Fard, Farzad Alavi; Doko Tchatoka, Firmin; Srianantha kumar, Sivagowry

**Article**

# Maximum entropy evaluation of asymptotic hedging error under a generalised jump-diffusion model

Journal of Risk and Financial Management

**Provided in Cooperation with:**
MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* Fard, Farzad Alavi; Doko Tchatoka, Firmin; Srianantha kumar, Sivagowry (2021) : Maximum entropy evaluation of asymptotic hedging error under a generalised jump-diffusion model, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 14, Iss. 3, pp. 1-19,
https://doi.org/10.3390/jrfm14030097

This Version is available at:
https://hdl.handle.net/10419/239513

Journal of
*Risk and Financial Management*

MDPI

*Article*

# Maximum Entropy Evaluation of Asymptotic Hedging Error under a Generalised Jump-Diffusion Model

**Farzad Alavi Fard [1], Firmin Doko Tchatoka [2,*] and Sivagowry Srianthakumar [3]**

[1] School of Business, The University of Sydney, Sydney, NSW 2006, Australia; farzad.alavifard@sydney.edu.au
[2] School of Economics, The University of Adelaide, Adelaide, SA 5005, Australia
[3] School of Economics, Finance and Marketing, RMIT University, Melbourne, VIC 3000, Australia; readergs@outlook.com
* Correspondence: firmin.dokotchatoka@adelaide.edu.au

**Abstract:** In this paper we propose a maximum entropy estimator for the asymptotic distribution of the hedging error for options. Perfect replication of financial derivatives is not possible, due to market incompleteness and discrete-time hedging. We derive the asymptotic hedging error for options under a generalised jump-diffusion model with kernel bias, which nests a number of very important processes in finance. We then obtain an estimation for the distribution of hedging error by maximising Shannon's entropy subject to a set of moment constraints, which in turn yields the value-at-risk and expected shortfall of the hedging error. The significance of this approach lies in the fact that the maximum entropy estimator allows us to obtain a consistent estimate of the asymptotic distribution of hedging error, despite the non-normality of the underlying distribution of returns.

**Keywords:** generalised jump; kernel biased; asymptotic hedging error; esscher transform; maximum entropy density; value-at-risk; expected shortfall

**JEL Classification:** C13; C51; G13



**Citation:** Fard, Farzad Alavi, Firmin Doko Tchatoka, and Sivagowry Srianthakumar. 2021. Maximum Entropy Evaluation of Asymptotic Hedging Error under a Generalised Jump-Diffusion Model. *Journal of Risk and Financial Management* 14: 97. https://doi.org/10.3390/jrfm14030097

Academic Editor: Tihana Škrinjarić

Received: 26 January 2021
Accepted: 24 February 2021
Published: 28 February 2021



## 1. Introduction

The theory of pricing and hedging options has been the centre of attention in modern mathematical finance since the seminal Black–Scholes model. It provides a theoretical value and hedging strategy for European options, under the key assumption that there exists a trading strategy that constructs a portfolio that perfectly replicates the pay-off of an option. Furthermore, Black and Scholes assume the underlying stock price follows a geometric Brownian motion, and trading may take place in continuous time. With these assumptions, they show that the initial value of the replicating portfolio provides the initial price of the option. Moreover, the Black–Scholes analysis demonstrates that an option can be created synthetically by dynamically trading in the underlying asset. Nevertheless, it is well accepted that the perfect replication of options by any self-financing strategy is impossible, due to market incompleteness as well as discrete-time hedging. These two sources of error are termed the jump error and the gamma error.

Many researchers, (see for example Hubalek et al. (2006), Cont et al. (2007), and Kennedy et al. (2009)) have studied the problem of hedging an option in an incomplete market, particularly where stock prices may jump. It is well understood that, except in very special cases, martingales with respect to the filtration of discontinuous processes cannot be represented in the form of a unique self-financing strategy, which leads to market incompleteness. At the jump time, both the model price of the option and value of the hedging portfolio jump. The former is a non-linear function of the stock price but the latter is a linear function of the stock price. Therefore, the jump induces a discrepancy between the value of the option and its replicating portfolio, and thus leads to a jump error. Furthermore, these researchers generally assume that the hedging portfolio can

be continuously rebalanced, which is only possible in the absence of transcription costs. In practice, this level of liquidity is not possible and market practitioners rebalance their hedging portfolio using discrete-time observations, just a few times per trading day. The discrete hedging of derivatives securities leads to the gamma error. This error is not easy to measure because the stochastic analysis techniques are not available in discrete time. Wang et al. (2015) provides a recent analysis of literature regarding hedging in a discrete time incomplete market. We base our research on the seminal work of Bertsimas et al. (2000), Hayashi and Mykland (2005) developed a methodology to analyse the discrete hedging error in a continuous-time framework using an asymptotic approach. This methodology was further developed by Tankov and Voltchkova (2009) who investigate the gamma risk via establishing a limit theorem for the renormalised error when the discretisation step tends to zero. Additionally, Rosenbaum and Tankov (2014) discuss the optimality conditions of discretised hedging strategies in the presence of jump.

In this paper we contribute to the literature by approaching the problem from a different angle. We characterise the risk in a dynamic hedge of options through the asymptotic distribution of hedging error. In particular, we investigate the case of conventional delta hedging for a European call option, although other types of options may be treated in a similar manner. Furthermore, we obtain an estimation for the distribution of hedging error by maximising Shannon's entropy (Shannon (1948)) subject to a set of moment constraints, which in turn yield the value-at-risk (VaR) and expected shortfall (ES) of the hedging error, two widely-used risk metrics in finance. In the literature there exist two dominant approaches for constructing the distribution of hedging error, namely, the parametric and non-parametric approaches. The new approach that we propose in this paper chooses the probability distribution with the most uncertainty, or maximum entropy, subject to what is known. This allows us to obtain a consistent estimate of the asymptotic distribution of hedging error, despite the non-normality of the underlying distribution of returns. As a result, we can drive a very generalised modelling framework, which can be applied in different areas of derivatives pricing.

We first extend the methodology introduced in Hayashi and Mykland (2005) to model the asymptotic hedging error for vanilla call options when the underlying asset is governed by a generalised jump-diffusion model with kernel bias. The class of kernel biased completely random measures is a wide class of jump-type processes that can nicely be represented by a generalised kernel-biased mixture of Poisson random measures. The importance of using a kernel biased completely random measure is to derive a variety of forms of distortion of jump sizes through the kernel link function. For example, Fard and Siu (2013) discuss the importance of this representation as it provides great flexibility in modelling different types of finite and infinite jump activities compared with existing studies.

Next, we estimate the probability density function of the hedging error using the maximum entropy (ME) methodology. It allows the agent to effectively combine the aforementioned risk factors dynamically over time to update its belief about the possible distribution governing the terminal hedging error. Maximum entropy is widely used in estimation and information theory, in which beliefs are updated so that the posterior coincides with the prior as closely as possible. Furthermore, ME methodology only updates those aspects of beliefs for which new evidence was gained (cf., Cover and Thomas (2012), Saghafian and Tomlin (2016) and references therein). Despite its widespread use in fields such as estimation theory, physics, statistical mechanics, and information theory among others, it is only recently that researchers have begun to appreciate its usefulness in econophysics. For instance, Xi et al. (2014) use the maximum entropy model to study business cycle synchronisation of the G7 economic system. Gzyl and Mayoral (2016) uses the maximum entropy principle to develop a non-parametric method of determining the prices of the zero coupon bonds, when the only information available consists of the prices of a few coupon bonds. Chan (2009) proposes a general modelling framework for the EM algorithm in approximating the distribution of financial returns in order to develop entropy-based risk metrics; Geman et al. (2014) and Xu et al. (2014) expand on Chan (2009) and develop a robust optimisation framework for portfolio construction;

Mistrulli (2011) and Zhou et al. (2013) applies the entropy maximisation principle to measure financial contagion and systemic risk. Geman et al. (2014), who use ME density in the VaR context, make an interesting observation that the real world is mostly ignorant about the importance of true probability distributions. They further point out that historically, finance theory has had a preference for parametric, less robust, methods. An approach that is based on distributional and parametric certainties may be useful for research purposes but does not accommodate responsible risk taking. Their study shows the importance of the use of true probability distributions in VaR calculations.

The remainder of this paper is structured as follows. In Section 2 we present the calculations for pricing European options under a generalised jump-diffusion model with kernel bias. Furthermore, we generalise the Hayashi and Mykland (2005) framework to derive the asymptotic hedging error, stemmed from the market incompleteness and discrete hedging. In Section 3, we obtain an estimation for the distribution of hedging error by maximising Shannon's entropy subject to a set of moment constraints, which in turn yield the value-at-risk and expected shortfall of the hedging error. Section 4 provides a numerical analysis to highlight the applicability of the method. Section 5 concludes the paper.

## 2. Modelling Framework

The inadequacy of the constant volatility of the Black–Scholes (henceforth, BS) option valuation model to replicate the characteristics of observed option prices is empirically evidenced in the finance literature. For example, the volatility smile for equity options exhibits a consistent pattern illustrating that implied volatilities are lower for options with higher strike prices. This is particularly the case particularly for short maturity options. Several methods are suggested to accommodate the volatility smile in option pricing, but, stochastic volatility models (see e.g., Heston 1993; Hull and White 1987) and jump diffusion models (see e.g., Bates 1991; Merton 1976; Naik and Lee 1990) have become popular alternatives to BS's constant volatility.

Many studies provide evidence that stochastic volatility models are a significant improvement (cf. Dotsis et al. (2007)). Nevertheless, their applicability in empirical studies suffers from two major drawbacks, namely, implausible correlation structure between returns and volatility, and excessively high "volatility of volatility" (cf. Bates (2000); Diavatopoulos et al. (2012); Dotsis et al. (2007)). Amongst these, the seminal work of Bakshi et al. (1997) reviews a range of alternative option pricing models, and conclude that stochastic volatility is of "first-order importance in improving upon the [Black–Scholes] formula", but adding jumps may provide further improvement based on an out-of-sample test.

As a result, a large body of the recent literature has focused on augmenting stochastic volatility models with jumps. For example, Bakshi et al. (2003), Eraker et al. (2003), and Bakshi et al. (2012) conduct a set of empirical investigations for the case of index options, evidencing the importance of adding the jump component in the returns process. Chang et al. (2013) finds that joint time series data of the underlying S&P 500 index and options on it strongly reject a stochastic volatility model without jumps. Shanahan et al. (2016) price long-maturity equity linked products, as a composition of three embedded options, using the Meixner process (to capture jumps) and a diffusion process to capture stochastic volatility.

On the other hand, option pricing with both jumps and stochastic volatility is rather cumbersome, as analytical solutions are rarely achievable. Even in the relatively simple context of vanilla American options, the studies of Bakshi et al. (2003) and Bakshi et al. (2012) uses European option calculations on a restricted set of out-of-the-money options.

Recently, some theoretical improvement have been offered in the literature, such as the polynomial approximation of stochastic volatility in Shanahan et al. (2016). However, due to the absence of supportive empirical evidence, we chose a more practical alternative, proposed by Andersen and Andreasen (2000). By so doing, we aim to extend the analysis of

Dupire (1994) to the case of jumps, whereby a model combining jumps with a deterministic volatility is developed. This type of model captures the observed behaviour of implied volatilities (see e.g., Andersen and Andreasen 2000). In addition, it is much easier to use for contracts where no analytic solution is readily available. Moreover, note that adopting a deterministic local volatility function approach has an advantage in calibrating European option prices. Indeed, instead of solving a set of backward equations, one for each option of a different strike and maturity, one only needs to solve a single one-dimensional forward equation for all strikes and maturities.

### 2.1. Financial Markets

We fix a complete probability space $(\Gamma, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ is the real-world probability measure. Let $\mathcal{T}$ denote the time index set $[0, T]$ of the economy. Let $\{r(t)\}_{t\in\mathcal{T}}$ be the instantaneous market interest rate of a money market account. Then, the dynamics of the value of the risk-free asset, $\{B_t\}_{t\in\mathcal{T}}$ would be:

$$\frac{dB_t}{B_t} = r(t)dt, \qquad B_0 = 1.$$

To model different types of finite and infinite jump activities, we adopt the kernel biased representation of completely random measures of James (2002, 2005).

To proceed, consider the measurable space $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$, where $\mathcal{B}(\mathcal{T})$ is the Borel $\sigma$-field (generated by the open subsets of $\mathcal{T}$). We denote by $\mathcal{B}_0$ the family of Borel sets $U \in \mathcal{R}^+$, with closure $\bar{U}$ not containing 0. Define $\mathcal{X} = \mathcal{T} \times \mathcal{R}^+$. Under these definitions, the measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is explicitly given by $(\mathcal{T} \times \mathcal{R}^+, \mathcal{B}(\mathcal{T}) \otimes \mathcal{B}_0)$.

Let $N(., U)$ define a Poisson random measure for all $U \in \mathcal{B}_0$. We denote by $N(dt, dz)$ the differential form of measure $N(t, U)$. In addition, define $\rho(dz|t)$ as a Lévy measure that depends on $t$. Let $\eta$ be a $\sigma$-finite (nonatomic) measure on $\mathcal{T}$. Following James (2005), we assume that there exists an arbitrary positive function $h(z)$ on $\mathcal{R}^+$, which along with $\rho$ and $\eta$ are chosen such that:

$$\sum_{i=1}^{N} \int_{\mathcal{B}} \int_{\mathcal{R}^+} \min(h(z), 1)\rho(dz|t)\eta(dt) < \infty.$$

and $h^2(z) \leq F_t\rho(z)$, where $F$ is a cádlág $\mathcal{F}_t$-adapted process. Define the intensity measure:

$$\nu(dt, dz) := \rho(dz|t)\eta(dt),$$

as well as the kernel biased completely random measure:

$$\mu(dt) := \int_{\mathcal{R}^+} h(z)N(dt, dz).$$

The latter is a kernel-biased Poisson random measure $N(dt, dz)$ over the state space of the jump size $\mathcal{R}^+$ with the mixing kernel function $h(z)$. We can replace the Poisson random measure with any random measure and choose some quite exotic functions for $h(z)$ to generate different types of finite and infinite jump activities. Let $\{W_t\}_{t\in\mathcal{T}}$ denote a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the $\mathbb{P}$-augmentation of its natural filtration $\mathcal{F}^W := \{\mathcal{F}_t^W\}_{t\in\mathcal{T}}$. Let $\widetilde{N}_{X_t}(dt, dz)$ denote the compensated Poisson random measure defined by:

$$\widetilde{N}(dt, dz) = N(dt, dz) - \rho(dz|t)\eta(dt).$$

Let $\mu_t$ and $\sigma_t$ denote the drift and volatility of the market value of the underlying asset, respectively. Consider a random jump process $A := \{A(t)|t \in \mathcal{T}\}$, such that:

$$dA_t = A_{t^-}\left[\mu_t dt + \sigma_t dW_t + \int_{\mathcal{R}^+} h(z)\widetilde{N}(dt, dz)\right], \tag{1}$$

where $A_0 = 0$. We assume under $\mathbb{P}$ the price process $\{S_t\}_{t\in\mathcal{T}}$ is defined as $S_t := \exp(A_t)$ so that:

$$dS_t = \left(\mu_t + \tfrac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t - \int_{\mathcal{R}^+}\left\{h(z) - e^{h(z)} + 1\right\}\rho d(z|t)\eta(dt) + \int_{\mathcal{R}^+}\left(e^{h(z)} - 1\right)\widetilde{N}(dt, dz), \quad (2)$$

with $S_0 = 1$.

*2.2. Esscher Transform*

It is well known that no arbitrage opportunities are necessary for the determination of a unique equivalent risk neutral martingale measure, thus ensuring the fair valuation of the option (Pilska 1997). In this paper, we emphasise on the possibility of having incomplete markets, and as such there may be more than one equivalent martingale measure, i.e., more than one no-arbitrage price. There are different methods to price and hedge derivative securities in incomplete financial markets. For example, one can choose an equivalent martingale measure by minimising the quadratic utility of the terminal hedging errors (see e.g., Follmer and Sondermann 1986; Follmer and Schweizer 1991; Schweizer 1995). One can also adopt an economic approach based on the marginal rate of substitution to select a pricing measure via a utility maximisation problem (see e.g., Davis 1997). Finally, one may employ the minimum entropy martingale measure method to determine the equivalent martingale measure (see e.g., Avellaneda 1998; Fard and Siu 2012; Frittelli 2000).

In this study we employ the Esscher transform to determine an equivalent martingale measure for the valuation of the option (see Gerber and Shiu 1994). The method provides market practitioners with a convenient and flexible way to value options. For example, it has been shown in Elliott et al. (2005) that for exponential Lévy models, the Esscher martingale transform for the linear process is also the minimal entropy martingale measure, i.e., the equivalent martingale measure which minimises the relative entropy, and this measure also has the property of preserving the Lévy structure of the model. In the framework of exponential Lévy models, the study of equivalent martingale measures the relationships and optimality properties, which has been developed in several directions, see Esche and Schweizer (2005), Hubalek and Sgarra (2006), and Tankov (2003) and the references therein.

Let $\mathcal{F}^A := \{\mathcal{F}_t^A\}_{t\in\mathcal{T}}$ and $\mathcal{F}^S := \{\mathcal{F}_t^S\}_{t\in\mathcal{T}}$ denote the $\mathbb{P}$-augmentation of the natural filtration generated by $A$ and $S$, respectively. Since, $\mathcal{F}^A$ and $\mathcal{F}^S$ are equivalent, we can use either one as an observed information structure. Write $B(\mathcal{T})$ for the Borel $\sigma$-field of $\mathcal{T}$ and let $BM(\mathcal{T})$ denote the collection of $B(\mathcal{T})$-measurable and nonnegative functions with compact support on $\mathcal{T}$. For each process $\theta \in BM(\mathcal{T})$, write:

$$(\theta.A)_t := \int_0^t \theta(u)dA(u), \qquad t \in \mathcal{T},$$

such that $\theta$ is integrable with respect to the return process.

Let $\{\Lambda_t\}_{t\in\mathcal{T}}$ denote a $\mathcal{G}$-adapted stochastic process:

$$\Lambda_t := \frac{e^{(\theta.A)_t}}{\mathcal{M}(\theta)_t}, \qquad t \in \mathcal{T},$$

where $\mathcal{M}(\theta)_t := E[e^{(\theta.A)_t}|\mathcal{F}_t^A]$ is a Laplace cumulant process and takes the following form:

$$\mathcal{M}(\theta)_t = \exp\left[\int_0^t \theta_s\left(\mu_s - \tfrac{1}{2}\sigma_s^2\right)ds + \tfrac{1}{2}\int_0^t \theta_s^2\sigma_s^2 ds + \int_0^t\int_{\mathcal{R}^+}\left(e^{\theta_s h(z)} - 1 - \theta h(z)\right)\rho(dz|s)\eta(ds)\right].$$

Therefore,

$$\Lambda_t = \exp\left[\int_0^t \theta_s \sigma_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 \sigma_s^2 ds + \int_0^t \int_{\mathcal{R}^+} \theta_s h(z) \widetilde{N}(dz, ds) \right. \tag{3}$$
$$\left. - \int_0^t \int_{\mathcal{R}^+} \left(e^{\theta_s h(z)} - 1 + \theta_s h(z)\right) \rho(dz|s)\eta(ds)\right].$$

The goal is to use $\Lambda_t$ in (3) as the Radon–Nikodym derivative to change the historical probability measure to the risk-neutral measure. Therefore, (3) is an essential part of our pricing formulation. An important characteristic of risk-neutral measure is that every discounted price process is a martingale under this measure. To establish this key property, it is paramount to show that (3) is $\mathcal{G}_t$-martingale.

**Lemma 1.** *$\Lambda_t$ is $\mathbb{P}$ martingale w.r.t $\mathcal{G}_t$.*

See Appendix A for proof.

For each $\theta \in L(A)$ define a new probability measure $\mathbb{P}^\theta \sim \mathbb{P}$ on $\mathcal{G}(T)$ by the Radon–Nikodym derivative:

$$\left.\frac{d\mathbb{P}^\theta}{d\mathbb{P}}\right|_{\mathcal{G}(T)} := \Lambda_T. \tag{4}$$

This new measure $d\mathbb{P}^\theta$ is defined by the Esscher transform $\Lambda_T$ associated with $\theta \in L(A)$.

The local-martingale condition, i.e., there exists an equivalent martingale measure under which discounted asset prices are local-martingales in the absence of arbitrage, is the foundation of asset pricing theory. Below, we state a necessary and sufficient condition for the local martingale condition in our framework.

**Proposition 1.** *For each $t \in \mathcal{T}$, let the discounted price of the risky asset at time t be:*

$$\widetilde{S}(t) := e^{-rt} S(t).$$

*Then the discounted price process $\widetilde{S} := \{\widetilde{S}(t) | t \in \mathcal{T}\}$ is an $\mathbb{P}^\theta$-local-martingale if and only if $\theta_t := \langle \theta, X_t \rangle$, $t \in \mathcal{T}$, is such that $\theta := (\theta_1, \theta_2, ..., \theta_N) \in \mathcal{R}^N$ satisfies the following equation:*

$$\theta_t \sigma_t^2 + \int_{\mathcal{R}^+} \{e^{\theta h(z)}(e^{h(z)} - 1) - h(z)\} \rho d(z|t) \eta'(t) = r_t - \mu_t. \tag{5}$$

See Appendix A for proof.

The results from the Lemma 1, Equation (4), and Proposition 1, allow us to use (3) to drive the risk-neutral dynamics of the return process.

**Proposition 2.** *Suppose $\widetilde{W}_t = W_t - \int_0^t \sigma_s \theta ds$ is a $\mathbb{P}^\theta$-Brownian motion, $\rho^\theta(dz|t) := e^{\theta h(z)} \rho(dz|t)$ is the $\mathbb{P}^\theta$ compensator of $N^\theta(dz, dt)$ then:*

$$dA_t = \left(\mu_t + 2\theta\sigma_s^2 - \tfrac{1}{2}\sigma_s^2\right)dt + \sigma_t d\widetilde{W}_t + \int_{\mathcal{R}^+} h(z)\left(1 - e^{-\theta h(z)}\right)\rho^\theta(dz|t)\eta(dt) + \int_{\mathcal{R}^+} h(z)\widetilde{N}^\theta(dz, dt). \tag{6}$$

See Appendix B for proof.

Similarly, we can derive the risk-neutral price process of the reference portfolio.

**Proposition 3.** *The price process of the reference portfolio S under $\mathbb{P}^\theta$ is:*

$$dS_t = (r_t - \frac{1}{2}\sigma_t^2)dt + \sigma_t d\widetilde{W}_t + \int_{\mathcal{R}^+} \left(e^{\theta h(z)}(e^{h(z)} - 1) - h(z)\right)\widetilde{N}^\theta(dt, dz). \tag{7}$$

**Proof.** Proof of Proposition 3. Recall $S_t := \exp(A_t)$. Then the proof can easily follow by applying Ito's Lemma and the martingale condition (5) to (6). □

We study the hedging of a European option with pay-off function $G$ using the popular delta hedging strategy. The option price is given by:

$$C(t,S) = E^{\theta}[G(S_T)|S_t = S], \tag{8}$$

where we assume $C \in C^{\infty}([0,T) \times \mathcal{R})$. Furthermore, the delta hedging strategy is $H_t := \frac{\partial C(t,S)}{\partial S}$, which is the most widely-used hedging strategy with a mathematically tractable structure. Detailed discussion about the hedging strategy is provided in the next subsection.

### 2.3. Continuous Hedging Strategy

We assume the existence of a continuous-time trading strategy $H$. If continuous-time hedging was possible, agents in the market would like to follow this strategy. This strategy may be chosen in several ways and may not lead to a perfect replication when markets are incomplete. In this study, we do not address the relative advantages of different choices of $H$. Rather, we suppose its existence from another generalised jump-diffusion process that satisfies the same set of assumptions as $S$. As such, we assume that:

$$dH_t = a_t dt + b_t dW_t + \int_{\mathcal{R}^+} \gamma_t \widetilde{N}(dt,dz), \tag{9}$$

under $\mathbb{P}$, where $a_t$, $b_t$, and $\gamma_t$ are the parameters of the process that will be determined below.

By applying the Itô Lemma to the definition of $H_t$, we can show the following decomposition:

$$\begin{aligned} dH_t := d\frac{\partial C(t,S)}{\partial S} =& \left\{ \frac{\partial^2 C}{\partial t \partial S}(t,S) + \left(\mu_t + \frac{1}{2}\sigma_t^2 + \int_{\mathcal{R}^+} h(z)\rho(dz|t)\eta'\right)\frac{\partial^2 C}{\partial S^2}(t,S) + \frac{1}{2}\sigma_t^2 \frac{\partial^3 C}{\partial S^3}(t,S) \right\} dt \\ &+ \sigma_t \frac{\partial^2 C}{\partial S^2}(t,S) dW_t + \int_{\mathcal{R}^+} \left\{ \frac{\partial C}{\partial S}(t, S + e^{h(z)} - 1) - \frac{\partial C}{\partial S}(t,S) \right\} \rho(dz|t)\eta' dt \\ &+ \int_{\mathcal{R}^+} \left\{ \frac{\partial C}{\partial S}(t, S + e^{h(z)} - 1) - \frac{\partial C}{\partial S}(t,S) \right\} \widetilde{N}(dt,dz). \end{aligned}$$

Then,

$$\begin{aligned} a_t =& \frac{\partial^2 C}{\partial t \partial S}(t,S) + \left(\mu_t + \frac{1}{2}\sigma_t^2 + \frac{1}{2}\sigma_t^2 \frac{\partial^3 C}{\partial S^3}(t,S) \right. \\ &\left. + \int_{\mathcal{R}^+} \left\{ \frac{\partial C}{\partial S}(t, S + e^{h(z)} - 1) - \frac{\partial C}{\partial S}(t,S) + h(z)\frac{\partial^2 C}{\partial S^2}(t,S) \right\} \rho(dz|t)\eta' dt \right. \\ b_t =& \sigma_t \frac{\partial^2 C}{\partial S^2}(t,S) \\ \gamma_t =& \frac{\partial C}{\partial S}(t, S + e^{h(z)} - 1) - \frac{\partial C}{\partial S}(t,S). \end{aligned}$$

### 2.4. Asymptotic Hedging Error

In what follows we drive the asymptotic distribution of the hedging error, generalising the methodology proposed in Hayashi and Mykland (2005) and Tankov and Voltchkova (2009).

Let,

$$\begin{aligned} &\mu_t^n := (-n) \vee \mu_t \wedge n; \quad \sigma_t^n := (-n) \vee \sigma_t \wedge n; \quad h^n(z) := -\sqrt{n\rho(n)} \vee h(z) \wedge \sqrt{n\rho(n)}; \\ &a_t^n := (-n) \vee a_t \wedge n; \quad b_t^n := (-n) \vee b_t \wedge n; \quad \gamma_t^n(z) := -\sqrt{n\rho(n)} \vee \gamma_t(z) \wedge \sqrt{n\rho(n)}. \end{aligned}$$

Then the processes $S^n$ and $H^n$ are Lévy–Itô processes with bounded coefficients and bounded jumps that coincide with $S$ and $H$ on the set:

$$\Omega_n := \{ \sup_{0\le t\le T} \max(|\mu_t|, |\sigma_t|, |F_t|) \le n; \sup_{0\le t\le T} \max(|a_t|, |b_t|, |F_t|) \le n; N([0,T], ((-\infty,-n)\cup(n,\infty)) = 0)\}.$$

Since all processes are supported cádlág, $P[\Omega_n] \to 1$.

The continuous re-balancing of a portfolio is practically unfeasible. Typically, holders of a position in an option $\Delta$-hedge in discrete time intervals of $t_i = iT/n$. Therefore, the trading strategy is piecewise constant and given by $F_{\phi_n(t)}$, where $\phi_n(t) = \sup\{t_i, t_i < t\}$. The value of the hedging portfolio at time $t$ is $V_0 + \int_0^t H_{s^-} dS_s$ with continuous hedging and $V_0 + \int_0^t H_{\phi_n(t)} dS_s$ with discrete hedging. Then the asymptotic distribution of the difference between discrete and continuous hedging is:

$$U_t^n = \int_0^t (H_s^- - H_{\phi_n(t)}) dS_s = \int_0^t H_{s^-}^n dS_s \tag{10}$$

where $n \to \infty$. For any process $A$ we set $A_t^n := A_t - A_{\phi_n(t)}$. Under the above conditions, Hayashi and Mykland (2005) provides a thorough discussion on the stable convergence of the bounded processes to their respective original processes.

Furthermore, define the renormalised hedging error process by:

$$Z_t^n = \sqrt{n} U_t^n = \sqrt{n} \int_0^t H_{s^-}^n dS_s. \tag{11}$$

Let $\{\widehat{W}_t\}_{t\in\mathcal{T}}$ be a standard Brownian motion independent of $W$ and $N$, and let $(\xi_k)_{k\ge1}$ and $(\xi'_k)_{k\ge1}$ be two sequences of a standard normal random variable and $(\zeta_k)_{k\ge1}$ sequence of independent uniform random variables on [0,1], such that the three sequences are independent from each other and other random elements. Let $(T_i)_{i\ge1}$ be an enumeration of the jump times of $N$, and define:

$$\Delta F_{T_i} = \frac{\partial C}{\partial S}(T_i, S_{T_i}) - \frac{\partial C}{\partial S}(T_i, S_{T_i^-}),\ \widetilde{\sigma}_{T_i^-} = \sigma_{T_i} \frac{\partial^2 C}{\partial S^{T_i}}(T_i, S_{T_i}),\ \widetilde{\sigma}_s = \sigma_s \frac{\partial^2 C}{\partial S^2}(s, S_s).$$

Then, applying Theorem 1 in Tankov and Voltchkova (2009) to the renormalised discrete delta hedging error $Z_t^n$ gives the following asymptotical convergence result in finite-dimensional laws:

$$\begin{aligned} Z_t &= \sqrt{\frac{T}{2}} \int_0^t \sigma_s \widetilde{\sigma}_s d\widehat{W}_s + \sqrt{T} \sum_{i:T_i\le t:\Delta S_t\ne0} \Delta F_{T_i} \sqrt{\zeta_i} \xi_i \sigma_{T_i} + \sqrt{T} \sum_{i:T_i\le t:\Delta S_t\ne0} \Delta S_{T_i} \sqrt{1-\zeta_i} \xi'_i \widetilde{\sigma}_{T_i^-} \qquad (12)\\ &\equiv \sqrt{\frac{T}{2}} \int_0^t \sigma_s^2 \frac{\partial^2 C}{\partial S^2} d\widehat{W}_s + \sqrt{T} \sum_{i:T_i\le t:\Delta S_t\ne0} \left( \frac{\partial C}{\partial S}(T_i, S_{T_i}) - \frac{\partial C}{\partial S}(T_i, S_{T_i^-}) \right) \sqrt{\zeta_i} \xi_i \sigma_{T_i} + \\ &\quad \sqrt{T} \sum_{i:T_i\le t:\Delta S_t\ne0} \Delta S_{T_i} \sqrt{1-\zeta_i} \xi'_i \sigma_{T_i} \frac{\partial^2 C}{\partial S^{T_i}}(T_i, S_{T_i}). \end{aligned}$$

## 3. Estimation of the Density of the Hedging Error

Let $Z_T$ be a realisation of the hedging error in (12) at date $T$ (say the maturity date of the option). Suppose that we have a random sample $\{Z(j) : j = 1, \dots, n\}$ of $n$ *i.i.d.* observations from $Z_T$, each with pdf $f_Z$ and cdf $F_Z$ on a support $\Omega \subseteq \mathbb{R}$. Note from (12) that $Z_T$ always has a continuous density almost everywhere (a.e). In this section, the VaR and the ES associated to $Z_T$ are of inferential interest. The VaR associated to $Z_T$ at level $\alpha$ is defined as:

$$\mathrm{VaR}_\alpha(Z_T) \quad = \quad \inf\{z \in \mathbb{R} : F_Z(z) \ge \alpha\} = F_Z^{-1}(\alpha), \tag{13}$$

while the expected shortfall is given as:

$$\begin{aligned} \mathrm{ES}_\alpha(Z_T) &= \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_\gamma(Z_T)d\gamma \\ &= -\frac{1}{\alpha}\Big(E[Z_T 1_{\{Z_T\le z_\alpha\}}] + z_\alpha[\alpha - F_Z(z_\alpha)]\Big), \end{aligned} \tag{14}$$

where $z_\alpha = \inf\{z \in \mathbb{R} : F_Z(z) \ge \alpha\}$ is the lowest $\alpha$-quantile and $1_A(z) = 1$ if $x \in A$ and $1_A(z) = 0$ else. The dual representation of (14) is given by:

$$\mathrm{ES}_\alpha(Z_T) = \inf_{Q\in\mathcal{Q}_\alpha} E^Q[Z_T], \tag{15}$$

where $Q$ is absolutely continuous such that $\frac{dQ}{dP} \le \alpha^{-1}$. If the density of $Z_T$ is continuous, then the expected shortfall is equivalent to the tail conditional expectation defined by $\mathrm{TCE}_\alpha(Z) = E[-Z_T | Z_T \le -\mathrm{VaR}_\alpha(Z_T)]$ . If $F_Z$ (or $f_Z$) were given, then the computation of $\mathrm{VaR}_\alpha(Z_T)$ and $\mathrm{ES}_\alpha(Z_T)$ is straightforward from (13)–(15). This is unfortunately not the case, and one has to resort to estimation techniques to approximate them. Suppose that we have a consistent estimate of $f_Z$, say $\hat{f}_Z$. So, we can also estimate its cdf and compute the estimate of the value-at-risk and the expected shortfall as:

$$\widehat{\mathrm{VaR}}_\alpha(Z_T) = \hat{F}_Z^{-1}(\alpha), \quad \widehat{\mathrm{ES}}_\alpha(Z_T) = \frac{1}{\alpha}\int_0^\alpha \widehat{\mathrm{VaR}}_\gamma(Z_T)d\gamma. \tag{16}$$

Our main goal is to find an estimator $\hat{f}_Z$ of $f_Z$ that captures the maximum uncertainty in $Z_T$. To achieve this goal, we will use the information entropy approach.

### *3.1. Information Entropy and Density Estimation*

The information entropy associated with $Z_T$ is defined as:

$$I_E(Z_T) = -\int_\Omega f_Z(z)\ln f_Z(z)dz, \tag{17}$$

where, by convention, we assume that $0 \times ln(0) = 0$. $I_E(Z_T)$ is a measure of the information carried by $Z_T$. As data are communicated more, they are corrupted with more noise so that the entropy increases, therefore they carry less information.

Let $g_Z^{(k)} : \Omega \to \mathbb{R}$ $(k = 0, 1, 2, \dots)$ be a moment function of $Z_T$. Then, the moment of $Z_T$ with respect to $g_Z^{(k)}$ is defined as:

$$\mu_k = \mathbb{E}[g_Z^{(k)}(z)] = \int_\Omega g_Z^{(k)}(z) f_Z(z)dx, \quad k = 0, 1, 2, \dots. \tag{18}$$

In practice, polynomial functions are often used for the moment function (for example, see Zellner and Highfield (1988)). $g_Z^{(k)}$ $(k = 0, 1, 2, \dots)$, i.e., $g_Z^{(k)}(z) = z^k$, $k = 0, 1, 2, \dots$, with the normalisation $g_Z^{(0)}(z) = 1$. In this case, we have:

$$\mu_0 = \int_\Omega f_Z(z)dz = 1, \ \mu_k = \int_\Omega z^k f_Z(z)dz, \quad k = 1, 2, \dots. \tag{19}$$

For the remainder of the paper, we assume the sequence $(\mu_k)_{k\ge1}$ satisfies the Carleman's condition [see Akhiezer (1965)], i.e.,

$$\sum_{k=1}^{\infty} \mu_{2k}^{-\frac{1}{2k}} = +\infty. \tag{20}$$

Note that when $\Omega = \mathbb{R}$ and $g_Z^{(k)}(z) = z^k$ for all $k = 0, 1, 2, \dots$, condition (20) is sufficient for the determinacy of the Hamburger moment problem.

Due to issues associated with using many moments, we only use $k = 0, \ldots, m$ moments in (19), and our goal is to find the density function $f_Z(z)$ such that:

$$\max_{f_Z \in \mathbf{L}^1(\Omega):\, f_Z \geq 0} \left\{ I_E(Z_T) = -\int_\Omega f_Z(z) \ln f_Z(z) dz \right\} \tag{21}$$

subject to

$$\bar{\mu}_k = \int_\Omega g_Z^{(k)}(z) f_Z(z) dz, \; k = 0, 1, 2, \ldots, m \tag{22}$$

where $\bar{\mu}_k = n^{-1} \sum_{j=1}^n g_Z^{(k)}(Z(j))$ is the empirical moment of $Z_T$ with respect to $g_Z^{(k)}$. For now, we assume that the number $m$ of selected moments of the pdf in (19) that have been chosen to match the empirical moments is fixed. Later on, we discuss how $m$ can be selected in a data-dependent manner based on a Bayesian- or Schwartz-type information criterion. Since $g_Z^{(0)} = 1$, the zeroth moment implies that $\mu_0 = \int_\Omega f_Z(z) dz = \bar{\mu}_0 = 1$ in (22). For example if $g_Z^{(k)}(z) = z^k, \; k = 0, 1, 2, 3, 4$ with $g_Z^{(0)}(z) = 1$, we have $m = 4$ and problem (22) becomes:

$$\begin{aligned} &\max_{f_Z \in \mathbf{L}^1(\Omega):\, f_Z \geq 0} I_E(Z_T) \quad \text{subject to} \\ &\bar{\mu}_0 = \int_\Omega f_Z(z) dz = 1, \quad \bar{\mu}_1 = \int_\Omega z f_Z(z) dz, \quad \bar{\mu}_2 = \int_\Omega z^2 f_Z(z) dz \\ &\bar{\mu}_3 = \int_\Omega z^3 f_Z(z) dz, \quad \bar{\mu}_4 = \int_\Omega z^4 f_Z(z) dz. \end{aligned} \tag{23}$$

For the remainder of the paper, we call the solution $\hat{f}_Z(z)$ of problem (22) the maximum-entropy estimator of $f_Z(z)$. We will now establish that such a solution exists and is unique. Let,

$$L(f_Z, \lambda_0, \ldots, \lambda_m) \;=\; -I_E(Z) + \sum_{k=0}^m \lambda_k \left( \int_\Omega g_Z^{(k)}(z) f_Z(z) dz - \bar{\mu}_k \right) \tag{24}$$

denote the Lagrange function associated with (22), where $\lambda_0, \ldots, \lambda_m$ are the Lagrange multipliers. By noting that solving (22) is equivalent to static problem:

$$\min_{f_Z \in \mathbf{L}^1(\Omega):\, f_Z \geq 0,\, \lambda_0, \ldots, \lambda_m} L(f_Z, \lambda_0, \ldots, \lambda_m), \tag{25}$$

it is sufficient to show that (25) has a unique solution with respect to $f_Z \in \mathbf{L}^1(\Omega) : \; f_Z \geq 0$. More often, the minimisation over $f_Z$ is carried out via the first variation of $L(f_Z, \lambda_0, \ldots, \lambda_m)$ with respect to $f_Z$. This type of calculation is misleading because $L(f_Z, \lambda_0, \ldots, \lambda_m)$ has support on the set $\{f_Z \in \mathbf{L}^1(\Omega) : \; f_Z \geq 0 \text{ a.e.}\}$, and the complement of this set is dense in $\mathbf{L}^1$, meaning that not only is $L(f_Z, \lambda_0, \ldots, \lambda_m)$ nowhere differentiable in $\mathbf{L}^1$, it is also nowhere continuous. Due to these reasons, we follow the convex duality approach in Borwein et al. (2003). From (Borwein et al. 2003, Theorems 1–2), there is a unique solution $f_Z$ of (25) satisfying:

$$\begin{aligned} f_Z(z; \lambda) \;&=\; \exp\left( -1 - \lambda_0 - \sum_{k=1}^m \lambda_k g_Z^{(k)}(z) \right) \\ &=\; \frac{1}{\theta} \exp\left( - \sum_{k=1}^m \lambda_k g_Z^{(k)}(z) \right) \end{aligned} \tag{26}$$

where $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)'$, $\theta = \exp(1 + \lambda_0)$, and the zeroth-moment equality implies that:

$$\theta = \int_\Omega \exp\left( - \sum_{k=1}^m \lambda_k g_Z^{(k)}(z) \right) dz.$$

To complete the closed form of $f_Z(z;\lambda)$, we must substitute $\lambda$ in (26) by an optimal value $\hat{\lambda}$. Zellner and Highfield (1988) consider the case in which $g_Z^{(k)}(z) = z^k$ where $k = 1,\ldots,m = 4$, and use an algorithm based on a Newton method to compute $\lambda_1,\ldots,\lambda_4$ from the restrictions in (22). This numerical approximation is cumbersome even for a moderate choice of $m = 4$. In this paper, we propose the maximum likelihood (ML) method to estimate $\lambda_1,\ldots,\lambda_m$, and $\theta$.

Since $\{Z(j) : j = 1,\ldots,n\}$ are *i.i.d.* with common pdf $\frac{1}{\theta}\exp\left(-\sum_{k=1}^m \lambda_k g_Z^{(k)}(z(j))\right)$, the likelihood function of the sample can be written as:

$$\begin{aligned}\mathcal{L}(Z(1),\ldots,Z(n);\lambda) &= \prod_{j=1}^n \frac{1}{\theta}\exp\left(-\sum_{k=1}^m \lambda_k g_Z^{(k)}(z_j)\right)\\ &= \frac{1}{\theta^n}\exp\left[-\sum_{j=1}^n\sum_{k=1}^m \lambda_k g_Z^{(k)}(z_j)\right]. \end{aligned} \tag{27}$$

The log-likelihood function from (27) is then given by $\ell(Z(1),\ldots,Z(n);\lambda) = -n\ln(\theta) - \sum_{j=1}^n\sum_{k=1}^m \lambda_k g_Z^{(k)}(z_j)$. The ML estimator of $\lambda = (\lambda_0,\lambda_1,\ldots,\lambda_m)'$ satisfies:

$$\max_{\lambda} \ell(Z(1),\ldots,Z(n);\lambda) \tag{28}$$

and we can prove the following on the existence of unique solutions for both problems (28) and (22).

**Proposition 4.** *Suppose that* $\bar{\mu}_k = n^{-1}\sum_{j=1}^n g_Z^{(k)}(Z(j)) < \infty \;\forall\; k = 1,\ldots,m$. *Then*:
(*a*) *Problem* (28) *has a unique solution with respect to* $\lambda$;
(*b*) *Problem* (22) *has a unique solution with respect to* $f_Z(\cdot)$.

**Proof.** First, observe first that $\ell(Z(1),\ldots,Z(n);\lambda)$ is a concave function in $\lambda$, and twice differentiable with respect to $\lambda$. The first order condition of (28) is given by:

$$\begin{aligned}\frac{\partial \ell(Z(1),\ldots,Z(n);\lambda)}{\partial \lambda_k} &= -n\frac{\partial \ln(\theta)}{\partial \lambda_k} - \sum_{j=1}^n g_Z^{(k)}(z_j) = 0\\ &\Leftrightarrow \quad \frac{\partial \ln(\theta)}{\partial \lambda_k} = -n^{-1}\sum_{j=1}^n g_Z^{(k)}(z_j) = -\bar{\mu}_k,\; k = 1,\ldots,m. \end{aligned} \tag{29}$$

Moreover, $\theta = \exp(1+\lambda_0) = \int_\Omega \exp\left(-\sum_{k=1}^m \lambda_k g_Z^{(k)}(z_j)\right)dz_j$ from the zeroth-moment equality. Thus,

$$\begin{aligned}\frac{\partial \ln(\theta)}{\partial \lambda_k} &= \frac{1}{\theta}\frac{\partial}{\partial \lambda_k}\int_\Omega \exp\left(-\sum_{k=1}^m \lambda_k g_Z^{(k)}(z_j)\right)dz_j = -\int_\Omega g_Z^{(k)}(z_j)\frac{1}{\theta}\exp\left(-\sum_{k=1}^m \lambda_k g_Z^{(k)}(z_j)\right)dz_j\\ &= -\int_\Omega g_Z^{(k)}(z_j) f_Z(z_j)dz_j = -\mu_k \equiv \mu_k(\lambda), \end{aligned} \tag{30}$$

where $\mu_k(\lambda)$ is a continuous function of $\lambda$. Therefore, (29) and (30) entail that:

$$\mu_k(\lambda) = \bar{\mu}_k,\; k = 1,\ldots,m. \tag{31}$$

So, as long as $\bar{\mu}_k < \infty$, (31) has a solution with respect to $\lambda$, i.e., the likelihood function $\ell(Z(1),\ldots,Z(n);\lambda)$ has a critical point $\hat{\lambda}$.

We will now show that this critical point, $\hat{\lambda}$, is the unique point where $\ell(Z(1),\ldots,Z(n);\lambda)$ is maximised. Let $H(Z(1),\ldots,Z(n);\lambda)$ denote the hessian matrix of the log-likelihood function $\ell(Z(1),\ldots,Z(n);\lambda)$. The $(k,p)$ entree of $H(Z(1),\ldots,Z(n);\lambda)$ is:

$$
\begin{aligned}
\frac{\partial}{\partial\lambda_p}\left(\frac{\partial\ell(Z(1),\ldots,Z(n);\lambda)}{\partial\lambda_k}\right) &= \frac{\partial}{\partial\lambda_p}\left(-n\frac{\partial\ln(\theta)}{\partial\lambda_k}-\sum_{j=1}^{n}g_Z^{(k)}(z_j)\right) = -n\frac{\partial}{\partial\lambda_p}\frac{\partial\ln(\theta)}{\partial\lambda_k}\\
&= n\frac{\partial}{\partial\lambda_p}\left(\frac{1}{\theta}\int_\Omega g_Z^{(k)}(z_j)\exp\left(-\sum_{k=1}^{m}\lambda_k g_Z^{(k)}(z_j)\right)dz_j\right)\\
&= -n\frac{1}{\theta}\int_\Omega g_Z^{(p)}(z_j)g_Z^{(k)}(z_j)\exp\left(-\sum_{k=1}^{m}\lambda_k g_Z^{(k)}(z_j)\right)dz_j +\\
&\quad n\left(\frac{1}{\theta^2}\int_\Omega g_Z^{(k)}(z_j)\exp\left(-\sum_{k=1}^{m}\lambda_k g_Z^{(k)}(z_j)\right)dz_j\right)\times\int_\Omega g_Z^{(p)}(z_j)\exp\left(-\sum_{k=1}^{m}\lambda_k g_Z^{(k)}(z_j)\right)dz_j\\
&= -n\left[\int_\Omega g_Z^{(p)}(z_j)g_Z^{(k)}(z_j)f_Z(z_j)dz_j-\int_\Omega g_Z^{(p)}(z_j)f_Z(z_j)dz_j\int_\Omega g_Z^{(k)}(z_j)f_Z(z_j)dz_j\right]\\
&= -n\left[\mathbb{E}(g_Z^{(p)}(z_j)g_Z^{(k)}(z_j))-\mathbb{E}(g_Z^{(p)}(z_j))\mathbb{E}(g_Z^{(k)}(z_j))\right]\\
&= -nCoV\left(g_Z^{(p)}(z_j),g_Z^{(k)}(z_j)\right),\ \forall\ p,k=1\ldots,m, \qquad (32)
\end{aligned}
$$

where $CoV(a,b)$ is the covariance of the two random variables $a$ and $b$. From (32), it is clear that $H(Z(1),\ldots,Z(n);\lambda)$ is a symmetric and strictly negative definite for all values of the vector of Lagrange multipliers. Thus, the critical point $\hat{\lambda}$ is a unique maximum, which completes the proof of Proposition 4(a). It follows immediately from (26) that $\hat{f}_Z(z;\hat{\lambda})$ is also a unique maximum, thus establishing Proposition 4(b). □

### *3.2. Choice of m and the Moment Functions in the Density Estimation*

To avoid the problems associated with using many moments, we suggest using a finite number $m$ of moments in the estimation of the density function $f_Z(z)$. In practice, the choice of $m$, as well as that of the moment functions $g_Z^{(k)}(\cdot),\ k=1,\ldots,m$, may not be obvious. In this section, we briefly discuss how both $m$ and $g_Z^{(k)}(\cdot),\ k=1,\ldots,m$ can be approximated in a data dependent manner. First, note from (26) that the density function $f_Z(z;\lambda)$ satisfies:

$$
f_Z(z;\lambda) = \frac{1}{\theta}\exp\left(-\sum_{k=1}^{m}\lambda_k g_Z^{(k)}(z)\right). \qquad (33)
$$

For the choice of moment functions, we replace each $g_Z^{(k)}(z)$ in (33) with its truncated Taylor series expansion of around the expectation of $z$. From this, the problem of looking for the proper $g_Z^{(k)}(z),\ k=0,1,\ldots,m$, moment functions is the same as finding an optimal order of truncation for each of the expansion:

$$
\begin{aligned}
\hat{\lambda} &= \arg\max_\lambda\prod_{j=1}^{n}\frac{1}{\theta}\exp\left(-\sum_{k=1}^{m}\lambda_k[a_{0k}+a_{1k}z_j+a_{2k}z_j^2+a_{3k}z_j^3+\ldots]\right)\\
&= \arg\max_\lambda\prod_{j=1}^{n}\frac{1}{\theta}\exp\left(\beta_0+\beta_1 z_j+\beta_2 z_j^2+\beta_3 z_j^3+\ldots\right), \qquad (34)
\end{aligned}
$$

where $\beta_0,\ \beta_1,\ \beta_1,\ \beta_3,\ \ldots$ are unknown coefficients to be estimated. We are interested in finding $m$ and the truncation order $l_0$ of the power series in (34). So, the search for $m$ and the moment functions is converted to a search for an optimal truncation orders $m$ and $l_0$, that yields the best fit of $\hat{f}_Z(z;\hat{\lambda})$ to the data. We suggest using Bayesian information criterion (BIC) or Schwartz information criterion (SIC) based on (34) to select the optimal $m$ and $l_0$. Let $\hat{\lambda}_{MLE}$ denote the ML estimator of $\lambda$ using the optimal order of truncation $l_0$. Then, the estimated pdf of $f_Z(z)$ is $\hat{f}_Z(z;\hat{\lambda}_{MLE})$, which can then be used to compute $\widehat{\mathrm{VaR}}_\alpha(Z)$ and $\widehat{\mathrm{ES}}_\alpha(Z)$ in (16).

## 4. Numerical Analysis

In this section we conduct a numerical experiment to analyse the sensitivity of the hedging error with respect to model parameters. In the previous sections we have defined a general jump-diffusion process with the jump component specified by a kernel biased completely random measure. This generalised framework nests a number of very important models in mathematical finance, including, but not limited to, the jump diffusion model of Merton (1976), the generalised gamma process discussed in Lo and Weng (1989), the variance gamma process by Madan et al. (1998), and the CGMY model of Carr et al. (2002). Here, for simplicity, we only use the generalised gamma (GG) process. The analysis can be easily extended to other classes of models or even their Markovian regime switching versions discussed in Fard and Siu (2013).

The GG process generalises several famous models in finance. For example, the inverse gamma (IG) and weighted gamma (WG) processes are special forms of the GG process. Let $\alpha \leq 1$ denote a constant shape parameter and $\delta(t)$ be the time-dependent scale parameter of the GG process. We can express the intensity process of the GG process as:

$$\rho(dz|t)\eta(dt) = \frac{1}{\Gamma(\alpha)\alpha e^{\delta(t)z}z^{\alpha}}dz\eta(dt),$$

where $\Gamma(\cdot)$ is the gamma function. We note that this process simplifies to a WG process when $\alpha = 1$ and to a IG process when $\alpha = \frac{1}{2}$.

Several observations are of order. First, to obtain the GG process, we must set the kernel function $h(z)$ to $cz^q$, for some constants $c$ and $q$, and choose a particular parametric form of the compensator measure. Second, different versions of GG processes are obtained for different sets of the kernel function parameters values. For example, if $c > 0$ and $q = 1$ (linear kernel function), then we obtained the scale-distorted version of the GG process. If $c > 1$, then the jump sizes are overstated, while they are understated if $0 < c < 1$. When $q > 0$ and $c = 1$, we obtain the power-distorted version of the GG process. When $q > 1$, small jump sizes (i.e., $0 < z < 1$) are understated and large jump sizes (i.e., $z > 1$) are overstated. Finally, when $0 < q < 1$, small jump sizes are overstated and large jump sizes are understated.

To simulate the GG process, we use the Poisson weighted algorithm by Lee and Kim (2004). Throughout the simulations, we set at $\alpha = 1/2$. It is worth mentioning that the Poisson weighted algorithm applies to a wide class of completely random measures, which are very difficult to simulate directly, see Lee and Kim (2004).

To implement the algorithm, divide the time interval to maturity $[0, T]$ into $nT$ equally spaced subintervals. Then for each $j = 0, 1, ..., n-1$, let $[t_j, t_{j+1}]$ be the $(j+1)$st subinterval. Let $M$ denote the number of jumps of the completely random measure over the term to maturity, such that, $M$ controls the accuracy of the approximation of the algorithm. To implement the Poisson weighted algorithm, we take the following steps:

Step 1. Let $\aleph$ denote a normalised density function defined by $\aleph^{-1} := \frac{1}{\eta'(t)}\int_0^T \eta'(s)ds$;

Step 2. Generate a set of *i.i.d* random variables $T_i \geq 0$ for each $i = 1, 2..., MT$, from $\aleph$;

Step 3. Generate the jump size $\Omega_i$ from the conditional density function $g_{T_i}$ (i.e., gamma);

Step 4. Evaluate whether $T_i \in [t_j, t_{j+1})$. If yes, calculate:

$$\lambda_i = \frac{\rho^{\theta}(\Omega_i|T_i)}{MTg_{t_i}(\Omega_i)}\int_0^T \eta'(s)ds.$$

$\lambda_i$ is the intensity of the Poisson distribution, used to generate the Poisson weights $\mathcal{W}$.

We provide a comparative analysis of the density function of the hedging error obtained by our ME estimation of the hedging error with that obtained by Monte Carlo simulation. In Monte Carlo simulation, we calculate the hedging error at $t_n$, $\Delta\Pi_n$, $n = 1, \ldots, N$, with hedging portfolio $\Pi_n := \Pi(t_n, S)$ defined by:

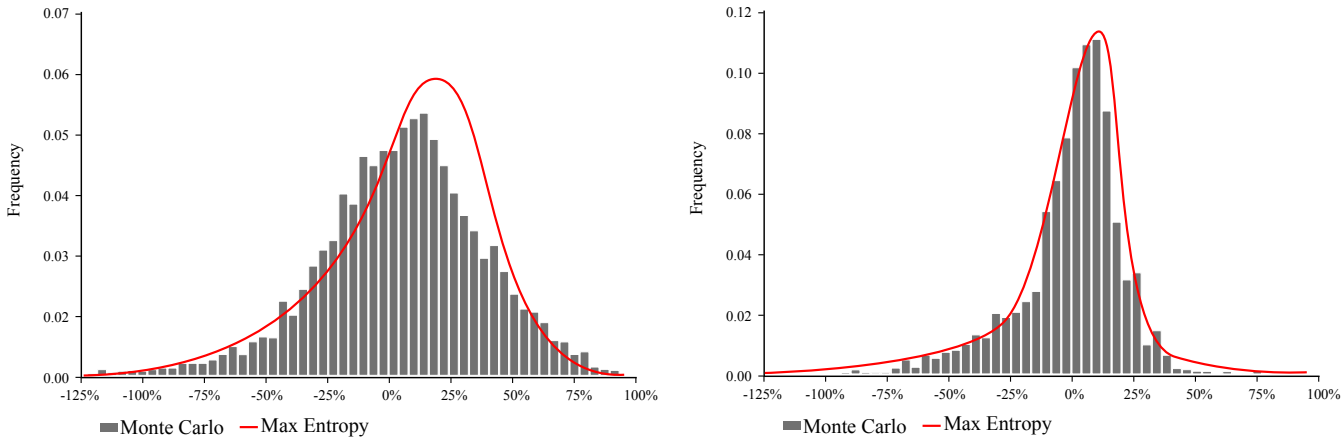
$$\Pi(t_n, s) = S(t_n)H(t_n, S) - C(t_n, S)$$

where the option price and delta are computed by evaluating the risk neutral expectation in Equation (8).

We consider hedging a one-month at the money call option with $S(0) = K = 1$. We analyse different hedging frequencies, namely, $N = 4$ (weekly hedging), $N = 20$ (daily hedging), $N = 80$ (hedging every hour and a half), and $N = 320$ (hedging every 20 min). For brevity, we suppose that the WG process with $c = q = 1$, equal model parameters under $\mathbb{P}$ and $\mathbb{P}^\theta$, zero interest rate, and zero drift for the underlying asset. Furthermore, we assume constant volatility, $\sigma = 18\%$. Then, the set of rebalancing times is $\{t_n\}, t_n = n\delta t, n = 1 \ldots, N$ with $\delta = 0.1$. All computations are fully vectorised in JULIA$^{\text{TM}}$.

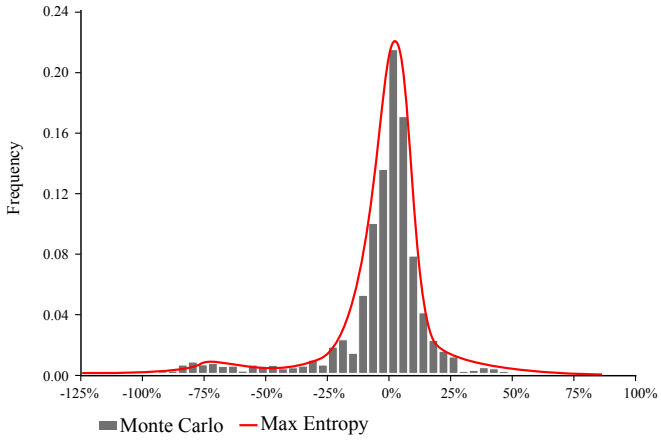
In this setting, the kernel-biased completely random measure on $\mathcal{T}$ is $\kappa(t) = \sum_{i=1}^{MT} h(\Omega_i \mathcal{W}_i) 1_{(T_i \leq t)}$. It can be shown that the process generated by the Poisson weighting algorithm converges in distribution to $\kappa(t)$ defined on $[0, T]$ with the Skorohod topology as $M \to \infty$ (see Lee and Kim 2004). Note that the true values of the truncation lag ($m$) and truncation order ($l_0$) in the entropy density are set at $m = l_0 = 4$. The optimal choice of these parameters with the BIC criterion varies within the pseudo-samples, but the minimum and maximum choice of each parameter are 2 and 5, respectively.

Figure 1 compares the distribution of hedging error obtained from the Monte Carlo Simulation and estimated from the ME method. In Table 1 we report the summary statistics for the distribution of hedging error, as well as the 99% and 95% VaR, obtained by the above two methods. From the numerical analysis, it follows that the distribution of hedging error for a small number of trades has a high volatility and is negatively skewed, which indicates that it is more likely to have large losses than large wins. It is noteworthy that the ME estimation of the distribution fails to sufficiently capture the right tail of the empirical distribution, nevertheless it adequately describes the left tail.

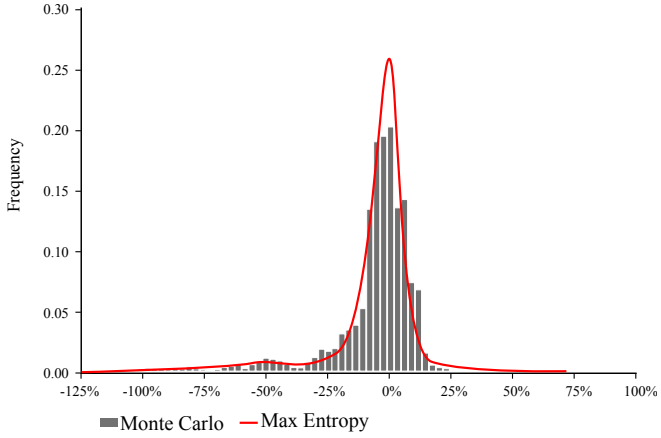



(**a**) Hedging every week ($N = 4$) (**b**) Hedging every day ($N = 20$)

**Figure 1.** *Cont.*




(**c**) Hedging every 1.5 h ($N = 80$)




(**d**) Hedging every 20 min ($N = 320$)

**Figure 1.** Distribution of hedging error of a one-month European call option under the Weighted Gamma, obtained using Monte Carlo simulation and computing Max Entropy estimation.

As we increase the frequency of the delta hedging trades, the distribution becomes more symmetric, and our ME estimation performs better. For $N = 20$, our distribution is strongly leptokurtic, indicating the option writers' large exposure to jumps. When trading frequency increases to 80 and 320, the distribution peaks around zero and the volatility significantly decreases, however, the left tail is still noticeable. From our analysis, it appears that the volatility is negatively related to the frequency of delta hedging trades. However, further research is required to conclusively establish the relationship between the number of trades and moments of the distributions.

**Table 1.** Summary statistics: Moments (Mj for j = 1, 2, 3, 4) of the distribution of hedging error under the GG (generalised gamma) process. VaR: Value-at-risk.

| | N | M1 (%) | M2 (%) | M3 (%) | M4 (%) | 99% VaR | 95% VaR |
|---|---|---|---|---|---|---|---|
| Max Entropy | 4 | 0.173 | 43.21 | −1.97 | 7.01 | 99.87 | 74.38 |
| Max Entropy | 20 | 0.154 | 28.44 | −2.41 | 16.01 | 88.31 | 72.11 |
| Max Entropy | 80 | −0.058 | 26.04 | −3.36 | 19.74 | 77.59 | 51.6 |
| Max Entropy | 320 | −0.041 | 22.14 | −3.99 | 14.32 | 61.33 | 46.31 |
| Monte Carlo | 4 | 0.131 | 42.52 | −1.03 | 6.51 | 101.73 | 75.12 |
| Monte Carlo | 20 | 0.143 | 29.37 | −2.33 | 15.36 | 76.12 | 62.49 |
| Monte Carlo | 80 | −0.072 | 25.41 | −3.21 | 19.85 | 74.06 | 49.33 |
| Monte Carlo | 320 | −0.074 | 23.33 | −4.01 | 18.94 | 64.83 | 48.99 |

## 5. Conclusions

Perfect replication of financial derivatives is not possible, given market incompleteness and discrete-time hedging. We characterised the risk in dynamic hedge of European options through the asymptotic distribution of hedging error. Furthermore, we obtained an estimation for the distribution of hedging error by maximising Shannon's entropy (Shannon (1948)) subject to a set of moment constraints, which in turn yielded the value-at-risk (VaR) and expected shortfall (ES) of the hedging error, two widely-used risk metrics in finance. This new approach chooses the probability distribution with the most uncertainty subject to what is known. Thus we obtained a consistent estimate of the asymptotic distribution of hedging error, despite the non-normality of the underlying distribution of returns. As a result, we derived a very generalised modelling framework, which can be applied in different areas of derivatives pricing. Some parametric specifications of this framework include, but are not limited to, the jump diffusion model of Merton (1976), the generalised gamma process discussed in Lo and Weng (1989), the variance gamma process by Madan et al. (1998), and the CGMY model of Carr et al. (2002).

Finally, we conducted a robust numerical simulation of the result, to highlight the practical applications of our model.

**Author Contributions:** Conceptualization, F.A.F., F.D.T.; methodology, F.A.F., F.D.T.; validation, F.A.F., F.D.T. and S.S.; formal analysis, F.A.F., F.D.T. and S.S.; investigation, F.A.F., F.D.T. and S.S.; resources, F.A.F., F.D.T. and S.S.; data curation, F.A.F., F.D.T. and S.S.; writing—original draft preparation, F.A.F., F.D.T. and S.S.; writing—review and editing, F.A.F., F.D.T. and S.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The project uses no real data other than Monte Carlo simulations.

**Acknowledgments:** The authors would like to thank Mervyn Silvapulle of the Department of Econometrics and Business Statistics at Monash University, for his insightful comments and helpful suggestions. Firmin Doko Tchatoka acknowledges the financial support from the School of Economics, the University of Adelaide.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Proof of Lemma 1

James (2002, 2005) shows that:

$$E\Big[\exp\Big(\int_0^t\int_{\mathcal{R}^+}\theta_s h(z)\widetilde{N}(ds,ds)\Big)\Big|\mathcal{G}_t\Big]=\exp\Big(\int_0^t\int_{\mathcal{R}^+}\Big(e^{\theta_s h(z)}-1+\theta_s h(z)\Big)\rho(dz|s)\eta(ds)\Big).$$

Then, by taking the conditional expectations of (3), the results follow.

**Proof.** Proof for Proposition 1 $\widetilde{S}$ is $\mathcal{F}^A$-adapted, $\widetilde{S}$ is an $(\mathcal{F}^Y,\mathbb{P}^\theta)$-local-martingale if and only if it is a $(\mathcal{G},\mathbb{P}^\theta)$-local-martingale. By Lemma 7.2.2 in Elliott and Kopp (2005), $\widetilde{S}$ is an $(\mathcal{G}^Y,\mathbb{P}^\theta)$-local-martingale if and only if $\Lambda\widetilde{S}:=\{\Lambda(t)\widetilde{S}(t)|t\in\mathcal{T}\}$ is a $(\mathcal{G},\mathbb{P}^\theta)$-local-martingale. First, by Bayes' rule:

$$\begin{aligned}
E^\theta\Big[\exp\Big(-\int_0^t r_s ds\Big)S_t\Big|\mathcal{G}_0\Big] &= \exp\Big(-\int_0^t r_s ds\Big)E\Big[\Lambda_t\exp\Big(\int_0^t dA_u\Big)|\mathcal{G}_0\Big]\\
&= \exp\Big(-\int_0^t r_s ds\Big)\frac{E\Big[\exp\Big(\int_0^t(\theta+1)dA_u\Big)|\mathcal{G}_0\Big]}{\mathcal{M}(\theta)_t}\\
&= \exp\Big(-\int_0^t r_s ds\Big)\frac{\mathcal{M}(\theta+1)_t}{\mathcal{M}(\theta)_t}\\
&= \exp\Big(\int_0^t(\mu_s-r_s-\frac{1}{2}\sigma_s^2)ds+\frac{1}{2}\int_0^t(2\theta+1)\sigma_s^2 ds\\
&+\int_0^t\int_{\mathcal{R}^+}\{e^{\theta h(z)}(e^{h(z)}-1)-h(z)\}\rho d(z|t)\eta(dt)\Big).
\end{aligned}$$

Then by setting time $s=0$, and applying the martingale condition we achieve:

$$\int_0^t(\mu_s-r_s-\frac{1}{2}\sigma_s^2)ds+\frac{1}{2}\int_0^t(1+2\theta)\sigma_s^2+\int_0^t\int_{\mathcal{R}^+}\{e^{\theta h(z)}(e^{h(z)}-1)-h(z)\}\rho d(z|t)\eta'(s)ds=0.$$

Hence, for each $t\in\mathcal{T}$, (5) must hold. □

**Appendix B. Proof for Proposition 2**

Assume that $\mathbb{P} \sim \mathbb{P}^\theta$ with density process $\Lambda_t$. Suppose $\mathcal{Z}_u \in BM(\mathcal{T})$. Then by Bayes' rule:

$$\begin{aligned}
\mathcal{M}_A^\theta(\mathcal{Z})_t :=& E^\theta[e^{(\mathcal{Z}.A)_t}|\mathcal{G}_0] = E[\Lambda_t.e^{(\mathcal{Z}.A)_t}|\mathcal{G}_0] \\
&= \exp\Bigg( \int_0^t \mathcal{Z}(\mu_s - \frac{1}{2}\sigma_s^2)ds + \int_0^t \frac{1}{2}(\mathcal{Z}+\theta)^2\sigma_s^2 ds \\
&\quad + \int_0^t \int_{\mathcal{R}^+} \Big\{ e^{(\mathcal{Z}+\theta)h(z)} - 1 - (\mathcal{Z}+\theta)h(z) \Big\} \rho(dz|s)\eta(ds) \\
&\quad - \int_0^t \frac{1}{2}(\theta.\sigma_s)^2 ds \\
&\quad - \int_0^t \int_{\mathcal{R}^+} \Big\{ e^{\theta h(z)} - 1 - \theta h(z) \Big\} \rho(dz|s)\eta(ds) \Bigg) \\
&= \exp\Bigg( \int_0^t \mathcal{Z}(\mu_s + 2\theta\sigma_s^2 - \frac{1}{2}\sigma_s^2)ds \\
&\quad + \int_0^t \int_{\mathcal{R}^+} (e^{\theta h(z)} - 1)\rho(dz|s)\eta(ds) + \frac{1}{2}\int_0^t \mathcal{Z}^2\sigma_s^2 ds \\
&\quad + \int_0^t \int_{\mathcal{R}^+} \Big\{ e^{\mathcal{Z}h(z)} - 1 - \mathcal{Z}h(z) \Big\} e^{\theta h(z)} \rho(dz|s)\eta(ds) \Bigg).
\end{aligned}$$

Then under $\mathbb{P}^\theta$, (6) holds.

## References


Akhiezer, Naum Il'ich. 1965. *The Classical Moment Problem and Some Related Questions in Analysis*. Edinburgh: Oliver & Boyd.

Andersen, Leif, and Jesper Andreasen. 2000. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research* 4: 231–62. [CrossRef]

Avellaneda, Marco. 1998. Minimum relative entropy calibration of asset pricing models. *International Journal of Theoretical and Applied Finance* 1: 447–72. [CrossRef]

Bakshi, Gurdip, Charles Cao, and Zhaodong Zhong. 2012. Assessing models of individual equity option prices. *SSRN*. [CrossRef]

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. 1997. Empirical performance of alternative option pricing models. *The Journal of Finance* 52: 2003–49. [CrossRef]

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan. 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16: 101–43. [CrossRef]

Bates, David S. 1991. The crash of'87: Was it expected? the evidence from options markets. *The Journal of Finance* 46: 1009–44. [CrossRef]

Bates, David S. 2000. Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics* 94: 181–238.

Bertsimas, Dimitris, Leonid Kogan, and Andrew W Lo. 2000. When is time continuous? *Journal of Financial Economics* 55: 173–204. [CrossRef]

Borwein, Jonathan, Rustum Choksi, and Pierre Marechal. 2003. Probability distributions of assets inferred from option prices via the principle of maximum entropy. *SIAM Journal on Optimization* 14: 464–78. [CrossRef]

Carr, Peter, Hélyette Geman, Dilip B Madan, and Marc Yor. 2002. The fine structure of asset returns: An empirical investigation. *The Journal of Business* 75: 305–33. [CrossRef]

Chan, Felix. 2009. Modelling time-varying higher moments with maximum entropy density. *Mathematics and Computers in Simulation* 79: 2767–78. [CrossRef]

Chang, Bo Young, Peter Christoffersen, and Kris Jacobs. 2013. Market skewness risk and the cross section of stock returns. *Journal of Financial Economics* 107: 46–68. [CrossRef]

Cont, Rama, Peter Tankov, and Ekaterina Voltchkova. 2007. Hedging with options in models with jumps. In *Stochastic Analysis and Applications*. Berlin/Heidelberg: Springer, pp. 197–217.

Cover, Thomas M., and Joy A. Thomas. 2012. *Elements of Information Theory*. Hoboken: John Wiley & Sons.

Davis, Mark H. A. 1997. Option pricing in incomplete market. In *Mathematics of Derivative Securities*. Edited by Michael Alan Howarth Dempster and Stanley R Pliska, Cambridge: Cambridge University Press, pp. 216–26.

Diavatopoulos, Dean, James S. Doran, Andy Fodor, and David R. Peterson. 2012. The information content of implied skewness and kurtosis changes prior to earnings announcements for stock and option returns. *Journal of Banking & Finance* 36: 786–802.

Dotsis, George, Dimitris Psychoyios, and George Skiadopoulos. 2007. An empirical comparison of continuous-time models of implied volatility indices. *Journal of Banking & Finance* 31: 3584–603.
Dupire, Bruno. 1994. Pricing with a smile. *Risk* 7: 18–20.
Elliott, Robert J., and P. Ekkehard Kopp. 2005. *Mathematics of Financial Market*. Berlin/Heidelberg: Springer.
Elliott, Robert J., Leunglung Chan, and Tak Kuen Siu. 2005. Option pricing and Esscher transform under regime switching. *Annals of Finance* 1: 423–32. [CrossRef]
Eraker, Bjørn, Michael Johannes, and Nicholas Polson. 2003. The impact of jumps in volatility and returns. *The Journal of Finance* 58: 1269–300. [CrossRef]
Esche, Felix, and Martin Schweizer. 2005. Minimal entropy preserves the lévy property: How and why. *Stochastic Processes and Their Applications* 115: 299–327. [CrossRef]
Fard, Farzad Alavi, and Tak Kuen Siu. 2012. Pricing and managing risks of European-style options in a Markovian regime-switching binomial model. *Annals of Finance* 1–18. doi:10.1007/s10436-012-0192-3. [CrossRef]
Fard, Farzad Alavi, and Tak Kuen Siu. 2013. Pricing participating products with Markov-modulated jump–diffusion process: An efficient numerical pide approach. *Insurance: Mathematics and Economics* 53: 712–21. [CrossRef]
Föllmer, Hans, and Dieter Sondermann. 1986. Hedging of contingent claims under incomplete information. In *Contributions to Mathematical Economics*. Edited by Werner Hildenbrand and Andreu Mas-Colell. Amsterdam: North-Holland, pp. 205–23.
Föllmer, Hans, and Martin Schweizer. 1991. *Hedging of Contingent Claims under Incomplete Information*. Edited by Mark H. A. Davis and Robert James Elliott. London: Gordon and Breach.
Frittelli, Marco. 2000. The Minimal Entropy Martingale Measure and the valuation problem in incomplete markets. *Mathematical Finance* 10: 39–52. [CrossRef]
Geman, Donald, Hélyette Geman, and Nassim Nicholas Taleb. 2014. Tail risk constraints and maximum entropy. *Entropy* 17: 3724–37. [CrossRef]
Gerber, Hans U., and Elias S.W. Shiu. 1994. Option pricing by Esscher transforms (with discussions). *Transactions of the Society of Actuaries* 46: 99–191.
Gzyl, Henryk, and Silvia Mayoral. 2016. Determination of zero-coupon and spot rates from treasury data by maximum entropy methods. *Physica A: Statistical Mechanics and Its Applications* 456: 38–50. [CrossRef]
Hayashi, Takaki, and Per A. Mykland. 2005. Evaluating hedging errors: An asymptotic approach. *Mathematical Finance* 15: 309–43. [CrossRef]
Heston, Steven L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies* 6: 327–43. [CrossRef]
Hubalek, Friedrich, and Carlo Sgarra. 2006. Esscher transforms and the minimal entropy martingale measure for exponential lévy models. *Quantitative Finance* 6: 125–45. [CrossRef]
Hubalek, Friedrich, Jan Kallsen, and Leszek Krawczyk. 2006. Variance-optimal hedging for processes with stationary independent increments. *The Annals of Applied Probability* 16: 853–85. [CrossRef]
Hull, John, and Alan White. 1987. The pricing of options on assets with stochastic volatilities. *The Journal of Finance* 42: 281–300. [CrossRef]
James, Lancelot. 2002. Poisson process partition calculus with applications to exchangeable models and Bayesian non-parametric. *arXiv* arXiv:math/0205093v1.
James, Lancelot. 2005. Bayesian Poisson process partition calculus with an application to Bayesian Levy moving averages. *The Annals of Statistics* 33: 1771–99. [CrossRef]
Kennedy, J. S., Peter A. Forsyth, and Kenneth R. Vetzal. 2009. Dynamic hedging under jump diffusion with transaction costs. *Operations Research* 57: 541–59. [CrossRef]
Lee, Jaeyong, and Y. Kim. 2004. Sampling methods for NTR processes. Paper presented at the International Workshop for Statistics (SRCCS'04), Lecture Note, Seoul, Korea, June..
Lo, Albert Y., and Chung-Sing Weng. 1989. On a class of Bayesian nonparametric estimates: Hazard rates estimates. *Annals of the Institute of Statistical Mathematics* 41: 227–45. [CrossRef]
Madan, Dilip B., Peter P. Carr, and Eric C. Chang. 1998. The variance gamma process and option pricing. *European Finance Review* 2: 79–105. [CrossRef]
Merton, Robert C. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3: 125–44. [CrossRef]
Mistrulli, Paolo Emilio. 2011. Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns. *Journal of Banking & Finance* 35: 1114–27.
Naik, Vasanttilak, and Moon Lee. 1990. General equilibrium pricing of options on the market portfolio with discontinuous returns. *Review of Financial Studies* 3: 493–521. [CrossRef]
Pilska, Stanley R. 1997. *Introduction to Mathematical Finance*. Malden: Blackwell.
Rosenbaum, Mathieu, and Peter Tankov. 2014. Asymptotically optimal discretization of hedging strategies with jumps. *The Annals of Applied Probability* 24: 1002–48. [CrossRef]
Saghafian, Soroush, and Brian Tomlin. 2016. The newsvendor under demand ambiguity: Combining data with moment and tail information. *Operations Research* 64: 167–85. [CrossRef]

Schweizer, Martin. 1995. On minimal martingale measure and and Follmer-Schweizer decomposition. *Stochastic Analysis and Applications* 13: 573–99. [CrossRef]
Shanahan, Brett, Farzad Alavi Fard, and John van der Hoek. 2016. Pricing participating policies under the meixner process and stochastic volatility. *Scandinavian Actuarial Journal* 2017: 559–583. [CrossRef]
Shannon, Claude Elwood. 1948. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review* 5: 3–55. [CrossRef]
Tankov, Peter. 2003. *Financial Modelling with Jump Processes*. Boca Raton: CRC Press, vol. 2.
Tankov, Peter, and Ekaterina Voltchkova. 2009. Asymptotic analysis of hedging errors in models with jumps. *Stochastic Processes and Their Applications* 119: 2004–27. [CrossRef]
Wang, Xiao-Tian, Zhong-Feng Zhao, and Xiao-Fen Fang. 2015. Option pricing and portfolio hedging under the mixed hedging strategy. *Physica A: Statistical Mechanics and Its Applications* 424: 194–206. [CrossRef]
Xi, Ning, Rachata Muneepeerakul, Sandro Azaele, and Yougui Wang. 2014. Maximum entropy model for business cycle synchronization. *Physica A: Statistical Mechanics and Its Applications* 413: 189–94. [CrossRef]
Xu, Yingying, Zhuwu Wu, Long Jiang, and Xuefeng Song. 2014. A maximum entropy method for a robust portfolio problem. *Entropy* 16: 3401–15. [CrossRef]
Zellner, Arnold, and Richard A. Highfield. 1988. Calculation of Maximum Entropy Distributions and Approximation of Marginal Posterior Distributions. *Journal of Econometrics* 37: 195–209. [CrossRef]
Zhou, Rongxi, Ru Cai, and Guanqun Tong. 2013. Applications of entropy in finance: A review. *Entropy* 15: 4909–31. [CrossRef]