*J. Appl. Probab.* **60,** 528–541 (2023)
*doi:*10.1017/jpr.2022.96

# ON THE CUMULANT TRANSFORMS FOR HAWKES PROCESSES

YOUNG LEE,* *Harvard University*
THORSTEN RHEINLÄNDER,** *Technische Universität Wien*

### Abstract

We consider the asset price as the weak solution to a stochastic differential equation driven by both a Brownian motion and the counting process martingale whose predictable compensator follows shot-noise and Hawkes processes. In this framework, we discuss the Esscher martingale measure where the conditions for its existence are detailed. This generalizes certain relationships not yet encountered in the literature.

*Keywords:* Hawkes process; stochastic logarithm; exponentially special semimartingale

2020 Mathematics Subject Classification: Primary 60G20
Secondary 60G55; 60E10

## 1. Introduction

We initiate the computation of the Esscher martingale measure for a class of general semimartingales whose predictable compensators have shot-noise and Hawkes process flavors.

Many different models to describe the evolution of financial assets have been proposed in the literature. These include the exponential Lévy processes (Merton, 1976; Carr and Geman, 2002) with their complements of stochastic volatility counterparts (cf. e.g. Hull and White, 1987; Barndorff-Nielsen and Shephard, 2001). One approach to modelling the stock price process using a class of pure-jump stochastic processes has been launched by Eberlein and Jacod (1997), with favourable arguments addressing the empirical realities and their implications of the no-arbitrage principle (Madan *et al.*, 1998). The price process we consider is capable of capturing contagion and jump clustering, a phenomenon that is empirically relevant when financial markets are in distress. Under these circumstances, stock price crashes occur more frequently than predicted by standard stochastic volatility models (Aït-Sahalia *et al.*, 2015). In particular, asset price crashes tend to cluster over short time spans, and standard models are unable to replicate this pattern of crash clustering. Therefore, related Hawkes models have been proposed to better replicate the empirical patterns found in asset returns and lead to an improved fit for option pricing models (Boswijk and Lalu, 2016).

In particular, we study the Esscher martingale measure when the asset price follows a stochastic differential equation driven by both a Brownian motion and the counting process martingale whose predictable compensator follows shot-noise and Hawkes processes (Brémaud and Massoulié, 2002; Dassios and Zhao, 2011; Boumezoued, 2016). Such models typically generate markets which are *incomplete*, wherein there exist infinitely many martingale measures which are equivalent to the physical measure describing the evolution

Received 7 April 2021; revision received 11 August 2022.
* Postal address: Harvard University, 1 Oxford Street, Cambridge, MA 02138, USA.
** Postal address: Wiedner Hauptstr. 8/E105-1 1040 Vienna, Austria.

of the underlying stock price process. Each equivalent martingale measure corresponds to a set of derivative prices respecting the no-arbitrage principle. The Esscher martingale measure for Lévy processes and stochastic volatility models is well studied (Hubalek and Sgarra, 2009). On another spectrum, Dassios and Zhao (2012) discussed the Esscher transform for a generalized contagion model for the purpose of importance sampling and ruin probability computations. The authors showed that under the Esscher transform, ruin becomes certain, and this facilitates simulation more efficiently than under the original measure, where the ruin is not certain and is even rare. Similar ideas for improving the simulation of rare events in the spirit of measure changes can be found in Asmussen (1985).

However, the Esscher martingale measure for a new class of *compound* shot-noise and Hawkes point processes has not been examined. Augmented with a Brownian component in a stochastic logarithm framework, we formulate the asset price as a general semimartingale whose drift is modulated by a one-layer neural network (Schmidhuber, 2015). This model may introduce flexibility and extend the standard log-Brownian paradigm behaviour, capturing unusual business cycles or other economic benchmarks. We determine the Esscher martingale measure in concert with the procedures outlined in Kallsen and Shiryaev (2002).

Section 2 summarizes the basic definitions and main results concerning that of shot-noise and Hawkes processes. In Section 3, we present our asset model as well as the basic properties of cumulant processes. In Section 4 we explicate our martingale approach to uncover the Esscher martingale measure. The conditions for absence of arbitrage and existence of solutions are discussed. Finally, the no-arbitrage valuation of derivatives is detailed in Section 5. We give some concluding remarks in Section 6. In the present text, we generally use the notation of Jacod and Shiryaev (2003), where $\mu_N$ and $\nu_N(dx, dt)$ respectively denote the jump measure associated to $N$ and its compensator.

## 2. Preliminaries and point processes

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space. We fix a finite time horizon $T$, and all stochastic processes are defined on $[0, T]$. Let $N$ be a *simple* point process on $\mathbb{R}$, that is, a family $\{N(W)\}_{W\in\mathcal{B}(\mathbb{R})}$ of random variables with values in $\{0, 1, 2, \ldots\} \cup \{\infty\}$ indexed by the Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R})$ of the real line $\mathbb{R}$, where $N(W) = \sum_{n\in\mathbb{Z}} \mathbb{1}_W(\tau_n)$ and $(\tau_n)_{n\in\mathbb{Z}}$ is a sequence of extended real-valued random variables such that almost surely (a.s.) $\tau_0 \le 0 < \tau_1$ and $\tau_n < \tau_{n+1}$ on $\{\tau_n < \infty\} \cap \{\tau_{n+1} > -\infty\}$ for every $n \in \mathbb{Z}$. We further assume that it is *nonexplosive*, wherein $\tau_n \to \infty$ a.s. as $n \to \infty$ (see p. 47 of Daley and Vere-Jones, 2003 and p. 8 of Last and Brandt, 1995, respectively). Let $\mathcal{H}_t = \sigma(N(W),\ W \in \mathcal{B}(\mathbb{R}),\ W \subset (-\infty, t))$. We also assume that our probability space supports a Brownian motion $B$, and denote its $\mathbb{P}$-augmented natural filtration by $(\mathcal{F}_t)$. Let $(\mathcal{G}_t) = (\mathcal{H}_t) \vee (\mathcal{F}_t)$. We assume the following martingale invariance property: all $(\mathcal{F}_t)$- and $(\mathcal{H}_t)$-martingales remain martingales in the larger filtration $(\mathcal{G}_t)$.

The process $\lambda_t$ is called the $\mathcal{G}_t$-conditional intensity of $N$ if for all intervals $(s,t]$, we have

$$\mathbb{E}[N((s, t]) \mid \mathcal{G}_s] = \mathbb{E}\left[\int_s^t \lambda_u\, du \,\middle|\, \mathcal{G}_s\right], \qquad \text{a.s.} \tag{2.1}$$

We use the notation $N_t := N(0, t]$ to denote the number of points in the interval $(0, t]$. The linear Hawkes process with *levels of excitation* $h : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a simple point process $N$ admitting the $\mathcal{G}_t$-intensity

$$\lambda_t = \bar{\lambda}_t + \int_{-\infty}^{t-} h(t-u)dN_u + \int_{-\infty}^{t-} \check{h}(t-\check{u})d\check{N}_{\check{u}}, \tag{2.2}$$

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

where $\bar{\lambda}_t > 0$, $h$ (resp. $\breve{h}$) is such that $h(\cdot): \mathbb{R}^+ \to \mathbb{R}^+$, and we always assume that $\|h\|_{L^1} = \int_0^\infty h(t)dt < \infty$, which ensures stationarity in the point process $N$ (Daley and Vere-Jones, 2003). $\breve{N}$ is an inhomogeneous Poisson process with non-random and bounded intensity $\sigma$. The sequence of event times $u$ gives the jump times of $N$, and this models the Hawkes or self-exciting property; they are endogenous by construction. By the same token, exogenous events occur at times $\breve{u}$; this is known as the *shot-noise* or commonly referred to as the 'Cox component' (Cox and Isham, 1980). Furthermore, the point process martingale $\widetilde{N}$ associated with the Hawkes process is given by

$$\widetilde{N}_\cdot := N_\cdot - \int_0^\cdot \lambda_u \, du. \tag{2.3}$$

**Characterizations of $h$ and $\breve{h}$.** One such function for $h$ (and that of $\breve{h}$) takes the following form: $h(t) = \alpha e^{-\delta t}$, $t \in [0, T]$, where $\alpha > 0$ and $\delta > 0$. Furthermore, we assume that $\delta > \alpha$ (resp. $\breve{\alpha}$), thus guaranteeing stationarity of $N$ (Brémaud and Massoulié, 2002; Dassios and Zhao, 2011). In the case of

$$\bar{\lambda}_t = \lambda^\flat + (\lambda_0 - \lambda^\flat)e^{-\delta t}, \tag{2.4}$$

where $\lambda_0 > 0$ is the initial intensity at time $t = 0$, and $\lambda^\flat \geq 0$ is a constant with $\lambda^\flat < \lambda_0$, we remark that $\lambda_t$ is Markovian and the pair $(N_t, \lambda_t)_{t \in [0,T]}$ forms a Markov process (Oakes, 1975). Another widely used function is the so-called power decay function, which is particularly used in the field of seismology; it takes the form $h(t) = \beta(t + \gamma)^{-\eta-1}$, where $\beta \geq 0$, $\gamma > 0$, $\eta > 0$, and $\eta\gamma^\eta > \beta$ for stationary of the point process to hold. When $\eta = 0$, we recover the Omori formula (Omori, 1894). When no constraints are placed on $\eta$, we obtain the modified Omori formula (Utsu *et al.*, 1995). For further discussion of this type of function, we refer the reader to Ogata (1998).

## 3. Structure condition and cumulant processes

We choose the money market account as numéraire, and the interest rate is taken to be null. We assume that the asset price dynamics $\tilde{S}$ under $\mathbb{P}$ is given by

$$\mathbb{P}: \ \frac{d\tilde{S}_t}{\tilde{S}_t} = a_t^{\boldsymbol{\theta}} \, dt + b_t \, dB_t + d(f_t^H(x) * (\mu_N - \nu_N))_t := dS_t, \tag{3.1}$$

where $\mu_N$ denotes the jump measure of $N$ while $\nu_N$ denotes its predictable compensator. In our framework of compound shot-noise and Hawkes point processes, we no end have

$$\nu_N := \nu_N(dx, dt) = \nu(dx)\lambda_t \, dt. \tag{3.2}$$

The deterministic function $a_t^{\boldsymbol{\theta}}$, representing the drift of the stock price process, is formulated to be a flexible family of functions in the form of a one-layer neural network,

$$a_t^{\boldsymbol{\theta}} := a(t, \boldsymbol{\theta}) = g_1 \left( \alpha + \sum_{k=1}^{K} \beta_k \cdot g_2(\gamma_k \cdot t + \pi_k) \right), \tag{3.3}$$

where $g_1(\cdot)$, $g_2(\cdot)$ are bounded activation functions, and $K$ is some number of hidden units (Zell, 1997). The quantity $\boldsymbol{\theta}$ denotes the set of parameters $\{\alpha, (\beta_k, \gamma_k, \pi_k)_{k=1}^K\}$, which modulate the drift of the stock price process. Each term $g_2(\gamma_k \cdot t + \pi_k)$ may be seen as a hidden

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

representation of the input time. By choosing $K$ sufficiently large, we can model complex nonlinear drift of the stock price process. Some choices of activation functions include the sigmoid $g_2(z) = (1 + e^{-z})^{-1}$, the inverse square $g_2(z) = (1 + z^2)^{-1}$, the hyperbolic tangent $g_2(z) = \tanh(z)$, and the bounded rectified linear activation function (or ReLU for short), given by $g_2(z) = \min(\max(0, z), D)$ for some predetermined $D > 0$ that defines the maximal output value that this function could produce (Schmidhuber, 2015). Furthermore, we assume that $b > \epsilon$ and $f^H > -1$ are deterministic and bounded.

To sum up, we work with the following restrictions on the parameters $(a^{\theta}, b, f^H)$.

**Assumption 1.** *The quantities $a_t$, $b_t > \epsilon$ and $f_t^H$ are deterministic, bounded functions with*

$$\int (f_t^H(x))^2 \nu(dx) < \infty. \tag{3.4}$$

We often suppress the time variable in the notation, in accordance with convention. Note that $\tilde{S}$ is the stochastic exponential of $S$, and thus solving the differential equation yields

$$\tilde{S}_t = \mathscr{E}\left(\int_0^{\cdot} a^{\theta}\, ds + b\, dB + d(f^H(x) * (\mu_N - \nu_N))\right)_t, \tag{3.5}$$

where $\mathscr{E}$ is the Doléans-Dade exponential. Our main goal in this manuscript is to calculate as explicitly as possible the Esscher martingale measure in the sense of Kallsen and Shiryaev (2002) for the stochastic logarithm $S = \int d\tilde{S}/\tilde{S}_-$, with probability measures $\mathbb{Q}$ such that $S$ is a local $\mathbb{Q}$-martingale. By Theorem III.33 of Protter (2005), the price process $\tilde{S} = \int \tilde{S}_- dS$ is a local $\mathbb{Q}$-martingale as well. Moreover we assume that the asset price process $S$ satisfies the following *structure condition*: the decomposition of $S$ can be uniquely written in the form $S = M + A$, where

$$A = \int \rho\, d\langle M, M\rangle \tag{3.6}$$

and $\rho$ is a predictable process satisfying $K_T := \int_0^T \rho_t^2\, d\langle M, M\rangle_t < \infty$, $\mathbb{P}$-a.s. (Ansel and Stricker, 1992). The notation $\langle M, M\rangle$ denotes the predictable compensator of the quadratic variation process $[M, M]$. In our Hawkes–Cox-driven price process from Equation (3.1), it is readily computed that

$$\rho = \frac{a^{\theta}}{b^2 + \lambda_- \int_{\mathbb{R}} (f^H(x))^2 \nu(dx)} \tag{3.7}$$

as well as that

$$K_T = \int_0^T \frac{(a_t^{\theta})^2}{b_t^2 + \lambda_t \int_{\mathbb{R}} (f_t^H(x))^2 \nu(dx)}\, dt < \infty, \tag{3.8}$$

since both the point processes $N$ and $\check{N}$ are nonexplosive and $\lambda_t$ is finite for $t \in [0, T]$. Hence $S$ is a special semimartingale (see p. 209 in He *et al.* (1998)).

One potential approach for the evaluation of a pricing measure is related to the Esscher martingale measure (Madan and Milne, 1991; Gerber and Shiu, 1994; Eberlein and Keller, 1995; Chan, 1999). In this section, we recall some definitions from Kallsen and Shiryaev (2002), which lays out an extended abstract theory for Esscher martingale transforms for general semimartingales.

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

We remark that if $X$ is a semimartingale, then $L(X)$ denotes the set of predictable $X$-integrable processes. Recall that a semimartingale $X$ is called *special* if it can be decomposed as $X = X_0 + M + A$ for some local martingale $M$ and some predictable process $A$ of finite variation, null at 0. We denote by $\mathcal{M}_{loc}$ the class of all local martingales.

**Definition 1.** Let $X$ be a real-valued semimartingale. $X$ is called *exponentially special* if $\exp(X - X_0)$ is a special semimartingale. A predictable process $A$ is called the exponential compensator of $X$ if $\exp(X - X_0 - A) \in \mathcal{M}_{loc}$.

**Definition 2.** The modified Laplace cumulant process $K^X(\vartheta)$ of $X$ in $\vartheta$ is defined to be the exponential compensator of $\int \vartheta\, dX$.

**Lemma 1.** *Let $\vartheta \in L(S)$ be such that $\int \vartheta\, dS$ is exponentially special. Then the modified Laplace cumulant process of $S$ in $\vartheta$ is given by*

$$K^S(\vartheta)_t = \int_0^t \tilde{\kappa}(\vartheta)_s\, ds, \tag{3.9}$$

*where*

$$\tilde{\kappa}(\vartheta)_t = \left( a_t^{\boldsymbol{\theta}} - \lambda_t \int_{\mathbb{R}} f_t^H(x)\nu(dx) \right) \vartheta_t + \frac{1}{2} b_t^2 \vartheta_t^2 + \left(e^{\int_{\mathbb{R}} f_t^H(x)\nu(dx)\vartheta_t} - 1\right)\lambda_t. \tag{3.10}$$

*Proof.* This follows immediately from Theorem 2.18 of Kallsen and Shiryaev (2002) following the semimartingale characteristics. □

The derivative of the cumulant process $DK^S(\vartheta) = \int D\tilde{\kappa}^S(\vartheta)\lambda dt$ in the sense of Definition 2.22 in Kallsen and Shiryaev (2002) is given by

$$D\tilde{\kappa}^S(\vartheta) = a^{\boldsymbol{\theta}} - \lambda \int_{\mathbb{R}} f^H(x)\nu(dx) + b^2\vartheta + f^H(x)\lambda e^{\vartheta f^H(x)}. \tag{3.11}$$

Assume that there exists a solution $\vartheta^\sharp$ to the martingale equation

$$DK^S(\vartheta) = 0. \tag{3.12}$$

In the case when

$$G_t^\sharp := \exp\left( \int_{0+}^t \vartheta_s^\sharp dS_s - K^S(\vartheta^\sharp)_t \right) \tag{3.13}$$

is a martingale, we can then define a probability measure $\mathbb{Q}^\sharp$ by

$$\frac{d\mathbb{Q}^\sharp}{d\mathbb{P}} = \exp\left( \int_{0+}^T \vartheta_t^\sharp dS_t - K^S(\vartheta^\sharp)_T \right). \tag{3.14}$$

The density process $(G_t^\sharp)_{t\in[0,T]}$ now defines a unique equivalent martingale measure $\mathbb{Q}^\sharp$ for $S$, known as the *Esscher martingale measure* for the process $S$ (Theorems 4.4 and 4.5 in Kallsen and Shiryaev, 2002).

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

## 4. Esscher martingale measure

In this section, we spell out the necessary and sufficient conditions for $(G_t^\sharp)_{t\in[0,T]}$ to be an Esscher martingale measure.

**Lemma 2.** *Define the following functions:*

$$m_1(\vartheta)=a^{\boldsymbol{\theta}},\quad m_2(\vartheta)=\lambda\int_{\mathbb{R}} f^H(x)\nu(dx)-b^2\vartheta-\lambda\int_{\mathbb{R}} f^H(x)e^{\vartheta f^H(x)}\nu(dx). \tag{4.1}$$

*Then there exists a bounded function* $\vartheta_t^\sharp:[0,T]\to\mathbb{R}$ *solving* $m_1(\vartheta_t^\sharp)=m_2(\vartheta_t^\sharp)$ *for all* $t\in[0,T]$ *and hence Equation* (3.12).

*Proof.* Observe that for a fixed $t\in[0,T]$, $m_2$ is decreasing in $\vartheta$, since

$$m_2'(\vartheta)=-b^2-\lambda\int_{\mathbb{R}}(f^H(x))^2e^{\vartheta f^H(x)}\nu(dx)<0, \tag{4.2}$$

using the fact that $\lambda>0$ is well defined for each $t\in[0,T]$ and $b$ is bounded. Also, note that $m_2(-\infty)=\infty$, $m_2(\infty)=-\infty$. Moreover, $m_1$ is bounded since $a^{\boldsymbol{\theta}}$ is bounded. Hence we conclude by the mean value theorem that there exists for all $t\in[0,T]$ a number $\vartheta_t^\sharp$ such that $m_1(\vartheta_t^\sharp)=m_2(\vartheta_t^\sharp)$. Note that $(m_2)^{-1}(\sup a^{\boldsymbol{\theta}})\le\vartheta^\sharp\le(m_2)^{-1}(\inf a^{\boldsymbol{\theta}})$, where $(m_2)^{-1}$ is the inverse function of $m_2$; hence $t\mapsto\vartheta_t^\sharp$ is a bounded function on $[0,T]$. □

We now turn our attention to proving that the process $G^\sharp$ is a martingale. For this we will need the following criterion, due to the criterion of Theorem 9 in Protter and Shimbo (2008).

**Proposition 1.** *Let* $\widetilde{M}$ *be a locally square-integrable martingale such that* $\Delta\widetilde{M}>-1$. *If*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\langle\widetilde{M}^c,\widetilde{M}^c\rangle_T+\langle\widetilde{M}^d,\widetilde{M}^d\rangle_T\right)\right]<\infty, \tag{4.3}$$

*where* $\widetilde{M}^c$ *and* $\widetilde{M}^d$ *are the continuous and purely discontinuous martingale parts of* $\widetilde{M}$, *then the Doléans-Dade* $\mathscr{E}(\widetilde{M})$ *is a strictly positive martingale on* $[0,T]$.

**Lemma 3.** *Define the following processes:*

$$G^\sharp=\mathscr{E}(R^\sharp) \tag{4.4}$$

*with*

$$R^\sharp=\int b\vartheta^\sharp\,dB+\left(e^{\vartheta^\sharp(e^{f^H(x)}-1)}-1\right)*(\mu_N-\nu_N). \tag{4.5}$$

*Then* $G^\sharp$ *is a martingale.*

*Proof.* We proceed to evaluate the following quantity:

$$\frac{1}{2}\langle R^\sharp,R^\sharp\rangle_T^c+\langle R^\sharp,R^\sharp\rangle_T^d=\frac{1}{2}\int_0^T b_t^2(\vartheta_t^\sharp)^2\,dt+\int_0^T\left(\int_{\mathbb{R}}e^{\vartheta_t^\sharp(e^{f_t^H(x)}-1)}\nu(dx)-1\right)^2\lambda_t\,dt<\infty$$

with probability 1, by the boundedness of the coefficients and that of $\vartheta^\sharp$, and using the fact that our point processes are nonexplosive. It follows that $\mathbb{E}[\exp(\frac{1}{2}\langle N,N\rangle_T^c+\langle N,N\rangle_T^d)]<\infty$. Hence the assertion follows from Proposition 1. □

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

Following the existence of solutions with the martingale property ensured, we state the main result with regard to the Esscher measure.

**Theorem 2.** *Let the conditions of Assumption 1 hold. Then there exists a function* $\vartheta^\sharp : [0, T] \to \mathbb{R}$ *with* $\vartheta^\sharp_t := \vartheta^\sharp(t)$ *which solves the martingale equation (3.12) for any* $t \in [0, T]$*, i.e.*

$$a^{\boldsymbol{\theta}} - \lambda \int_{\mathbb{R}} f^H(x)\nu(dx) + b^2\vartheta + f^H(x)\lambda e^{\vartheta f^H(x)} = 0.$$

*Furthermore, the derivative*

$$\frac{d\mathbb{Q}^\sharp}{d\mathbb{P}} = \mathscr{E}\left(\int_0^\cdot \vartheta_t b_t\, dB_t + \left(e^{\vartheta^\sharp_t(e^{f^H_t(x)}-1)} - 1\right) * (\mu_N - \nu_N)\right)_T \tag{4.6}$$

*defines a probability measure* $\mathbb{Q}^\sharp \sim \mathbb{P}$ *on* $\mathcal{G}_T$*, which is called the Esscher martingale measure for S. Moreover, under* $\mathbb{Q}^\sharp$*, we have that*

$$B^{\mathbb{Q}^\sharp}_t := B_t - \int_0^t \vartheta^\sharp_s b_s ds \tag{4.7}$$

*is a* $\mathbb{Q}^\sharp$*-Brownian motion, and the predictable compensator* $\nu^\sharp_N$ *under* $\mathbb{Q}^\sharp$ *is given by*

$$\nu^\sharp_N := \nu^\sharp_N(dx, dt) = e^{\vartheta^\sharp_t(e^{f^H_t(x)}-1)}\nu(dx)\lambda_t\, dt. \tag{4.8}$$

*Proof.* From Lemma 2, there exists a function $\vartheta^\sharp$ solving Equation (3.12). To complete the proof, we apply Theorem 4.4 in Kallsen and Shiryaev (2002) to conclude that the density in Equation (4.6) defines an equivalent local martingale measure for $\mathscr{E}(S)$. By Lemma 3, we have that $G^\sharp$ is a proper martingale and thus defines a density process. They dynamics under $\mathbb{Q}^\sharp$ follows from Girsanov's theorem (cf. Theorem III.40 in Protter (2005)). □

**Remark 1.** (1) *The role of* $\vartheta^\sharp$ and $a^{\boldsymbol{\theta}}$. For the Esscher martingale measure to exist, the choice of $\boldsymbol{\theta}$ and the concrete specification of the drift $a^{\boldsymbol{\theta}}$ must be such that there exists a solution $\vartheta^\sharp$ that solves the martingale equation (3.12); otherwise we say that the Esscher martingale measure does not exist.

(2) *Family of drifts and existence of* $\vartheta^\sharp$. It is well known that the dynamics of asset returns cannot be adequately described with constant drift and volatility (Mandelbrot and Taylor, 1967; Clark, 1973). There have been many attempts to extend the model by describing the evolution of drift and volatility beyond constants (cf. e.g. Masi *et al.*, 1995, and references therein).

(3) *Special cases.* In our case, the formulation of the flexible family of drifts in Equation (3.1) is an interesting one provided that $\vartheta^\sharp$ exists. This *one-layer* neural network drift naturally subsumes the standard constant drift as a special case if we let $\beta_k = \gamma_k = \pi_k = 0$. Second, for identity activation of $g_1(\cdot)$ and $g_2(\cdot)$ a suitable indicator function, we get an $a^{\boldsymbol{\theta}}_t$ that is piecewise constant on intervals and could be utilized as part of modelling a regime switching; see for example Çetin and Verschuere (2009) and Elliott and Siu (2010).

**Example** (**Hawkes with neural network drift**). Suppose the stock price process takes the form

$$\mathbb{P}: \frac{d\tilde{S}_t}{\tilde{S}_t} = a^{\boldsymbol{\theta}}_t\, dt + b\ dB_t + d(f^H(x) * (\mu_N - \nu_N))_t, \qquad t \in [0, T], \tag{4.9}$$

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

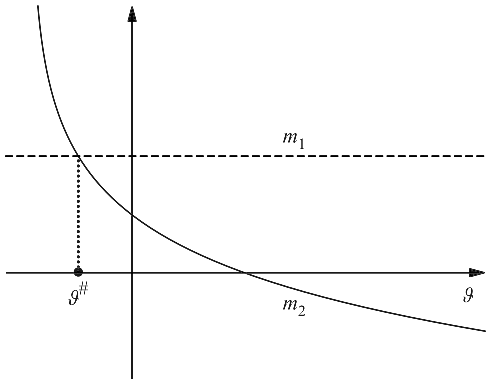


FIGURE 1. An illustration of the existence of $\vartheta^{\sharp}$ solving Equation (4.11), where $m_1$ and $m_2$ are given in Equations (4.12) and (4.13), respectively.

where $f$ is a bounded function independent of time $t$ and $b$ is a constant. We set $a_t^{\boldsymbol{\theta}}$ to be

$$a_t^{\boldsymbol{\theta}} = \alpha + \frac{\beta_1}{1 + e^{-(\gamma_1 t + \pi_1)}} + \frac{\beta_2}{1 + e^{-(\gamma_2 t + \pi_2)}}, \quad t \in [0, T], \tag{4.10}$$

where $\boldsymbol{\theta} = (\alpha, \beta_1, \beta_2, \gamma_1, \gamma_2, \pi_1, \pi_2)$, all of which are taken to be positive constants. We remark that Assumption 1 is satisfied for these choices of $a^{\boldsymbol{\theta}}$, $b$, and $f$. We set $g_1(\cdot)$ and $g_2(\cdot)$ to be the identity and the sigmoid activation functions, respectively, while setting the number of hidden units to $K = 2$. We first check for existence of a solution $\vartheta^{\sharp}$ to the martingale equation (3.12), which in this example returns

$$\alpha + \frac{\beta_1}{1 + e^{-(\gamma_1 t + \pi_1)}} + \frac{\beta_2}{1 + e^{-(\gamma_2 t + \pi_2)}} - \lambda \int_{\mathbb{R}} f^H(x)\nu(dx) + b^2\vartheta + f^H(x)\lambda e^{\vartheta f^H(x)} = 0. \tag{4.11}$$

Define the quantities

$$m_1 = \alpha + \frac{\beta_1}{1 + e^{-(\gamma_1 t + \pi_1)}} + \frac{\beta_2}{1 + e^{-(\gamma_2 t + \pi_2)}}, \tag{4.12}$$

$$m_2 = \lambda \int_{\mathbb{R}} f^H(x)\nu(dx) - b^2\vartheta - f^H(x)\lambda e^{\vartheta f^H(x)}. \tag{4.13}$$

Invoke Lemma 2 to ensure the existence of a solution $\vartheta^{\sharp}$ to Equation (4.11). An illustration is given in Figure 1.

From Theorem 2, we see that the process

$$B_t - \int_0^t \sigma_s^{\sharp} ds \tag{4.14}$$

is a $\mathbb{Q}^{\sharp}$-Brownian motion, where we have defined

$$\sigma^{\sharp} := \vartheta^{\sharp} b. \tag{4.15}$$

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

Furthermore, the quantity

$$\nu_N^{\sharp}(dx, dt) = e^{\vartheta_t^{\sharp}(e^{f^H(x)}-1)}\nu(dx)\lambda_t \, dt \tag{4.16}$$

is the predictable compensator $\nu_N^{\sharp}$ of $\mu_N$ under $\mathbb{Q}^{\sharp}$. The process $\sigma^{\sharp}$ represents the risk premium associated with the continuous martingale part $B$, and the predictable function $e^{\vartheta_t^{\sharp}(e^{f^H(x)}-1)} - 1$ is interpreted as the risk premium associated with the jumps of the discontinuous part of $\mu_N$ (Theorem III.3.23 in Jacod and Shiryaev, 2003), where $\vartheta^{\sharp}$ solves Equation (4.11).

## 5. Valuation of contingent claims

Having computed the Esscher martingale measure, we now detail the optimal hedging strategy as well as the differential equations for the prices of derivatives under this measure.

### 5.1. Mean-variance hedging

One criterion for hedging is mean-variance hedging (Bouleau and Lamberton, 1989). Consider a contingent claim with maturity $T > 0$, defined by a $\mathcal{G}_T$-measurable random variable $H$. Define the initial capital $V_0$ and a trading strategy which will be defined by a portfolio $(\psi_t^0, \psi_t)_{t\in[0,T]}$ taking values in $\mathbb{R}^2$. As before, we choose the savings account as numéraire and work with a zero interest rate. In mean-variance hedging, we look for a self-financing trading strategy given by an initial capital $V_0$ and a portfolio over the lifetime of the contingent claim which minimizes the shortfall at the terminal date $T$ in a mean square sense:

$$\inf_{\psi, V_0} \mathbb{E}_{\mathbb{Q}^{\sharp}} |H - V_T|^2, \quad V_T = V_0 + \int_0^T \psi \, d\tilde{S}_t$$

where

$$H - V_T = H - V_0 - \int_0^T \psi_t d\tilde{S}_t.$$

In particular, we have chosen a pricing rule given by the Esscher martingale measure $\mathbb{Q}^{\sharp}$. Let $H \in \mathcal{L}^2(\Omega, \mathbb{G}, \mathbb{Q}^{\sharp})$; $H$ has finite variance and $(\tilde{S}_t)_{t\in[0,T]}$ is a square-integrable $\mathbb{Q}^{\sharp}$-martingale. Let us consider those $(\psi_t)_{t\in[0,T]}$ whose terminal values satisfy

$$\mathcal{X} := \left\{ \psi : \mathbb{E}_{\mathbb{Q}^{\sharp}}\left( \int_0^T \psi_t^2 d[\tilde{S}, \tilde{S}]_t \right) < \infty \right\}. \tag{5.1}$$

Define $\mathcal{L}^2(\tilde{S})$ as the set of portfolios $\psi$ satisfying (5.1). Since $\psi \in \mathcal{L}^2(\tilde{S})$, and using the fact that $\tilde{S}$ is a martingale under $\mathbb{Q}^{\sharp}$, the gains process $G(\psi) = \int_0^{\cdot} \psi \, d\tilde{S}$ is also a square-integrable martingale. The mean-variance hedging problem can now be recast as

$$\inf_{\psi \in L^2(\tilde{S}), V_0} \mathbb{E}_{\mathbb{Q}^{\sharp}} \left| H - V_0 - \int_0^T \psi_t d\tilde{S}_t \right|^2,$$

or equivalently

$$\inf_{\psi \in L^2(\tilde{S}), V_0} \mathbb{E}_{\mathbb{Q}^{\sharp}} \left| H - V_0 - G_T(\psi) \right|^2.$$

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

The condition (5.1) implies that the value process $V$ is a square-integrable martingale, and we have $\mathbb{E}_{\mathbb{Q}^\sharp}[V_T] = V_0$. Applying the identity $\mathbb{E}(Z^2) = (\mathbb{E}(Z))^2 + \mathbb{V}(Z)$ to the random variable $Z := H - V_0 - G_T(\psi)$, we obtain

$$\begin{aligned} J_0(V_0, \psi) &:= \mathbb{E}_{\mathbb{Q}^\sharp} \left| H - V_0 - \int_0^T \psi_t d\tilde{S}_t \right|^2 \\ &= |\mathbb{E}_{\mathbb{Q}^\sharp}[H] - V_0|^2 + \mathbb{V}(H - V_0 - G_T(\psi)). \end{aligned}$$

If the writer of the contingent claim tries to minimize the residual risk $J_0(V_0, \psi)$, the optimal value that he/she will ask for is a premium

$$V_0 = \mathbb{E}_{\mathbb{Q}^\sharp}[H].$$

We see that $\mathbb{E}_{\mathbb{Q}^\sharp}[H]$ is the initial value of any strategy $\psi \in \mathcal{L}^2(\tilde{S})$ designed to minimize the shortfall at maturity, and we take this as the definition of the *price* associated with our contingent claim $H$ at time 0. By the same token, if the writer sells the option at time $t > 0$ and intends to minimize the remaining risk $J_t(V_t, \psi) := \mathbb{E}_{\mathbb{Q}^\sharp}\left( \left| H - V_0 - \int_0^T \psi_t d\tilde{S}_t \right|^2 | \mathcal{G}_t \right)$, he/she will ask a premium $V_t = \mathbb{E}_{\mathbb{Q}^\sharp}(H \mid \mathcal{G}_t)$. We will take this quantity to define the price of the contingent claim $H$ at time $t$ under the Esscher measure $\mathbb{Q}^\sharp$.

For illustrative purposes, in the framework of pricing and hedging derivatives, we work with the asset price process in Equation (3.1), setting $a_t^{\boldsymbol{\theta}}$ and $b_t$ to be constants $a > 0$ and $b > 0$, respectively, and setting $f_t^H(x) = f^H(x)$ to be a bounded function independent of time. Furthermore, we assume that $h(t) = \alpha e^{-\delta t}$ and, without loss of generality, that $\check{h} = 0$ in Equation (2.2). With this, we see that $(N, \lambda)$ is jointly Markov and $\lambda_t$ can be recorded with an alternative expression

$$\lambda = \int -\delta(\lambda - \lambda^\flat)dt + \int \alpha dN \tag{5.2}$$

to Equation (2.2) when $\bar{\lambda}$ takes the form in Equation (2.4). From Theorem 2, the Esscher martingale measure $\mathbb{Q}^\sharp$ exists and the dynamics of $\tilde{S}$ takes the form

$$\mathbb{Q}^\sharp: \qquad \frac{d\tilde{S}}{S_-} = b\, dB^{\mathbb{Q}^\sharp} + \int f^H(x)(\mu_N - \nu_N^\sharp), \tag{5.3}$$

where the quantities $\mathbb{Q}^\sharp$ and $\nu^\sharp$ are defined in Theorem 2, and $\vartheta^\sharp$ solves Equation (3.12).

5.1.1. *Pricing.* Consider a contingent claim of European type with maturity $T$ and payoff $H(\tilde{S}_T)$. The payoff $H$ is assumed to satisfy

$$|H(z_1) - H(z_2)| \leq k|z_1 - z_2| \tag{5.4}$$

for some $k > 0$. For put and call options, this assumption is satisfied with the choice of $k = 1$. The value $C_t$ of such a derivative claim is given by $C_t = C(t, \tilde{S}_t, \lambda_t)$, where

$$C(t, s, \lambda) = e^{-r(T-t)}E[H(\tilde{S}_T) \mid \tilde{S}_t = s, \lambda_t = \lambda].$$

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

Let $H$ be a contingent claim with terminal payoff $H = H(\tilde{S}_T)$. Then the price of $H$ at time $t$ is given by $C(t, s, \lambda)$ where

$$\begin{aligned} C : [0, T] \times [0, \infty) \times (0, \infty) &\rightarrow \mathbb{R}, \\ (t, s, \lambda) &\mapsto C(t, s, \lambda) = \mathbb{E}_{\mathbb{Q}^\sharp}\left[H(\tilde{S}_T) \mid \tilde{S}_t = s, \lambda_t = \lambda\right]. \end{aligned}$$

Assume that $C \in \mathcal{C}^{1,2}$, i.e., the functions $(t, s, \lambda) \mapsto C(t, \tilde{S}, \lambda)$ are continuously differentiable with respect to $t$ and twice continuously differentiable with respect to $s$. Hence the Itô formula can be applied to $C_t = C(t, \tilde{S}_t, \lambda_t)$ between 0 and $T$. Define

$$\widetilde{C}_t^{\mathbb{Q}^\sharp,*} := \sup_{t\in[0,T]} |C_t| = \sup_{t\in[0,T]} C_t,$$

since $H \geq 0$. Applying the Doob inequality to the martingale $C$ yields

$$\mathbb{E}_{\mathbb{Q}^\sharp}(\widetilde{C}_t^{\mathbb{Q}^\sharp,*})^2 \leq 4\mathbb{E}_{\mathbb{Q}^\sharp} H^2 < \infty$$

since $H$ has finite variance. By the Burkholder–Davis–Gundy inequality, there exists a constant $c > 0$ such that

$$c\mathbb{E}_{\mathbb{Q}^\sharp}\{[C, C]_T\} \leq \mathbb{E}_{\mathbb{Q}^\sharp}(\widetilde{C}_T^{\mathbb{Q}^\sharp,*})^2 < \infty.$$

Hence we obtain

$$\mathbb{E}_{\mathbb{Q}^\sharp}\{[C, C]_T\} < \infty.$$

This implies that $C$ is a square-integrable $\mathbb{Q}^\sharp$-martingale. With the expression of (5.2) and by Corollary I-3.16 of Jacod and Shiryaev (2003), we conclude that the finite-variation term vanishes, giving us the PDE

$$\begin{aligned} &\frac{\partial C}{\partial t} - s\lambda \int_{\mathbb{R}} f^H(x) e^{\vartheta^\sharp(e^{f^H(x)}-1)} \nu(dx) \frac{\partial C}{\partial s} + \frac{1}{2} s^2 b^2 \frac{\partial^2 C}{\partial s^2} - \delta(\lambda - a)\frac{\partial C}{\partial \lambda} \\ &\quad + \lambda \int_{\mathbb{R}} (C(t, s(1 + f^H(x)), \lambda + \alpha) - C(t, s, \lambda)) e^{\vartheta^\sharp(e^{f^H(x)}-1)} \nu(dx) = 0 \end{aligned} \tag{5.5}$$

with terminal condition

$$C(T, s, \lambda) = H(s), \qquad \lambda > 0,\ s \geq 0. \tag{5.6}$$

Having derived the Esscher price for European-type contracts, we now characterize the optimal hedging strategy.

5.1.2. *Optimal hedging strategy.* Recall that under $\mathbb{Q}^\sharp$, the stock price process $\tilde{S}$ is a square-integrable martingale. Consider now a self-financing trading strategy with $\psi \in L^2(\tilde{S})$; the value process $V$ of the portfolio is also a martingale whose value at maturity $T > 0$ is

$$\mathbb{Q}^\sharp: \qquad V_T = V_0 + \int_0^T \psi_t d\tilde{S}_t = V_0 + G_T(\psi).$$

Also, $C(T, \tilde{S}_T, \lambda_T) = H(\tilde{S}_T) = H$ and $C(0, \tilde{S}_0, \lambda_0) = \mathbb{E}_{\mathbb{Q}^\sharp}[H(\tilde{S}_T)] = V_0$, so that

$$H - V_0 - G_T(\psi) = C(T, \tilde{S}_T, \lambda_T) - C(0, \tilde{S}_0, \lambda_0) - G_T(\psi).$$

Taking into consideration Equation (5.5), we evaluate the quantity

$$C(T, \tilde{S}_T, \lambda_T) - C(0, \tilde{S}_0, \lambda_0) - G_T(\psi)$$
$$= \int_0^T \left(\psi_t \tilde{S}_{t-} - \tilde{S}_{t-} \frac{\partial}{\partial s} C(t, \tilde{S}_{t-}, \lambda_{t-})\right) b\, dB_t$$
$$+ \int_0^T \int_{\mathbb{R}} \big(\psi_t \tilde{S}_{t-} f^H(x) - (C(t, \tilde{S}_{t-}(1+f^H(x)), \lambda_{t-} + \alpha) - C(t, \tilde{S}_{t-}, \lambda_{t-})\big)(\mu_N - \nu_N^\sharp)_t,$$

so that

$$J_0(V_0, \psi) = \mathbb{E}_{\mathbb{Q}^\sharp}\left[\int_0^T \tilde{S}_{t-}^2 \left(\psi_t - \frac{\partial}{\partial s}(t, \tilde{S}_{t-}, \lambda_{t-})dt\right)^2 b^2 dt\right]$$
$$+ \mathbb{E}_{\mathbb{Q}^\sharp}\left[\int_0^T \int_{\mathbb{R}} (\psi_t \tilde{S}_{t-} f^H(x)\right.$$
$$\left. - (C(t, \tilde{S}_{t-}(1+f^H(x)), \lambda_{t-} + \alpha) - C(t, \tilde{S}_{t-}, \lambda_{t-})^2))\nu_N^\sharp(dx, dt)\right].$$

To obtain the optimal risk-minimizing hedge, we minimize the above expression with respect to $\psi$:

$$\tilde{S}_{t-}^2 \sigma^2 \left(\psi_t - \frac{\partial}{\partial s}(t, \tilde{S}_{t-}, \lambda_{t-})\right) +$$
$$\lambda \int_{\mathbb{R}} (\tilde{S}_{t-} \psi_t f^H(x) - (C(t, \tilde{S}_{t-}(1+f^H(x)), \lambda_{t-} + \alpha) - C(t, \tilde{S}_{t-}, \lambda_{t-}))$$
$$\cdot \tilde{S}_{t-} f^H(x) e^{\vartheta^\sharp (e^{f^H(x)} - 1)} \nu(dx) \stackrel{!}{=} 0.$$

Hence we have that the minimal-risk hedge amounts to holding a position in the underlying $(S_t)_{t>0}$ equal to $\psi_t = \Delta(t, \tilde{S}_{t-}, \lambda_{t-})$, where

$$\Delta(t, s, \lambda)$$
$$= \frac{b^2 \frac{\partial}{\partial s} C(t, s, \lambda) + \frac{\lambda}{s} \int_{\mathbb{R}} (C(t, s(1+f^H(x)), \lambda + \alpha) - C(t, s, \lambda)) e^{\vartheta^\sharp (e^{f^H(x)} - 1)} \nu(dx)}{b^2 + \lambda \int_{\mathbb{R}} (f^H(x))^2 e^{\vartheta^\sharp (e^{f^H(x)} - 1)} \nu(dx)}.$$

A sanity check for convexity yields that

$$\tilde{S}_{t-}^2 b^2 + \lambda_{t-} \int_{\mathbb{R}} \tilde{S}_{t-}^2 (f^H(x))^2 e^{\vartheta^\sharp (e^{f^H(x)} - 1)} \nu(dx) > 0.$$

## 6. Postlude

(1) *Integro-differential equations.* The assumption that the kernel is of exponential form, i.e., $h(t) = \alpha e^{-\delta t}$, in Equation (2.2) is instrumental to the results derived in Section 5, owing to the joint Markovian structure of $N$ and $\lambda$. For general kernels $h$, it is unclear whether explicit

https://doi.org/10.1017/jpr.2022.96 Published online by Cambridge University Press

expressions can be obtained. This is due to the non-Markovian nature of the pair $(N, \lambda)$ induced by such kernel functions (cf. discussions in Boumezoued, 2016, and Gao *et al.*, 2018).

Our derivations and heuristics suggest that the same result should hold for a wide class of parameter specifications satisfying Assumption 1, provided that Lemma 2 is respected, guaranteeing the existence of $\vartheta^\sharp$ to Equation (3.12). In Section 5, we stick to simple specifications for the sake of illustration. Furthermore, since the payoff functions $H$ satisfy the Lipschitz condition in Equation (5.4), these results hold for put and call options, as well as any combinations of these, such as butterfly spreads and strangles, to name a few.

(2) *Other martingale measures.* Given the emergence of new financial markets where an established market for exchange-traded derivatives has yet to emerge, it is often not possible to identify the *statistical martingale measure* (see Chapter 6 in Schoutens, 2003) for the purpose of derivative valuation. The statistical martingale measure is a pricing measure 'chosen by the market' through the minimization of some distance between the observed and theoretical option prices. It may be tricky to compute this measure in the absence of a liquid market with actively traded vanilla options written on the underlying asset.

To overcome this problem, we propose the use of the Esscher martingale measure, which, under general conditions, minimizes the entropy-Hellinger distance. This observation has been documented; see Remark 10 in Hubalek and Sgarra, 2009, which explicates the relationships (cf. Theorem 4.3 in Choulli and Stricker, 2005, and Theorem 4.4 in Kallsen and Shiryaev, 2002). Other martingale measures, such as the minimal entropy martingale measure (Delbaen *et al.*, 2002), are often chosen as well. That said, the identification of this measure is quite involved and goes beyond the scope of this paper, so it is left to future research.

(3) *Open problems.* We have studied the Esscher martingale measure for a general stochastic differential equation augmented by a counting process whose predictable compensator has shot-noise and Hawkes elements. The techniques of proof are fairly general and may be applied to various extensions of the proposed framework. One possibility may be to generalize the Cox element $\breve{N}$ to a Lévy subordinator, which is *a.s.* increasing.


## Acknowledgements

We are thankful to three anonymous referees for their helpful comments on improving the manuscript. This work was conducted while Y. L. was at Harvard University.


## Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

## References


AÏT-SAHALIA, Y., CACHO-DIAZ, J. AND LAEVEN, R. J. A. (2015). Modeling financial contagion using mutually exciting jump processes. *J. Financial Econom.* **117**, 585–606.

ANSEL, J.-P. AND STRICKER, C. (1992). Lois de martingale, densités et décomposition de Föllmer Schweizer. *Ann. Inst. H. Poincaré Prob. Statist.* **28**, 375–392.

ASMUSSEN, S. (1985). Conjugate processes and the simulation of ruin problems. *Stoch. Process. Appl.* **20**, 213–229.

BARNDORFF-NIELSEN, O. AND SHEPHARD, N. (2001). Modelling by Lévy processes for financial econometrics. In *Lévy Processes*, Birkhäuser, Boston, pp. 283–318.

BOSWIJK, H., PETER, R. J. A. L. AND LALU, A. (2016). Asset returns with self-exciting jumps: option pricing and estimation with a continuum of moments. Tech. Rep., University of Amsterdam.

BOULEAU, N. AND LAMBERTON, D. (1989). Residual risks and hedging strategies in Markovian markets. *Stoch. Process. Appl.* **33**, 131–150.
BOUMEZOUED, A. (2016). Population viewpoint on Hawkes processes. *Adv. Appl. Prob.* **48**, 463–480.
BRÉMAUD, P. AND MASSOULIÉ, L. (2002). Power spectra of general shot noises and Hawkes point processes with a random excitation. *Adv. Appl. Prob.* **34**, 205–222.
CARR, P. AND GEMAN, H. (2002). The fine structure of asset returns: an empirical investigation. *J. Business* **75**, 305–332.
CHAN, T. (1999). Pricing contingent claims on stocks driven by Lévy processes. *Ann. Appl. Prob.* **9**, 504–528.
CHOULLI, T. AND STRICKER, C. (2005). Minimal entropy-Hellinger martingale measure in incomplete markets. *Math. Finance* **15**, 465–490.
CLARK, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**, 135–155.
COX, D. R. AND ISHAM, V. (1980). *Point Processes*. Chapman and Hall/CRC, Boca Raton.
DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*, 2nd edn. Springer, New York.
DASSIOS, A. AND ZHAO, H. (2011). A dynamic contagion process. *Adv. Appl. Prob.* **43**, 814–846.
DASSIOS, A. AND ZHAO, H. (2012). Ruin by dynamic contagion claims. *Insurance Math. Econom.* **51**, 93–106.
DELBAEN, F. *et al.*et al. (2002). Exponential hedging and entropic penalties. *Math. Finance* **12**, 99–123.
EBERLEIN, E. AND JACOD, J. (1997). On the range of options prices. *Finance Stoch.* **1**, 131–140.
EBERLEIN, E. AND KELLER, U. (1995). Hyperbolic distributions in finance. *Bernoulli* **1**, 281–299.
ELLIOTT, R. AND SIU, T. (2010). On risk minimizing portfolios under a Markovian regime-switching Black–Scholes economy. *Ann. Operat. Res.* **176**, 271–291.
GAO, X., ZHOU, X. AND ZHU, L. (2018). Transform analysis for Hawkes processes with applications in dark pool trading. *Quant. Finance* **18**, 265–282.
GERBER, H. AND SHIU, E. (1994). Option pricing by Esscher transforms. *Trans. Soc. Actuaries* **46**, 99–140.
HE, S., WANG, J. AND YAN, J. (1998). *Semimartingale Theory and Stochastic Calculus*. CRC Press, Boca Raton.
HUBALEK, F. AND SGARRA, C. (2009). On the Esscher transforms and other equivalent martingale measures for Barndorff-Nielsen and Shephard stochastic volatility models with jumps. *Stoch. Process. Appl.* **119**, 2137–2157.
HULL, J. AND WHITE, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finance* **42**, 281–300.
JACOD, J. AND SHIRYAEV, A. (2003). *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin.
KALLSEN, J. AND SHIRYAEV, A. N. (2002). The cumulant process and Esscher's change of measure. *Finance Stoch.* **6**, 397–428.
LAST, G. AND BRANDT, A. (1995). *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer, New York.
MADAN, D. B., CARR, P. AND CHANG, E. C. (1998). The variance gamma process and option pricing. *Europ. Finance Rev.* **2**, 79–105.
MADAN, D. B. AND MILNE, F. (1991). Option pricing with v. g. martingale components. *Math. Finance* **1**, 39–55.
MANDELBROT, B. AND TAYLOR, H. M. (1967). On the distribution of stock price differences. *Operat. Res.* **15**, 1057–1062.
MASI, G. B. D., KABANOV, Y. AND RUNGGALDIER, W. J. (1995). Mean-variance hedging of options on stocks with Markov volatilities. *Theory Prob. Appl.* **39**, 172–182.
MERTON, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *J. Financial Econom.* **3**, 125–144.
OAKES, D. (1975). The Markovian self-exciting process. *J. Appl. Prob.* **12**, 69–77.
OGATA, Y. (1998). Space-time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50**, 379–402.
OMORI, F. (1894). On the aftershocks of earthquakes. *J. College Sci. Imperial Univ. Tokyo* **7**, 111–120.
PROTTER, P. (2005). *Stochastic Integration and Differential Equations*. Springer, Berlin, Heidelberg.
PROTTER, P. AND SHIMBO, K. (2008). No arbitrage and general semimartingales. In *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz* (Inst. Math. Statist. Collect. **4**), Institute of Mathematical Statistics, Beachwood, OH, pp. 267–283.
SCHMIDHUBER, J. (2015). Deep learning in neural networks: an overview. *Neural Networks* **61**, 85–117.
SCHOUTENS, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley, Chichester.
UTSU, T., OGATA, Y. AND MATSU'URA, R. S. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth* **43**, 1–33.
ZELL, A. (1997). *Simulation neuronaler Netze*. Oldenbourg, Munich.
ÇETIN, U. AND VERSCHUERE, M. (2009). Pricing and hedging in carbon emissions markets. *Internat. J. Theoret. Appl. Finance* **12**, 949–967.