0 \leq t \leq T$ is a càdlàg process, that is a process, whose paths are right continuous paths and have existing left limits.

By the Lévy-Itô theorem the Lévy process $L_t = (L_t^{(1)}, ..., L_t^{(m)}), 0 \leq t \leq T$ can be uniquely decomposed as

$$\begin{aligned} L_t^{(i)} = &\sum_{k=1}^{l} a_{ik} B_t^{(k)} + b_i t + \int_0^{t+} \int_{\mathbb{R}_0^m} z_i \mathbf{1}_{\{\|z\| \geq 1\}} N(ds, dz) \\ &+ \int_0^{t+} \int_{\mathbb{R}_0^m} z_i \mathbf{1}_{\{\|z\| < 1\}} \widetilde{N}(ds, dz), \end{aligned}$$

for $0 \leq t \leq T$, $i = 1, ..., m$, where $B_t = (B_t^{(k)})_{1 \leq k \leq l} \in \mathbb{R}^l, 0 \leq t \leq T$ is a Brownian motion, $(a_{ik})_{1 \leq i \leq m, 1 \leq k \leq l} \in \mathbb{R}^{m \times l}, (b_i)_{1 \leq i \leq m} \in \mathbb{R}^m$ and $\widetilde{N}(ds, dz) = N(ds, dz) - ds\nu(dz)$ the compensated Poisson random measure associated with the Lévy process $L_.$. Here ν is a $\sigma-$finite measure on the Borel sets $\mathcal{B}(\mathbb{R}_0^m)$, $\mathbb{R}_0^m := \mathbb{R}^m \setminus \{0\}$, referred to as Lévy measure, which satisfies the integrability condition

$$\int_{\mathbb{R}_0^d} 1 \wedge \|z\|^2 \, \nu(dz) < \infty$$

for the Euclidean norm $\|\cdot\|$. See e.g. [**?**] or [**?**] for more information on Lévy processes.

In what follows we want to estimate the unknown "parametrization" process $X_t, 0 \leq t \leq T$ from the observed insurance data (**??**) by analyzing the non-linear filtering problem

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X, \tag{8}$$

$$dY_t = h(t, X_t)dt + dB_t^Y + \int_{\mathbb{R}^m} \varsigma N_\lambda(dt, d\varsigma), \tag{9}$$

for the *signal process* $X_t \in \mathbb{R}^n$ and the *observation process* $Y_t \in \mathbb{R}^m, 0 \leq t \leq T, n, m \in \mathbb{N}$ on a complete probability space $(\Omega, \mathcal{F}, \mu)$, where the Brownian motion $B_t^Y \in \mathbb{R}^n$ is independent of the Brownian motion $B_t^X \in \mathbb{R}^m$ and the integer valued random measure N_λ, whose predictable compensator $\widehat{\mu}$ with respect to a augmented filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$ (generated by $B_.^X, B_.^Y, N_\lambda$) is given by

$$\widehat{\mu}(dt, d\varsigma, \omega) = \lambda(t, X_t, \varsigma)dt\nu(d\varsigma) \tag{10}$$

for the Lévy measure ν of $L_t \in \mathbb{R}^m$ and a Borel function λ. Further the initial condition X_0 in (**??**) is a random variable, which is independent of B_t^X, B_t^Y and N_λ.

In order to guarantee a unique strong solution to the system (**??**) and (**??**), we require for the time being that the continuous coefficients $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^{n \times n}, h : [0, T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and $\lambda : [0, T] \times \mathbb{R}^n \times \mathbb{R}_0^m \longrightarrow \mathbb{R}$ fulfill a linear growth and Lipschitz condition, that is

$$\|b(x)\| + \|\sigma(x)\| + \|h(t, x)\| + \int\limits_{\mathbb{R}_0^m} |\lambda(t, x, \varsigma)| \, \nu(d\varsigma) \leq C(1 + \|x\|) \tag{11}$$

and

$$\begin{aligned} &\|b(x)-b(y)\|+\|\sigma(x)-\sigma(y)\|+\|h(t,x)-h(t,y)\| \\ &+\int\limits_{\mathbb{R}_0^m}|\lambda(t,x,\varsigma)-\lambda(t,y,\varsigma)|\,\nu(d\varsigma) \\ &\leq C\,\|x-y\| \end{aligned} \tag{12}$$

for all x, y, t and a constant $C < \infty$, where $\|\cdot\|$ stands for a vector or matrix norm.

For the convenience of the reader we now want to give a derivation of the Zakai equation for the unnormalized filter of the non-linear filtering problem (**??**), (**??**). See e.g. [**?**] or [**?**] in the case of Wiener noise driven obervation processes.

For this purpose denote by $\pi_t : \Omega \times \mathcal{B}(\mathbb{R}^n) \longrightarrow [0,\infty)$ the regular conditional probability measure of the signal process X_t given the $\sigma-$algebra $\mathcal{F}_t^Y$, generated by $\{Y_s, 0 \leq s \leq t\}$ and the null sets $\mathcal{N}$. Then

$$E[f(X_t)\,\big|\mathcal{F}_t^Y\,] = \langle \pi_t, f\rangle$$

for all $f \in C_b(\mathbb{R}^n)$ (space of bounded continuous functions), where $\langle \pi_t, f\rangle := \int_{\mathbb{R}^n} f(x)\pi_t(\omega, dx)$.

Suppose that the function $\lambda : [0,T]\times\mathbb{R}^n\times\mathbb{R}_0^m \longrightarrow \mathbb{R}$ is strictly positive and consider the density process

$$\begin{aligned} \Lambda_t := \exp\{&\sum_{i=1}^m \int_0^t -h_i(s,X_s)dB_s^{Y,i} - \frac{1}{2}\int_0^t \|h(s,X_s)\|^2\,ds \\ &+\int_0^t\int_{\mathbb{R}_0^m} -\log\lambda(s,X_s,\varsigma)N_\lambda(ds,d\varsigma) + \int_0^t\int_{\mathbb{R}_0^m}(\lambda(s,X_s,\varsigma)-1)ds\nu(d\varsigma)\}, \end{aligned} \tag{13}$$

for $0 \leq t \leq T$, where $B_t^Y = (B_s^{Y,1}, ..., B_s^{Y,m})^*$ and $h(t,x) = (h_1(t,x), ..., h_m(t,x))^*$ ($*$ transposition). Further, assume that

$$E[\Lambda_T] = 1. \tag{14}$$

Remark 1 *Using stopping time localization of Doleans-Dade exponentials, one obtains e.g. the following sufficient conditions for (**??**):*

$$\sup_{0\leq t\leq T} E\Bigg[\exp(6\int_0^t \|h(s,X_s)\|^2\,ds + 4\int_0^t\int_{\mathbb{R}_0^m}(1-\lambda^{-1}(s,X_s,\varsigma))\lambda(s,X_s,\varsigma)ds\nu(d\varsigma) - \int_0^t\int_{\mathbb{R}_0^m}(1-\lambda^{-4}(s,X_s,\varsigma))\lambda(s,X_s,\varsigma)ds\nu(d\varsigma)\Bigg] < \infty \tag{15}$$

$$E\left[\int_0^T\int_{\mathbb{R}_0^m}\left|(\lambda^{-4}(s,X_s,\varsigma)-1)\lambda(s,X_s,\varsigma)\right|\nu(d\varsigma)ds\right] + E\left[\int_0^T(\int_{\mathbb{R}_0^m}|(\lambda(s,X_s,\varsigma)-1)|\,\nu(d\varsigma))^2 ds\right] < \infty \tag{16}$$

$$E\left[\int_0^T\int_{\mathbb{R}_0^m}|\lambda(s,X_s,\varsigma)\log\lambda(s,X_s,\varsigma)|\,ds\nu(d\varsigma)\right] < \infty \tag{17}$$

*An example which satisfies the conditions (**??**), (**??**) and (**??**) in the case $m=1$ is given by*

$$\nu(d\varsigma) = \varphi(\varsigma)d\varsigma, \tag{18}$$

where

$$\varphi(\varsigma) = \begin{cases} \frac{1}{|\varsigma|^{1+\alpha}} & ,\ if\ |\varsigma|\leq 1 \\ 0 & else \end{cases}$$

for $\alpha\in(0,1)$ as well as h is a bounded Borel measurable function and

$$\lambda(s,x,\varsigma) = \exp(\Psi(x)\,|\varsigma|) \tag{19}$$

for a bounded and continuous function $\Psi:\mathbb{R}\longrightarrow\mathbb{R}$.

Define now the probability measure π with Radon-Nikodym derivative on $(\Omega,\mathcal{F}_t\,)$ given by

$$\left.\frac{d\pi}{d\mu}\right|_{\mathcal{F}_t} = \Lambda_t.$$

and require that

$$\int_{\mathbb{R}_0^d}\|z\|\,\nu(dz) < \infty \tag{20}$$

Then by Girsanov's theorem and the uniqueness of semimartigale characteristics (see e.g. [**?**]), the observation process $Y_t, 0 \leq t \leq T$ becomes a Lévy process being independent of the signal process under the new probability measure π. More precisely, the system (**??**), (**??**) has the following representation under π :

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dB_t^X \\ dY_t &= dB_t + dL_t, \end{aligned} \tag{21}$$

where $Y.$ is a Lévy process independent of $X.$ with

$$B_t := B_t^Y - \int_0^t (-h(s, X_s))ds, 0 \leq t \leq T$$

the Gaussian part and

$$L_t = \int_0^t \int_{\mathbb{R}_0^m} \varsigma N(ds, d\varsigma)$$

the jump component with respect the Poisson random measure $N(ds, d\varsigma) := N_\lambda(ds, d\varsigma)$ with compensator $ds\nu(d\varsigma)$.

Since $Y.$ is a Lévy process under π, we also observe that the (augmented) filtration $\mathcal{F}_t^Y$, $0 \leq t \leq T$ is right-continuous.

The so called *unnormalized filter* $\langle \Psi_t, \cdot \rangle, 0 \leq t \leq T$ is a stochastic process taking values in the space of finite Borel measures on $\mathbb{R}^n$, and is given by the Kallianpur-Striebel-formula, which is a consequence of Bayes' rule:

Theorem 2 *The optimal filter π_t has the representation*

$$\langle \pi_t, f \rangle = \frac{\langle \Psi_t, f \rangle}{\langle \Psi_t, 1 \rangle}$$

with

$$\langle \Psi_t, f \rangle := E_\pi[Z_t f(X_t) \,|\, \mathcal{F}_t^Y]$$

for all $f \in C_b(\mathbb{R}^n)$, where E_π denotes the espectation with respect to π and

where

$$\begin{aligned} Z_t &:= \Lambda_t^{-1} \\ &= \exp\{\sum_{i=1}^{m} \int_0^t h_i(s, X_s) dB_s^i - \frac{1}{2} \int_0^t \|h(s, X_s)\|^2 ds \\ &\quad + \int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, X_s, \varsigma) N(ds, d\varsigma) \\ &\quad + \int_0^t \int_{\mathbb{R}_0^m} (1 - \lambda(s, X_s, \varsigma)) ds \nu(d\varsigma)\}, \end{aligned} \tag{22}$$

for $0 \leq t \leq T$ *under* π*.*

Remark 3 *We mention the fact that*

$$E_\pi[\xi \,|\mathcal{F}_t^Y] = E_\pi[\xi \,|\mathcal{A}]$$

for all $\mathcal{F}_t-$*measurable* ξ *with* $E_\pi[|\xi|] < \infty$*, where*

$$\mathcal{A} := \bigvee_{0 \leq t \leq T} \mathcal{F}_t^Y.$$

*See Proposition 3.15 in [**?**].*

We also need the following Lemmata for the derivation of the Zakai equation:

Lemma 4 *Let* $f \in C_b^\infty(\mathbb{R}^n)$ *(space of smooth functions on* $\mathbb{R}^n$ *with bounded partial derivatives). Assume that the coefficients* b, σ *in (**??**), (**??**) are bounded and that*

$$\begin{aligned} &E\Bigg[\exp\Bigg(496 \int_0^T \|h(s, X_s)\|^2 ds + \int_0^T \left| \int_{\mathbb{R}_0^m} (1 - \lambda^{32}(s, X_s(\theta), \varsigma)) \nu(d\varsigma) \right| ds \\ &+ 32 \int_0^T \left| \int_{\mathbb{R}_0^m} (1 - \lambda(s, X_s(\theta), \varsigma)) \nu(d\varsigma) \right| ds \Bigg)\Bigg] < \infty, \end{aligned} \tag{23}$$

$$E\left[\left(\int_0^T \Big(\int_{\mathbb{R}_0^m} |\log \lambda(r, X_r(\theta), \varsigma)|^j \nu(d\varsigma) \Big)^k dr \right)^4\right] < \infty, \tag{24}$$

for all $j = 1, 2, 4, 8$ *,*$k = 1, 2, 3$ *and*

$$\begin{aligned} & E[\int_0^T \int_{\mathbb{R}_0^m} \left|1 - \lambda^{32}(s, X_s(\theta), \varsigma)\right| \nu(d\varsigma)ds] \\ & \quad + E[(\int_0^T (\int_{\mathbb{R}_0^m} (1 - \lambda(r, X_r(\theta), \varsigma))\nu(d\varsigma))^2 ds)^4] \\ & \quad < \infty. \end{aligned} \tag{25}$$

Then there exists a càdlàg modification of the unnormalized filter $\langle \Psi_\cdot, f \rangle$ *.*

Proof. See Appendix. ■

Remark 5 *An example satisfying the assumptions (??)-(??) in Lemma* **??** *is given by Remark* **??***.*

Lemma 6 *Consider* $\mathbb{F}-$*predictable processes* $\alpha_t, \beta_t, \gamma_t(\cdot), 0 \leq t \leq T$ *such that*

$$\begin{aligned} E_\pi[\int_0^T (|\alpha_s| + |\beta_s|^2)ds] &< \infty, \\ E_\pi[\int_0^T \int_{\mathbb{R}_0^m} |\gamma_s(\varsigma)|^2 \, ds\nu(d\varsigma)] &< \infty. \end{aligned}$$

Then

$$E_\pi[\int_0^t \alpha_s ds \left|\mathcal{F}_t^Y\right.] = \int_0^t E_\pi[\alpha_s \left|\mathcal{F}_s^Y\right.] \, ds,$$

$$E_\pi[\int_0^t \beta_s dB_s \left|\mathcal{F}_t^Y\right.] = \int_0^t E_\pi[\beta_s \left|\mathcal{F}_s^Y\right.] \, dB_s$$

$$E_\pi[\int_0^T \int_{\mathbb{R}_0^m} \gamma_s(\varsigma)\widetilde{N}(ds, d\varsigma) \left|\mathcal{F}_t^Y\right.] = \int_0^t \int_{\mathbb{R}_0^m} E_\pi[\gamma_s(\varsigma) \left|\mathcal{F}_s^Y\right.] \, \widetilde{N}(ds, d\varsigma)$$

and

$$E_\pi[\int_0^t \beta_s dB_s^X \left|\mathcal{F}_t^Y\right.] = 0.$$

Proof. The proof is essentially based on the independence of the increments of the process $Y_t, 0 \leq t \leq T$ under π and can be e.g. found in [**?**] or [**?**] in the case of Brownian motion. ■

Using the latter auxiliary result, we obtain the following Zakai equation for the unnormalized filter of the non-linear filtering problem (**??**), (**??**):

Theorem 7 *Assume the conditions of Lemma* **??** *and require that*

$$\sup_{0\leq s\leq T} E_\pi[|Z_s(\lambda(s,X_s,\varsigma)-1)|^p] < \infty \tag{26}$$

for all ς *and some* $p>1$. *Then the unnormalized filter* $\langle \Psi_t,\cdot\rangle, 0\leq t\leq T$ *is a càdlàg* $\mathcal{F}_t^Y-$*adapted solution to the Zakai equation, that is to the SPDE*

$$\begin{aligned}\langle \Psi_t, f\rangle &= \langle \Psi_0, f\rangle + \int_0^t \langle \Psi_s, \mathcal{L}f\rangle\, ds + \int_0^t \langle \Psi_s, f\cdot h^*(s,\cdot)\rangle\, dB_s \\ &\quad + \int_0^t \langle \Psi_{s-}, f\cdot(\lambda(s,\cdot,\varsigma)-1)\rangle\, \widetilde{N}(ds,d\varsigma)\end{aligned} \tag{27}$$

for all $f\in D\subset C_c^\infty((\mathbb{R}^n)$ *(space of compactly supported infinitely often differentiable functions of* $\mathbb{R}^n$*), where* D *is a (countable) dense subset of* $L^2(\mathbb{R}^n)$ *and* $\mathcal{L}$ *the generator of the diffusion process* X. *given by*

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^n \sigma_{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}f(x) + \sum_{i=1}^n b_i(x)\frac{\partial}{\partial x_i}f(x) \tag{28}$$

with $\sigma(x)=(\sigma_{ij}(x))_{1\leq i,j\leq n}$ *and* $b(x)=(b_1(x),...,b_n(x))^*$ *and where* $\widetilde{N}(ds,d\varsigma)$

is the compensated Poisson random measure associated with the Lévy process $Y_t, 0\leq t\leq T$ *under* π.

Proof. It follows from Itô's Lemma for $f\in C_c^\infty((\mathbb{R}^n)$ that

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L}f(X_s)ds + \int_0^t \nabla^* f(X_s)\sigma(X_s)dB_s^X,$$

where ∇^* denotes the transposed gradient. On the other hand we know that the process $Z_t, 0\leq t\leq T$ in Theorem **??** satisfies the SDE

$$Z_t = 1+\sum_{i=1}^m \int_0^t Z_s h_i(s,X_s)dB_s^i + \int_0^t\int_{\mathbb{R}_0^m} Z_{s-}(\lambda(s,X_s,\varsigma)-1)\widetilde{N}(ds,d\varsigma). \tag{29}$$

So using integration by parts, we obtain that

$$\begin{aligned}Z_t f(X_t) &= f(X_0) + \int_0^t Z_s\mathcal{L}f(X_s)ds + \int_0^t Z_s\nabla^* f(X_s)\sigma(X_s)dB_s^X \\ &\quad + \sum_{i=1}^m \int_0^t Z_s f(X_s)h_i(s,X_s)dB_s^i \\ &\quad + \int_0^t\int_{\mathbb{R}_0^m} Z_{s-}f(X_s)(\lambda(s,X_s,\varsigma)-1)\widetilde{N}(ds,d\varsigma).\end{aligned}$$

The conditional expectation with respect to $\mathcal{F}_t^Y$ applied to the latter equation combined with Lemma **??** then gives

$$\begin{aligned}
\langle \Psi_t, f\rangle &= \langle \Psi_0, f\rangle + E_\pi[\int_0^t Z_s \mathcal{L} f(X_s) ds \,|\mathcal{F}_t^Y] \\
&\quad + E_\pi[\int_0^t Z_s \nabla^* f(X_s)\sigma(X_s) dB_s^X \,|\mathcal{F}_t^Y] \\
&\quad + \sum_{i=1}^m E_\pi[\int_0^t Z_s f(X_s) h_i(s, X_s) dB_s^i \,|\mathcal{F}_t^Y] \\
&\quad + E_\pi[\int_0^t \int_{\mathbb{R}_0^m} Z_{s-} f(X_s)(\lambda(s, X_s, \varsigma) - 1)\widetilde{N}(ds, d\varsigma) \,|\mathcal{F}_t^Y] \\
&= \langle \Psi_0, f\rangle + \int_0^t \langle \Psi_s, \mathcal{L} f\rangle \, ds + \sum_{i=1}^m \int_0^t \langle \Psi_t, f \cdot h_i(s, \cdot)\rangle \, dB_s^i \\
&\quad + \int_0^t \int_{\mathbb{R}_0^m} \langle \Psi_{s-}, f \cdot (\lambda(s, \cdot, \varsigma) - 1)\rangle \, \widetilde{N}(ds, d\varsigma),
\end{aligned}$$

where we used Lemma **??**, Remark **??**, the continuity of the paths of $X_t, 0 \leq t \leq T$, the continuity of λ and (**??**) in connection with uniform integrability under the measure π. ■

Remark 8 *The condition (**??**) in Theorem **??** holds, if e.g.*

$$\sup_{0 \leq s \leq T} E_\pi[Z_s^{pr}] < \infty$$

and

$$\sup_{0 \leq s \leq T} E_\pi[|\lambda(s, X_s, \varsigma) - 1|^{pq}] < C$$

for all ς, some constant C with $pr < 2$, $\frac{1}{r} + \frac{1}{q} = 1, r, q > 1$ are satisfied. The latter conditions are e.g. covered by the conditions $B1 - B6$ in the paper later on.

In addition to the conditions (**??**), (**??**) let us from now on also require that the drift coefficient b is bounded and $\sigma = Id$ (identity).

Using the independence of the increments of the observation process $Y.$ under π and the probability density of the signal process X_t, which exists in this case, our assumptions on b, h, λ and ν imply that there is an $\mathcal{F}_t^Y-$adapted process $\Phi(t, \cdot), 0 \leq t \leq T$, called *unnormalized conditional density*, such that

$$\langle \Psi_t, f\rangle = \int_{\mathbb{R}^n} f(x)\Phi(t, x) dx, 0 \leq t \leq T$$

for all $f \in C_b(\mathbb{R}^n)$. Hence we can recast the Zakai equation (**??**) in terms of the unnormalized density and find that $\Phi(t,\cdot), 0 \leq t \leq T$ satisfies a stochastic Fokker-Planck-equation, that is the SPDE

$$\begin{aligned} d_t\Phi(t,x) &= \mathcal{L}^*\Phi(s,x)dt + \\ &\quad \Phi(t,x)h^*(s,x)dB_t + \int_{\mathbb{R}_0^m} \Phi(t^-,x)(\lambda(s,x,\varsigma)-1)\widetilde{N}(dt,d\varsigma) \\ \Phi(0,x) &= p_0(x), \end{aligned} \tag{30}$$

where $\mathcal{L}^*$ is the adjoint operator of the generator $\mathcal{L}$ of X_t and where $p_0(x)$ is the probability density of X_0, in a *weak sense*, that is $\Phi \in L^2_{loc}([0,T] \times \mathbb{R}^n; L^2(\Omega))$ is $\mathcal{F}_t^Y-$adapted process, which solves the equation

$$\begin{aligned} &\int_{\mathbb{R}^n} \Phi(t,x)f(x)dx \\ = &\int_{\mathbb{R}^n} p_0(x)f(x)dx + \int_0^t \int_{\mathbb{R}^n} \Phi(s,x)\mathcal{L}f(x)dxds \\ &+ \int_0^t \int_{\mathbb{R}^n} \Phi(s,x)h^*(s,x)f(x)dxdB_s \\ &+ \int_0^t \int_{\mathbb{R}_0^m} \int_{\mathbb{R}^n} \Phi(s^-,x)(\lambda(s,x,\varsigma)-1)f(x)dx\widetilde{N}(ds,d\varsigma), 0 \leq t \leq T \end{aligned} \tag{31}$$

for all $f \in C_c^\infty((\mathbb{R}^n)$.

In fact, it was shown in [**?**] that the Zakai equation for the unnormalized density (**??**) has a unique strong solution $\Phi(t,x)$ to (**??**) in $L^p(\mu), p \geq 1$, which is twice continuously differentiable in x, under the following conditions

- $A1$: The Lévy measure ν is bounded.
- $A2$: The drift coefficient b is contained in $C_b^{2+\beta}(\mathbb{R}^n)$.
- $A3$: The initial condition p_0 in (**??**) is positive and belongs to $C_b^{2+\beta}(\mathbb{R}^n)$.
- $A4$: The intensity rate λ is strictly positive and $\lambda(\cdot,\cdot,\varsigma) \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^m) \cap C^{2+\beta}(\mathbb{R}_+ \times \mathbb{R}^m)$ uniformly in ς.
- $A5$: $\sum_{i=1}^n \frac{\partial}{\partial x_i} b_i \in C_b^2(\mathbb{R}^n) \cap C^{2+\beta}(\mathbb{R}^n)$.
- $A6$: The observation function h is contained in $C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n) \cap C^{2+\beta}(\mathbb{R}_+ \times \mathbb{R}^n)$.
- $A7$: $\Lambda_t, 0 \leq t \leq T$ in (**??**) is a martingale,

where $C_b^{l,r}(\mathbb{R}_+ \times \mathbb{R}^d)$ is the space of l-times in $t \in (0,\infty)$ and r-times in $x \in \mathbb{R}^d$ continuously differentiable, whose partial derivatives are bounded and have continuous extensions to $\mathbb{R}_+ \times \mathbb{R}^d$ ($\mathbb{R}_+ := [0,\infty)$). The space $C^{r+\beta}(\mathcal{U})$ denotes the space of functions in $C^r(\mathcal{U})$ with all partial derivatives up to order r being Hölder continuous of order $\beta \in (0,1)$.

Moreover, the strong solution Φ to (**??**) has the following explicit representation:

$$
\begin{aligned}
& \Phi(t,x,\omega) \qquad (32)\\
= \;& E_\vartheta^x[p_0(X_t^*(\theta))\exp(-\sum_{i=1}^n \int_0^t \frac{\partial}{\partial x_i} b_i(X_s^*(\theta))ds)\\
& \exp\{\int_{T-t}^T h^*(s, X^*_{s-(T-t)}(\theta))dB_s(\omega) - \frac{1}{2}\int_{T-t}^T \left\| h(s, X^*_{s-(T-t)}(\theta))\right\|^2 ds\\
& + \int_{T-t}^T \int_{\mathbb{R}_0^m} \log(\lambda(s, X^*_{s-(T-t)}(\theta),\varsigma))\widetilde{N}(ds,d\varsigma,\omega)\\
& + \int_{T-t}^T \int_{\mathbb{R}_0^m} (\log(\lambda(s, X^*_{s-(T-t)}(\theta),\varsigma)) - (\lambda(s, X^*_{s-(T-t)}(\theta),\varsigma)-1))ds\nu(d\varsigma)\}]
\end{aligned}
$$

where $X_s^*(\theta) = X_s^{*,x}(\theta), 0 \le s \le T$, is the solution to the time-homogeneous SDE

$$dX_t^* = -b(X_t^*)dt + dB_t^*, X_0^* = x \in \mathbb{R}^d \qquad (33)$$

for a Brownian motion $B_\cdot^*$, defined on an auxiliary probability space $(\Theta, \mathcal{K}, \vartheta)$.

In order to capture "regime switching effects" in the framework of our model for the stochastic transition rates $\mu_{ik}(t,x), 0 \le t \le T$ in (**??**) and in view of Monte Carlo simulation techniques with respect to such transition rates, we now want to extend the representation of Φ (**??**) under the conditions $A1 - A7$ to the case, when the drift coefficent b of the signal process is merely bounded and measurable. In addition, we aim at relaxing the condition $A1$ of compound Poisson Lévy measures ν in (**??**) to that of finite-variation Lévy measures ν satisfying (**??**). Furthermore, we will show that such a Φ solves the Zakai equation (**??**) in the weak sense.

To this end we need to recall the concept of stochastic integration over the plane with respect to Brownian local time. See [**?**]:

Consider elementary functions $f_\Delta : [0,1] \times \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f_\Delta(s,x) = \sum_{(s_j,x_i)\in\Delta} f_{ij}\chi_{(s_j,s_{j+1}]}(s).\chi_{(x_i,x_{i+1}]}(x)\,, \qquad (34)$$

where $(x_i)_{1\le i\le n}$, $(f_{ij})_{1\le i\le n,1\le j\le m}$ are finite sequences of real numbers, $(s_j)_{1\le j\le m}$ a partition of $[0,1]$ and $\Delta = \{(s_j,x_i), 1\le i\le n, 1\le j\le m\}$. Denote by $\{L(t,x)\}_{0\le t\le 1,x\in\mathbb{R}}$ the local time of a $1-$dimensional Brownian motion B. Then the integral of integration of f_Δ with respect to L is defined as

$$\begin{aligned} &\int_0^1\int_{\mathbb{R}} f_\Delta(s,x)L(ds,dx) \qquad (35)\\ = &\sum_{(s_j,x_i)\in\Delta} f_{ij}(L(s_{j+1},x_{i+1}) - L(s_j,x_{i+1}) - L(s_{j+1},x_i) + L(s_j,x_i)). \end{aligned}$$

The latter integral can be generalized to integrands of the Banach space $(\mathcal{H},\|\cdot\|)$ of measurable functions f endowed with the norm

$$\begin{aligned} \|f\| &= 2\left(\int_0^1\int_{\mathbb{R}}(f(s,x))^2\exp(-\frac{x^2}{2s})\frac{dsdx}{\sqrt{2\pi s}}\right)^{1/2} \qquad (36)\\ &\quad + \int_0^1\int_{\mathbb{R}}|xf(s,x)|\exp(-\frac{x^2}{2s})\frac{dsdx}{s\sqrt{2\pi s}}. \end{aligned}$$

If f is such that $f(t,\cdot)$ is locally square integrable and $f(t,\cdot)$ continuous in t as a map from $[0,T]$ to $L^2_{loc}(\mathbb{R})$, then $f\in\mathcal{H}$ and

$$\int_0^t\int_{\mathbb{R}} f(s,x)L(ds,dx),\ 0\le t\le T$$

exists as well as

$$E\left[\left|\int_0^t\int_{\mathbb{R}} f(s,x)L(ds,dx)\right|\right] \le \|f\|$$

for $0\le t\le T$. Further, if $f(t,x)$ is differentiable in x, then

$$\int_0^t\int_{\mathbb{R}} f(s,x)L(ds,dx) = -\int_0^t f'(s,B_s)\,ds\,,\quad 0\le t\le T\,,$$

where $f'(s,x)$ denotes the derivative in x. See [**?**].

Assume now that $\bar{B}_t = \left(\bar{B}_t^{(1)},...,\bar{B}_t^{(n)}\right), 0\le t\le T$ is a Brownian motion, whose components $\bar{B}_t^{(i)}$ are defined on probability spaces $(\Omega_i,\mathcal{F}_i,\mu_i)$, $i = 1,...,n$. In what follows we denote by

$$\widehat{B}_t := (\widehat{B}_t^{(1)},...,\widehat{B}_t^{(d)}) := \bar{B}_{T-t}\,,\quad 0\le t\le T\,, \qquad (37)$$

the time-reversed Brownian motion. The process $\widehat{B}_t^{(i)}$ satisfies for each $i = 1, ..., d$ the equation

$$\widehat{B}_t^{(i)} = \bar{B}_1^{(i)} + \widetilde{W}_t^{(i)} - \int_0^t \frac{\widehat{B}_s^{(i)}}{T-s} ds\,, \quad 0 \leq t \leq T\,, \quad \text{a.e.,} \tag{38}$$

where $\widetilde{W}_t^{(i)}$, $0 \leq t \leq T$ are independent μ_i-Brownian motions with respect to the filtrations $\mathcal{F}_t^{\widehat{B}^{(i)}}$ generated by $\widehat{B}_\cdot^{(i)}$, $i = 1, ..., n$. See [**?**].

Using the relation (**??**) one obtains the following decomposition of local time-space integrals (see [**?**]):

$$\begin{aligned} & \int_0^t \int_{\mathbb{R}} f_i(s,x) L_i(ds,dx) \qquad (39) \\ = & \int_0^t f_i(s, \bar{B}_s^{(i)}) d\bar{B}_s^{(i)} + \int_{T-t}^T f_i(T-s, \widehat{B}_s^{(i)}) d\widetilde{W}_s^{(i)} \\ & - \int_{T-t}^T f_i(T-s, \widehat{B}_s^{(i)}) \frac{\widehat{B}_s^{(i)}}{T-s} ds, \end{aligned}$$

$0 \leq t \leq T$, a.e. for $f_i \in \mathcal{H}$, $i = 1, ..., n$. Here $L_i(t,x)$ is the local time of $\bar{B}_\cdot^{(i)}$ on (Ω_i, μ_i), $i = 1, ..., n$.

In the sequel we also need the following auxiliary result (see also [**?**]):

Lemma 9 *Let $B_t, 0 \leq t \leq T$ be a $1-$dimensional Brownian motion. Then*

$$E\left[\exp\left(k \int_0^T \frac{|B_t|}{t} dt\right)\right] < \infty$$

for all $k \geq 0$.

Proof. See Appendix. ■

Let us now assume that the following conditions are satisfied

$B1$: The Lévy measure ν fulfills condition (**??**).

$B2$: The drift coefficient b is Borel measurable and bounded.

$B3$: The initial condition p_0 in (**??**) is positive and belongs to $C_b^{2+\beta}(\mathbb{R}^n)$.

$B4$: The intensity rate λ is strictly positive and $\lambda(\cdot,\cdot,\varsigma) \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n) \cap C^{2+\beta}(\mathbb{R}_+ \times \mathbb{R}^n)$ uniformly in ς.

$B5$: The observation function h is contained in $C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n) \cap C^{2+\beta}(\mathbb{R}_+ \times \mathbb{R}^n)$.

$B6$: λ satisfies (**??**)-(**??**), (**??**)-(**??**) and the following integrability conditions

$$\begin{aligned} & \sup_{x\in\mathcal{U}} E[\exp(200\{\int_0^T \int_{\mathbb{R}_0^m} \left|(\lambda(s,\bar{B}_s^x,\varsigma)-1\right| \\ & + \max_{n=1}^{100} \int_0^T \int_{\mathbb{R}_0^m} \left|(\lambda^{2n}(s,\bar{B}_s^x,\varsigma)-1\right| ds\nu(d\varsigma)\})] \\ < \ & \infty \end{aligned} \tag{40}$$

and

$$\sup_{x\in\mathcal{U}} E\left[\left(\int_0^T \int_{\mathbb{R}_0^m} \left|\log\lambda(s,\bar{B}_s^x,\varsigma)\right|^i ds\nu(d\varsigma)\right)^8\right] < \infty, \tag{41}$$

for all $i = 1, 2, 4, 8$ and all bounded $\mathcal{U} \subset \mathbb{R}^n$.

We mention that condition $B2$ implies the the existence of a unique strong solution $X_\cdot^*$ to the the SDE (**??**). See e.g. [**?**].

We obtain the following existence result for weak solutions of a singular stochastic Fokker-Planck equation driven by Lévy noise, that is the SPDE (**??**) with the adjoint operator $\mathcal{L}^*$ of the generator $\mathcal{L}$ of X_t for merely bounded and measurable drift coefficients $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$:

Theorem 10 *Suppose that the conditions $B1-B6$ hold. Then there exists a weak solution Φ to the SPDE (**??**), which is given in law by the unnormalized*

density and takes the explicit form

$$
\begin{aligned}
& \Phi(t,x,\omega) \qquad (42)\\
= \;& E_\vartheta[p_0(\bar{B}_t^x(\theta))\exp(\sum_{i=1}^n\{\int_0^t b_i(\bar{B}_s^x(\theta))d\bar{B}_s^{(i)}+\int_{T-t}^T b_i(\widehat{B}_s^x(\theta))d\widetilde{W}_s^{(i)}\\
& -\int_{T-t}^T b_i(\widehat{B}_s^x(\theta))\frac{\widehat{B}_s^{(i)}}{T-s}ds\})\\
& \exp\{\int_{T-t}^T h^*(s,\bar{B}_{s-(T-t)}^x(\theta))dB_s(\omega)-\frac{1}{2}\int_{T-t}^T\left\|h(s,\bar{B}_{s-(T-t)}^x(\theta))\right\|^2 ds\\
& +\int_{T-t}^T\int_{\mathbb{R}_0^m}\log(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma))\widetilde{N}(ds,d\varsigma,\omega)\\
& +\int_{T-t}^T\int_{\mathbb{R}_0^m}(\log(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma))\\
& -(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma)-1))ds\nu(d\varsigma)\}\mathcal{E}(\int_0^T -b^*(\bar{B}_s^x(\theta))d\bar{B}_s)],
\end{aligned}
$$

where E_ϑ denotes the expectation with respect to the product measure $\vartheta = \mu_1\times...\times\mu_n$ with $\bar{B}_\cdot^{(i)}$ is a Brownian motion on (Ω_i,μ_i), $i=1,...,n$, $\bar{B}_t^x(\theta) := x+\bar{B}_t(\theta)$ and $\widehat{B}_t^x(\theta) := x+\widehat{B}_t(\theta)$. Further,

$$
\begin{aligned}
& \mathcal{E}(\int_0^t -b^*(\bar{B}_s^x(\theta)d\bar{B}_s)\\
= \;& \exp(\int_0^t -b^*(\bar{B}_s^x(\theta))d\bar{B}_s(\theta)-\frac{1}{2}\int_0^t\left\|b(\bar{B}_s^x(\theta)\right\|^2 ds), 0\le t\le T
\end{aligned}
$$

is the Doleans-Dade exponential.

Proof. The proof is based on the explicit representation for the unnormalized density Φ in (**??**) and an approximation argument with respect to the function b and the Lévy measure ν.

Consider a sequence of Borel sets $U_r, r\ge 1$ of $\mathbb{R}_0^m$ with $U_r\nearrow\mathbb{R}_0^m$ such that $\nu(U_r)<\infty$ for all r. Define the compound Poisson Lévy measures ν_r by

$$\nu_r(B)=\int_B 1_{U_r}(\varsigma)\nu(d\varsigma),$$

where 1_A is the indicator function of a set A. In the sequel we denote by $N_r(ds,d\varsigma)$ the Poisson random measure associated with the Lévy measure ν_r, $r\ge 1$.

Let us also choose functions $b^r \in C_c^\infty(\mathbb{R}^n)$, $r \geq 1$ such that

$$\|b^r(x)\| \leq M < \infty$$

for a constant M and all x, r as well as

$$b^r(x) \longrightarrow b(x) \text{ a.e.}$$

for $r \longrightarrow \infty$.

In the following let us denote by Φ_r the unique (strong) solution to the SPDE (**??**) with respect to the drift coefficent b^r and by $X_t^{*,r}$, $0 \leq t \leq T$ the strong solution to the SDE

$$dX_t^{*,r} = -b^r(X_t^{*,r})dt + dB_t^*, X_0^{*,r} = x \in \mathbb{R}^n \tag{43}$$

for all r.

In what follows let $x \in U$ for a bounded set $U \subset \mathbb{R}^n$.

Using Girsanov's theorem and the explicit representation of Φ_r in (**??**) for $b = b^r$ and $\nu = \nu_r$ based on the condition $B6$ we find that

$$\begin{aligned}
& \Phi_r(t,x,\omega) \qquad (44)\\
= \; & E_\vartheta[p_0(\bar{B}_t^x(\theta))\exp(\sum_{i=1}^n\{\int_0^t \frac{\partial}{\partial x_i} b_i^r(\bar{B}_s^x(\theta))ds)\\
& \exp\{\int_{T-t}^T h^*(s,\bar{B}_{s-(T-t)}^x(\theta))dB_s(\omega) - \frac{1}{2}\int_{T-t}^T \left\|h(s,\bar{B}_{s-(T-t)}^x(\theta))\right\|^2 ds\\
& + \int_{T-t}^T \int_{\mathbb{R}_0^m} \log(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma))\widetilde{N}_r(ds,d\varsigma,\omega)\\
& + \int_{T-t}^T \int_{\mathbb{R}_0^m} (\log(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma))\\
& -(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma)-1))ds\nu_r(d\varsigma)\}\mathcal{E}(\int_0^T -(b^r(\bar{B}_s^x(\theta)))^* d\bar{B}_s)].
\end{aligned}$$

If we apply (**??**) to $\bar{B}_\cdot^{(i)}$ on (Ω_i, μ_i) for

$$\begin{aligned}
& f_i(s,z)\\
= \; & b_i^r(x_1 + \bar{B}_s^{(1)}(\omega_1), ...,\\
& x_{i-1} + \bar{B}_s^{(i-1)}(\omega_{i-1}), z, x_i + \bar{B}_s^{(i+1)}(\omega_{i+1}), ..., x_n + \bar{B}_s^{(n)}(\omega_n))
\end{aligned}$$

then we get that

$$
\begin{aligned}
& \Phi_r(t,x,\omega) \qquad (45)\\
= \;& E_\vartheta[p_0(\bar{B}_t^x(\theta))\exp(\sum_{i=1}^n\{\int_0^t b_i^r(\bar{B}_s^x(\theta))d\bar{B}_s^{(i)}+\int_{T-t}^T b_i^r(\widehat{B}_s^x(\theta))d\widetilde{W}_s^{(i)}\\
& -\int_{T-t}^T b_i^r(\widehat{B}_s^x(\theta))\frac{\widehat{B}_s^{(i)}}{T-s}ds\})\\
& \exp\{\int_{T-t}^T h^*(s,\bar{B}_{s-(T-t)}^x(\theta))dB_s(\omega)-\frac{1}{2}\int_{T-t}^T\left\|h(s,\bar{B}_{s-(T-t)}^x(\theta))\right\|^2 ds\\
& +\int_{T-t}^T\int_{\mathbb{R}_0^m}\log(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma))\widetilde{N}_r(ds,d\varsigma,\omega)\\
& +\int_{T-t}^T\int_{\mathbb{R}_0^m}(\log(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma))\\
& -(\lambda(s,\bar{B}_{s-(T-t)}^x(\theta),\varsigma)-1))ds\nu_r(d\varsigma)\}\mathcal{E}(\int_0^T -(b^r(\bar{B}_s^x(\theta)))^*d\bar{B}_s)].
\end{aligned}
$$

Then it follows from the mean value theorem that

$$
\begin{aligned}
& E[(\Phi_r(t,x,\omega)-\Phi(t,x,\omega))^2]\\
& \quad = E[(p_0(\bar{B}_t^x(\theta)))^2(I_1^r+I_2^r+I_3^r)^2\Big(\int_0^1\exp(I_0^r+\tau(I_1^r+I_2^r+I_3^r))d\tau\Big)^2],
\end{aligned}
$$

where E is an expectation with respect to a probability measure under which $Y.$ associated with the bounded and measurable drift coefficient b is the Lévy process of the type in (**??**) and where

$$I_0^r := I_{0,1}^r + I_{0,2}^r + I_{0,3}^r$$

with

$$
\begin{aligned}
I_{0,1}^r := \sum_{i=1}^n\{&\int_0^t b_i^r(\bar{B}_s^x(\theta))d\bar{B}_s^{(i)}+\int_{T-t}^T b_i^r(\widehat{B}_s^x(\theta))d\widetilde{W}_s^{(i)}\\
&-\int_{T-t}^T b_i^r(\widehat{B}_s^x(\theta))\frac{\widehat{B}_s^{(i)}}{T-s}ds\},
\end{aligned}
$$

$$
\begin{aligned}
& I^r_{0,2} := \int_{T-t}^{T} h^*(s, \bar{B}^x_{s-(T-t)}(\theta)) dB_s(\omega) - \frac{1}{2}\int_{T-t}^{T} \left\| h(s, \bar{B}^x_{s-(T-t)}(\theta)) \right\|^2 ds \\
& + \int_{T-t}^{T}\int_{\mathbb{R}^m_0} \log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \widetilde{N}_r(ds, d\varsigma, \omega) \\
& + \int_{T-t}^{T}\int_{\mathbb{R}^m_0} (\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \\
& - (\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma) - 1)) ds \nu_r(d\varsigma) \\
= & \int_{T-t}^{T} h^*(s, \bar{B}^x_{s-(T-t)}(\theta)) dB_s(\omega) - \frac{1}{2}\int_{T-t}^{T} \left\| h(s, \bar{B}^x_{s-(T-t)}(\theta)) \right\|^2 ds \\
& + \int_{T-t}^{T}\int_{\mathbb{R}^m_0} \log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) 1_{U_r}(\varsigma) \widetilde{N}(ds, d\varsigma, \omega) \\
& + \int_{T-t}^{T}\int_{\mathbb{R}^m_0} (\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \\
& - (\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma) - 1)) 1_{U_r}(\varsigma) ds \nu(d\varsigma)
\end{aligned}
$$

$$
I^r_{0,3} := \int_0^T -(b^r(\bar{B}^x_s(\theta)))^* d\bar{B}_s - \frac{1}{2}\int_0^T \left\| b^r(\bar{B}^x_s(\theta)) \right\|^2 ds
$$

and

$$
\begin{aligned}
I^r_1 := & \sum_{i=1}^{n} \{ \int_0^t (b^r_i(\bar{B}^x_s(\theta)) - b_i(\bar{B}^x_s(\theta))) d\bar{B}^{(i)}_s + \int_{T-t}^{T} (b^r_i(\widehat{B}^x_s(\theta)) - b_i(\widehat{B}^x_s(\theta))) d\widetilde{W}^{(i)}_s \\
& - \int_{T-t}^{T} (b^r_i(\widehat{B}^x_s(\theta)) - b_i(\widehat{B}^x_s(\theta))) \frac{\widehat{B}^{(i)}_s}{T-s} ds \},
\end{aligned}
$$

$$
\begin{aligned}
& I^r_2 \\
: = & \int_{T-t}^{T}\int_{\mathbb{R}^m_0} \log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) (\widetilde{N}_r(ds, d\varsigma, \omega) - \widetilde{N}(ds, d\varsigma, \omega)) \\
& + \int_{T-t}^{T}\int_{\mathbb{R}^m_0} (\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \\
& - (\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma) - 1)) ds (\nu_r(d\varsigma) - \nu(d\varsigma)) \\
= & \int_{T-t}^{T}\int_{\mathbb{R}^m_0} \log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) (1_{U_r}(\varsigma) - 1) \widetilde{N}(ds, d\varsigma, \omega) \\
& + \int_{T-t}^{T}\int_{\mathbb{R}^m_0} (\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \\
& - (\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma) - 1)) (1_{U_r}(\varsigma) - 1) ds \nu(d\varsigma),
\end{aligned}
$$

$$
\begin{aligned}
I_3^r &:= \int_0^T -((b^r(\bar{B}_s^x(\theta)))^* - (b(\bar{B}_s^x(\theta)))^*)d\bar{B}_s \\
&\quad -\frac{1}{2}\int_0^T (\left\|b^r(\bar{B}_s^x(\theta))\right\|^2 - \left\|b(\bar{B}_s^x(\theta))\right\|^2)ds.
\end{aligned}
$$

It follows from the boundedness of the probability density p_0 and Hölder's inequality that

$$
E[(\Phi_r(t,x,\omega) - \Phi(t,x,\omega))^2] \leq C J_1^r J_2^r
$$

for a constant $C < \infty$, where

$$
J_1^r := E[(I_1^r + I_2^r + I_3^r)^4]^{1/2},
$$

$$
J_2^r := (\int_0^1 E[\exp(4(I_0^r + \tau(I_1^r + I_2^r + I_3^r)))]d\tau)^{1/2}.
$$

Using Burkholder's inequality, we find that

$$
\begin{aligned}
& J_1^r \\
\leq\ & K_1(E[|I_1^r|^4]^{1/2} + E[|I_2^r|^4]^{1/2} + E[|I_3^r|^4]^{1/2}) \\
\leq\ & K_2(\sum_{i=1}^n \{E[\int_0^T (b_i^r(\bar{B}_s^x(\theta)) - b_i(\bar{B}_s^x(\theta)))^4 ds]^{1/2} \\
& + E[\int_{T-t}^T (b_i^r(\widehat{B}_s^x(\theta)) - b_i(\widehat{B}_s^x(\theta)))^4 ds]^{1/2} \\
& + E[(\int_{T-t}^T \left| b_i^r(\widehat{B}_s^x(\theta)) - b_i(\widehat{B}_s^x(\theta)) \right| (\left| \widehat{B}_s^{(i)} \right| /(T-s))ds)^4]^{1/2}\} \\
& + E[(\int_{T-t}^T \int_{\mathbb{R}_0^m} ((\log(\lambda(s, \bar{B}_{s-(T-t)}^x(\theta), \varsigma)))^2 (1_{U_r}(\varsigma) - 1))^2 ds\nu(d\varsigma))^2]^{1/2} \\
& + E[\int_{T-t}^T \int_{\mathbb{R}_0^m} ((\log(\lambda(s, \bar{B}_{s-(T-t)}^x(\theta), \varsigma)))^4 (1_{U_r}(\varsigma) - 1))^4 ds\nu(d\varsigma)]^{1/2} \\
& + E[(\int_{T-t}^T \int_{\mathbb{R}_0^m} \left| (\log(\lambda(s, \bar{B}_{s-(T-t)}^x(\theta), \varsigma)) - (\lambda(s, \bar{B}_{s-(T-t)}^x(\theta), \varsigma) - 1)) \right| \\
& \cdot |(1_{U_r}(\varsigma) - 1)|\, ds\nu(d\varsigma))^4]^{1/2} \\
& + E[\int_0^T \left\| b^r(\bar{B}_s^x(\theta)) - b(\bar{B}_s^x(\theta)) \right\|^4 ds]^{1/2} \\
& + E[\int_0^T (\left\|b^r(\bar{B}_s^x(\theta))\right\|^2 - \left\|b(\bar{B}_s^x(\theta))\right\|^2)^4 ds]^{1/2})
\end{aligned}
$$

for constants $K_1, K_2 < \infty$. Since

$$E[(\int_{T-t}^{T} \left|\widehat{B}_s^{(i)}\right| /(T-s))ds)^l] < \infty$$

for all $l \geq 1$ (see Lemma **??**) and since by assumption

$$E\left[\Big(\int_{T-t}^{T}\int_{\mathbb{R}_0^m} (\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)))^2 ds\nu(d\varsigma)\Big)^2\right] < \infty,$$

$$E\left[\Big(\int_{T-t}^{T}\int_{\mathbb{R}_0^m} \left|\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma))\right| + \left|\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma) - 1\right| ds\nu(d\varsigma)\Big)^4\right] < \infty,$$

it follows from dominated convergence that

$$J_1^r \longrightarrow 0 \text{ for } r \longrightarrow \infty.$$

Further, we obtain by Hölder's inequality that

$$\begin{aligned} J_2^r &= (\int_0^1 E[\exp(4(I_0^r + \tau(I_1^r + I_2^r + I_3^r)))]d\tau)^{1/2} \\ &\leq (\int_0^1 E[\exp(24I_{0,1}^r)]^{1/6} E[\exp(24I_{0,2}^r)]^{1/6} E[\exp(24I_{0,3}^r)]^{1/6} \\ &\quad E[\exp(6\tau I_1^r)]^{1/6} E[\exp(6\tau I_2^r)]^{1/6} E[\exp(6\tau I_3^r)]^{1/6} d\tau)^{1/2}. \end{aligned}$$

Using localization applied to Doleans-Dade exponentials combined with Lemma **??** once more, we get that

$$\begin{aligned} & E[\exp(24I_{0,1}^r)] \\ \leq\ & E[\exp(K_1 \int_0^T \left\|b^r(\bar{B}_s^x(\theta))\right\|^2 ds)]^{2/3} \\ & \cdot E[\exp(K_2 \int_{T-t}^{T} \left|\widehat{B}_s^{(i)}\right| /(T-s))ds)]^{1/3} \\ \leq\ & K_3 E[\exp(K_2 \int_{T-t}^{T} \left|\widehat{B}_s^{(i)}\right| /(T-s))ds)]^{1/3} \\ \leq\ & K_4 < \infty \end{aligned}$$

for constants $K_i, i = 1, ..., 4$. On the other hand, repeated use of a localization argument with respect to Doleans-Dade exponentials yields

$$\begin{aligned} & E[\exp(24I^r_{0,2})] \qquad (46)\\ \leq\ & E[\exp(1128\int_0^T \left\|h(s,\bar{B}^x_s)\right\|^2 ds \\ & +48\int_0^T\int_{\mathbb{R}^m_0}\left|1-\lambda(s,\bar{B}^x_s,\varsigma)\right| ds\nu(d\varsigma) \\ & +\int_0^T\int_{\mathbb{R}^m_0}\left|1-\lambda^{48}(s,\bar{B}^x_s,\varsigma)\right| ds\nu(d\varsigma))]^{1/2} \\ <\ & \infty \end{aligned}$$

Similarly to the above estimates we see from our assumptions that

$$\begin{aligned} & E[\exp(24I^r_{0,3})] \\ \leq\ & E[\exp(K\int_0^T \left\|b^r(\bar{B}^x_s)\right\|^2 ds)]^{1/2} \\ \leq\ & M<\infty, \end{aligned}$$

$$\begin{aligned} & E[\exp(6\tau I^r_1)] \\ \leq\ & E[\exp(K_1\int_0^T \left\|b^r(\bar{B}^x_s(\theta))-b(\bar{B}^x_s(\theta))\right\|^2 ds)]^{2/3} \\ & \cdot E[\exp(K_2\int_{T-t}^T \left|\widehat{B}^{(i)}_s\right|/(T-s))ds)]^{1/3} \\ \leq\ & C<\infty, \end{aligned}$$

$$\begin{aligned} & E[\exp(6\tau I^r_2)] \qquad (47)\\ \leq\ & E[\exp(24\int_0^T\int_{\mathbb{R}^m_0}\left|1-\lambda(s,\bar{B}^x_s,\varsigma)\right| ds\nu(d\varsigma) \\ & +\int_0^T\int_{\mathbb{R}^m_0}\left|1-\lambda^{12}(s,\bar{B}^x_s,\varsigma)\right| ds\nu(d\varsigma))]^{1/2} \\ <\ & R<\infty \end{aligned}$$

and

$$\begin{aligned} & E[\exp(6\tau I^r_3)] \\ \leq\ & E[\exp(K\int_0^T (\left\|b^r(\bar{B}^x_s\right\|^2+\left\|b(\bar{B}^x_s)\right\|^2)ds)]^{1/2} \\ \leq\ & H<\infty. \end{aligned}$$

Altogether, we obtain that for all bounded sets $U \subset \mathbb{R}^n$:

$$\sup_{0\leq t\leq T, x\in U} E[(\Phi_r(t,x) - \Phi(t,x))^2] \longrightarrow 0 \tag{48}$$

for $r \longrightarrow \infty$.

Denote by $\mathcal{L}_r$ the differential operator in (**??**) for $b = b^r$. Then, using the Itô-isometry, relation (**??**) implies that

$$\int_{\mathbb{R}^n} \Phi_r(t,x) f(x) dx \longrightarrow \int_{\mathbb{R}^n} \Phi(t,x) f(x) dx,$$

$$\int_0^t \int_{\mathbb{R}^n} \Phi_r(s,x) \mathcal{L}_r f(x) dx ds \longrightarrow \int_0^t \int_{\mathbb{R}^n} \Phi(s,x) \mathcal{L} f(x) dx ds,$$

$$\int_0^t \int_{\mathbb{R}^n} \Phi_r(s,x) h^*(s,x) f(x) dx dB_s \longrightarrow \int_0^t \int_{\mathbb{R}^n} \Phi(s,x) h^*(s,x) f(x) dx dB_s$$

and

$$\begin{aligned} & \int_0^t \int_{\mathbb{R}_0^m} \int_{\mathbb{R}^n} \Phi_m(s^-,x)(\lambda(s,x,\varsigma) - 1) f(x) dx \widetilde{N}_r(ds, d\varsigma) \\ = & \int_0^t \int_{\mathbb{R}_0^m} \int_{\mathbb{R}^n} \Phi_m(s^-,x)(\lambda(s,x,\varsigma) - 1) f(x) 1_{U_r}(\varsigma) dx \widetilde{N}(ds, d\varsigma) \\ \longrightarrow & \int_0^t \int_{\mathbb{R}_0^m} \int_{\mathbb{R}^n} \Phi(s^-,x)(\lambda(s,x,\varsigma) - 1) f(x) dx \widetilde{N}(ds, d\varsigma) \end{aligned}$$

for $r \longrightarrow \infty$ in $L^2(\Omega)$ uniformly in t for all $f \in C_c^\infty(\mathbb{R}^n)$. Thus $\Phi \in L^2_{loc}([0,T] \times \mathbb{R}^n; L^2(\Omega))$ is an adapted process, which solves the SPDE (**??**) in a weak sense.

Consider now the unique strong solutions $X_t^r, 0 \leq t \leq T$, $r \geq 1$ of the SDE for the signal process (**??**)

$$dX_t^r = b^r(X_t^r)dt + dB_t^X, X_0^r = x$$

and denote by $Y_t^r, 0 \leq t \leq T$ the corresponding oberservation process.

It is known that

$$X_t^r \longrightarrow X_t \text{ for } r \longrightarrow \infty$$

in $L^2(\Omega)$ for all t. See e.g. [**?**].

Denote by $Z_\cdot^r$, π_r the Doleans-Dade exponentials and probability measures in (**??**) with respect to b^r and ν_r, $r \geq 0$, where we set $b_0 := b$, $\nu_0 := \nu$,

$Z_\cdot := Z^0_\cdot$, $\pi_r := \pi_0$. Since $X^r_\cdot$ is independent of $Y^r_\cdot$ under π_r, we find by means of Girsanov's theorem applied to the signal process that

$$\begin{aligned}
& E_{\pi_r}[Z^r_t f(X^r_t)\,|\mathcal{F}^{Y^r}_t]\,(\omega) \\
= \;& E_\vartheta[\exp\{\sum_{i=1}^m \int_0^t h_i(s,\bar{B}^x_s(\theta))dB^{i,r}_s(\omega) - \frac{1}{2}\int_0^t \left\|h(s,\bar{B}^x_s(\theta))\right\|^2 ds \\
& + \int_0^t \int_{\mathbb{R}^m_0} \log\lambda(s,\bar{B}^x_s(\theta),\varsigma)1_{U_r}(\varsigma)N^r(ds,d\varsigma,\omega) \\
& + \int_0^t \int_{\mathbb{R}^m_0} (1-\lambda(s,\bar{B}^x_s(\theta),\varsigma))1_{U_r}(\varsigma)ds\nu(d\varsigma)\} \\
& f(\bar{B}^x_t(\theta))\mathcal{E}(\int_0^T (b^r(\bar{B}^x_s(\theta)))^* d\bar{B}_s)]
\end{aligned}$$

π_r-a.e. and therefore $\mu-$a.e., where $B^r_\cdot$ and N^r are the Brownian motion and Poisson random measure under π_r, respectively.

Let g be a bounded Lipschitz function on $\mathbb{R}$. Then

$$\begin{aligned}
& E_\mu[(g(E_{\pi_r}[Z^r_t f(X^r_t)\,|\mathcal{F}^{Y^r}_t])] \\
= \;& E_{\pi^*}[g(E_\vartheta[\exp\{\sum_{i=1}^m \int_0^t h_i(s,\bar{B}^x_s(\theta))dB^i_s - \frac{1}{2}\int_0^t \left\|h(s,\bar{B}^x_s(\theta))\right\|^2 ds \\
& + \int_0^t \int_{\mathbb{R}^m_0} \log\lambda(s,\bar{B}^x_s(\theta),\varsigma)1_{U_r}(\varsigma)N(ds,d\varsigma) \\
& + \int_0^t \int_{\mathbb{R}^m_0} (1-\lambda(s,\bar{B}^x_s(\theta),\varsigma))1_{U_r}(\varsigma)ds\nu(d\varsigma)\} \\
& f(\bar{B}^x_t(\theta))\mathcal{E}(\int_0^T (b^r(\bar{B}^x_s(\theta)))^* d\bar{B}_s)])]
\end{aligned}$$

where π^* is a probability measure under which $B_\cdot$ is a Brownian motion independent of a Poisson random measure N associated with the Lévy measure ν. By using the same reasoning as above, one sees that

$$E_\mu[(g(E_{\pi_r}[Z^r_t f(X^r_t)\,|\mathcal{F}^{Y^r}_t])] \longrightarrow E_\mu[(g(E_\pi[Z_t f(X_t)\,|\mathcal{F}^Y_t])]$$

for $r \longrightarrow \infty$. Hence

$$E_{\pi_r}[Z^r_t f(X^r_t)\,|\mathcal{F}^{Y^r}_t] \longrightarrow E_\pi[Z_t f(X_t)\,|\mathcal{F}^Y_t]$$

for $r \longrightarrow \infty$ in distribution. Similarly, by employing the representation (**??**), we have that

$$E_{\pi_r}[Z^r_t f(X^r_t)\,|\mathcal{F}^{Y^r}_t] \longrightarrow \int_{\mathbb{R}^n} f(x)\Phi(t,x)dx$$

for $r \longrightarrow \infty$ in distribution for all $f \in C_c^\infty(\mathbb{R}^n)$.

On the other hand, since X_t possesses a probability density, we also know that there exists an unnormalized density $\widetilde{\Phi}$ of the corresponding filter problem. So we obtain that

$$\begin{aligned} & E_\mu[(g(E_\pi[Z_t f(X_t)\,\big|\mathcal{F}_t^Y\,])] \\ = \;& E_\mu[(g(\int_{\mathbb{R}^n} f(x)\widetilde{\Phi}(t,x)dx)] = E_\mu[(g(\int_{\mathbb{R}^n} f(x)\Phi(t,x)dx)] \end{aligned}$$

for all bounded Lipschitz functions g and $f \in C_c^\infty(\mathbb{R}^n)$. If we now choose f such that

$$f(x) = \frac{1}{\varepsilon^n}\eta(\frac{y-x}{\varepsilon})$$

for a standard mollifier and $\varepsilon > 0$, then we find for $\varepsilon \searrow 0$ that

$$E_\mu[(g(\widetilde{\Phi}(t,x))] = E_\mu[(g(\Phi(t,x))]$$

$x-$a.e. Hence, separability implies that for all t

$$\widetilde{\Phi}(t,x) \stackrel{\text{law}}{=} \Phi(t,x)$$

$x-$a.e. ■

Our next result, which pertains to the regularity of solutions Φ given by (**??**) in the case of discontinuous drift coefficients $b : \mathbb{R} \longrightarrow \mathbb{R}$, requires the following additional condition:

$$\begin{aligned} & E_\vartheta[\int_0^T \left\|h(s,\bar{B}_s^{x_1}(\theta)) - h(s,\bar{B}_s^{x_2}(\theta))\right\|^{16} ds \qquad (49) \\ & + \int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s,\bar{B}_s^{x_1}(\theta),\varsigma)) - \log(\lambda(s,\bar{B}_s^{x_2}(\theta),\varsigma))\right|^8 \nu(d\varsigma)ds \\ & +(\int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s,\bar{B}_s^{x_1}(\theta),\varsigma)) - \log(\lambda(s,\bar{B}_s^{x_2}(\theta),\varsigma))\right|^4 \nu(d\varsigma)ds)^2 \\ & \int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s,\bar{B}_s^{x_1}(\theta),\varsigma)) - \log(\lambda(s,\bar{B}_s^{x_2}(\theta),\varsigma))\right|^2 \nu(d\varsigma)ds)^4 \\ & +(\int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s,\bar{B}_s^{x_1}(\theta),\varsigma)) - \log(\lambda(s,\bar{B}_{s-(T-t)}^{x_2}(\theta),\varsigma))\right| ds\nu(d\varsigma))^8 \\ & (\int_0^T \int_{\mathbb{R}_0^m} \left|\lambda(s,\bar{B}_s^{x_1}(\theta),\varsigma) - \lambda(s,\bar{B}_s^{x_2}(\theta),\varsigma)\right| ds\nu(d\varsigma))^8] \\ \leq \;& C(|x_1 - x_2|^{16} + |x_1 - x_2|^8) \end{aligned}$$

for all $x_1, x_2 \in \mathbb{R}$, where C is a constant.

We obtain the following regularity result:

Theorem 11 *Retain the conditions of Theorem* **??** *and suppose that (***??***) holds. Further, assume that the drift coefficient b in (***??***) is a step function of the form*

$$b(x) = \sum_{i=1}^{r} \xi_i 1_{(a_i, b_i]}(x), x \in \mathbb{R},$$

where $\xi_i, a_i, b_i \in \mathbb{R}, i = 1, ..., r$.
*Then for all t a modification of the weak solution $\Phi(t, \cdot)$ to the SPDE (***??***) in Theorem* **??** *is locally Hölder continuous with exponent α for all $\alpha \in (0, 1/4)$.*

Proof. Using relation (**??**), we see that Φ can be written as

$$\Phi(t, x, \omega) = E_\vartheta[p_0(\bar{B}_t^x(\theta))I(x)],$$

where

$$\begin{aligned}
&I(x) := \exp\{\int_0^t \int_{\mathbb{R}} b(y) L^{\bar{B}^x}(ds, dy) \\
&\int_{T-t}^{T} h^*(s, \bar{B}^x_{s-(T-t)}(\theta)) dB_s(\omega) - \frac{1}{2}\int_{T-t}^{T} \left\| h(s, \bar{B}^x_{s-(T-t)}(\theta)) \right\|^2 ds \\
&+ \int_{T-t}^{T} \int_{\mathbb{R}_0^m} \log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \widetilde{N}(ds, d\varsigma, \omega) \\
&+ \int_{T-t}^{T} \int_{\mathbb{R}_0^m} (\log(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma)) \\
&-(\lambda(s, \bar{B}^x_{s-(T-t)}(\theta), \varsigma) - 1)) ds \nu(d\varsigma)\} \mathcal{E}(\int_0^T -b(\bar{B}_s^x(\theta)) d\bar{B}_s).
\end{aligned}$$

Hence,

$$\begin{aligned}
& E[(\Phi(t, x_1, \omega) - \Phi(t, x_2, \omega))^4] \\
\leq\ & CE[E_\vartheta[(p_0(\bar{B}_t^{x_1}(\theta)) - p_0(\bar{B}_t^{x_2}(\theta)))^4 (I(x_1))^4 \\
& + (I(x_1) - I(x_2))^4]]
\end{aligned} \tag{50}$$

for a constant C. On the other hand, it follows from the mean value theorem and Hölder's inequality that

$$\begin{aligned}
& E[E_\vartheta[(I(x_1)-I(x_2))^4]] \\
\leq\ & CE[E_\vartheta[((I_1(x_1)-I_1(x_2))^8+(I_2(x_1)-I_2(x_2))^8+(I_3(x_1)-I_3(x_2))^8)]]^{\frac{1}{2}} \\
& E[E_\vartheta[\int_0^1 \exp(8(I_1(x_1)+I_2(x_1)+I_3(x_1) \\
& +\tau(I_1(x_1)-I_1(x_2)+I_2(x_1)-I_2(x_2)+I_3(x_1)-I_3(x_2))))d\tau)]]^{\frac{1}{2}},
\end{aligned}$$

where

$$I_1(x) := \int_0^t \int_{\mathbb{R}} b(y) L^{\bar{B}^x}(ds,dy),$$

$$\begin{aligned}
& I_2(x) \\
:\ =\ & \int_{T-t}^T h^*(s,\bar{B}^x_{s-(T-t)}(\theta))dB_s(\omega) - \frac{1}{2}\int_{T-t}^T \left\| h(s,\bar{B}^x_{s-(T-t)}(\theta))\right\|^2 ds \\
& + \int_{T-t}^T \int_{\mathbb{R}_0^m} \log(\lambda(s,\bar{B}^x_{s-(T-t)}(\theta),\varsigma))\widetilde{N}(ds,d\varsigma,\omega) \\
& + \int_{T-t}^T \int_{\mathbb{R}_0^m} \log(\lambda(s,\bar{B}^x_{s-(T-t)}(\theta),\varsigma)) - (\lambda(s,\bar{B}^x_{s-(T-t)}(\theta),\varsigma)-1))ds\nu(d\varsigma)
\end{aligned}$$

and

$$I_3(x) := \int_0^T -b(\bar{B}^x_s(\theta))d\bar{B}_s - \frac{1}{2}\int_0^T (b(\bar{B}^x_s(\theta)))^2 ds.$$

Using the above notation, we have that

$$\begin{aligned}
& E[E_\vartheta[((I_3(x_1)-I_3(x_2))^8]] \\
\leq\ & C\{E_\vartheta[(\int_0^T b(\bar{B}^{x_1}_s(\theta))d\bar{B}_s - \int_0^T b(\bar{B}^{x_2}_s(\theta))d\bar{B}_s)^8] \\
& + E_\vartheta[(\int_0^T (b(\bar{B}^{x_1}_s(\theta)))^2 ds - \int_0^T (b(\bar{B}^{x_2}_s(\theta)))^2 ds)^8].
\end{aligned}$$

Further , it follows from the Tanaka formula and our assumptions on the drift coefficient b that

$$\begin{aligned}
& \int_0^T b(\bar{B}^x_s(\theta))d\bar{B}_s \\
& = \sum_{i=1}^r \xi_i(L(T,b_i-x)-L(T,a_i-x)) - \sum_{i=1}^r \xi_i\{(\bar{B}^x_s-b_i)^- - (\bar{B}^x_s-a_i)^-\} \text{ a.e.}
\end{aligned}$$

where $(a)^- := \min(0, a)$, $a \in \mathbb{R}$.So

$$\begin{aligned} & E_\vartheta[(\int_0^T b(\bar{B}_s^{x_1}(\theta))d\bar{B}_s - \int_0^T b(\bar{B}_s^{x_2}(\theta))d\bar{B}_s)^8] \\ \leq\ & C\sum_{i=1}^r |\xi_i|^8 \{E_\vartheta[(L(T, b_i - x_1) - L(T, b_i - x_2))^8] \\ & + E_\vartheta[(L(T, a_i - x_1) - L(T, a_i - x_2))^8] \\ & + E_\vartheta[((\bar{B}_s^{x_1} - b_i)^- - (\bar{B}_s^{x_2} - b_i)^-)^8] \\ & + E_\vartheta[((\bar{B}_s^{x_1} - a_i)^- - (\bar{B}_s^{x_2} - a_i)^-)^8]\}. \end{aligned}$$

On the other hand, it is well known that

$$E_\vartheta[(L(t_1, z_1) - L(t_2, z_2))^{2l}] \leq C_{n,T}\{|t_1 - t_2|^l + (|z_1 - z_2|^l\} \qquad (51)$$

for a constant $C_{n,T}$. See e.g. [**?**]. Thus

$$\begin{aligned} & E_\vartheta[(\int_0^T b(\bar{B}_s^{x_1}(\theta))d\bar{B}_s - \int_0^T b(\bar{B}_s^{x_2}(\theta))d\bar{B}_s)^8]^{\frac{1}{2}} \\ \leq\ & C(\sum_{i=1}^r |\xi_i|^4)(|x_1 - x_2|^2 + |x_1 - x_2|^4). \end{aligned}$$

Further, we also see from the occupation time formula that

$$\begin{aligned} & E_\vartheta[(\int_0^T (b(\bar{B}_s^{x_1}(\theta)))^2 ds - \int_0^T (b(\bar{B}_s^{x_2}(\theta)))^2 ds)^8] \\ =\ & E_\vartheta[(\int_{\mathbb{R}} (b(y))^2 L(T, y - x_1)dy - \int_{\mathbb{R}} (b(y))^2 L(T, y - x_2)dy)^8] \\ =\ & E_\vartheta[(\int_{\mathbb{R}} (b(y))^2 (L(T, y - x_1) - L(T, y - x_2))dy)^8] \\ & C_b E_\vartheta[\int_{\mathbb{R}} (b(y))^{16} (L(T, y - x_1) - L(T, y - x_2))^8 dy] \end{aligned}$$

for a constant C_b depending on the compact support of the function b. Therefore, we see from (**??**) that

$$\begin{aligned} & E_\vartheta[(\int_0^T (b(\bar{B}_s^{x_1}(\theta)))^2 ds - \int_0^T (b(\bar{B}_s^{x_2}(\theta)))^2 ds)^4]^{\frac{1}{2}} \\ \leq\ & C_b(\int_{\mathbb{R}} (b(y))^{16}) E_\vartheta[(L(T, y - x_1) - L(T, y - x_2))^8]dy)^{\frac{1}{2}} \\ \leq\ & C(\int_{\mathbb{R}} (b(y))^{16} dy)^{\frac{1}{2}} |x_1 - x_2|^4 . \end{aligned}$$

The latter yields

$$E[E_\vartheta[((I_3(x_1) - I_3(x_2))^8]]^{\frac{1}{2}} \leq C\{|x_1 - x_2|^2 + |x_1 - x_2|^4\}.$$

Since

$$I_1(x) = \int_0^t \int_{\mathbb{R}} b(y) L^{\bar{B}^x}(ds, dy) = \sum_{i=1}^r \xi_i (L(t, b_i - x) - L(t, a_i - x))$$

by definition (see **??**), we can employ the same reasoning as above and obtain that

$$E[E_\vartheta[((I_1(x_1) - I_1(x_2))^8]]^{\frac{1}{2}} \leq C\,|x_1 - x_2|^2\,.$$

Further, it follows from Burkholder's inequality in connection with the assumptions (**??**), (**??**) and the inequality (**??**) that

$$\begin{aligned}
& E[E_\vartheta[(I_2(x_1) - I_2(x_2))^8]] \\
\leq\ & CE_\vartheta[\int_0^T \left\|h(s, \bar{B}_s^{x_1}(\theta)) - h(s, \bar{B}_s^{x_2}(\theta))\right\|^{16} ds \\
& + \int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s, \bar{B}_s^{x_1}(\theta), \varsigma)) - \log(\lambda(s, \bar{B}_s^{x_2}(\theta), \varsigma))\right|^8 \nu(d\varsigma) ds \\
& + (\int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s, \bar{B}_s^{x_1}(\theta), \varsigma)) - \log(\lambda(s, \bar{B}_s^{x_2}(\theta), \varsigma))\right|^4 \nu(d\varsigma) ds)^2 \\
& \int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s, \bar{B}_s^{x_1}(\theta), \varsigma)) - \log(\lambda(s, \bar{B}_s^{x_2}(\theta), \varsigma))\right|^2 \nu(d\varsigma) ds)^4 \\
& + (\int_0^T \int_{\mathbb{R}_0^m} \left|\log(\lambda(s, \bar{B}_s^{x_1}(\theta), \varsigma)) - \log(\lambda(s, \bar{B}_{s-(T-t)}^{x_2}(\theta), \varsigma))\right| ds \nu(d\varsigma))^8 \\
& (\int_0^T \int_{\mathbb{R}_0^m} \left|\lambda(s, \bar{B}_s^{x_1}(\theta), \varsigma) - \lambda(s, \bar{B}_s^{x_2}(\theta), \varsigma)\right| ds \nu(d\varsigma))^8 \\
& + \int_0^T \left|b(\bar{B}_s^{x_1}(\theta)) - b(\bar{B}_s^{x_2}(\theta))\right|^{16} ds] \\
\leq\ & C(|x_1 - x_2|^{16} + |x_1 - x_2|^8)
\end{aligned}$$

for a constant C.

Finally, using the same arguments as in the proof of Theorem **??**, our

assumptions imply that

$$\begin{aligned} & \sup_{x_1,x_2\in\mathcal{U}\times\mathcal{U}} E[E_\vartheta[\int_0^1 \exp(8(I_1(x_1)+I_2(x_1)+I_3(x_1) \\ & +\tau(I_1(x_1)-I_1(x_2)+I_2(x_1)-I_2(x_2)+I_3(x_1)-I_3(x_2))))d\tau)]]^{\frac{1}{2}} \\ \leq \quad & M<\infty \end{aligned}$$

for bounded sets $\mathcal{U}\subset\mathbb{R}$, where M is a constant depending on $\mathcal{U}$.

Altogether, we see from the above estimates that

$$E[(\Phi(t,x_1,\omega)-\Phi(t,x_2,\omega))^4]\leq C\left|x_1-x_2\right|^2$$

for all x_1,x_2 on bounded intervals $\mathcal{U}\subset\mathbb{R}$ with a constant C depending on $\mathcal{U}$.

So it follows from Kolmogorov's Lemma that for all t there is a continuous modification of $\Phi(t,\cdot)$, which is locally Hölder continuous with exponent α for all $\alpha\in(0,1/4)$. ■

Remark 12 *An example which fulfills the assumptions of Theorem* **??** *is given by (***??***), (***??***) in the case of a truncated $\alpha-$stable Lévy process with $\alpha\in(0,1)$, when $\Psi\in C_b^3(\mathbb{R})$ and $h=0$.*

3 The Model

As mentioned in the introduction of this paper, we aim at modeling stochastic transition rates $\mu_{ik}(t,x), 0\leq t\leq T$ for states $i,k\in S$ of the insured by the following stochastic Gompertz-Makeham model $GM(r,s)$ given by

$$\mu_{ik}(t,x)=h_{ik}^{1,r}(t,x)+\exp(h_{ik}^{2,s}(t,x)), \tag{52}$$

where $h_{ik}^{1,r}(t,x)$, $h_{ik}^{2,s}(t,x)$ are time-dependent stochastic polynomials of degree r and s, respectively, that is

$$h_{ik}^{1,r}(t,x)=\sum_{l=0}^{r}Y_t^{(l)}x^l$$

and

$$h_{ik}^{2,s}(t,x) = \sum_{l=0}^{s} Y_t^{(r+1+l)} x^l$$

for all $i, k \in S$.

Here the coefficients of the polynomials

$$Y_t = (Y_t^{(0)}, ..., Y_t^{(r)}, Y_t^{(r+1)}, ..., Y_t^{(r+s)})^*, 0 \leq t \leq T \tag{53}$$

are described by a generalized Cox process given by

$$dY_t = h(t, X_t)dt + dB_t^Y + \int_{\mathbb{R}^m} \varsigma N_\lambda(dt, d\varsigma), \tag{54}$$

where $m = r + s + 2$ and where the integer valued random measure N_λ has a $\mathbb{F}-$predictable compensator of the form

$$\widehat{\mu}(dt, d\varsigma, \omega) = \lambda(t, X_t, \varsigma)dt\nu(d\varsigma)$$

for a Lévy measure ν associated with a Lévy process $L_t \in \mathbb{R}^m$ and a Borel function λ. Further, the process $X_t, 0 \leq t \leq T$ is the strong solution to the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X, \tag{55}$$

where the Brownian motion $B_t^Y \in \mathbb{R}^n$ is independent of the Brownian motion $B_t^X \in \mathbb{R}^m$ and the integer valued random measure N_λ. Here the initial value X_0 is supposed to be square integrable and to be independent of $B_\cdot^X, B_\cdot^Y$, N_λ.

An important feature of our model is the unknown "parametrization" process $X_t, 0 \leq t \leq T$, which we use to describe the occurrence of changing types of jumps or types of heavy-tailedness of distributions of real data. The phenomenon of (semi-) heavy tailedness- as mentioned- may arise from data with cohort effects in mortality modeling or short term catastrophic events as e.g. earthquakes. However, the "regime switch" itself between different types of jumps, which may be of long-term nature and due to regulatory changes in the insurance branch or political decisions with a long-term impact on the economy, is modeled by the process $X_t, 0 \leq t \leq T$. In order to capture the "regime switching" effects of data, we may assume that $X_t, 0 \leq t \leq T$ is the strong solution to a SDE with singular drift coefficient

given by

$$\begin{aligned} dX_t &= \begin{cases} d\widetilde{X}_t = b(X_t)dt + dB_t^X & \\ d\widetilde{Y}_t = 0 & , \widetilde{Y}_0 = a_1 \\ d\widetilde{Z}_t = 0 & , \widetilde{Z}_0 = a_2 \end{cases} , \\ X_0 &= (\widetilde{X}_0, \widetilde{Y}_0, \widetilde{Z}_0)^* \in \mathbb{R}^{3l} \end{aligned}$$

where

$$\begin{aligned} b(x) &= \\ b(x_1, ..., x_l, a_1, a_2) &= \begin{cases} a_1 & , \text{ if } \|(x_1, ..., x_l)^*\| \geq \tau \\ a_2 & \text{else} \end{cases} \end{aligned} \tag{56}$$

for $\widetilde{X}_t := (X_t^{(1)}, ..., X_t^{(l)})^*$ and a "critical" threshold $\tau > 0$. Here, the vectors $a_1, a_2 \in \mathbb{R}^l$ can be interpreted as the different "regime switching" states of the jump intensity of the generalized Cox process $Y_t, 0 \leq t \leq T$ in (**??**), depending on whether $\left\|\widetilde{X}_t\right\| \geq \tau$ or not.

A natural generalization of the model (**??**) to the case of multiple "regime switching" states $a_1, ..., a_r \in \mathbb{R}^l$ is the following

$$b(x) = b(x_1, ..., x_l, a_1, ..., a_r) = \sum_{i=1}^{r} a_i 1_{\Gamma_i}(x_1, ..., x_l), \tag{57}$$

where $\{\Gamma_i\}_{i=1,...,r}$ is a partition of $\mathbb{R}^l$.

An alternative model to the above ones, which is able capture long-term effects of data, is given by the following "regime switching" mean-reversion model

$$dX_t = \begin{cases} dX_t^{(1)} = b(X_t)dt + dB_t^X & \\ dX_t^{(2)} = 0, X_0^{(2)} = a & \\ dX_t^{(3)} = 0, X_0^{(3)} = b_1 & , \\ dX_t^{(4)} = 0, X_0^{(4)} = b_2 & \\ dX_t^{(5)} = 0, X_0^{(5)} = \tau & \end{cases} \tag{58}$$

where

$$b(x) = \begin{cases} a(b_1 - x) & , \text{if } x \geq \tau \\ a(b_2 - x) & \text{else} \end{cases} .$$

for a mean reversion coefficient $a \geq 0$ and long-run average levels $b_1, b_2 \geq 0$, depending on a "critical" threshold $\tau > 0$.

We mention that a unique strong solution to (**??**) exists. See also [**?**], where the authors consider the latter model in connection with a regime switching short rate model in finance.

In our model (**??**), (**??**) we may e.g. choose the observation function h in (**??**) to be a constant. In this case the process $Y_t, 0 \leq t \leq T$ can be regarded as a Lévy process with a Lévy measure ν "parametrized" by the process $X_t, 0 \leq t \leq T$.

The unknown process $X_t, 0 \leq t \leq T$ or more generally $f(X_t), 0 \leq t \leq T$ for Borel functions f can be estimated from (indirectly and under optimization constraints unique) observed insurance data $Y_t, 0 \leq t \leq T$ by means of the optimal filter

$$\langle \pi_t, f \rangle = E[f(X_t) \left| \mathcal{F}_t^Y \right.], 0 \leq t \leq T. \tag{59}$$

Using the Kallianpur-Striebel formula in Theorem **??** under the condition (**??**) , we may also write (**??**) as

$$\langle \pi_t, f \rangle = \frac{\langle \Psi_t, f \rangle}{\langle \Psi_t, 1 \rangle} \tag{60}$$

with

$$\langle \Psi_t, f \rangle := E_\pi[Z_t f(X_t) \left| \mathcal{F}_t^Y \right.]$$

for all $f \in C_b(\mathbb{R}^n)$, where $Z_t, 0 \leq t \leq T$ is the Doleans-Dade exponential (**??**).

Principally, we could now use Monte-Carlo techniques in connection with Theorem **??** to simulate the unnormalized conditional density (**??**) and to compute the optimal filter by means of (**??**).

An alternative method to the latter one, which we want to discuss in this Section, is the Monte-Carlo method directly applied to the unnormalized filter $\langle \Psi_t, \cdot \rangle, 0 \leq t \leq T$.

In fact, we have the following result in the case of Lipschitz continuous coefficients b, σ, h and λ :

Proposition 13 *Assume that the functions b, σ, h and λ are bounded and satisfy the conditions (**??**), (**??**). In addition, require (**??**), (**??**) and (**??**) hold. Let $X_t^i, 0 \leq t \leq T, i \geq 1$ be a sequence of i.i.d. copies of the solution $X_t, 0 \leq t \leq T$ to (**??**) on our probability space, being independent of $Y_t, 0 \leq$*

*$t \leq T$, and denote by Z_t^i,$0 \leq t \leq T$ the stochastic exponential in (**??**) based on $X_t^i, 0 \leq t \leq T$ for all $i \geq 1$. Let $f \in C_b(\mathbb{R}^n)$. Then*

$$M^l(f) := \frac{1}{l}\sum_{i=1}^{l} Z_t^i f(X_t^i) \underset{l \longrightarrow \infty}{\longrightarrow} \langle \Psi_t, f \rangle = E_\pi[Z_t f(X_t) \left| \mathcal{F}_t^Y \right.] \quad a.e. \tag{61}$$

for all t. Moreover, for all t there exists a constant $C < \infty$ such that

$$E_\pi[(M^l(f) - \langle \Psi_t, f \rangle)^2] \leq \frac{1}{l} C \left\| f \right\|_\infty^2 \tag{62}$$

for all $l \geq 1$.

Proof. Since $X_\cdot$ is independent of $Y_\cdot$ under π, we can represent $\langle \Psi_t, f \rangle$ as

$$\begin{aligned} & \langle \Psi_t, f \rangle (\omega) \\ = \; & E_\vartheta[\exp\{\sum_{i=1}^{m} \int_0^t h_i(s, X_s(\theta)) dB_s^i(\omega) - \frac{1}{2}\int_0^t \left\| h(s, X_s) \right\|^2 ds \\ & + \int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, X_s(\theta), \varsigma) N(ds, d\varsigma, \omega) \\ & + \int_0^t \int_{\mathbb{R}_0^m} (1 - \lambda(s, X_s(\theta), \varsigma)) ds \nu(d\varsigma)\} f(X_t(\theta))], \end{aligned} \tag{63}$$

where E_ϑ denotes the expectation with respect to $X_s(\theta), 0 \leq s \leq T$ on a separate probability space. Using the latter in connection with an expectation E_ω in the direction of the other probability space, we get that

$$\begin{aligned} & E_\pi[(M^l(f) - \langle \Psi_t, f \rangle)^2] \\ = \; & E_\omega[E_\vartheta[(M^l(f) - \langle \Psi_t, f \rangle)^2]] \\ = \; & E_\omega[\frac{1}{l^2} E_\vartheta[(\sum_{i=1}^{l} (Z_t^i f(X_t^i) - \left\langle \Psi_t^i, f \right\rangle))^2] \\ = \; & E_\omega[\frac{1}{l^2} \sum_{i=1}^{l} E_\vartheta[((Z_t^i f(X_t^i) - \left\langle \Psi_t^i, f \right\rangle))^2] \\ = \; & \frac{1}{l} E_\omega[E_\vartheta[((Z_t^1 f(X_t^1) - \left\langle \Psi_t^1, f \right\rangle))^2] \\ \leq \; & \frac{4 \left\| f \right\|_\infty^2}{l} E_\pi[(Z_t)^2]. \end{aligned}$$

It follows from the conditions (**??**), (**??**) and (**??**) that there is a constant depending on the sizes of h and λ such that

$$E_\pi[(Z_t)^2] \leq C.$$

Relation (**??**) is a consequence of the strong law of large numbers applied to (**??**). ■

Since we are interested to apply Proposition **??** to our model for stochastic transition rates in the case of discontinuous coefficients b in (**??**), (**??**) or (**??**) and $\sigma = Id$, we may approximate the drift coefficient b by a bounded Lipschitz continuous function $\widetilde{b}$. To be more precise, using the notation of the previous Section, we mention that under the conditions (**??**)-(**??**) we get for bounded coefficients b that

$$\begin{aligned}
& E_\pi[Z_t f(X_t) \left| \mathcal{F}_t^Y \right.](\omega) \\
= \; & E_\vartheta[\exp\{\sum_{i=1}^m \int_0^t h_i(s, \bar{B}_s^x(\theta)) dB_s^i(\omega) - \frac{1}{2}\int_0^t \left\| h(s, \bar{B}_s^x(\theta)) \right\|^2 ds \\
& + \int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, \bar{B}_s^x(\theta), \varsigma) N(ds, d\varsigma, \omega) \\
& + \int_0^t \int_{\mathbb{R}_0^m} (1 - \lambda(s, \bar{B}_s^x(\theta), \varsigma)) ds \nu(d\varsigma)\} \\
& f(\bar{B}_t^x(\theta)) \mathcal{E}(\int_0^T (b(\bar{B}_s^x(\theta)))^* d\bar{B}_s)].
\end{aligned}$$

Denote by $\left\langle \Psi_t^b, \cdot \right\rangle, 0 \leq t \leq T$ the unnormalized filter associated with b. Then, using the mean value theorem, we obtain just as in the proof of Theorem **??** that

$$\begin{aligned}
& \sup_{0 \leq t \leq T} E_\pi[(\left\langle \Psi_t^b, \cdot \right\rangle - \left\langle \Psi_t^{\widetilde{b}}, \cdot \right\rangle)^2] \qquad (64) \\
\leq \; & C E_\pi[\int_0^T \left\| b(\bar{B}_s^x) - \widetilde{b}(\bar{B}_s^x) \right\|^2 ds \\
& + (\int_0^T \left\| b(\bar{B}_s^x) \right\|^2 - \left\| \widetilde{b}(\bar{B}_s^x) \right\|^2 ds)^2]^{\frac{1}{2}}
\end{aligned}$$

for a constant C depending on the sizes of b, h and λ. So, if we approximate b by $\widetilde{b}$ in the sense that

$$\int_0^T \int_{\mathbb{R}^n} \left\| b(y) - \widetilde{b}(y) \right\|^4 \frac{1}{(2\pi t)^{\frac{n}{2}}} \exp(- \left\| y - x \right\|^2 / 2t) dy dt < \frac{1}{r^2}$$

for $r \geq 1$ sufficiently large, then we see from (**??**) and (**??**) that

$$E_\pi[(M^{l,\widetilde{b}}(f) - \langle \Psi_t, f\rangle)^2] \leq C(\frac{1}{r} + \frac{1}{l}) \tag{65}$$

for all $l \geq 1$ and a constant C, where $M^{l,\widetilde{b}}(f)$ is the sum in (**??**) with respect to $\widetilde{b}$.

We remark that the estimate (**??**) can also be established in the case of coefficients b in (**??**), since in this case one can still apply Girsanov's theorem.

Finally, we mention that the paths of $Z_t^i, 0 \leq t \leq T$ and $X_t^i, 0 \leq t \leq T$ with respect to $\widetilde{b}$ can be simulated by using Euler-approximation scheme applied to the SDE's (**??**) and (**??**).

4 Appendix

We collect in this Section some results and proofs which we need in the maintext of the article.

Lemma 14 *Let $M_t, 0 \leq t \leq T$ be stochastically continuous process on $(\Omega, \mathcal{F}, \mu)$. Suppose that*

$$E[|M_t - M_s|^a |M_s - M_u|^b] \leq C |t-u|^{1+\gamma}$$

for some constants $a, b, C, \gamma > 0$ and all $0 \leq u \leq s \leq t \leq T$. Then $M_t, 0 \leq t \leq T$ has a càdlàg modification.

Proof. See e.g. Theorem 6.4.1 in [**?**] for a proof. ■

Proof of Lemma ??. Without loss of generality, we consider the case, when $\sigma = Id$.

Since $X_\cdot$ is independent of $Y_\cdot$ under π, we can represent $\langle \Psi_t, f\rangle$ as

$$\begin{aligned} & \langle \Psi_t, f\rangle(\omega) \\ = \; & E_\vartheta[\exp\{\sum_{i=1}^n \int_0^t h_i(s, X_s(\theta)) dB_s^i(\omega) - \frac{1}{2}\int_0^t \|h(s, X_s)\|^2 ds \\ & + \int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, X_s(\theta), \varsigma) N(ds, d\varsigma, \omega) \\ & + \int_0^t \int_{\mathbb{R}_0^m} (1 - \lambda(s, X_s(\theta), \varsigma)) ds \nu(d\varsigma)\} f(X_t(\theta))], \end{aligned} \tag{66}$$

where E_ϑ denotes the expectation with respect to $X_s(\theta), 0 \leq s \leq T$ on a separate probability space.

Set

$$\begin{aligned} I_t &= \sum_{i=1}^{n} \int_0^t h_i(s, X_s(\theta)) dB_s^i(\omega) - \frac{1}{2} \int_0^t \|h(s, X_s)\|^2 \, ds \\ &\int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, X_s(\theta), \varsigma) \widetilde{N}(ds, d\varsigma, \omega) \\ &+ \int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, X_s(\theta), \varsigma) ds \nu(d\varsigma) \\ &+ \int_0^t \int_{\mathbb{R}_0^m} (1 - \lambda(s, X_s(\theta), \varsigma)) ds \nu(d\varsigma). \end{aligned}$$

Using Hölder's inequality we see that

$$\begin{aligned} & E[|\langle \Psi_t, f \rangle - \langle \Psi_s, f \rangle|^2 |\langle \Psi_s, f \rangle - \langle \Psi_u, f \rangle|^2] \\ = \; & E[E_\vartheta[|\exp(I_t)(f(X_t(\theta)) - f(X_s(\theta))) + f(X_s(\theta))(\exp(I_t) - \exp(I_s))|^2 \\ & |\exp(I_s)(f(X_s(\theta)) - f(X_u(\theta))) + f(X_u(\theta))(\exp(I_s) - \exp(I_u))|^2]] \\ \leq \; & C(J_1 + J_2 + J_3 + J_4), \end{aligned}$$

where

$$\begin{aligned} J_1 &: = E[E_\vartheta[\exp(2I_t) \exp(2I_s) \\ &\quad (f(X_t(\theta)) - f(X_s(\theta)))^2 (f(X_s(\theta)) - f(X_u(\theta)))^2]], \\ J_2 &: = E[E_\vartheta[\exp(2I_t) f(X_u(\theta))^2 \\ &\quad (f(X_t(\theta)) - f(X_s(\theta)))^2 (\exp(I_s) - \exp(I_u))^2]], \\ J_3 &: = E[E_\vartheta[f(X_s(\theta))^2 \exp(2I_s) \\ &\quad (f(X_s(\theta)) - f(X_u(\theta)))^2 (\exp(I_t) - \exp(I_s))^2]], \\ J_4 &: = E[E_\vartheta[f(X_s(\theta))^2 f(X_u(\theta))^2 \\ &\quad (\exp(I_t) - \exp(I_s))^2 (\exp(I_s) - \exp(I_u))^2]]. \end{aligned}$$

Using our assumptions we obtain by Hölder's inequality in connection with the independence of Brownian increments that

$$\begin{aligned} J_1 &\leq C E_\vartheta[(X_t(\theta) - X_s(\theta))^4 (X_s(\theta) - X_u(\theta))^4]^{1/2} \\ &\leq C \, |t - u|^2 . \end{aligned}$$

On the other hand, using the mean value theorem and Burkholder's inequality we see that

$$\begin{aligned}
& E[(\exp(I_s)-\exp(I_u))^4] \\
= \;& E[(\int_0^1 (I_s-I_u)\exp(I_s+\theta(I_s-I_u))d\theta)^4] \\
\leq \;& CE[(I_s-I_u)^8]^{1/2} \\
\leq \;& C\{E[(\int_u^s \|h(r,X_r)\|^2\,dr)^4]^{1/2}+E[(\int_u^s\int_{\mathbb{R}_0^m}\log\lambda(r,X_r(\theta),\varsigma)\nu(d\varsigma)dr)^8]^{1/2} \\
& +E[(\int_u^s\int_{\mathbb{R}_0^m}(1-\lambda(r,X_r(\theta),\varsigma))\nu(d\varsigma)ds)^8]^{1/2} \\
& +\sum_{i=1}^n E[(\int_u^s (h_i(r,X_r(\theta)))^2dr)^4]^{1/2} \\
& +E[\int_u^s\int_{\mathbb{R}_0^m}(\log\lambda(r,X_r(\theta),\varsigma))^8\nu(d\varsigma)dr]^{1/2} \\
& +E[(\int_u^s\int_{\mathbb{R}_0^m}(\log\lambda(r,X_r(\theta),\varsigma))^4\nu(d\varsigma)dr)^2]^{1/2} \\
& +E[(\int_u^s\int_{\mathbb{R}_0^m}(\log\lambda(r,X_r(\theta),\varsigma))^2\nu(d\varsigma)dr)^4]^{1/2}\}.
\end{aligned}$$

So

$$\begin{aligned}
& E[(\exp(I_s)-\exp(I_u))^4] \\
\leq \;& C\{|s-u|^4+|s-u|^2\,E[(\int_0^T(\int_{\mathbb{R}_0^m}\log\lambda(r,X_r(\theta),\varsigma)\nu(d\varsigma))^2dr)^4]^{1/2} \\
& +|s-u|^2\,E[(\int_0^T(\int_{\mathbb{R}_0^m}(1-\lambda(r,X_r(\theta),\varsigma))\nu(d\varsigma))^2ds)^4]^{1/2} \\
& +|s-u|^2+|s-u|^{\frac{1}{4}}\,E[(\int_0^T(\int_{\mathbb{R}_0^m}(\log\lambda(r,X_r(\theta),\varsigma))^8\nu(d\varsigma))^2dr)^{1/2}]^{1/2} \\
& +|s-u|^{\frac{1}{2}}\,E[(\int_0^T(\int_{\mathbb{R}_0^m}(\log\lambda(r,X_r(\theta),\varsigma))^4\nu(d\varsigma))^2dr)]^{1/2} \\
& +|s-u|\,E[(\int_0^T(\int_{\mathbb{R}_0^m}(\log\lambda(r,X_r(\theta),\varsigma))^2\nu(d\varsigma))^2dr)^2]^{1/2}\} \\
\leq \;& C\,|t-u|^{\frac{1}{4}}\,.
\end{aligned}$$

We also see by using Lipschitzianity that

$$E_\vartheta[(f(X_t(\theta)) - f(X_s(\theta)))^8]^{1/4} \leq C\,|t-u|\,.$$

Hence by our assumptions we get that

$$J_2 \leq C\,|t-u|^{1+\frac{1}{8}}\,.$$

Similarly, we see that

$$J_3 \leq C\,|t-u|^{1+\frac{1}{8}}\,.$$

Because of the mean value theorem and the independence of increments of Lévy processes we obtain from our assumptions that

$$\begin{aligned}
& E[E_\vartheta[(\exp(I_t) - \exp(I_s))^2(\exp(I_s) - \exp(I_u))^2]] \\
= \;& E[E_\vartheta[(\int_0^1 (I_t - I_s)\exp(I_s + \theta_1(I_s - I_u))d\theta_1)^2 \\
& (\int_0^1 (I_s - I_u)\exp(I_u + \theta_1(I_s - I_u))d\theta_1)^2]] \\
\leq \;& CE_\vartheta[E[(I_t - I_s)^4(I_s - I_u)^4]] = CE_\vartheta[E[(I_t - I_s)^4]E[(I_s - I_u)^4]].
\end{aligned}$$

Further, Hölder's and Burkholder's inequality implies that

$$
\begin{aligned}
& E[(I_s - I_u)^4] \\
\leq\ & C\{E[(\int_u^s \|h(r, X_r)\|^2 \, dr)^2] + E[(\int_u^s \int_{\mathbb{R}_0^m} \log \lambda(r, X_r(\theta), \varsigma)\nu(d\varsigma)dr)^4] \\
& + E[(\int_u^s \int_{\mathbb{R}_0^m} (1 - \lambda(r, X_r(\theta), \varsigma))\nu(d\varsigma)ds)^4] \\
& + \sum_{i=1}^n E[(\int_u^s (h_i(r, X_r(\theta)))^2 dr)^2] \\
& + E[\int_u^s \int_{\mathbb{R}_0^m} (\log \lambda(r, X_r(\theta), \varsigma))^4 \nu(d\varsigma)dr] \\
& + E[(\int_u^s \int_{\mathbb{R}_0^m} (\log \lambda(r, X_r(\theta), \varsigma))^2 \nu(d\varsigma)dr)^2]\} \\
\leq\ & C\{|s-u|^2 + |s-u|^2 \, E[(\int_0^T (\int_{\mathbb{R}_0^m} \log \lambda(r, X_r(\theta), \varsigma)\nu(d\varsigma))^2 dr)^2] \\
& + |s-u|^2 \, E[(\int_0^T (\int_{\mathbb{R}_0^m} (1 - \lambda(r, X_r(\theta), \varsigma))\nu(d\varsigma))^2 ds)^2] + |s-u|^2 \\
& + |s-u|^{\frac{2}{3}} \, E[(\int_0^T (\int_{\mathbb{R}_0^m} (\log \lambda(r, X_r(\theta), \varsigma))^4 \nu(d\varsigma))^3 dr)^{1/3}] \\
& + |s-u| \, E[(\int_0^T (\int_{\mathbb{R}_0^m} (\log \lambda(r, X_r(\theta), \varsigma))^2 \nu(d\varsigma))^2 dr)]\} \\
\leq\ & |t-u|^{\frac{2}{3}} \, (K + CL(\theta))
\end{aligned}
$$

$\theta-$a.e., where

$$
\begin{aligned}
L(\theta) \ :\ & = E[(\int_0^T (\int_{\mathbb{R}_0^m} \log \lambda(r, X_r(\theta), \varsigma)\nu(d\varsigma))^2 dr)^2] \\
& + E[(\int_0^T (\int_{\mathbb{R}_0^m} (1 - \lambda(r, X_r(\theta), \varsigma))\nu(d\varsigma))^2 ds)^2] \\
& + E[(\int_0^T (\int_{\mathbb{R}_0^m} (\log \lambda(r, X_r(\theta), \varsigma))^4 \nu(d\varsigma))^3 dr)^{1/3}] \\
& + E[(\int_0^T (\int_{\mathbb{R}_0^m} (\log \lambda(r, X_r(\theta), \varsigma))^2 \nu(d\varsigma))^2 dr)].
\end{aligned}
$$

Hence

$$\begin{aligned}
& E_\vartheta[E[(I_t - I_s)^4]E[(I_s - I_u)^4]] \\
\leq\ & |t-u|^{\frac{4}{3}} E_\vartheta[(K + CL(\theta))^2] \\
\leq\ & C\,|t-u|^{\frac{4}{3}} \cdot \\
& \cdot(K + E_\vartheta[E[(\int_0^T (\int_{\mathbb{R}_0^m} \log\lambda(r, X_r(\theta), \varsigma)\nu(d\varsigma))^2 dr)^4]] \\
& + E_\vartheta[E[(\int_0^T (\int_{\mathbb{R}_0^m} (1 - \lambda(r, X_r(\theta), \varsigma))\nu(d\varsigma))^2 ds)^4]] \\
& + E_\vartheta[E[(\int_0^T (\int_{\mathbb{R}_0^m} (\log\lambda(r, X_r(\theta), \varsigma))^4 \nu(d\varsigma))^3 dr)^{2/3}]] \\
& + E_\vartheta[E[(\int_0^T (\int_{\mathbb{R}_0^m} (\log\lambda(r, X_r(\theta), \varsigma))^2 \nu(d\varsigma))^2 dr)^2]]).
\end{aligned}$$

Altogether, it follows that

$$E[|\langle \Psi_t, f\rangle - \langle \Psi_s, f\rangle|^2 \, |\langle \Psi_s, f\rangle - \langle \Psi_u, f\rangle|^2] \leq C\,|t-u|^{1+\frac{1}{8}}$$

for a constant $C < \infty$ depending on f, which gives the proof in connection with Lemma **??**. ■

Proof of Lemma ??. It follows from (**??**) that

$$\begin{aligned}
& \int_0^t \int_{\mathbb{R}} sgn(x) L(ds, dx) \\
=\ & \int_0^t sgn(B_s) dB_s + \int_{T-t}^T sgn(\widehat{B}_s) d\widetilde{W}_s - \int_{T-t}^T sgn(\widehat{B}_s) \frac{\widehat{B}_s}{T-s} ds \\
=\ & \int_0^t sgn(B_s) dB_s + \int_{T-t}^T sgn(\widehat{B}_s) d\widetilde{W}_s - \int_{T-t}^T \frac{\left|\widehat{B}_s\right|}{T-s} ds,
\end{aligned}$$

where

$$sgn(x) := \begin{cases} 1 & \text{, if } x > 0 \\ -1 & \text{else} \end{cases}.$$

Since

$$\int_0^t \int_{\mathbb{R}} sgn(x) L(ds, dx) = -2\,|B_t| + 2 \int_0^t sgn(B_s) dB_s$$

by means of Tanaka's formula, we find that

$$\begin{aligned}\int_0^t \frac{|B_s|}{s} ds &= \int_{T-t}^T \frac{\left|\widehat{B}_s\right|}{T-s} ds \\ &= -\int_0^t sgn(B_s) dB_s + \int_{T-t}^T sgn(\widehat{B}_s) d\widetilde{W}_s + 2\,|B_t|\,.\end{aligned}$$

Using the latter combined with the supermartingale property of Doleans-Dade exponentials and Hölder's inequality, we get that

$$\begin{aligned}& E[\exp(k\int_0^T \frac{|B_t|}{t} dt)] \\ \leq\ & E[\exp(-3k\int_0^t sgn(B_s) dB_s)]^{1/3} E[\exp(3k\int_{T-t}^T sgn(\widehat{B}_s) d\widetilde{W}_s)]^{1/3}\,\cdot \\ & \cdot E[\exp(6k\,|B_t|)]^{1/3} \\ \leq\ & C_{T,k} < \infty\end{aligned}$$

for a constant $C_{T,k}$ depending on T and k. ■

References

[1] Bain, A., Crişan, D.: Fundamentals of Stochastic Filtering. Applications of Mathematics. Springer (2008).

[2] Benssoussan, A.: Stochastic Control of Partially Observable Systems. Cambridge University Press (2004).

[3] Baños, D., Duedahl, S., Meyer-Brandis, T., Proske, F.: Computing greeks without derivatives. Manuscript, University of Oslo (2015).

[4] Benth, F.E., Deck, T., Potthoff, J., Vaage, G.: Explicit strong solutions of SPDE's with applications to non-linear filtering. Acta Appl. Math. 51, 215-242 (1998).

[5] Bertoin, J.: Lévy Processes. Cambridge University Press, Cambridge 1996.

[6] Borkar, V.S.: Probability Theory. Universitext, Springer (1995).

[7] Cairns, A.J.G., Blake, D., Dowd, K.: A two factor model for stochastic mortality with uncertainty: theory and calibration. Journal of Risk and Insurance, 73(4), pp. 687-718 (2006).

[8] Cox, S.H., Lin, Y., Wang, S.S.: Multivariate exponential tilting and pricing implications for mortality securitization. The Journal of Risk and Insurance, 73(4), 719-736 (2006).

[9] Davis, M.H.A.: Lectures on Stochastic Control and Nonlinear Filtering. Tata Institute of Fundamental Research 75. Springer-Verlag (1984).

[10] Davis, M.H.A., Marcus, S.I.: An introduction to non-linear filtering. In Hazewinkel, M., Willems, J.C. (eds.): Stochastic Systems: the mathematics of filtering and identification and applications, pp 53-75, Reidel (1981).

[11] Di Masi, B., Runggaldier, W.: On approximation methods for nonlinear filtering. In Lecture Notes Math. 972, eds Mitter, S:K:, Moro, A., Springer-Verlag, Berlin, pp. 249-259 (1982).

[12] Eisenbaum, N.: Integration with respect to local time. Potential Anal. 13, 303-328 (2000).

[13] Fleming, W.H., Rishel, R.W.: Deterministic and Stochastic Optimal Control. Springer Verlag (1975).

[14] Frey, R., Runggaldier, W: A non-linear filtering approach to volatility estimation with a view towards high frequency data. Int. J. of Th. and Appl. Finance, Vol. 4, No.2, 199-210 (2001).

[15] Giacometti, R., Ortobelli, S., Bertocchi, M.I.: Impact of different distributional assumptions in forecasting Italian mortality rates. Investment Management and Financial Innovations, Vol. 6, No. 3 (2009).

[16] Grigelionis, B.: Stochastic non-linear filtering equations and semimartingales. In Lecture Notes Math. 972, eds Mitter, S.K., Moro, A., Springer-Verlag, Berlin (1982).

[17] Gyöngy, I., Krylov, N.V.: SPDE's with unbounded coefficients and applications I. Stochastics 32, 53-91 (1990).

[18] Gyöngy, I., Krylov, N.V.: SPDE's with unbounded coefficients and applications II. Stochastics 32, 165-180, (1990).

[19] Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer, Berlin Heidelberg New York (1987).

[20] Kallianpur, G.: Stochastic Filtering Theory. Springer-Verlag, New York (1980).

[21] Kalman, R. E., Bucy, R.S.: New results in filtering and prediction theory. Trans. ASME J. Basic Eng., 83, 95-108 (1961).

[22] Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus. Springer (1988).

[23] Koller, M: Stochastic Models in Life Insurance. EAA Series. Springer (2012).

[24] Kunita, H.: Stochastic Flows and Stochastic Differential Equations. Cambridge University Press, Cambridge (1990).

[25] Le Bris, C., Lions, P.-L.: Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients. Rapport de recherche du CERMICS 349 (2007).

[26] Lee, R.D., Carter, L.R.: Modeling and forecasting U.S. Mortality. J. American Stat. Association, Vol. 87, No. 419, 659-671 (1992).

[27] Lipster, R.S., Shiryayev, A.N.: Statistics of Random Processes. Springer Verlag (1977).

[28] Mandrekar, V., Meyer-Brandis, T., Proske, F.: A Bayes formula for non-linear filtering with Gaussian and Cocx noise. Probability and Statistics (2011).

[29] Menoukeu-Pamen, O., Meyer-Brandis, T., Nilssen, T., Proske, F., Zhang, T.: Mathematische Annalen, 357(2), pp. 761-799 (2013).

[30] Meyer-Brandis, T., Proske, F.: Explicit solution to a non-linear filtering problem for Lévy processes with applications to finance. Applied Mathematics and Optimization, Vol. 50, No. 2, pp. 119-134 (2004).

[31] Meyer-Brandis, T., Proske, F.: Construction of strong solutions to SDE's via Malliavin calculus. Journal of Functional Analysis, Vol. 258, No. 11, 3922-3953 (2009).

[32] Milidonis, A., Lin, Y., Cox, S.: Mortality regimes and pricing. North American Actuarial Journal, Vol. 15, No. 2, pp. 266-289 (2011).

[33] Pardoux, E.: Stochastic differential equations and filtering of diffusion processes. Stochastics 3, 127-167 (1979).

[34] Pardoux, E.: Filtrage non lineaire et equations aux derivees partielle stochastiques associees. In: D:L: Burkholder, E. Pardoux and Snitzman (eds). École d' Été de Probabilité de Saint-Flour XIX-1989. Lecture Notes in Math. 1464, Springer-Verlag, New York, pp. 67-163 (1991).

[35] Renshaw, A.E., Haberman, S.: A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics, 38, 556-570 (2006).

[36] Sato, K.: Lévy Processes and Infinitely Divisible Distributions, Cambridge University Studies in Advanced Mathematics, Vol. 68, Cambridge University Press, Cambridge 1999.

[37] Wang, C.-W., Huang, H.-C., Liu, I.-C.: Mortality modeling with non-Gaussian innovations and applications to the valuation of longevity swaps. Journal of Risk and Insurance, Vol. 80, No. 3 (2013).

[38] Xiong, J.: An Introduction to Stochastic Filtering Theory. Oxford University Press (2008).

[39] Zakai, M.: On the optimal filtering of diffusion processes. Z. Wahrsch. verw. Geb., No. 11, 230-243, (1969).

[40] Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. Mat. Sb. (N.S.), Vol. 93(135), No. 1, 129-149 (1974).