

Quadratic hedging of weather and catastrophe risk by using short term climate predictions

Stefan Ankirchner and Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany

February 12, 2008

Abstract

The extent to which catastrophic weather events occur strongly depends on global climate conditions such as average sea surface temperatures (SST) or sea level pressures. Some of the factors can be predicted up to a year in advance, and should therefore be taken into account in any reasonable management of weather related risk. In this paper we first set up a risk model that integrates climate factors. Then we show how variance minimizing hedging strategies explicitly depend on the factors' prediction. Our analysis is based on a detailed study of the predictable representation property on the combined Poisson and Wiener spaces. Using tools of the stochastic calculus of variations we derive a representation formula of the Clark-Ocone type. Finally, we exemplify the theory developed in a case study of US hurricane risk. We derive hedging strategies taking into account that US hurricane activity strongly depends on the SST of the Pacific Ocean.

2000 AMS subject classifications: primary 91B28, 60H07; secondary 37H10, 60H07.

Key words and phrases: stochastic climate model; mean variance hedging; Poisson random measure; predictable representation; Malliavin calculus; stochastic calculus of variations; difference operator.

Introduction

The intensity with which many weather related events and catastrophes occur is known to be strongly influenced by climate factors. For instance,

the average number of hurricanes observed in the Atlantic basin possesses a very high correlation with climate parameters such as Pacific and Atlantic sea surface temperatures (SST) (see [22]). Rainfall in the African Sahel zone is linked to Atlantic SST. Similarly, there exists a strong connection between precipitation patterns in North and South America and Pacific sea level pressures and SST. Any reasonable assessment of weather and catastrophe risk therefore has to take into account the evolution of correlated climate factors, and will be enhanced by possibly available *predictions* of their values. Climate variables are usually meant to describe long term trends in the environmental conditions on earth. There is a broad variety of characteristic time scales for climate variables ranging from several years to many millennia. For example, the ENSO SST anomaly is observed every 3-8 years, while glaciation cycles are known to take 20 to 100 millennia. Thinking of their immediate economic relevance, due to the fact that lifetimes of financial contracts or budget planning horizons of insurance companies are of the order of a couple of years, only short scale climate variables should matter in our considerations. And forecast intervals should be of the same *short term* order of magnitude, from several months to a couple of years (see [31]). By a *prediction* we understand any forecast with a mean square error that is considerably below the inherent variance of the system.

Climatologists have developed different methods in order to make short term predictions, depending on the factors to be forecasted. For an overview on established prediction methods we refer to the book [31]. One standard method, developed in Penland [26] with particular significance given to the ENSO South Pacific SST anomaly, is based on a low dimensional linear model of the most relevant factors in the framework of random dynamical systems. We follow this approach and therefore use a description of the factor process by a multidimensional linear SDE.

The risk generated by events or even catastrophes based on weather and climate is exogenous from the perspective of usual financial markets. Contracts such as catastrophe bonds or catastrophe (*cat*) derivatives that are written on non-tradable risk of this type therefore can be seen as instruments of *securitization*, serving the task of transferring exogenous risk to capital markets, in which illiquidity is another major issue. Dealing with their pricing and hedging mathematically, one faces an archetypical model of an incomplete financial market. Developing techniques for this purpose has long become a major topic in finance and insurance. They are based on utility indifference methods (Davis [11], [12], Becherer [8], Delbaen et. al. [14]), super-replication (El Karoui, Quenez [16]), quantile hedging (Föllmer, Leukert [17], [18]), or risk measures and optimal design of derivatives (Bar-

riou [5], Barrieu and El Karoui [6], [7]). The problem of hedging a claim written on illiquid assets with a liquid traded proxy, already closer to the setting of this paper, has been studied by several authors. Davis [13], Musiela and Zariphopoulou [24], and Ankirchner, Imkeller and Popier [3] study the HJB equation of the associated optimization problem, discuss numerical and semi-closed form solutions, identify indifference prices in terms of diversification pressure, which depends in a sensitive manner on the correlation between illiquid asset and its liquid proxy, but without using the power of predictions. More recently, Ankirchner, Imkeller and dos Reis [2] follow the BSDE interpretation of the martingale optimality principle to introduce the truly stochastic counterpart of the HJB approach, and describe indifference prices by a system of forward-backward stochastic differential equations (FBSDE).

The main aim of this paper is to include the benefits short term predictions of relevant climate factors can bring to strategies of optimal investment into cat derivatives designed to make climate risk tradable. So the main ingredients of the model to be studied are a *risk process* that, along with the *price dynamics* of the financial market to which risk is supposed to be transferred, contain an essential jump component, described as the integral processes of Poisson type random measures. The third important ingredient is given by the *prediction process* of the relevant climate factors, described by a finite dimensional *linear stochastic differential equation*. It interacts with the risk process through an explicit functional dependence. The technique we use here to deal with the incompleteness of the market model due to non-hedgeable residual risk (*basis risk*), usually called *variance hedging*, looks for those strategies that minimize the mean square distance from the agent's liability. In practice, it is almost impossible to estimate drift coefficients of asset price processes. We will therefore solve the problem from the perspective of a risk neutral measure, under which the cat derivative's price process is a martingale. This has the additional advantage that we do not have to discuss *variance optimal martingale measures* that prove to be very hard to find. The method of mean variance hedging for financial models driven by Lévy noise has been investigated in many papers. We refer to Chapter 10 in [10] for a survey of this issue.

Here is an outline of the paper. We choose an approach of mean variance hedging that is based on the predictable representation property (PRP). In Section 1 we introduce the model and recall how one can derive abstract representations of the optimal variance hedges by means of integral representations of the risk process. In order to make the representations more explicit, in Section 2 we will then study in more detail the PRP on the Pois-

son and Wiener space. We will derive an explicit version of the PRP on the combined Poisson and Wiener spaces, and prove an intrinsic formula of the Clark-Ocone type. With the mathematical tools of the stochastic calculus of variations at hand, in Section 3 we will then proceed to derive explicit formulas for the mean variance hedging strategies. Section 4 is devoted to the example of US hurricane risk, where predictions of pacific SST by linear models are used.

1 The model

Let $(L_t)_{t \geq 0}$ be a risk process that represents the losses or, more generally, determines the financial obligations of an economic agent resulting from catastrophic events. For example L_t may be the aggregate storm claim amount of an insurer between times 0 and t . Alternatively, L may measure the total economic losses due to catastrophes. Moreover, one may think of L as an index that is the underlying of a catastrophe derivative, for example a European call option $(L_T - K)^+$ with maturity $T > 0$ and strike K . Our aim is to derive dynamic strategies that optimally hedge the risky obligations derived from L . For this we want to take into account the predictions of climate indices, such as Pacific and Atlantic SST temperatures, that are heavily correlated with the development of the risk process L . We will show how the hedging strategies explicitly depend on the climate predictions.

We first specify the *prediction model* of the climate factors that influence the catastrophe risk process L .

Throughout let R be a vector climate process. For example R may be the SST anomaly at points of a grid covering an area of an ocean. In order to reduce the dimension of such a system often a 'principal components' analysis (PCA) is used (see [31]). To this end every state is written as a linear combination of orthogonal eigenvectors, the so-called *empirical orthogonal functions (EOF)*. By restricting the system to the eigenvectors with the largest eigenvalues one can often considerably reduce the dimension without losing the basic pattern of the climate process. Note that in this case the process R consists of the coefficients of a linear combination of EOFs.

We assume that the process R is an m -dimensional diffusion. Let us fix an initial time $t \in [0, T)$, as well as an initial state $r \in \mathbb{R}^m$ to be taken by our factor process at this time. Then, conditioned on taking the value r at time t , the factor process is supposed to satisfy the SDE

$$R_s^{t,r} = r + \int_t^s b(u, R_u^{t,r}) du + \int_t^s \gamma(u, R_u^{t,r}) dW_u, \quad (1)$$

where W is a d -dimensional Brownian motion, $s \in [t, \infty)$, $b : \mathbb{R}_+ \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $\gamma : \mathbb{R}_+ \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$. We will assume that the coefficients satisfy a growth and a Lipschitz condition. More precisely, assume that there exists a constant $C \in \mathbb{R}_+$ such that for all $r, r' \in \mathbb{R}^m$ and $t \in \mathbb{R}_+$

$$\begin{aligned} |b(t, r) - b(t, r')| + |\gamma(t, r) - \gamma(t, r')| &\leq C(|r - r'|), \\ |b(t, r)| + |\gamma(t, r)| &\leq C(1 + |r|). \end{aligned} \quad (2)$$

Condition (2) guarantees that there exists a unique solution of (1). It moreover implies that all moments of $R^{t,r}$ are bounded and satisfy the following estimates: For any $p \geq 2$ there exists a constant $C \in \mathbb{R}$ such that for all $r, r' \in \mathbb{R}^m$ we have

$$E \left[\sup_{s \in [t, T]} (1 + |R_s^{t,r}|^p) \right] \leq C(1 + |r|)^p \quad (3)$$

$$E \left[\sup_{s \in [t, T]} |R_s^{t,r} - R_s^{t,r'}|^p \right] \leq C|r - r'|^p \quad (4)$$

(see f.ex. [21], Thm 3.2, p. 340).

A fundamental example: The linear prediction model

Frequently climatologists use linear stochastic models for their predictions (see [26], or Ch.8.7.3 in [31]). Linear models are often able to capture the basic features of climate phenomena, for instance the *principal oscillation patterns (POP)* (see [26]). For example in [27] the SST in the tropical Indian-Pacific ocean is described as a linear dynamical system disturbed by white noise.

We briefly recall how one can use linear models for short-term climate forecasting. For more details and background information we refer to [26]. Let the dynamics of R satisfy a linear SDE of the form

$$dR_s^{t,r} = BR_s^{t,r} ds + \Sigma dW_s, \quad (5)$$

where $B \in \mathbb{R}^{m \times m}$ is the feedback matrix, $\Sigma \in \mathbb{R}^{m \times d}$ the noise amplitude. All the eigenvalues of B are supposed to have negative real parts, which guarantees that $R^{t,r}$ is asymptotically stable.

Let e^B denote the usual matrix exponential. Then the solution of the SDE (5) is given by

$$R_s^{t,r} = e^{B(s-t)} \left(r + \int_t^s e^{-B(u-t)} \Sigma dW_u \right)$$

(see f.ex. Thm 8.2.2 in [4]). This implies that $(R_s^{t,r})_{s \geq t}$ is a Gaussian process, where $R_s^{t,r}$ has the mean $m_s = e^{B(s-t)}r$. Notice that m_s solves the deterministic version of Equation (5),

$$dm_t = Bm_t dt, \quad m_s = r.$$

Moreover, the covariance matrix $K(s) = E[(R_s^{t,r} - m_s)(R_s^{t,r} - m_s)^*]$ satisfies

$$K(s) = \int_t^s e^{B(s-u)} \Sigma \Sigma^* (e^{B(s-u)})^* du$$

(see f.ex. Thm 8.2.6 in [4]). As $s \rightarrow \infty$, the covariance $K(s)$ converges to $\bar{K} = \int_0^\infty e^{Bu} \Sigma \Sigma^* (e^{Bu})^* du$, which coincides with the covariance matrix of R with respect to its stationary distribution.

For a given lime lag $\tau > 0$ observe that

$$E[(R_{\tau+s}^{t,r} - m_{\tau+s})(R_s^{t,r} - m_s)^*] = e^{B\tau} \int_0^s e^{B(s-u)} \Sigma \Sigma^* (e^{B(s-u)})^* du,$$

and hence

$$\lim_{s \rightarrow \infty} E[(R_{\tau+s}^{t,r} - m_{\tau+s})(R_s^{t,r} - m_s)^*] = e^{B\tau} \bar{K}.$$

Let $0 \leq t_1 \leq t_2 \leq \dots$ be a sequence of times such that $\lim_n t_n = \infty$. Then $\lim_n m_{t_n} = 0$, and Birkhoff's ergodic theorem implies that a.s.

$$\lim_n \frac{1}{n} \sum_{i=1}^n R_{t_i+\tau}^{0,r} (R_{t_i}^{0,r})^* = e^{B\tau} \bar{K}. \quad (6)$$

In practice (6) is used as follows to estimate the matrix B . Let $f(j, t_i)$, $i \in \{1, \dots, N\}$, $j \in \{1, \dots, m\}$, be an ensemble of N observations of R at uniform time intervals, i.e. $f(j, t_i)$ is the SST temperature observed at grid point j at time t_i . Let τ be a time lag, with a length of $l \in \mathbb{N}$ time intervals, and let $C(\tau) = (c_{jk}^\tau)$ be the covariance matrix with coefficients $c_{jk}^\tau = \frac{1}{N-l} \sum_{i=1}^{N-l} f(j, t_i) f(k, t_i + \tau)$. Then an estimate of B is given by $\frac{1}{\tau} \log(C(\tau)C(0)^{-1})$.

Provided that the system is in state r at time t , the Maximum Likelihood estimator of R at time s is given by $m_s = e^{B(s-t)}r$. The mean square error of this estimator satisfies $E|R_s^{t,r} - m_s|^2 = \text{tr}(K(s)) = \sum_{i=1}^m K_{ii}(s)$. The fraction

$$\delta(s) = \frac{\text{tr}(K(s))}{\text{tr}(\bar{K})},$$

is called *relative discrepancy* in [26], and measures the forecast skill of the estimator.

The financial model

Next we model the aggregate financial losses that an economic agent faces due to natural catastrophes. We will suppose that the dynamics is an integral process with respect to a random measure.

Let (U, \mathcal{U}) be a standard Borel space and let ν be a σ -finite measure on U . Denote by λ the Lebesgue measure, and let μ be a homogeneous Poisson random measure with compensator $\pi = \lambda \otimes \nu$. By this we mean, in accordance with [19], that μ is an integer-valued random measure such that

- for all $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{U}$ with $\pi(A) < \infty$, the random variable $\mu(\omega; A)$ is Poisson distributed with intensity $\pi(A)$,
- for all A and $B \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{U}$ with $A \cap B = \emptyset$, the random variables $\mu(\omega; A)$ and $\mu(\omega; B)$ are independent.

We denote by (\mathcal{F}_t) the smallest filtration satisfying the usual conditions and containing the filtration generated by μ and the Brownian motion W . The predictable σ -algebra on $\mathbb{R}_+ \times \Omega$ will be denoted by \mathcal{P} . As usual we denote by $L^2(P \otimes \pi)$ the set of all measurable processes which are square integrable relative to the product measure $P \otimes \pi$. If $\varphi \in L^2(P \otimes \pi)$ is $\mathcal{P} \otimes \mathcal{U}$ -measurable, then we may define the stochastic integral of φ relative to the compensated random measure $(\mu - \pi)$,

$$(\varphi * (\mu - \pi))_t = \int_0^t \int_U \varphi(s, z) (\mu - \pi)(ds, dz),$$

as in [19].

Let $\eta : [0, T] \times U \times \mathbb{R}^m \rightarrow \mathbb{R}$ be a measurable mapping, and suppose that the risk process L evolves according to the equation

$$L_t = \int_0^t \int_U \eta(s, z, R_s^{0,r}) \mu(ds, dz). \quad (7)$$

We need to make assumptions such that for almost all ω , the integral in (7) is defined. To this end we suppose that there exists a $p \geq 1$ and a function $a : U \rightarrow \mathbb{R}_+$ with $\int_U a(z) d\nu(z) < \infty$ such that for all $s \in [0, T]$, $z \in U$ and $r \in \mathbb{R}^m$

$$|\eta(s, z, r)| \leq a(z)(1 + |r|^p). \quad (8)$$

Condition (8) and the moment estimate (3) yield

$$E \int_0^T \int_U |\eta(s, z, R_s^{0,r})| \nu(dz) ds \leq \left(\int_U a(z) d\nu(z) \right) E \int_0^T (1 + |R_s^{0,r}|^p) ds < \infty.$$

This implies that $E \int_0^T \int_U |\eta(s, z, R_s^{0,r})| \mu(ds, dz) < \infty$ (see f.ex. Theorem 1.8, Chapter II, [19]), and hence, for almost all ω , the integral in (7) is defined and finite for all $t \in [0, T]$.

Throughout we suppose that the financial obligations of the agent are of the form $g(L_T)$, where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a measurable mapping. In the remainder we will frequently use the abbreviation $F = g(L_T)$.

One may think of different interpretations of g . For example L_T may represent the losses reported to an insurance company, and $g(L_T)$ the losses minus the part that is carried by a reinsurance company. Alternatively, we may interpret L_t as an underlying of a catastrophe derivative. In this case $g(L_T)$ is a European type option with maturity T , underlying L and payoff function g .

Finally we suppose that there exists an exchange traded and liquid financial security, for example a catastrophe bond, that is highly correlated with the risk process L . To simplify the analysis we assume that our agent can borrow money without interest. Moreover we suppose that the price of the security evolves according to the integral equation

$$S_t = S_0 + \int_0^t \rho(s) dW_s + \int_0^t \int_U \sigma(s, z) (\mu - \pi)(ds, dz), \quad (9)$$

where $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ is \mathcal{P} -measurable with $E \int_0^T |\rho|^2(s) ds < \infty$, and $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ is $\mathcal{P} \otimes \mathcal{U}$ -measurable such that $E \int_0^T \int_{\mathbb{R}_0} \sigma^2(s, z) d\nu(z) ds < \infty$. Note that W is the same Brownian motion that also drives the factor process in (1) and μ is the Poisson random measure that determines the risk process in (7).

Let \mathcal{A} be the set of predictable processes $\vartheta : [0, T] \times \Omega \rightarrow \mathbb{R}$ such that $E \int_0^T \vartheta^2(s) (|\rho|^2(s) + \int_U \sigma^2(s, z) d\nu(z)) ds < \infty$. The elements of \mathcal{A} will be called *strategies*. For any $\vartheta \in \mathcal{A}$ we interpret ϑ_t as the number of units of the risky asset that are held in an investment portfolio at time t . Note that for any strategy the stochastic integral $(\vartheta \cdot S)$ with respect to the semimartingale S is defined. Let x be the initial capital, and denote the wealth at time $t \in [0, T]$ by $X_t^\vartheta = x + (\vartheta \cdot S)_t$.

Among all strategies we aim at finding the strategy $\vartheta^* \in \mathcal{A}$ that minimizes the quadratic error between the wealth at maturity T and the pay-off $g(L_T)$, i.e. that satisfies

$$E[(F - X_T^{\vartheta^*})^2] = \min\{E[(F - X_T^\vartheta)^2] : \vartheta \in \mathcal{A}\}.$$

We will refer to ϑ^* as the *optimal mean variance hedging strategy* or simply optimal variance hedge.

We start by deriving an abstract representation of the optimal variance hedge by using the predictable representation property of the Brownian

motion and the Poisson random measure with respect to (\mathcal{F}_t) . If $g(L_T)$ is square integrable, then it is known that there exists a \mathcal{P} -measurable process $\zeta \in L^2(P \otimes \lambda; \mathbb{R}^d)$ and a $\mathcal{P} \otimes \mathcal{U}$ -measurable process $\psi \in L^2(P \otimes \pi)$ such that

$$g(L_T) = E(F) + \int_0^T \zeta(s) dW_s + \int_0^T \int_U \psi(s, z) (\mu - \pi)(ds, dz). \quad (10)$$

As shown in the next result, the optimal variance hedge ϑ^* has a simple representation in terms of integrands appearing in (10).

Theorem 1.1. *Let $g(L_T)$ be square integrable, and ζ and ψ as in (10). Then the optimal variance hedge is given by*

$$\vartheta^*(s) = \frac{\langle \rho(s), \zeta(s) \rangle + \int \sigma(s, z) \psi(s, z) d\nu(z)}{|\rho|^2(s) + \int \sigma^2(s, z) d\nu(z)}.$$

Proof. We first show that $\vartheta^* \in \mathcal{A}$. By using $(a + b)^2 \leq 2a^2 + 2b^2$, and the Cauchy-Schwarz inequality we get

$$\begin{aligned} & E \int_0^T \vartheta^{*2}(s) \left(|\rho|^2(s) + \int_U \sigma^2(s, z) d\nu(z) \right) ds \\ & \leq E \int_0^T \frac{2\langle \rho(s), \zeta(s) \rangle^2 + 2 \left(\int \sigma(s, z) \psi(s, z) d\nu(z) \right)^2}{\left(|\rho|^2(s) + \int \sigma^2(s, z) d\nu(z) \right)^2} \left(|\rho|^2(s) + \int_U \sigma^2(s, z) d\nu(z) \right) ds \\ & \leq 2E \int_0^T \frac{|\rho|^2(s) |\zeta|^2(s)}{|\rho|^2(s) + \int \sigma^2(s, z) d\nu(z)} ds + 2E \int_0^T \frac{\left(\int \sigma(s, z) \psi(s, z) d\nu(z) \right)^2}{|\rho|^2(s) + \int \sigma^2(s, z) d\nu(z)} ds \\ & \leq 2E \int_0^T |\zeta|^2(s) ds + 2E \int_0^T \int_U \psi^2(s, z) d\nu(z) ds \\ & < \infty. \end{aligned}$$

Recalling the abbreviation $F = g(L_T)$, and using the predictable representation formula (10), we have for all $\vartheta \in \mathcal{A}$

$$\begin{aligned} & E \left(X_T^\vartheta - F \right)^2 \\ & = E \left((X_T^\vartheta - x) - (F - E(F)) \right)^2 + (E(F) - x)^2 \\ & = E \left(\int_0^T \vartheta(s) \rho(s) - \zeta(s) dW_s + \int_0^T \int_U \vartheta(s) \sigma(s, z) - \psi(s, z) (\mu - \pi)(dt, dz) \right)^2 \\ & \quad + (E(F) - x)^2 \\ & = E \int_0^T \left[|\vartheta(s) \rho(s) - \zeta(s)|^2 + \int (\vartheta(s) \sigma(s, z) - \psi(s, z))^2 d\nu(z) \right] ds + (E(F) - x)^2. \end{aligned}$$

Among all $\vartheta \in \mathcal{A}$ the strategy ϑ^* minimizes the square error of

$$|\vartheta(s) \rho(s) - \zeta(s)|^2 + \int_U (\vartheta(s) \sigma(s, z) - \psi(s, z))^2 d\nu(z), \text{ for } P \otimes \lambda \text{ a.a. } (\omega, s),$$

and hence ϑ^* satisfies $E(X_T^{\vartheta^*} - F)^2 = \min\{E(X_T^\vartheta - F)^2 : \vartheta \in \mathcal{A}\}$. \square

Our ultimate aim is to find more explicit descriptions of the representing integrands ζ and ψ and thus of the optimal variance hedge. For this purpose we will show in the next section that one can interpret both integrands as variational derivatives. But before, we want to remark that the optimal variance hedge can be seen as a generalisation of the classical Delta hedge. For instance, suppose that the derivative is measurable relative to μ , that $\rho = 0$ and the jump measure ν is a Dirac measure. Then the variance hedge, given by $\vartheta^* = \frac{\psi(s)}{\sigma(s)}$, resembles very much the replicating strategy of the binomial asset pricing model.

2 Representing martingales with Picard's difference operator

Stochastic calculus with respect to Poisson random measures allows for predictable representation. Any square-integrable martingale adapted to the filtration generated by a Poisson random measure can be written as a stochastic integral of a predictable process φ with respect to the compensated random measure. We begin by recalling that the representing integrand φ can be obtained via an annihilation and creation operation introduced by Picard [28], [29]. This provides a Poisson representation formula corresponding to the 'Clark-Ocone' formula on the Wiener space. We then proceed by combining the representation properties for the Poisson and the Wiener space. We provide an explicit representation formula with respect to filtrations that are generated by a Poisson random measure and the Wiener process.

2.1 The pure jump case

We start by recalling some of the definitions and notation introduced in [28], [29]. First we define the canonical space which we will have to work on if we want to use Picard's difference operator.

Let (U, \mathcal{U}) be a standard Borel space and ν a σ -finite measure on U . Let Ω be the set of all measures ω on $\mathbb{R}_+ \times U$ with values in $\mathbb{Z}_+ \cup \{\infty\}$ such that $\omega(\{(t, u)\}) \leq 1$ for all $(t, u) \in \mathbb{R}_+ \times \mathcal{U}$, and $\omega([0, t] \times B) < \infty$ if $t \in \mathbb{R}_+$ and $\nu(B) < \infty$. Ω will be referred to as the *canonical Poisson space*. Let the canonical random measure be defined by

$$\mu(\omega; A) = \omega(A), \quad A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{U},$$

and let P be a measure on Ω such that μ is Poisson random measure with compensator $\pi = \lambda \otimes \nu$. Throughout let (\mathcal{F}_t) be the filtration generated

by μ and, in this subsection here, denote by \mathcal{P} the associated predictable σ -algebra on $\Omega \times \mathbb{R}_+$.

For any pair $(t, z) \in \mathbb{R}_+ \otimes U$ we consider two operations on Ω defined by

$$\varepsilon_{(t,z)}^-(A) = \omega(A \cap \{(t, z)\}^c), \quad \varepsilon_{(t,z)}^+(A) = \varepsilon_{(t,z)}^-(A) + 1_A(t, z).$$

For any function $F : \Omega \rightarrow \mathbb{R}$ let $\Delta F : \Omega \times \mathbb{R}_+ \otimes U \rightarrow \mathbb{R}$ be the mapping defined by

$$\Delta_{(t,z)}F = F \circ \varepsilon_{(t,z)}^+ - F.$$

We will call ΔF *Picard difference* of F (see Def. 1.11 in [29]).

Let F be a bounded random variable with representation $F = E(F) + (\psi * (\mu - \pi))_\infty$, where $\psi \in L^2(P \otimes \pi)$ is predictable. We claim that for fixed $z \in U$, $\psi(t, z)$ is equal to the predictable projection of $\Delta_{(\cdot, z)}F$. For the convenience of the reader we recall the definition of the predictable projection of a stochastic process.

Definition 2.1. (see Theorem 43, Ch. VI, [15]) Let $H : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}$ be a positive or bounded measurable process. It is known that there exists a unique (up to indistinguishability) \mathcal{P} -predictable process H° , called *predictable projection* of H , such that for any predictable stopping time τ we have

$$E[H_\tau 1_{\{\tau < \infty\}} | \mathcal{F}_{\tau-}] = H_\tau^\circ 1_{\{\tau < \infty\}} \quad \text{a.s.}$$

If H is neither bounded nor positive, but if $|H|$ is indistinguishable from a finite process, then one may define the predictable projection of H by setting $H^\circ = (H^+)^\circ - (H^-)^\circ$ (see Remark 44 (f), Ch. VI, [15]).

The next result guarantees that we may always choose a nice version of the Picard difference $\Delta_{(t,z)}F$.

Lemma 2.2. *Let F be a bounded or positive random variable. Then $\Delta F : \Omega \times \mathbb{R}_+ \times U \rightarrow \mathbb{R}$ is bounded and measurable also; and there exists a $\mathcal{P} \otimes \mathcal{U}$ measurable map $\Omega \times \mathbb{R}_+ \times U \rightarrow \mathbb{R}$, denoted by $[\Delta_{(t,z)}F]^{\hat{\circ}}(\omega)$, such that for all $z \in U$, $[\Delta_{(\cdot, z)}F]^{\hat{\circ}}$ is a version of the predictable projection of $\Delta F_{(\cdot, z)}$.*

Proof. This follows immediately from Proposition 3 in [30]. □

The next theorem establishes the Picard difference as the variational derivative of a \mathcal{F}_T -measurable random variable with respect to Poisson random measures.

Theorem 2.3. (*Predictable representation property*) *Let F be bounded and \mathcal{F}_T -measurable. Then*

$$F = E(F) + \int_0^T \int_U [\Delta_{(t,z)}F]^{\hat{\circ}} (\mu - \pi)(dt, dz). \quad (11)$$

In a special case the predictable representation formula (11) is first established in [23] with a proof based on an Ito chaos expansion for Levy processes and a derivative operation defined as a weighted shift of the expansion. It is shown in [1] that (11) holds for arbitrary bounded variables and that it can be proved with very *elementary* methods.

2.2 The general case

Let $\Omega_1 = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ be the d -dimensional Wiener space, W the coordinate process, (\mathcal{F}_t^W) the Wiener filtration, and denote by P_1 the Wiener measure. Besides let Ω_2 be the canonical Poisson space with a measure P_2 making the canonical random measure μ a Poisson random measure with compensator $\pi = \lambda \otimes \nu$. We will work on the product space $\Omega = \Omega_1 \times \Omega_2$, endowed with the filtration $(\mathcal{F}_t) = (\mathcal{F}_t^W \otimes \mathcal{F}_t^\mu)$ and the product measure $P = P_1 \otimes P_2$.

Next we briefly recall the definition of the Malliavin derivative operator. Let $\mathcal{C}_p^\infty(\mathbb{R}^k)$ be the set of all infinitely continuously differentiable functions $f : \mathbb{R}^k \rightarrow \mathbb{R}$ that have derivatives of at most polynomial growth. We denote by \mathcal{S}_p the set of random variables of the form $F = f(\int_0^T h_1 dW, \dots, \int_0^T h_k dW)$, where $k \in \mathbb{N}$, $f \in \mathcal{C}_p^\infty(\mathbb{R}^k)$ and $h_i \in L^2([0, T]^d)$ for all $1 \leq i \leq k$. The *Malliavin derivative* of F is the element in $L^2(\Omega_1 \times [0, T]^d)$ defined by

$$DF = \sum_{i=1}^k h_i \frac{\partial}{\partial x_i} f\left(\int_0^T h_1 dW, \dots, \int_0^T h_k dW\right).$$

One can show that the class of random variables \mathcal{S}_p is a dense set in $L^2(\Omega_1)$ and that D is a closable operator (see Ch. 1 in [25]). As usual, we denote by $\mathbb{D}^{1,2}(\Omega_1)$ the closure of \mathcal{S}_p with respect to the norm

$$\|F\|_{1,2} = \left(E(F^2) + E \int_{[0,T]^d} (DF)^2(s_1, \dots, s_d) ds_1 \dots ds_n \right)^{\frac{1}{2}}.$$

We now address the predictable representation property on the product space Ω . We first show it for products of random variables in $\mathbb{D}^{1,2}(\Omega_1)$ with elements of $L^\infty(\Omega_2)$, i.e. the set of bounded \mathcal{F}_T^μ -measurable functions $\Omega_2 \rightarrow \mathbb{R}$.

Lemma 2.4. *Let $F \in \mathbb{D}^{1,2}(\Omega_1)$ and $G \in L^\infty(\Omega_2)$. Then*

$$FG = E(FG) + \int_0^T (G D_s F)^\diamond dW_s + \int_0^T \int_U (F \Delta_{s,z} G)^\diamond d(\mu - \pi)(s, z).$$

Proof. By the Clark-Ocone formula and Theorem 2.3 we have $F = E(F) + \int_0^T (D_s F)^\varphi dW_s$ and $G = E(G) + \int_0^T \int_U [\Delta_{(t,z)} G]^{\tilde{\varphi}} (\mu - \pi)(dt, dz)$. The product formula implies

$$\begin{aligned} FG &= E(F)E(G) + \int_0^T \left(E(G) + \int_0^s \int_U [\Delta_{(r,z)} G]^{\tilde{\varphi}} (\mu - \pi)(dr, dz) \right) (D_s F)^\varphi dW_s \\ &\quad + \int_0^T \int_U \left(E(F) + \int_0^s (D_r F)^\varphi dW_r \right) [\Delta_{(s,z)} G]^{\tilde{\varphi}} (\mu - \pi)(ds, dz) \\ &= E(F)E(G) + \int_0^T (G1_{[0,T]})^\varphi (D_s F)^\varphi dW_s \\ &\quad + \int_0^T \int_U (F1_{[0,T]})^\varphi [\Delta_{(s,z)} G]^{\tilde{\varphi}} (\mu - \pi)(ds, dz). \end{aligned}$$

Since F and G are independent, and predictable projections and multiplications can be interchanged we obtain

$$FG = E(FG) + \int_0^T (G D_s F)^\varphi dW_s + \int_0^T \int_U (F \Delta_{(s,z)} G)^{\tilde{\varphi}} (\mu - \pi)(ds, dz).$$

□

Let $\widehat{\mathcal{L}}$ be the set of linear combinations of products FG where $F \in \mathbb{D}^{1,2}(\Omega_1)$ and $G \in L^\infty(\Omega_2)$. We next extend the Malliavin derivative and the Picard difference operator to the product space Ω in the following way. For any random variable $F : \Omega \rightarrow \mathbb{R}$, and $\omega = (\omega_1, \omega_2) \in \Omega$ let $\Delta_{(t,z)} F(\omega_1, \omega_2) = F(\omega_1, \varepsilon_{(t,z)}^+ \omega_2) - F(\omega_1, \omega_2)$.

For any $H \in \widehat{\mathcal{L}}$ let

$$\|H\| = \left(E(H^2) + E \int_{[0,T]^d} (DH)^2 d\lambda^d \right)^{\frac{1}{2}}, \quad (12)$$

and observe that this norm does not include the ΔH term. The reason is that the variational differentiability on the Poisson component of our space does not require additional smoothness beyond square integrability. We denote by \mathcal{L} the set of all random variables H for which there exists a sequence (H^n) in $\widehat{\mathcal{L}}$ such that H^n converges to H in $L^2(P)$, and the derivatives DH^n converge to a limit X in $L^2(\Omega \times [0, T]^d)$. For any $H \in \mathcal{L}$ we define $DH = X$ if X is in $L^2(\Omega \times [0, T]^d)$.

Lemma 2.5. *D is a uniquely defined operator on \mathcal{L} .*

Proof. Let $C_b^\infty(\mathbb{R}^k)$ be the set of all infinitely continuously differentiable functions $f : \mathbb{R}^k \rightarrow \mathbb{R}$ such that f and all its derivatives are bounded. Let \mathcal{S}_b

denote the set of random variables of the form $f(\int_0^T h_1 dW, \dots, \int_0^T h_k dW)$, where $k \geq 1$, $f \in \mathcal{C}_b^\infty(\mathbb{R}^k)$ and $h_i \in L^2([0, T]^d)$. Note that \mathcal{S}_b is dense in $L^2(\Omega_1)$. Let (H^n) be a sequence in $\widehat{\mathcal{L}}$ such that $\lim_n E(H^n)^2 = 0$ and DH^n has an $L^2(\Omega \times [0, T]^d)$ limit X . We have to show that $X = 0$. To this end we prove for any bounded and adapted $\phi \in L^2(\Omega \times [0, T]^d)$, that $\lim_n E(\int_{[0, T]^d} \phi DH^n d\lambda^d) = 0$. This implies that $E(\int_{[0, T]^d} \phi X d\lambda^d) = 0$ for all adapted $\phi \in L^2(\Omega \times [0, T]^d)$, and hence $X = 0$.

Integration by parts (see f.ex. Lemma 1.2.2, [25]) yields

$$E\left(\int_{[0, T]^d} \phi DH^n d\lambda^d\right) = E\left(H^n \int_0^T \phi_s dW_s\right).$$

From boundedness of ϕ , together with the Cauchy-Schwarz inequality we get $\lim_n E\left(H^n \int_0^T \phi_s dW_s\right) = 0$. Hence the result follows. \square

For any $H \in \mathcal{L}$ we define $\|H\|$ as in Equation (12). It is straightforward to show that $(\mathcal{L}, \|\cdot\|)$ is a Banach space.

Theorem 2.6. *Let $H \in \mathcal{L}$ be \mathcal{F}_T -measurable. Then*

$$H = E(H) + \int_0^T (DH)^\varrho dW + \int_0^T \int_U (\Delta_{s,z} H)^{\widehat{\varphi}} (\mu - \pi)(ds, dz).$$

Proof. Let H^n be a sequence in $\widehat{\mathcal{L}}$ such that $\lim_n \|H^n - H\| = 0$. By Lemma 2.4 the result is true for every H^n . Convergence of each summand implies the statement also for H . \square

3 Making hedging strategies explicit

With the representation property shown in the previous section we are now in a position to refine the optimal variance hedge formulas obtained in Section 1. In the remainder we will always work on the canonical space introduced in Section 2.2. On this space we will consider the model introduced in Section 1. So again let $R^{t,r}$ denote the factor process solving a SDE with coefficients satisfying (1) and (2), let the risk process L satisfy (7), and let S be the price of a correlated risky asset with dynamics (9).

As an immediate corollary of Theorem 2.6 we obtain the following refinement of Theorem 1.1.

Theorem 3.1. *Let $g(L_T) \in \mathcal{L}$. Then the optimal variance hedge satisfies*

$$\vartheta^*(s) = \frac{\rho(s)[D_s g(L_T)]^\varrho + \int \sigma(s, z)[\Delta_{(s,z)} g(L_T)]^{\widehat{\varphi}} d\nu(z)}{\rho^2(s) + \int \sigma^2(s, z) d\nu(z)}.$$

In the next two sections we are going to derive explicit representations of the predictable projection of the Picard difference $\Delta_{(t,z)}g(L_T)$ and the Malliavin derivative $D_tg(L_T)$. To this end we define for all $t \in [0, T]$, $x \in \mathbb{R}$ and $r \in \mathbb{R}$ the auxiliary function

$$h(t, x, r) = E \left[g \left(x + \int_t^T \int_U \eta(s, z, R_s^{t,r}) \mu(ds, dz) \right) \right]. \quad (13)$$

The function h will be called *value function* since $h(t, x, r)$ is the expected derivative pay-off conditioned on $L_t = x$ and the factor process to be in r at time t . Notice that Condition (8) guarantees that h is defined and finite, and that the Markov property of (L, R) implies that $E[g(L_T)|\mathcal{F}_t] = h(t, L_t, R_t)$.

Before we can simplify the predictable projections of $\Delta_{(t,z)}g(L_T)$ and $D_tg(L_T)$, we first have to show some regularity properties of the value function h . This will indeed take most of our remaining efforts.

We remark that if h satisfies sufficiently strong smoothness properties, then the optimal quadratic hedge can be determined without using the representation in terms of the Picard difference and the Malliavin trace as provided by Theorem 2.6. More precisely assume that h is continuously differentiable in t and twice continuously differentiable in r ; and that $\kappa(t, z) = h(t, L_{t-} + \eta(t, z, R_t), R_t) - h(t, L_{t-}, R_t)$ satisfies $\int_0^T \int_U |\kappa(s, z)| d\nu(z) ds < \infty$ almost surely. Then one can derive from Ito's formula that $g(L_T) = E(F) + \int_0^T \frac{\partial h^2}{\partial r^2}(s, L_{s-}, R_s) \rho(s, R_s) dW_s + \int_0^T \int_U \kappa(s, z) (\mu - \pi)(ds, dz)$, from where one can deduce the optimal variance hedge in a similar way as in Chapter 10.4, [10]. In the following we will not use Ito's formula and therefore we need only weaker assumptions on h .

3.1 Calculating the Picard difference

In this subsection we are going to derive an explicit representation of the Picard difference $\Delta_{(t,z)}g(L_T)$.

We first have to show some continuity properties of the value function h . This will be done with the help of (3) and (4), and further moment estimates of the factor process $R^{0,r}$. Recall that the estimates (3), (4) and the Markov property of the factor process $R^{t,r}$ imply that for all $p \geq 2$ there exists a constant C' such that for all $r \in \mathbb{R}^m$ and $t, t' \in [0, T]$

$$E \left[\sup_{s \in [t \vee t', T]} |R_s^{t',r} - R_s^{t,r}|^p \right] \leq C'(1 + |r|)^p |t - t'| \quad (14)$$

(for a rigorous proof of this estimate we refer to Lemma 4.5.6 in [20]).

We next give sufficient conditions for the value function h , defined in (13), to be continuous.

Lemma 3.2. *Let g be Lipschitz continuous. Then h is Lipschitz continuous in x . If in addition there exists a constant $L \in \mathbb{R}_+$ such that for all $s \in [0, T]$, $z \in U$ and $r, r' \in \mathbb{R}^m$ we have*

$$\int_U |\eta(s, z, r) - \eta(s, z, r')| d\nu(z) \leq L|r - r'|, \quad (15)$$

then h is also Lipschitz continuous in r , and continuous in t .

Proof. Let $t \in [0, T]$, $x \in \mathbb{R}$ and $r, r' \in \mathbb{R}^m$. Since g is Lipschitz, there exists an $M \in \mathbb{R}_+$ with

$$\begin{aligned} |h(t, x, r) - h(t, x, r')| &\leq ME \left[\int_t^T \int_U |\eta(s, z, R_s^{t,r}) - \eta(s, z, R_s^{t,r'})| \mu(ds, dz) \right] \\ &\leq ME \left[\int_t^T \int_U |\eta(s, z, R_s^{t,r}) - \eta(s, z, R_s^{t,r'})| d\nu(z) ds \right]. \end{aligned}$$

Let C be as in Equation (4) for $p = 2$. Then Condition (15) and Hölder's inequality imply

$$\begin{aligned} |h(t, x, r) - h(t, x, r')| &\leq ML E \left[\int_t^T |R_s^{t,r} - R_s^{t,r'}| ds \right] \\ &\leq M m(T-t)^{\frac{1}{2}} \left(E \int_t^T (R_s^{t,r} - R_s^{t,r'})^2 ds \right)^{\frac{1}{2}} \\ &\leq M m T C |r - r'|, \end{aligned}$$

and hence the Lipschitz continuity in r .

Finally let $\delta \neq 0$ such that $t + \delta \in [0, T]$. Let $a : U \rightarrow \mathbb{R}_+$ be as in Condition (8), $m = \int_U a(z) d\nu(z)$, and observe that

$$\begin{aligned} &|h(t + \delta, x, r) - h(t, x, r)| \\ &\leq ME \left| \int_{t+\delta}^T \int_U \eta(s, z, R_s^{t+\delta,r}) \mu(ds, dz) - \int_t^T \int_U \eta(s, z, R_s^{t,r}) \mu(ds, dz) \right| \\ &\leq ME \left[\int_{t+\delta}^T \int_U |\eta(s, z, R_s^{t+\delta,r}) - \eta(s, z, R_s^{t,r})| \mu(ds, dz) + \int_t^{t+\delta} \int_U |\eta(s, z, R_s^{t,r})| \mu(ds, dz) \right] \\ &\leq ME \left[L \int_0^T |R_s^{t+\delta,r} - R_s^{t,r}| ds + \int_t^{t+\delta} \int_U a(z) (1 + |R_s^{t,r}|^p) d\nu(z) ds \right] \\ &\leq M(L + m) \left\{ \int_0^T E |R_s^{t+\delta,r} - R_s^{t,r}| ds + \int_t^{t+\delta} E (1 + |R_s^{t,r}|^p) ds \right\}. \end{aligned}$$

The estimates (3) and (14) further imply

$$|h(t + \delta, x, r) - h(t, x, r)| \leq M(L + m)(C' + C)T(1 + |r|)^p |\delta|,$$

which shows that h is continuous in t . □

Theorem 3.3. *Let g be bounded and Lipschitz continuous, and suppose that η satisfies (15) and (8). Then the predictable projection of the Picard difference $\Delta_{(t,z)}g(L_T)$ is given by*

$$[\Delta_{(t,z)}g(L_T)]^{\tilde{\phi}} = h(t, L_{t-} + \eta(t, z, R_t), R_t) - h(t, L_{t-}, R_t).$$

Proof. Let $\zeta(s)$ and $\psi(s, z)$ be the predictable processes such that $F - E(F) = \int_0^T \zeta(s) dW_s + \int_0^T \int_{\mathbb{R}_0} \psi(s, z) (\mu - \pi)(ds, dz)$. Since $h(\cdot, L, R)$ is a martingale, for all $t \leq T$ we have $\psi(s, z) = (\Delta_{(s,z)}h(t, L_t, R_t))^{\tilde{\phi}}$, $P \otimes \pi$ -a.s. on $\Omega \times [0, t] \times U$.

Now let for any partition $\Gamma : 0 = t_0 < t_1 < \dots < t_n = T$,

$$Z_{(t,z)}^\Gamma = \sum_{i=1}^n 1_{]t_{i-1}, t_i]}(t) \Delta_{(t,z)}h(t_i, L_{t_i}, R_{t_i}).$$

It follows that $\psi(t, z) = \left(Z_{(t,z)}^\Gamma \right)^{\tilde{\phi}}$. Moreover observe that

$$\begin{aligned} \Delta_{(t,z)}h(r, L_r, R_r) &= h(r, L_r \circ \varepsilon_{(t,z)}^+, R_r) - h(r, L_r, R_r) \\ &= 1_{[0,r]}(t) [h(r, L_r + \eta(t, z, R_t), R_r) - h(r, L_r, R_r)], \end{aligned}$$

$P \otimes \pi$ -a.s. By Lemma 3.2, h is continuous in t, r and x , and consequently, for any sequence of partitions $(\Gamma^n)_n$ of $[0, T]$ with meshsizes $|\Gamma^n|$ tending to 0, the sequence $(Z^{\Gamma^n})_n$ converges to $Z_{(t,z)} = h(t, L_t + \eta(t, z, R_t), R_t) - h(t, L_t, R_t)$. For all $t \in [0, T]$, we have almost surely $L_t = L_{t-}$, and hence, for every fixed $z \in U$, the predictable projection of the process $Z_{(t,z)}$ is given by the left continuous version $h(t, L_{t-} + \eta(t, z, R_t), R_t) - h(t, L_{t-}, R_t)$. Finally, since Z^{Γ^n} is uniformly bounded, we may interchange predictable projections and limits (see f.ex. p. 104, [15]), which implies

$$\begin{aligned} \psi(s, z) &= \lim_n \left(Z_{(t,z)}^{\Gamma^n} \right)^{\tilde{\phi}} = \left(\lim_n Z_{(t,z)}^{\Gamma^n} \right)^{\tilde{\phi}} \\ &= h(t, L_{t-} + \eta(t, z, R_t), R_t) - h(t, L_{t-}, R_t), \end{aligned}$$

and hence the result. \square

3.2 Sufficient criteria for the Malliavin differentiability of the derivative

In this section we will give sufficient conditions for $g(L_T)$ to belong to \mathcal{L} and we will show how to simplify the predictable projection of the Malliavin derivative $D_t g(L_T)$.

First we remark that the growth and Lipschitz condition (2) guarantee that the factor process $R^{t,r}$, defined in (1), is Malliavin differentiable. Moreover, the Malliavin gradient has a representation involving, for (t, r) fixed, the global flow on the space of nonsingular linear operators $\Phi^{t,r}$ on \mathbb{R}^m defined by the equation

$$\Phi_s^{t,r} = I_m + \int_t^s \nabla_r b(u, R_u^{t,r}) \Phi_u^{t,r} du + \int_t^s \nabla_r \gamma(u, R_u^{t,r}) \Phi_u^{t,r} dW_u, \quad s \geq t.$$

Here $\nabla_r b$ and $\nabla_r \gamma$ describe the gradients of b resp. γ existing in the weak sense under (2), I_m the $m \times m$ unit matrix. The Malliavin gradient is then given by the formula (see Nualart [25], p. 126)

$$D_\vartheta R_s^{t,r} = \Phi_s^{t,r} (\Phi_\vartheta^{t,r})^{-1} \gamma(\vartheta, R_\vartheta^{t,r}), \quad t \leq \vartheta \leq s. \quad (16)$$

Theorem 3.4. *Let η be continuously differentiable in r and suppose that there exists a real $p \geq 1$ and a function $b : U \rightarrow \mathbb{R}_+$ with $\int_U (b(z) \vee b^2(z)) d\nu(z) < \infty$ such that for all $s \in [0, T]$, $z \in U$, and $r, r' \in \mathbb{R}^m$ we have*

$$|\eta(s, z, r) - \eta(s, z, r')| \leq b(z) |r - r'|. \quad (17)$$

Then $L_t \in \mathcal{L}$ for all $t \in [0, T]$ and

$$D_\vartheta L_t = \int_0^t \int_U \nabla_r \eta(t, z, R_s^{0,r}) D_\vartheta R_s^{0,r} \mu(ds, dz).$$

Proof. Let $t \in [0, T]$. First note that the Lipschitz property (17) implies that $\eta(t, z, R_t^{0,r})$ is Malliavin differentiable (see Proposition 1.2.4 in [25]).

For any partition $\Gamma : 0 = t_0 \leq t_1 \leq \dots \leq t_n = t$ let

$$A_t^\Gamma = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_U \eta(s, z, R_{t_i}^{0,r}) \mu(\omega; ds, dz).$$

Then it is straightforward to show that A_t^Γ belongs to \mathcal{L} and $D_\vartheta A_t^\Gamma = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_U \nabla_r \eta(s, z, R_{t_i}^{0,r}) D_\vartheta R_{t_i}^{0,r} \mu(\omega; ds, dz)$. Moreover, with $m = \int_U (b(z) \vee$

$b^2(z)d\nu(z)$,

$$\begin{aligned}
E \int_0^T (D_\vartheta A_t^\Gamma)^2 d\vartheta &\leq E \int_0^T \left(\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_U b(z) |D_\vartheta R_{t_i}^{0,r}| \mu(\omega; ds, dz) \right)^2 d\vartheta \\
&\leq E \int_0^T \left(\int_0^t \int_U b(z) \sup_{s \geq \vartheta} |D_\vartheta R_s^{0,r}| \mu(\omega; ds, dz) \right)^2 d\vartheta \\
&\leq E \int_0^T 2 \left(\int_0^t \int_U b(z) \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}| (\mu - \pi)(ds, dz) \right)^2 d\vartheta \\
&\quad + E \int_0^T 2 \left(\int_0^t \int_U b(z) \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}| \pi(ds, dz) \right)^2 d\vartheta \\
&\leq 2E \int_0^T \int_0^t \int_U b^2(z) \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}|^2 d\nu(z) ds d\vartheta \\
&\quad + 2E \int_0^T \left(\int_0^t \int_U b(z) \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}| d\nu(z) ds \right)^2 d\vartheta \\
&\leq 2(mT + m^2 T^2) E \int_0^T \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}|^2 d\vartheta.
\end{aligned}$$

It is known that $\sup_{\vartheta \in [0, T]} E \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}|^2 < \infty$ (see Theorem 2.2.1 in [25]), and hence the collection $(D_\vartheta A_t^\Gamma)$, Γ a finite partition of $[0, t]$, is bounded in $L^2(\Omega \times [0, T])$.

Observe that for any partition $\Gamma : 0 = t_0 \leq t_1 \leq \dots \leq t_n = t$ we have

$$\left| \sum_{i=1}^n 1_{[t_{i-1}, t_i]}(s) \nabla_r \eta(s, z, R_{t_i}^{0,r}) D_\vartheta R_{t_i}^{0,r} \right| \leq b(z) \sup_{v \geq \vartheta} |D_\vartheta R_v^{0,r}|. \quad (18)$$

Now let Γ_n be a sequence of partitions of $[0, t]$ with mesh sizes converging to 0. It is straightforward to show that the sequence $A_t^{\Gamma_n}$ converges to L_t in $L^2(\Omega)$. Inequality (18), the continuity of ∇_r in r and an application of the dominated convergence theorem imply that, for almost all ϑ , ω and t , we have $\lim_n D_\vartheta A_t^{\Gamma_n} = \int_0^t \int_U \nabla_r \eta(s, z, R_s^{0,r}) D_\vartheta R_s^{0,r} \mu(\omega; ds, dz)$. Since the sequence $(D_\vartheta A_t^{\Gamma_n})$ is bounded in $L^2(\Omega \times [0, T])$, it also converges in $L^2(\Omega \times [0, T])$ and thus, by the very definition, L_t belongs to \mathcal{L} and $D_\vartheta L_t = \int_0^t \int_U \nabla_r \eta(t, z, R_s^{0,r}) D_\vartheta R_s^{0,r} \mu(ds, dz)$. \square

Corollary 3.5. *If g is Lischitz continuous and if the assumptions of Theorem 3.4 are satisfied, then $g(L_T)$ belongs to \mathcal{L} .*

Proof. This follows from Proposition 1.2.4 in [25]. \square

We next give sufficient conditions for h to be differentiable in r . First we have to impose stronger conditions on the coefficients b and γ of the SDE (1). Besides (2) we assume that there exists a constant $C \in \mathbb{R}_+$ such that for all $r, r' \in \mathbb{R}^m$ and $t \in [0, T]$

$$\begin{aligned} |\nabla_r b(t, r) - \nabla_r b(t, r')| + |\nabla_r \gamma(t, r) - \nabla_r \gamma(t, r')| &\leq C(|r - r'|), \\ |\nabla_r b(t, r)| + |\nabla_r \gamma(t, r)| &\leq C. \end{aligned} \quad (19)$$

Condition (19) implies that there exists a version of $R^{t,r}$ that is differentiable in r , and the derivative $\nabla_r R^{t,r}$ is a solution of the SDE

$$\nabla_r R_s^{t,r} = I_m + \int_t^s \nabla_r b(u, R_u^{t,r}) \nabla_r R_u^{t,r} du + \int_t^s \nabla_r \gamma(u, R_u^{t,r}) \nabla_r R_u^{t,r} dW_u, \quad s \in [t, T],$$

(see f.ex [21], Thm3.4, p. 346). Moreover, from standard moments estimates (see f.ex. [21], Thm 3.2, p. 340) it follows that for any $p \geq 2$ there exists a constant $C \in \mathbb{R}$ such that for all $r, r' \in \mathbb{R}^m$ we have

$$E \left[\sup_{s \in [t, T]} (1 + |\nabla_r R_s^{t,r}|^p) \right] \leq C(1 + |r|)^p \quad (20)$$

$$E \left[\sup_{s \in [t, T]} |\nabla_r R_s^{t,r} - \nabla_r R_s^{t,r'}|^p \right] \leq C|r - r'|^p. \quad (21)$$

It is known that not only the derivatives, but also the difference quotients have bounded moments. In the following lemma we collect some estimates we will need later.

Lemma 3.6. *Suppose (2) and (19) hold. Let $1 \leq i \leq m$ and denote by e_i the unit vector in \mathbb{R}^m in the direction of the i th coordinate. For any $\delta \in \mathbb{R} \setminus \{0\}$ let $\xi^{t,r,\delta} = \frac{1}{\delta}(R^{t,r} - R^{t,r+\delta e_i})$. Then for all $p \geq 2$ there exists a constant $C \in \mathbb{R}_+$ such that for any $\delta \in \mathbb{R} \setminus \{0\}$ and $r, r' \in \mathbb{R}^m$ we have*

$$E \left[\sup_{s \in [t, T]} |\xi^{t,r,\delta}|^p \right] \leq C, \quad (22)$$

$$E \left[\sup_{s \in [t, T]} |\xi^{t,r,\delta} - \xi^{t,r',\delta}|^p \right] \leq C|r - r'|^p. \quad (23)$$

Proof. See for example [21], Theorem 3.3, p. 342. □

The estimates of $R^{t,r}$ and its derivatives allow to show that the auxiliary function h is differentiable in r .

Lemma 3.7. *Suppose (2) and (19) hold. Let g be Lipschitz continuous and differentiable, and assume that ∇g is Lipschitz continuous also. Suppose that there exists a measurable function $b : U \rightarrow \mathbb{R}_+$ with $\int_U (b(z) \vee b^2(z)) d\nu(z) < \infty$ such that for all $s \in [0, T]$, $z \in U$ and $r, r' \in \mathbb{R}$ we have*

$$|\eta(s, z, r) - \eta(s, z, r')| \leq b(z)|r - r'| \quad (24)$$

$$|\nabla_r \eta(s, z, r) - \nabla_r \eta(s, z, r')| \leq b(z)|r - r'|. \quad (25)$$

Then h is differentiable in r and

$$\begin{aligned} \nabla_r h(t, x, r) &= E \left[\nabla g \left(x + \int_t^T \int_U \eta(s, z, R_s^{t,r}) \mu(ds, dz) \right) \right. \\ &\quad \left. \times \int_t^T \int_U \nabla_r \eta(s, z, R_s^{t,r}) \nabla_r R_s^{t,r} \mu(ds, dz) \right]. \end{aligned}$$

Proof. Fix t and x , and set $A^r = x + \int_t^T \int_U \eta(s, z, R_s^{t,r}) \mu(ds, dz)$ and $B^r = \int_t^T \int_U \nabla_r \eta(s, z, R_s^{t,r}) \nabla_r R_s^{t,r} \mu(ds, dz)$. Then, the Lipschitz continuity and boundedness of ∇g imply that there exist constants C_1, C_2, \dots such that for all $\delta \neq 0$ we have

$$\begin{aligned} & \left| \frac{h(t, x, r + \delta) - h(t, x, r)}{\delta} - E(\nabla g(A^r)B^r) \right| \\ & \leq E \left| \frac{A^{r+\delta} - A^r}{\delta} \int_0^1 \nabla g(A^r + \vartheta(A^{r+\delta} - A^r)) d\vartheta - \nabla g(A^r)B^r \right| \\ & \leq E \left| \left(\frac{A^{r+\delta} - A^r}{\delta} - B^r \right) \int_0^1 \nabla g(A^r + \vartheta(A^{r+\delta} - A^r)) d\vartheta \right| \\ & \quad + E \left| B^r \int_0^1 [\nabla g(A^r + \vartheta(A^{r+\delta} - A^r)) - \nabla g(A^r)] d\vartheta \right| \\ & \leq C_1 \left\{ E \left| \frac{A^{r+\delta} - A^r}{\delta} - B^r \right| + E \left| B^r (A^{r+\delta} - A^r) \right| \right\} \quad (26) \end{aligned}$$

We have to show that the two summands in (26) converge to 0 as $\delta \rightarrow 0$. We start with the first one. To this end let $\xi^{t,r,\delta} = \frac{1}{\delta}(R^{t,r} - R^{t,r+\delta})$ and observe that

$$\frac{\eta(s, z, R_s^{t,r+\delta}) - \eta(s, z, R_s^{t,r})}{\delta} = \xi_s^{t,r,\delta} \int_0^1 \nabla_r \eta(s, z, R_s^{t,r} + \vartheta(R_s^{t,r+\delta} - R_s^{t,r})) d\vartheta.$$

Then

$$\begin{aligned}
& E \left| \frac{A^{r+\delta} - A^r}{\delta} - B^r \right| \\
&= E \left| \int_t^T \int_U \left(\frac{\eta(s, z, R_s^{t, r+\delta}) - \eta(s, z, R_s^{t, r})}{\delta} - \nabla_r \eta(s, z, R_s^{t, r}) \nabla_r R_s^{t, r} \right) \mu(ds, dz) \right| \\
&\leq E \int_t^T \int_U \int_0^1 \left| \xi_s^{t, r, \delta} \left(\nabla_r \eta(s, z, R_s^{t, r} + \vartheta(R_s^{t, r+\delta} - R_s^{t, r})) - \nabla_r \eta(s, z, R_s^{t, r}) \right) \right| d\vartheta \mu(ds, dz) \\
&\quad + E \int_t^T \int_U \left| (\xi_s^{t, r, \delta} - \nabla_r R_s^{t, r}) \nabla_r \eta(s, z, R_s^{t, r}) \right| \mu(ds, dz)
\end{aligned}$$

We show separately that the two summands on the RHS of the preceding inequality converge to 0 as $\delta \rightarrow 0$. Due to (25), and the estimates (22) and (4), the first summand is bounded by

$$\begin{aligned}
& C_2 E \left[\int_t^T |\xi_s^{t, r, \delta}| |R_s^{t, r+\delta} - R_s^{t, r}| ds \right] \\
&\leq C_3 \left(E \left[\int_t^T (\xi_s^{t, r, \delta})^2 ds \right] \right)^{\frac{1}{2}} \left(E \left[\int_t^T |R_s^{t, r+\delta} - R_s^{t, r}|^2 ds \right] \right)^{\frac{1}{2}} \\
&\leq C_4 |\delta|,
\end{aligned}$$

and, since $|\nabla_r \eta(s, z, r)| \leq b(z)$, the second summand is bounded by

$$\begin{aligned}
& C_5 E \int_t^T \int_U b(z) |\xi_s^{t, r, \delta} - \nabla_r R_s^{t, r}| d\nu(z) ds \\
&\leq C_6 E \int_t^T |\xi_s^{t, r, \delta} - \nabla_r R_s^{t, r}| ds \\
&\leq C_7 |\delta|.
\end{aligned}$$

By letting δ converge to 0 we obtain that $\lim_{\delta \rightarrow 0} E \left| \frac{A^{r+\delta} - A^r}{\delta} - B^r \right| = 0$.

It remains to show that the second summand in (26) vanishes also as $\delta \rightarrow 0$. Notice that Hölder's inequality implies

$$E \left| B^r (A^{r+\delta} - A^r) \right| \leq (E |B^r|^2)^{\frac{1}{2}} \left(E |A^{r+\delta} - A^r|^2 \right)^{\frac{1}{2}}$$

By (24) we have

$$\begin{aligned}
E|B^r|^2 &\leq 4E \left(\int_t^T \int_U \nabla_r \eta(s, z, R_s^{t,r}) \nabla_r R_s^{t,r} (\mu - \pi)(ds, dz) \right)^2 \\
&\quad + 4E \left(\int_t^T \int_U |\nabla_r \eta(s, z, R_s^{t,r}) \nabla_r R_s^{t,r}| \pi(ds, dz) \right)^2 \\
&\leq 4E \int_t^T \int_U (\nabla_r \eta(s, z, R_s^{t,r}) \nabla_r R_s^{t,r})^2 d\nu(z) ds \\
&\quad + 4E \left(\int_t^T \int_U b(z) |\nabla_r R_s^{t,r}| d\nu(z) ds \right)^2 \\
&\leq 4E \int_t^T \int_U b^2(z) |\nabla_r R_s^{t,r}|^2 d\nu(z) ds + C_8 E \left(\int_t^T |\nabla_r R_s^{t,r}| ds \right)^2 \\
&\leq C_9 E \int_t^T |\nabla_r R_s^{t,r}|^2 ds < \infty.
\end{aligned}$$

Similarly, with (24) and (4),

$$\begin{aligned}
E|A^{r+\delta} - A^r|^2 &= E \left(\int_t^T \int_U \eta(s, z, R_s^{t,r+\delta}) - \eta(s, z, R_s^{t,r}) \mu(ds, dz) \right)^2 \\
&\leq E \int_t^T \int_U (\eta(s, z, R_s^{t,r+\delta}) - \eta(s, z, R_s^{t,r}))^2 d\nu(z) ds \\
&\quad + E \left(\int_t^T \int_U \eta(s, z, R_s^{t,r+\delta}) - \eta(s, z, R_s^{t,r}) d\nu(z) ds \right)^2 \\
&\leq E \int_t^T \int_U b^2(z) |R_s^{t,r+\delta} - R_s^{t,r}|^2 d\nu(z) ds \\
&\quad + E \left(\int_t^T \int_U b(z) |R_s^{t,r+\delta} - R_s^{t,r}| d\nu(z) ds \right)^2 \\
&\leq C_{10} E \int_t^T |R_s^{t,r+\delta} - R_s^{t,r}|^2 ds \\
&\leq C_{11} |\delta|,
\end{aligned}$$

which implies that $\lim_{\delta \rightarrow 0} E|B^r(A^{r+\delta} - A^r)| = 0$. \square

Theorem 3.8. *Let the assumptions of Lemma 3.7 be fulfilled. Then the predictable projection of $D_t g(L_T)$ satisfies $[D_t g(L_T)]^\varphi = \nabla_r h(t, L_{t-}, R_t^{0,r}) \gamma(t, R_t^{0,r})$.*

Proof. Let $0 \leq u < T$. It follows from Lemma 1.9 in [25] and the chain rule that for $P \otimes \lambda$ -a.a. (ω, s)

$$\begin{aligned}
1_{[0,u]}(s) E[D_s g(L_T) | \mathcal{F}_u] &= D_s(E[g(L_T) | \mathcal{F}_u]) = D_s(h(u, L_u, R_u^{0,r})) \\
&= \nabla_r h(u, L_u, R_u^{0,r}) D_s R_u^{0,r}. \tag{27}
\end{aligned}$$

For any partition $\Gamma : 0 = t_0 < t_1 < t_n \dots < t_n < T$ let

$$X_s^\Gamma = \sum_{i=1}^n 1_{]t_{i-1}, t_i]}(s) E[D_s g(L_T) | \mathcal{F}_{t_i}].$$

The optional projection of $D_s g(L_T)$ coincides with the optional projection of X_s^Γ . Note that (27) and (16) yield $\lim_{|\Gamma| \rightarrow 0} X_s^\Gamma = \nabla_r h(s, L_s, R_s^{0,r}) \gamma(s, R_s^{0,r})$. Since $\nabla_r h$ is continuous, we obtain the predictable projection of $D_s g(L_T)$ by taking left limits. \square

4 Case study: Hedging US hurricane risk by predicting Pacific sea surface temperatures

We aim at applying the results to hedging of risk due to hurricanes in the Atlantic basin that landfall in the US. Many environmental factors are correlated with the hurricane activity in the Atlantic (see [22]). For example, it is known since long that the sea surface temperature (SST) in the tropical Pacific ocean considerably influences the number of hurricanes that develop in the Atlantic basin every year in the period between June and November. A rough statistical analysis shows that the number of hurricanes is approximately Poisson distributed with an intensity depending on the average Pacific SST during the hurricane season (see [9]). A Poisson regression, the results of which will be shown in a forthcoming paper, seems to reveal that this dependence is linear.

Consider a grid of N points covering the Pacific ocean. Let $e_1, \dots, e_N \in \mathbb{R}^N$ be the EOFs of the Pacific SST anomaly field. At any time the SST anomaly grid can be decomposed into a linear combination $\sum_{i=1}^N r_i e_i$. Now let e_1, \dots, e_m be the EOFs of the $m \ll N$ largest eigenvalues, so that $\sum_{i=1}^m r_i e_i$ is a good approximation of the SST. Let $K \subset \{1, \dots, N\}$ be the set containing the grid points in the *tropical* Pacific ocean. Let $\bar{e}_i = \frac{1}{K} \sum_{j \in K} e_i^j$, where e_i^j denotes the j th component of eigenvector e_i . Then the average tropical SST anomaly is given by $\bar{r} = \langle r, \bar{e} \rangle = \sum_{i=1}^m r_i \bar{e}_i$.

Let R_t denote the vector of the m first coefficients of the EOF linear combination describing the SST anomaly grid at time t , and assume that the process R_t solves the linear SDE (5). We assume that the Atlantic hurricanes occur with an intensity at time t

$$\lambda_t = \alpha + \beta \bar{R}_t,$$

for some real constants α and β . Thus the number of hurricanes occurring during a season is a non homogeneous Poisson process $(N_t)_{0 \leq t \leq T}$, with intensity λ_t and with T being the end of a hurricane season.

We consider an economic agent facing losses that by time t amount to $L_t = \sum_{k=1}^{N_t} Y_k$, where Y_k are i.i.d. distributed random variables. Suppose that there exists a liquid security with price (S_t) , and assume that the joint jump size distribution of L and the price S has a density $j(x, y)$ independent of k , i.e. for any $t \geq 0$, and for all $A \in \mathcal{B}(\mathbb{R}_+)$, $B \in \mathcal{B}(\mathbb{R})$ we have

$$P(\Delta L_t \in A, \Delta S_t \in B | \Delta N_t \neq 0) = \int_A \int_B j(x, y) dx dy.$$

In the remainder we always assume that the jump distribution satisfies the integrability condition $\int_0^\infty \int_{\mathbb{R}} (|x| + y^2 + |xy|) j(x, y) dx dy < \infty$.

Note that the process L_t is an *inhomogeneous* compound Poisson process. Nevertheless, we can write L_t as an integral with respect to a *homogeneous* Poisson random measure on the product space $U = \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}$. For this purpose let $\nu(dx, dy, dz) = j(x, y) dx dy dz$, and μ a Poisson random measure with compensator $\pi = \lambda \otimes \nu$. We put $\eta(x, y, z, r) = x 1_{[0, \alpha + \beta \bar{r}]}(z)$. Then the aggregate losses of our agent may be rewritten as

$$L_t = \int_0^t \int_U \eta(x, y, z, R_s) \mu(ds, dx, dy, dz) = \int_0^t \int_U x 1_{[0, \alpha + \beta \bar{R}_s]}(z) \mu(ds, dx, dy, dz)$$

First note that Condition (15) is satisfied, since with $M := \int_0^\infty \int_{\mathbb{R}} |x| j(x, y) dx dy$

$$\begin{aligned} \int_U |\eta(x, y, z, r) - \eta(x, y, z, r')| dz &\leq M \beta |\bar{r} - \bar{r}'| = M \beta |\langle r - r', \bar{e} \rangle| \\ &\leq M \beta |\bar{e}| |r - r'|. \end{aligned}$$

We suppose that the security price evolves according to (9). In addition we assume that $\sigma(s, x, y, z) = y 1_{[0, \alpha + \beta \bar{r}]}(z)$, which implies the price of the security to jump whenever L jumps.

By Theorem 1.1 the optimal variance hedge of the losses L_T is a function of ζ and ψ appearing in the integral representation

$$L_T = E(L_T) + \int_0^T \zeta(s) dW_s + \int_0^T \int_U \psi(s, z) (\mu - \pi)(ds, dz).$$

Note that Theorem 3.3 implies that $\psi(t, x, y, z) = \eta(x, y, z, R_t^{0,r})$. This, together with Theorem 1.1, yields the following proposition in case the white noise of the security is independent of the noise driving the SST anomaly process R , i.e. $\Sigma \rho^{tr}(t) = 0$.

Proposition 4.1. *If $\Sigma \rho^{tr}(t) = 0$, then the optimal variance hedge satisfies*

$$\begin{aligned} \vartheta^*(t) &= \frac{\int_U xy 1_{[0, \alpha + \beta \bar{R}_t^{0,r}]}(z) \nu(dx, dy, dz)}{|\rho|^2(t) + \int_U y^2 1_{[0, \alpha + \beta \bar{R}_t^{0,r}]}(z) \nu(dx, dy, dz)} \\ &= \frac{(\alpha + \beta \bar{R}_t^{0,r}) \int_{\mathbb{R}_+ \times \mathbb{R}} xy j(x, y) dx dy}{|\rho|^2(t) + (\alpha + \beta \bar{R}_t^{0,r}) \int_{\mathbb{R}_+ \times \mathbb{R}} y^2 j(y, x) dx dy}. \end{aligned}$$

The proposition shows that if $\Sigma\rho^{tr}(t) = 0$, then the more has to be invested in the security the higher the SST and hence the higher the intensity of occurring losses.

In order to obtain the optimal variance hedge in the case where $\Sigma\rho^{tr}(t)$ does not vanish, we first compute explicitly the integrand ζ .

Lemma 4.2. *We have $\zeta(s) = \nabla_r h(t, L_{t-}, R_t^{0,r})\Sigma$ on the set where $\inf_{0 \leq t \leq T} \bar{R}_t^{0,r} \geq -\frac{\alpha}{\beta}$.*

Proof. We prove the result by approximating η with functions that satisfy the assumptions of Lemma 3.7 and Theorem 3.8. Note that Theorem 3.8 requires $\nu(U) < \infty$. By standard stopping arguments we may assume that $R^{0,r}$ is bounded from above by a constant $C \in \mathbb{R}_+$ and hence we may replace U by $[0, C] \times \mathbb{R} \times \mathbb{R}$ which has finite measure relative to ν .

Let $\Phi : \mathbb{R} \rightarrow [0, 1]$ be a twice continuously differentiable function such that $\Phi(x) = 0$ for all $x \geq 1$, $\Phi(x) = 1$ for all $x \leq 0$, and $\Phi'(0) = \Phi'(1) = 0$. Define

$$\eta^n(z, r) = \Phi(n[z - \alpha - \beta\bar{r}]),$$

and observe that $\lim_n \eta^n(t, r) = \eta(t, r)$ for all $t \in [0, T]$ and $r \in \mathbb{R}^m$. Let $\kappa^n(r) = \int_U \eta^n(z, r) dz$ for all $r \in \mathbb{R}^m$. Since $\nu(U) < \infty$, dominated convergence implies that $\kappa^n(r)$ converges to $\kappa(r) = \int_0^\infty \eta(r, z) dz = (\alpha + \beta\bar{r}) \vee 0$ as $n \rightarrow \infty$.

The functions η^n and their derivatives $\nabla_r \eta^n$ are Lipschitz continuous, and the assumptions (24) and (25) are satisfied. It follows that $h^n(t, x, r) = x + E \int_t^T \int_U \eta^n(z, R_s^{t,r}) \mu(ds, dz)$ is differentiable with

$$\nabla_r h^n(t, x, r) = E \int_t^T \nabla R_s^{t,r} \nabla_r \kappa^n(R_s^{t,r}) ds.$$

Now let $L_T^n = \int_0^T \int_U \eta^n(z, R_s^{0,r}) \mu(ds, dz)$, and notice that L_T^n converges to L_T in $L^2(\Omega)$. By Theorem 3.8 and the Ito Isometry the processes $\nabla_r h^n(t, L_{t-}, R_t^{0,r})\Sigma$ converge to $\zeta(s)$ in $L^2(\Omega \times [0, T])$. On the set $\{r : \bar{r} > -\frac{\alpha}{\beta}\}$ we have $\nabla_r \kappa^n \rightarrow \nabla_r \kappa$ pointwise, whence the result. \square

The lemma immediately yields:

Proposition 4.3. *The optimal variance hedge satisfies*

$$\vartheta^*(t) = \frac{\nabla_r h(t, L_{t-}, R_t^{0,r})\Sigma\rho^{tr}(t) + (\alpha + \beta\bar{R}_t^{0,r}) \int_{\mathbb{R}_+ \times \mathbb{R}} xyj(x, y) dx dy}{|\rho|^2(t) + (\alpha + \beta\bar{R}_t^{0,r}) \int_{\mathbb{R}_+ \times \mathbb{R}} y^2 j(y, x) dx dy}.$$

on the set where $\inf_{0 \leq t \leq T} \bar{R}_t^{0,r} \geq -\frac{\alpha}{\beta}$.

References

- [1] S. Ankirchner. On filtration enlargements and purely discontinuous martingales. *to appear in: Stoch. Proc. Appl.*
- [2] S. Ankirchner, P. Imkeller, and G. Dos Reis. Pricing and hedging of derivatives based on non-tradable underlyings. Preprint, arXiv:0712.3746v1 [math.PR], 2007.
- [3] S. Ankirchner, A. Popier, and P. Imkeller. Optimal cross hedging of insurance derivatives, 2007.
- [4] L. Arnold. *Stochastische Differentialgleichungen*. R. Oldenbourg Verlag, Munich, 1973.
- [5] P. Barrieu. Structuration optimale de produits financiers en marché illiquide et trois excursions dans d'autres domaines des probabilités. Thèse de doctorat, Paris VI, 2002.
- [6] P. Barrieu and N. El Karoui. Optimal design of derivatives in illiquid markets. *Quant. Finance*, 2(3):181–188, 2002.
- [7] P. Barrieu and N. El Karoui. Optimal derivatives design under dynamic risk measures. In *Mathematics of finance*, volume 351 of *Contemp. Math.*, pages 13–25. Amer. Math. Soc., Providence, RI, 2004.
- [8] D. Becherer. Rational hedging and valuation with utility-based preferences. Ph.D. thesis, Technische Universität Berlin, 2001.
- [9] M.C. Bove, J.J. O'Brien, J.B. Elsner, C.W. Landsea, and X. Niu. Effect of El Niño on U.S. Landfalling Hurricanes, Revisited. *Bulletin of the American Meteorological Society*, 79(11).
- [10] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [11] M. Davis. Option valuation and hedging with basis risk. In *System theory: modeling, analysis and control (Cambridge, MA, 1999)*, volume 518 of *Kluwer Internat. Ser. Engrg. Comput. Sci.*, pages 245–254. Kluwer Acad. Publ., Boston, MA, 2000.
- [12] M. Davis. Pricing weather derivatives by marginal value. *Quant. Finance*, 1(3):305–308, 2001.

- [13] M. Davis. Optimal hedging with basis risk. In R. Liptser Y. Kabanov and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance*. Springer, 2006.
- [14] F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker. Exponential hedging and entropic penalties. *Math. Finance*, 12(2):99–123, 2002.
- [15] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.
- [16] N. El Karoui and M.-C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.*, 33(1):29–66, 1995.
- [17] H. Föllmer and P. Leukert. Quantile hedging. *Finance Stoch.*, 3(3):251–273, 1999.
- [18] H. Föllmer and P. Leukert. Efficient hedging: cost versus shortfall risk. *Finance Stoch.*, 4(2):117–146, 2000.
- [19] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.
- [20] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.
- [21] H. Kunita. Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and stochastic analysis*, Trends Math., pages 305–373. Birkhäuser Boston, Boston, MA, 2004.
- [22] C.W. Landsea, Jr.R.A. Pielke, A.M. Mestas-Nuñez, and J.A. Knaff. Atlantic basin hurricanes: indices of climatic changes. *Climatic change*, 42, 1999.
- [23] A. Løkka. Martingale representation of functionals of Lévy processes. *Stochastic Anal. Appl.*, 22(4):867–892, 2004.
- [24] M. Musiela and T. Zariphopoulou. An example of indifference prices under exponential preferences. *Finance Stoch.*, 8(2):229–239, 2004.

- [25] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.
- [26] C. Penland. Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Wea. Rev.*, 117:2165–2185, 1989.
- [27] C. Penland and P. D. Sardeshmukh. The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, 8(8):1999–2024, 1999.
- [28] J. Picard. Formules de dualité sur l'espace de Poisson. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(4):509–548, 1996.
- [29] J. Picard. On the existence of smooth densities for jump processes. *Probab. Theory Related Fields*, 105(4):481–511, 1996.
- [30] C. Stricker and M. Yor. Calcul stochastique dépendant d'un paramètre. *Z. Wahrsch. Verw. Gebiete*, 45(2):109–133, 1978.
- [31] H. van den Dool. *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press, Oxford, 2007.