Discretely Sampled Variance and Volatility Swaps versus their Continuous Approximations

Robert Jarrow^{*}

Younes Kchia[†]

a[†] Martin Larsson[‡]

Philip Protter^{§¶}

August 9, 2018

Abstract

Discretely sampled variance and volatility swaps trade actively in OTC markets. To price these swaps, the continuously sampled approximation is often used to simplify the computations. The purpose of this paper is to study the conditions under which this approximation is valid. Our first set of theorems characterize the conditions under which the discretely sampled swap values are finite, given the values of the continuous approximations exist. Surprisingly, for some otherwise reasonable price processes, the discretely sampled swap prices do not exist, thereby invalidating the approximation. Examples are provided. Assuming further that both swap values exist, we study sufficient conditions under which the discretely sampled values converge to their continuous counterparts. Because of its popularity in the literature, we apply our theorems to the 3/2 stochastic volatility model. Although we can show finiteness of all swap values, we can prove convergence of the approximation only for some parameter values.

KEY WORDS: variance swaps, volatility swaps, NFLVR, semimartingales

1 Introduction

Although variance and volatility swaps only started trading in the mid-1900s, they have since become a standard financial instrument useful for managing volatility risk (see Carr and Lee [7] for the history of volatility derivative markets). In the pricing and hedging of variance and volatility swaps, a distinction is made between payoffs that are discretely or continuously sampled. Discretely sampled variance and volatility swaps trade in the over-the counter (OTC) markets. In contrast, continuously sampled variance and volatility swaps are only an abstract construct, often used to approximate the values of their discretely sampled counterparts (see Broadie and Jain [4],[5], Chan and Platen [10], and

 $^{^*}$ Johnson Graduate School of Management, Cornell University, Ithaca, NY, 14853 and Kamakura Corporation

[†]Centre de Mathématiques Appliquées, Ecole Polytechnique, Paris

[‡]School of Operations Research, Cornell University, Ithaca, NY, 14853

[§]Statistics Department, Columbia University, New York, NY, 10027

[¶]Supported in part by NSF grant DMS-0906995

Carr and Lee [8]). These approximations depend on an exchange of a limit operator (as the discrete sampling period goes to zero) with an expectation operator. This operator exchange is often invoked without adequate justification. The purpose of this paper is to characterize the conditions under which this operator exchange is valid.

For our investigation we utilize the martingale pricing methodology where we take as given the asset's price process assuming markets are arbitrage free (in the sense of No Free Lunch with Vanishing Risk (NFLVR)). This evolution is taken to be very general. We only assume that the price process is a strictly positive semimartingale with possibly discontinuous sample paths. We also assume, as a standing hypothesis, that the continuously sampled variance and volatility swaps have finite values. Otherwise, before the analysis begins, the approximation would not make sense.

Our first two Theorems (1 and 2) characterize the *additional conditions* needed on the price process such that the discretely sampled variance and volatility swaps have finite values. Of course, when the discretely sampled variance and volatility swap values do not exist, the approximation is again nonsensical. Surprisingly, we provide examples of otherwise reasonable price processes, with stochastic volatility of the volatility, where these discretely sample variance swap values do not exist.

Next, assuming both the continuous and discretely sampled variance swap values are finite, we study conditions justifying an exchange of the limit and expectation operators. In this regard, under no additional hypotheses, Theorem 3 provides an upper bound for the maximum difference between these two values. Theorem 4 proves, under some additional moment conditions, that the exchange of the two operators is valid. Furthermore, this theorem also provides infomation on the rate of convergence (1/n where n is the number of discretely sampled prices).

Lastly, given the recent interest in the 3/2 stochastic volatility model for pricing volatility derivatives (see Carr and Sun [9], Chan and Platen [10]), we explore its consistency with valuing discretely sampled variance swaps with their continuously sampled counterparts. Here we show that both the discrete and continuously sampled variance swaps have finite values. Unfortunately, we can only prove convergence of the discrete to the continuously sampled variance swap values for some parameter ranges, but not all. A complete characterization of this convergence for the 3/2 stochastic volatility model remains an open question.

An outline for this paper is as follows. Section 2 gives the framework underlying the model. Section 3 studies the finiteness and convergence of the discretely sampled variance and volatility swap values. Finally, Section 4 provides examples to illustrate the theorems proved.

2 Framework

Let a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be given, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ satisfies the usual conditions and T is a fixed time horizon. We suppose that there is a liquidly traded asset paying no dividends, whose market price process is modeled by a semimartingale $S = (S_t)_{t \in [0,T]}$ such that S > 0 and $S_- > 0$. The value of the money market account is chosen as numeraire, or, viewed differently, the interest rate is zero. The price process is assumed to be arbitrage free in the sense of No Free Lunch with Vanishing Risk (NFLVR), see Delbaen and Schachermayer [13] and [14], which guarantees the existence of at least one equivalent probability measure under which S is a local martingale. We assume P is such a measure, and that the modeler prices future payoffs by taking expectations with respect to P.

2.1 Variance and volatility swaps

A variance swap (with strike zero) and maturity T is a contract which pays the "realized variance," i.e. the square of the logarithm returns up to time T, namely

$$\frac{252}{n} \sum_{i=1}^{n} \left(\ln \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2,$$

where $0 = t_0 < ... < t_n = T$ is a regular sampling of the time interval [0, T], i.e. $t_i - t_{i-1} = \frac{1}{n}$ for i = 1, ..., n. Finally, 252 is the number of trading days per year. The maturity T is approximately $\frac{n}{252}$.

Rather than considering the quantity

$$P^{n}(T) = \sum_{i=1}^{n} \left(\ln \frac{S_{t_{i}}}{S_{t_{i-1}}} \right)^{2},$$

practitioners¹ often use its limit

$$P(T) = [\ln S, \ln S]_T$$

in the pricing of variance swaps (see Carr and Lee [7]). That is, one approximates the quantity $E(P^n(T))$ by E(P(T)). The continuous approximation P(T) to the variance swap's true payoff $P^n(T)$ is justified by the fact that

$$[\ln S, \ln S]_T = \lim_{n \to \infty} \sum_{i=1}^n \left(\ln \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2,$$

where the limit is in probability and taken over a sequence of subdivisions whose mesh size tends to zero.

Analogous to a variance swap, a *volatility swap* is a security written on the square root of the variance swap's payoff. We will use the notation

$$V^{n}(T) = \sqrt{\sum_{i=0}^{n-1} \left(\ln \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2} \quad \text{and} \quad V(T) = \sqrt{\langle \ln S \rangle_T}.$$

¹Confirmed in a private discussion with Peter Carr.

Again, $V^n(T)$ is the volatility swap payoff up to multiplicative and additive constants.

Our aim is to investigate the validity of these approximations, motivated by the fact that the convergence of $P^n(T)$ to P(T), respectively $V^n(T)$ to V(T), is only in probability, which a priori does not guarantee convergence of their expectations. It is the latter convergence one needs in order to justify the use of this approximation in the context of pricing.

To simplify the notation later on, we introduce the following convention.

Notation. For a process $X = (X_t)_{t \in [0,T]}$ we define

$$\delta_i X = X_{t_i} - X_{t_{i-1}}$$
 $(i = 1, \dots, n)$

In particular, $P^n(T) = \sum_{i=1}^n (\delta_i \ln S)^2$.

2.2 Mathematical preliminaries

Since S > 0 and $S_{-} > 0$ there is a semimartingale M such that

$$S = S_0 \mathcal{E}(M)$$

where $\mathcal{E}(\cdot)$ denotes stochastic exponential; see [18], Theorem II.8.3. In fact, since S is a local martingale and hence a special semimartingale, it follows that from Theorem II.8.21 in [18] that M is a stochastic integral with respect to the local martingale part of S. Furthermore, since S > 0, the jumps of M satisfy $\Delta M > -1$, so by the Ansel-Stricker Theorem (see [2]), M is a local martingale.

The following notation is standard. We direct the reader to [18] for details. The random measure μ^M associated with the jumps of M is given by

$$\mu^{M}(dt, dx) = \sum_{s} \mathbf{1}_{\{\Delta M_{s} \neq 0\}} \varepsilon_{(s, \Delta M_{s})}(dt, dx),$$

and its predictable compensator is $\nu(dt, dx)$. We may then write

$$S = S_0 \exp\left\{M - \frac{1}{2} \langle M^c, M^c \rangle - (x - \ln(1+x)) * \mu^M\right\},$$
(1)

and furthermore decompose M as

$$M = M^{c} + M^{d} = M^{c} + x * (\mu^{M} - \nu), \qquad (2)$$

where M^c is the continuous local martingale part of M and $M^d = x * (\mu^M - \nu)$ is the jump part. Both these processes are local martingales. The following elementary result will be useful later, so we state it as a lemma.

Lemma 1 The quadratic variation of M and $\ln S$ are given by

(i) $[M, M] = \langle M^c, M^c \rangle + x^2 * \mu^M$

(ii) $[\ln S, \ln S] = \langle M^c, M^c \rangle + (\ln(1+x))^2 * \mu^M$

Proof. Part (i) is a standard fact. For part (ii), write $X = (x - \ln(1 + x)) * \mu^M$ and notice that $\ln S - \ln S_0 = M - X - \frac{1}{2} \langle M^c, M^c \rangle$. Hence

$$\begin{aligned} [\ln S, \ln S] &= [M - X, M - X] \\ &= [M, M] + [X, X] - 2[M, X] \\ &= \langle M^c, M^c \rangle + x^2 * \mu^M + (x - \ln(1 + x))^2 * \mu^M - 2x(x - \ln(1 + x)) * \mu^M. \end{aligned}$$

By the Cauchy-Schwartz inequality,

$$|x(x - \ln(1 + x))| * \mu^{M} = \sum_{s \leq \cdot} |\Delta M_{s}(\Delta M_{s} - \ln(1 + \Delta M_{s}))|$$

$$\leq \sqrt{\sum_{s \leq \cdot} (\Delta M_{s})^{2}} \sqrt{\sum_{s \leq \cdot} (\Delta M_{s} - \ln(1 + \Delta M_{s}))^{2}} < \infty,$$

so $x(x - \ln(1+x)) * \mu^M$ also converges absolutely, a.s. Termwise manipulation is therefore allowed, so since $x^2 + (x - \ln(1+x))^2 - 2x(x - \ln(1+x)) = (\ln(1+x))^2$, we get

$$x^{2} * \mu^{M} + (x - \ln(1+x))^{2} * \mu^{M} - 2x(x - \ln(1+x)) * \mu^{M} = (\ln(1+x))^{2} * \mu^{M}.$$

The result follows. \blacksquare

We also have the following lemma, which gives the semimartingale decomposition of $\ln S$.

Lemma 2 Assume that $\ln S$ is locally integrable. Then

$$\ln S - \ln S_0 = M^c + \ln(1+x) * (\mu^M - \nu) - \frac{1}{2} \langle M^c, M^c \rangle - (x - \ln(1+x)) * \nu.$$

Proof. From (1) and (2) we have

$$\ln S - \ln S_0 = M^c - \frac{1}{2} \langle M^c, M^c \rangle + x * (\mu^M - \nu) - (x - \ln(1 + x)) * \mu^M.$$

Our assumption together with the fact that M^c and $x * (\mu^M - \nu)$ are local martingales, hence locally integrable, implies that $\langle M^c, M^c \rangle$ and $(x-\ln(1+x))*\mu^M$ are locally integrable (notice that both are nonnegative). Hence $(x - \ln(1+x))*\nu$ is locally integrable (see [18], Proposition II.1.8), so we may add and subtract this quantity to the right side of the previous display to obtain

$$\ln S - \ln S_0 = M^c - \frac{1}{2} \langle M^c, M^c \rangle + \ln(1+x) * (\mu^M - \nu) - (x - \ln(1+x)) * \nu.$$

This is the desired expression. \blacksquare

3 Approximation using the quadratic variation

3.1 Finiteness of expectations

In order for there to be any hope that E(P(T)) accurately approximates $E(P^n(T))$, a minimal requirement is that both these quantities be finite. Perhaps somewhat surprisingly, they need not be finite or infinite simultaneously. This is of course a potentially serious issue, since the value of E(P(T)), when finite, is nonsensical as an approximation of $E(P^n(T))$ if this is infinite.

The following result gives necessary and sufficient conditions for $E(P^n(T))$ to be finite, given that the approximation P(T) is known to have finite expectation.

Theorem 1 Assume that $P(T) \in L^1$. The following statements are equivalent.

(i)
$$P^n(T) \in L^1$$
 for at least one $n \ge 1$
(ii)
$$\begin{cases} \langle M^c, M^c \rangle_T \in L^2\\ (x - \ln(1+x)) * \nu_T \in L^2 \end{cases}$$

Proof. First of all, note that since $P^n(T) = \sum_{i=1}^n (\delta_i \ln S)^2$, we have $P^n(T) \in L^1$ if and only if $\delta_i \ln S \in L^2$ for each *i*, which is equivalent to having $\ln S_{t_i} - \ln S_0 \in L^2$ for each *i*. We thus need to show that this is equivalent to condition (*ii*) in the statement of the theorem.

By Lemma 1, our basic assumption $P(T) = [\ln S, \ln S]_T \in L^1$ implies that $\langle M^c, M^c \rangle_T \in L^1$ and $(\ln(1+x))^2 * \mu_T^M \in L^1$. Hence both M_t^c and $\ln(1+x) * (\mu^M - \nu)_t$ are in L^2 for every $t \leq T$. Therefore, using the representation from Lemma 2, namely

$$\ln S - \ln S_0 = M^c - \frac{1}{2} \langle M^c, M^c \rangle + \ln(1+x) * (\mu^M - \nu) - (x - \ln(1+x)) * \nu,$$

we deduce that $\ln S_{t_i} - \ln S_0 \in L^2$ for each *i* holds if and only if

$$\frac{1}{2}\langle M^c, M^c \rangle_{t_i} + (x - \ln(1+x)) * \nu_{t_i} \in L^2$$

for each *i*. Since both $\langle M^c, M^c \rangle$ and $(x - \ln(1+x)) * \nu$ are nonnegative and nondecreasing, this is equivalent to condition (*ii*). The proof is finished.

We note that the above theorem implies that if $P^n(T) \in L^1$ for some *n* then $P^n(T) \in L^1$ for all $n \ge 1$. An analogous result holds for volatility swaps.

Theorem 2 Assume that $V(T) \in L^1$. The following statements are equivalent.

(i)
$$V^n(T) \in L^1$$
 for some $n \ge 1$
(ii)
$$\begin{cases} \langle M^c, M^c \rangle_T \in L^1 \\ (x - \ln(1+x)) * \nu_T \in L^1 \end{cases}$$

Proof. Using for instance the fact that all norms on \mathbb{R}^n are equivalent, there are constants $0 < c \leq C < \infty$ such that

$$cV^n(T) \le \sum_{i=1}^n \left|\delta_i \ln S\right| \le CV^n(T)$$
 a.s.

Thus $V^n(T) \in L^1$ is equivalent to $\delta_i \ln S \in L^1$ for each *i*, which is equivalent to $\ln S_{t_i} - \ln S_0 \in L^1$ for each *i*.

Now the proof is similar to that of Theorem 1. By Lemma 1, $V(T) = \sqrt{[\ln S, \ln S]_T} \in L^1$ implies that $\sqrt{\langle M^c, M^c \rangle_T}$ and $\sqrt{(\ln(1+x))^2 * \mu_T^M}$ are in L^1 , which via the Burkholder-Davis-Gundy inequalities implies that M_t^c and $\ln(1+x) * (\mu^M - \nu)_t$ are in L^1 for each $t \leq T$. By Lemma 2, therefore, $\ln S_{t_i} - \ln S_0 \in L^1$ for each *i* if and only if

$$\frac{1}{2}\langle M^c, M^c \rangle_{t_i} + (x - \ln(1+x)) * \nu_{t_i} \in L^1$$

for each *i*. This is equivalent to condition (ii) of the theorem.

Again, we note that Theorem 2 implies that if $V^n(T) \in L^1$ for some n then $V^n(T) \in L^1$ for all $n \ge 1$.

When S is continuous we have the following variation of Theorem 1, which indicates that in many cases, $E(P^n(T)) < \infty$ is a stronger requirement than $E(P(T)) < \infty$. This will be used in Section 4, where we discuss specific examples. Notice that when S is continuous, $P(T) = [\ln S, \ln S] = \langle M^c, M^c \rangle = \langle M, M \rangle.$

Proposition 1 Assume that S is continuous. Then the following are equivalent.

- (i) $M_t \in L^1$ and $P^n(T) \in L^1$
- (*ii*) $P(T) \in L^2$

Proof. In the continuous case, $\ln S - \ln S_0 = M - \frac{1}{2} \langle M, M \rangle$. First assume (*ii*), i.e. that $P(T) = \langle M, M \rangle_T \in L^2$. Then certainly $M_t \in L^2$ for every $t \leq T$, hence $\ln S_t \in L^2$ for every $t \leq T$. Therefore $\delta_i \ln S \in L^2$ for every *i*, and as in the beginning of the proof of Theorem 1, this implies that $P^n(T) \in L^1$. Hence (*i*) holds. Conversely, if (*i*) is satisfied, then $\delta_i \ln S \in L^2$ for every *i*, so $\ln S_T \in L^2$. We also have $M_T \in L^1$ by assumption, so $\langle M, M \rangle_T = 2(M_T - \ln S_T + \ln S_0)$ is in L^1 , implying that M_T is in fact in L^2 . This lets us strengthen the previous conclusion to $\langle M, M \rangle_T \in L^2$, which is (*ii*).

Notice that the implication $(ii) \Rightarrow (i)$ also follows from Theorem 1. We can prove an analogous result for volatility swaps.

Proposition 2 Assume that S is continuous. Then the following are equivalent.

- (i) $M_t \in L^1$ and $V^n(T) \in L^1$
- (ii) $V(T) \in L^2$, i.e. $\langle M, M \rangle_T \in L^1$.

Proof. As in the proof of Theorem 2, $V^n(T) \in L^1$ is equivalent to $\delta_i \ln S \in L^1$ for each *i*. The rest of the proof is similar to that of Proposition 1.

In the case of continuous price processes, Proposition 1 makes it clear how one can proceed to construct examples where the approximation P(T) has finite expectation, but the true payoff $P^n(T)$ does not. Indeed, any process S of the form $S = S_0 \mathcal{E}(M)$ will do, where M is a continuous local martingale that satisfies

$$\begin{cases} \langle M, M \rangle_T \in L^1 \\ \langle M, M \rangle_T \notin L^2. \end{cases}$$
(3)

It is clear that such processes exist; what is less clear is to what extent examples can be found among models that appear in applications. In Section 4 we provide examples demonstrating that the condition (3) can appear in models that may appear innocuous at first sight.

3.2 Bounds on the approximation error

In this section we assume that both E(P(T)) and $E(P^n(T))$ are finite, and study conditions under which they are close for large n; as already mentioned, although $P^n(T) \to P(T)$ in probability, the expectations need not converge. We start by showing that under general conditions, the two expectations at least cannot be too far apart. We then impose additional structure on the model and give conditions that guarantee convergence. The focus of this section is on variance swaps. The analysis of volatility swaps is more complicated and is omitted.

Theorem 3 Assume that $P^n(T)$ and P(T) both are in L^1 . Then there is a constant C > 0, independent of n, such that

$$|E(P(T)) - E(P^n(T))| \le C \quad \text{for all } n.$$

The proof of Theorem 3 is straightforward once the following lemma has been established.

Lemma 3 Assume that $P^n(T)$ and P(T) both are in L^1 and define

$$N = \ln(1+x) * (\mu^{M} - \nu)$$

$$A = \frac{1}{2} \langle M^{c}, M^{c} \rangle + (x - \ln(1+x)) * \nu.$$

Then $\langle M^c, M^c \rangle_T$ and $[N, N]_T$ are in L^1 . Moreover,

$$\left|E\left(P(T)\right) - E\left(P^{n}(T)\right)\right| \le E\left(\sum_{i=1}^{n} (\delta_{i}A)^{2}\right) + 2E\left(\sum_{i=1}^{n} |\delta_{i}M^{c}|\delta_{i}A\right) + 2E\left(\sum_{i=1}^{n} |\delta_{i}N|\delta_{i}A\right),$$

and we have

$$E\Big(\sum_{i=1}^{n} |\delta_{i}M^{c}|\delta_{i}A\Big) \leq \sqrt{E\big(\langle M^{c}, M^{c}\rangle_{T}\big)} \sqrt{E\big(\sum_{i=1}^{n} (\delta_{i}A)^{2}\big)}$$

and

$$E\Big(\sum_{i=1}^{n} |\delta_i N| \delta_i A\Big) \le \sqrt{E\big([N,N]_T\big)} \sqrt{E\big(\sum_{i=1}^{n} (\delta_i A)^2\big)}.$$

Proof. Observe that N is a purely discontinuous local martingale with $[N, N] = (\ln(1 + x))^2 * \mu^M$, and A is a nondecreasing process. Since by assumption $P(T) = [\ln S, \ln S]_T \in L^1$, Lemma 1 implies that both M^c and N are L^2 martingales, which is the first assertion. Since also $P^n(T) \in L^1$, Theorem 1 shows that $A_T \in L^2$. Again by Lemma 1,

$$E\left([\ln S, \ln S]_T\right) = \sum_{i=1}^n E\left(\delta_i \langle M^c, M^c \rangle\right) + \sum_{i=1}^n E\left(\delta_i[N, N]\right)$$
$$= \sum_{i=1}^n E\left((\delta_i M^c)^2 + (\delta_i N)^2\right).$$

Lemma 2 shows that with N and A as above, we have $\ln S - \ln S_0 = M^c + N - A$. Hence $\delta_i \ln S = \delta_i M^c + \delta_i N - \delta_i A$, and therefore

$$E(P^{n}(T)) - E(P(T)) = \sum_{i=1}^{n} E((\delta_{i}A)^{2} + 2(\delta_{i}M^{c})(\delta_{i}N) - 2(\delta_{i}M^{c})(\delta_{i}A) - 2(\delta_{i}N)(\delta_{i}A))$$
$$= \sum_{i=1}^{n} E((\delta_{i}A)^{2} - 2(\delta_{i}M^{c})(\delta_{i}A) - 2(\delta_{i}N)(\delta_{i}A)),$$

where the second equality holds because M^c and N are orthogonal L^2 martingales. The triangle inequality and Jensen's inequality yield

$$\left| E(P(T)) - E(P^n(T)) \right| \le E\left(\sum_{i=1}^n (\delta_i A)^2\right) + 2E\left(\sum_{i=1}^n |\delta_i M^c| \delta_i A\right) + 2E\left(\sum_{i=1}^n |\delta_i N| \delta_i A\right)$$
$$= (I) + (II) + (III).$$

This settles the first inequality in the statement of the lemma. For (II), applying the Cauchy-Schwartz inequality twice, first for the sum and then the expectation, yields

$$(\mathrm{II}) \le 2E\left(\sqrt{\sum_{i=1}^{n} (\delta_i M^c)^2} \sqrt{\sum_{i=1}^{n} (\delta_i A)^2}\right) \le 2\sqrt{E\left(\sum_{i=1}^{n} (\delta_i M^c)^2\right)} \sqrt{E\left(\sum_{i=1}^{n} (\delta_i A)^2\right)}.$$

It only remains to notice that $E\left(\sum_{i=1}^{n} (\delta_i M^c)^2\right) = \sum_{i=1}^{n} E\left(\delta_i \langle M^c, M^c \rangle\right) = E\left(\langle M^c, M^c \rangle_T\right).$ An analogous calculation yields

(III)
$$\leq 2\sqrt{E([N,N]_T)} \sqrt{E\left(\sum_{i=1}^n (\delta_i A)^2\right)},$$

thus proving the lemma. \blacksquare

Proof of Theorem 3. By Lemma 3 it suffices to bound $E\left(\sum_{i=1}^{n} (\delta_i A)^2\right)$. Since $\delta_i A \ge 0$ for each *i*, we get

$$E\left(\sum_{i=1}^{n} (\delta_i A)^2\right) \le E\left(\left\{\sum_{i=1}^{n} \delta_i A\right\}^2\right) = E(A_T^2) < \infty.$$

The proof is complete. \blacksquare

Notice that constant C can be taken to be

$$C = E(A_T^2) + 2\sqrt{E(\langle M^c, M^c \rangle_T)}\sqrt{E(A_T^2)} + 2\sqrt{E([N,N]_T)}\sqrt{E(A_T^2)},$$

where N and A are as in Lemma 3. It is worth pointing out that since C is independent of n, it is in particular valid for n = 1. It can therefore not be expected to provide a tight bound for large n. Such results can be obtained under additional structure on the model, which we now introduce.

For the remainder of this section we assume that our probability space supports an mdimensional Brownian motion $W = (W^1, \ldots, W^n)$ and a Poisson random measure $\mu = \mu(dt, dz)$ on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $dt \otimes F(dz)$, where $\int (z^2 \wedge 1)F(dz) < \infty$. Moreover, we assume that M^c is a stochastic integral with respect to W, and that M^d is a stochastic integral with respect to $\mu - dt \otimes F(dz)$. That is, we assume that there are predictable processes a^1, \ldots, a^m and a predictable function $\psi > -1$ such that

$$M_t^c = \sum_{k=1}^m \int_0^t a_s^k dW_s^k \qquad and \qquad M_t^d = \psi * (\mu - dt \otimes F(dz))_t.$$

In this case the compensator ν of μ^M satisfies

$$\int_0^T \int_{(-1,\infty)} G(s,x) \nu(ds,dx) = \int_0^T \int_{\mathbb{R}} G(s,\psi(s,z)) F(dz) ds$$

for every nonnegative predictable function G. Moreover, it is a classical result that there exists a one-dimensional Brownian motion B and a nonnegative predictable process σ such that

$$M_t^c = \int_0^t \sigma_s dB_s.$$

Under this structure we can formulate conditions on σ and ψ under which the expectation of $P^n(T)$ converges to the expectation of P(T) as $n \to \infty$.

Theorem 4 With the notation and assumptions just described above, assume that the following conditions hold:

$$\begin{cases} E\left\{\int_0^T \sigma_s^4 ds\right\} < \infty \\ E\left\{\int_0^T \int_{\mathbb{R}} (1 \lor |z|^{-p}) \left\{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\right\} F(dz) ds\right\} < \infty, \end{cases}$$

for some $p \ge 0$ such that $\int_{\mathbb{R}} (1 \wedge |z|^p) F(dz) < \infty$. Then

$$\lim_{n \to \infty} |E(P(T)) - E(P^n(T))| = 0.$$

More specifically, there are constants C and D such that

$$|E(P(T)) - E(P^{n}(T))| \le Cn^{-1} + Dn^{-1/2}.$$

Notice that $p \leq 2$ always works, in the sense that $\int_{\mathbb{R}} (1 \wedge z^2) F(dz) < \infty$. However, depending on F(dx), smaller values of p may also work, imposing less stringent restrictions on ψ . In particular, if the Poisson random measure only has finitely many jumps, so that $\int_{\mathbb{R}} F(dx) < \infty$, we may take p = 0, and the condition on ψ reduces to

$$E\Big\{\int_0^T \int_{\mathbb{R}} \big\{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\big\}F(dz)ds\Big\} < \infty.$$

We remark that Theorem 4 generalizes results in [4], where the authors study the Black-Scholes model, the Heston stochastic volatility model, the Merton jump-diffusion model, and the stochastic volatility with jumps model by Bates and Scott.

The proof of Theorem 4 requires two lemmas.

Lemma 4 For $0 \le s < t \le T$ we have, a.s.,

$$\left\{ \int_{s}^{t} \int_{(-1,\infty)} (x - \ln(1+x))\nu(du, dx) \right\}^{2} \le (t-s)C \int_{0}^{T} \int_{\mathbb{R}} (1 \vee |z|^{-p}) \left\{ \psi(s,z)^{2} + (\ln(1+\psi(s,z)))^{2} \right\} F(dz) ds$$

for any $p \ge 0$ such that $\int_{\mathbb{R}} (1 \wedge |z|^p) F(dz) < \infty$, and a finite constant C that does not depend on s, t or p.

Proof. Jensen's inequality yields

$$\left\{\int_{s}^{t}\int_{(-1,\infty)}(x-\ln(1+x))\nu(du,dx)\right\}^{2}$$
$$=\left\{\int_{s}^{t}\int_{\mathbb{R}}(\psi(s,z)-\ln(1+\psi(s,z)))F(dz)ds\right\}^{2}$$
$$\leq (t-s)\int_{s}^{t}\left\{\int_{\mathbb{R}}G(s,z)F(dz)\right\}^{2}ds,$$

where we have defined $G(s, z) = \psi(s, z) - \ln(1 + \psi(s, z))$. Splitting up the integral over \mathbb{R} as the sum of the integrals over $\{|z| \leq 1\}$ and $\{|z| > 1\}$, and applying the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, we obtain

$$\Big\{\int_{\mathbb{R}} G(s,z)F(dz)\Big\}^2 \le 2\Big\{\int_{\{|z|\le 1\}} G(s,z)F(dz)\Big\}^2 + 2\Big\{\int_{\{|z|>1\}} G(s,z)F(dz)\Big\}^2.$$

Since $C_1 = F(\{|z| > 1\}) < \infty$, Jensen's inequality applied to the second term yields

$$\left\{\int_{\{|z|>1\}} G(s,z)F(dz)\right\}^2 \le C_1 \int_{\{|z|>1\}} G(s,z)^2 F(dz).$$

For the first term, let $\widetilde{F}(dz) = |z|^p F(dz)$. Then $C_2 = \widetilde{F}(\{|z| \le 1\}) < \infty$ by assumption, and using once again Jensen's inequality we get

$$\begin{split} \left\{ \int_{\{|z| \le 1\}} G(s,z) F(dz) \right\}^2 &= \left\{ \int_{\{|z| \le 1\}} |z|^{-p} G(s,z) \widetilde{F}(dz) \right\}^2 \\ &\le C_2 \int_{\{|z| \le 1\}} |z|^{-2p} G(s,z)^2 \widetilde{F}(dz) \\ &= C_2 \int_{\{|z| \le 1\}} |z|^{-p} G(s,z)^2 F(dz). \end{split}$$

Assembling the different terms and noting that

$$G(s,z)^{2} \leq 2\left\{\psi(s,z)^{2} + (\ln(1+\psi(s,z)))^{2}\right\}$$

yields the result with $C = 4(C_1 \vee C_2)$.

Lemma 5 Under the assumptions of Theorem 4, we have $P(T) \in L^1$, $\langle M^c, M^c \rangle_T \in L^2$, and $(x - \ln(1 + x)) * \nu_T \in L^2$.

Proof. By Jensen's inequality, $E\{(\int_0^T \sigma_s^2 ds)^2\} \leq TE\{\int_0^T \sigma_s^4 ds\}$, which is finite by hypothesis. So $\langle M^c, M^c \rangle_T \in L^2$. To prove that $P(T) \in L^1$ it therefore suffices to note that

$$E\Big\{(\ln(1+x))^2 * \mu_T^M\Big\} = E\Big\{(\ln(1+\psi))^2 * \mu_T\Big\} = E\Big\{(\ln(1+\psi))^2 * (ds \otimes F(dz))\Big\},\$$

which is finite by assumption, since $1 \vee |z|^{-p} \ge 1$. Finally, an application of Lemma 4 with s = 0 and t = T gives a bound on $E\{((x - \ln(1 + x)) * \nu_T)^2\}$ that is finite by assumption. The lemma is proved.

Proof of Theorem 4. By Lemma 5, the hypotheses of Lemma 3 are satisfied, so we have the bound

$$\begin{split} \left| E(P(T)) - E(P^n(T)) \right| &\leq E\left(\sum_{i=1}^n (\delta_i A)^2\right) + 2E\left(\sum_{i=1}^n |\delta_i M^c| \delta_i A\right) + 2E\left(\sum_{i=1}^n |\delta_i N| \delta_i A\right) \\ &= (\mathbf{I}) + (\mathbf{II}) + (\mathbf{III}), \end{split}$$

where $N = \ln(1+x) * (\mu^M - \nu)$ and $A = \frac{1}{2} \langle M^c, M^c \rangle + (x - \ln(1+x)) * \nu$. We first deal with (I). Using the inequality $(x+y)^2 \le 2x^2 + 2y^2$ we obtain

$$\sum_{i=1}^{n} E\Big((\delta_i A)^2\Big) \le \sum_{i=1}^{n} \left\{ \frac{1}{2} E\Big((\delta_i \langle M^c, M^c \rangle)^2\Big) + 2E\Big((\delta_i (x - \ln(1+x)) * \nu)^2\Big) \right\}.$$

Now, Jensen's inequality yields

$$E\Big\{(\delta_i \langle M^c, M^c \rangle)^2\Big\} = E\Big\{\Big(\int_{t_{i-1}}^{t_i} \sigma_s^2 ds\Big)^2\Big\} \le \frac{T}{n} E\Big\{\int_{t_{i-1}}^{t_i} \sigma_s^4 ds\Big\}.$$

Furthermore, Lemma 4 with $s = t_{i-1}$ and $t = t_i$ yields

$$E\Big\{(\delta_i(x - \ln(1+x)) * \nu)^2\Big\}$$

$$\leq \frac{C_1}{n} \int_{t_{i-1}}^{t_i} \int_{\mathbb{R}} (1 \vee |z|^{-p}) \{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\} F(dz) ds$$

for some constant C_1 independent of i. Summing over i shows that

$$\begin{split} (\mathbf{I}) &\leq \frac{T}{2n} E \Big\{ \int_0^T \sigma_s^4 ds \Big\} \\ &\quad + \frac{2C_1}{n} E \Big\{ \int_0^T \int_{\mathbb{R}} (1 \vee |z|^{-p}) \big\{ \psi(s,z)^2 + (\ln(1+\psi(s,z)))^2 \big\} F(dz) ds \Big\}, \end{split}$$

which is equal to n^{-1} times a constant C that is finite by assumption.

Concerning the two remaining terms (II) and (III), Lemma 3 gives

(II)
$$\leq 2\sqrt{E(\langle M^c, M^c \rangle_T)} \sqrt{E(\sum_{i=1}^n (\delta_i A)^2)}$$

and

(III)
$$\leq 2\sqrt{E([N,N]_T)}\sqrt{E(\sum_{i=1}^n (\delta_i A)^2)},$$

so that (II) + (III) $\leq C_2 \sqrt{(I)} \leq C_2 \sqrt{C} n^{-1/2}$ for some constant C_2 . The claim now follows with C as above and $D = C_2 \sqrt{C}$.

4 Examples

4.1 Strict local martingales

The literature on asset price bubbles centers around the phenomenon that S can be a strict local martingale under the risk neutral measure P, see [12], [19], [20]. Moreover, in [16] this issue has been noted to cause complications for pricing using PDE techniques. On the other hand, alternative criteria of no arbitrage type have been proposed by various authors to guarantee the existence of a *true* martingale measure, for instance [22], [6], [17]. It is therefore natural to ask about the relationship between our previous results and the true martingale (or strict local martingale) property of S. We give two examples, in the continuous case, showing that the two are not connected in general. More specifically, the

examples show that the martingale property of S has little to do with the integrability of $\langle M, M \rangle_T$.

Our first example uses the following criterion, which is well-known [6]. As in Section 3.2, B is standard Brownian motion.

Lemma 6 Assume S > 0 satisfies the stochastic differential equation $dS_t = \sigma(S_t)dB_t$. It is then a true martingale if and only if for some a > 0,

$$\int_{a}^{\infty} \frac{x}{\sigma^{2}(x)} dx = \infty.$$
(4)

Example 1 (S a strict local martingale and $\langle M, M \rangle_T \in L^2$) Consider the Constant Elasticity of Variance (CEV) models $dS_t = S_t^{\alpha} dB_t$. By Lemma 6, S is a strict local martingale if and only if $\alpha > 1$. We would like to choose $\alpha > 1$ such that $M_t = \int_0^t S_s^{\alpha-1} dB_s$ is a martingale with an integrable quadratic variation, i.e. $E\{(\int_0^t S_s^{2(\alpha-1)})^2\} < \infty$. This can be achieved with $\varepsilon \in (0, 1)$ and $\alpha = 1 + \frac{\varepsilon}{4} > 1$. Indeed,

$$E\left\{\left(\int_0^T S_s^{2(\alpha-1)} ds\right)^2\right\} \le TE\left(\int_0^T S_s^{4(\alpha-1)} ds\right) = TE\left(\int_0^T S_s^{\varepsilon} ds\right) = T\int_0^T E(S_s^{\varepsilon}) ds.$$

Since $\varepsilon \in (0,1)$, $x \mapsto x^{\varepsilon}$ is concave. Jensen's inequality thus implies that the right side above is dominated by $T \int_0^T E(S_s)^{\varepsilon} ds \leq T^2 S_0^{\varepsilon} < \infty$, where $E(S_s) \leq S_0$ because S is a positive local martingale, hence a supermartingale. This shows that S can be a strict local martingale, even if the quadratic variation $\langle M, M \rangle_T = \langle \ln S, \ln S \rangle_T$ is in L^2 .

For completeness, we also give a simple example showing that the reverse situation is also possible: that S can be well-behaved (a bounded martingale), while M is not.

Example 2 (S a bounded martingale and M a strict local martingale) Let X be the reciprocal of a Bessel(3) process. It is well-known that X is a strict local martingale, see e.g. [11, p. 20-21]. Set $M_t = -X_t$, so that M is a strict local martingale with $M_t \in L^1$ and $M_t < 0$ a.s. for all $t \ge 0$. Now, $P(T) = \langle M, M \rangle_T$ is not in L^1 (otherwise M would be a true martingale), and also $P^n(T)$ fails to be in L^1 by Proposition 1.

However, since $S_t = \mathcal{E}(M)_t = \exp\{-X_t - \frac{1}{2}\langle X, X \rangle_t\} \leq 1$, it is a bounded local martingale, hence a true martingale. In this example, the "bad" behavior of M is caused by its ability to take on very large negative values. This does not carry over to S, since it is obtained through exponentiation.

4.2 Stochastic volatility of volatility

We now proceed to give an example of a class of continuous stock price models that look innocuous, but where the conditions (3) at the end of Section 3.1 are satisfied for certain parameter values. In those cases, Proposition 1 implies that $E(P(T)) < \infty$ but $E(P^n(T)) = \infty$.

We use stochastic volatility models with stochastic volatility of volatility. Let B, W and Z be three Brownian motions, and let ρ denote the correlation between W and Z,

i.e. $d\langle W, Z \rangle_t = \rho dt$. No restrictions will be imposed on the correlation structure of (B, W, Z), other than through the parameter ρ . We consider the following model for the stock price S, its volatility v and the volatility of volatility w:

$$dS_t = S_t \sqrt{v_t} dB_t \tag{5}$$

$$dv_t = v_t \sqrt{w_t} dW_t \tag{6}$$

$$dw_t = \kappa(\theta - w_t)dt + \eta\sqrt{w_t}dZ_t,\tag{7}$$

where κ , θ , η are positive constants. Maintaining our previous notation, where M is the stochastic logarithm of S, we have that $M_t = \int_0^t \sqrt{v_s} dB_s$, so $\langle M, M \rangle_t = \int_0^t v_s ds$. Recall condition (3) from Section 3.1:

$$\left\{ \begin{array}{l} \langle M,M\rangle_T\in L^1\\\\ \langle M,M\rangle_T\notin L^2. \end{array} \right.$$

Note that v is a nonnegative local martingale and hence a supermartingale. Together with Fubini's theorem, this yields

$$E(\langle M, M \rangle_T) = E\left\{\int_0^T v_s ds\right\} = \int_0^T E(v_s) ds \le Tv_0,$$

so $\langle M, M \rangle_T \in L^1$. Now we wish to find conditions such that $\langle M, M \rangle_T = \int_0^T v_s ds \notin L^2$. To this end, define

$$T^* = \sup\{t : E(v_t^2) < \infty\},\$$

and let $\chi = 2\rho\eta - \kappa$ and $\Delta = \chi^2 - 2\eta^2$. It is proven in [1] that

$$T^* = \begin{cases} \frac{1}{\sqrt{\Delta}} \ln(\frac{\chi + \sqrt{\Delta}}{\chi - \sqrt{\Delta}}) & \text{if } \Delta \ge 0 \text{ and } \chi > 0, \\\\ \frac{2}{\sqrt{-\Delta}} \left(\arctan(\frac{\sqrt{-\Delta}}{\chi}) + \pi \mathbf{1}_{\{\chi < 0\}} \right) & \text{if } \Delta < 0, \\\\ +\infty & \text{otherwise.} \end{cases}$$

The next step is to establish that v is in fact a martingale.

Lemma 7 The process v defined above is a true martingale.

Proof. Using Feller's test of explosion (see e.g. [21]), a straightforward calculation shows that w does not explode under P. Therefore, using the same techniques as in [1], it suffices to establish that the auxiliary process \hat{w} , defined as the solution to

$$d\hat{w}_t = \left(\kappa(\theta - \hat{w}_t) + \rho\eta\hat{w}_t\right)dt + \eta\sqrt{\hat{w}_t}dZ_t, \qquad \hat{w}_0 = w_0,$$

is non-explosive. This can again be verified using Feller's criterion.

We may now conclude our construction by choosing the parameters ρ , η and κ such that $T^* < \infty$, and then choose $T > T^*$. In this case, Fubini's theorem implies that

$$E\left\{\left(\int_0^T v_s ds\right)^2\right\} = E\left\{\int_0^T \int_0^T v_s v_t ds dt\right\} = \int_0^T \int_0^T E(v_s v_t) ds dt.$$

Moreover, $E(v_s v_t) = E(v_s E(v_t | \mathcal{F}_s)) = E(v_s^2)$, so we get

$$E\left\{\left(\int_{0}^{T} v_{s} ds\right)^{2}\right\} = \int_{0}^{T} \int_{0}^{T} E(v_{s}^{2}) ds dt \ge \int_{T^{*}}^{T} \int_{T^{*}}^{T} E(v_{s}^{2}) ds dt = \infty.$$

We conclude that $\langle M, M \rangle_T \notin L^2$, and can summarize our findings as follows.

Example 3 Suppose in the stock price model with stochastic volatility of volatility described above, the parameters are such that $T > T^*$. Then the preceding discussion shows that

$$E(P(T)) < \infty$$
 but $E(P^n(T)) = \infty$.

That is, the approximation to the variance swap payoff has finite expectation, whereas the true payoff does not.

It is interesting to note that it is sometimes possible to change to an equivalent measure $Q \sim P$, under which the price process is still a local martingale, and such that both P(T) and $P^n(T)$ become integrable. The authors would like to thank Kerry Back [3] for posing the question of whether or not this can happen. To carry out the construction, let us continue to consider the stochastic volatility of volatility model described above.

Proposition 3 Assume that we are in the framework of a doubly stochastic volatility model as described in (5), (6), and (7). Suppose that $\Delta \ge 0$ and $\chi > 0$, so that $T^* < \infty$, and assume also that B is independent of (W, Z). Then there is an equivalent measure Q such that S is a local martingale under Q, and $\tilde{T}^* = \sup\{t : E_Q(v_t^2) < \infty\} = \infty$. As a consequence,

$$E_Q(P(T)) < \infty$$
 and $E_Q(P^n(T)) < \infty$

and we have $\lim_{n\to\infty} P^n(T) = P(T)$.

Proof. We can find a Brownian motion W', independent of W and B, such that

$$Z_t = \rho W_t + \sqrt{1 - \rho^2} W_t'$$

Let Q be the measure whose density process $Y_t = E_P(\frac{dQ}{dP} | \mathcal{F}_t)$ is given by $dY_t = -Y_t\gamma\sqrt{w_t}dW'_t$, where $\gamma > 0$ is a constant to be determined. To show that Y is indeed a martingale on [0, T], it suffices to verify, as in Lemma 7, that the auxiliary process

$$d\hat{w}_t = \left(\kappa(\theta - \hat{w}_t) - \gamma \rho \eta \hat{w}_t\right) dt + \eta \sqrt{\hat{w}_t} dZ_t, \qquad \hat{w}_0 = w_0,$$

is non-explosive. This can again be done using Feller's criteria. Next, it follows from Girsanov's theorem that the dynamics of w under Q is given by

$$dw_t = \tilde{\kappa}(\tilde{\theta} - w_t)dt + \eta\sqrt{w_t}d\tilde{Z}_t,$$

where

$$\tilde{\kappa} = \kappa + \gamma \eta \sqrt{1 - \rho^2}, \qquad \tilde{\theta} = \frac{\kappa \theta}{\kappa + \gamma \eta \sqrt{1 - \rho^2}},$$

and $d\tilde{Z}_t = dZ_t + \gamma \sqrt{1 - \rho^2} \sqrt{w_t} dt$ is Brownian motion under Q. Hence, if we define

$$\tilde{\chi} = 2\rho\eta - \tilde{\kappa}$$
 and $\tilde{\Delta} = \tilde{\chi}^2 - 2\eta^2$,

we have that $\tilde{T}^* = \infty$ if $\tilde{\chi} \leq 0$ and $\tilde{\Delta} \geq 0$. But

$$\tilde{\chi} = \chi - \gamma \eta \sqrt{1 - \rho^2}$$

and

$$\tilde{\Delta} = \Delta + \gamma \eta \sqrt{1 - \rho^2} (\gamma \eta \sqrt{1 - \rho^2} - 2\chi),$$

so it suffices to choose $\gamma \geq \frac{2\chi}{\eta\sqrt{1-\rho^2}}$.

The verification of the last assertion is straightforward: $E_Q(P(T)) < \infty$ is proved in the same way as under the measure P. To show that $E_Q(P^n(T)) < \infty$ and $\lim_{n\to\infty} P^n(T) = P(T)$, note that $\int_0^T E_Q(v_t^2) dt < \infty$ due to the continuity and finiteness of $E_Q(v_t^2)$ on the compact interval [0, T]. An application of Theorem 4 concludes the proof.

4.3 The 3/2-stochastic volatility model

A model that has received considerable attention both in the theoretical and empirical literature is the 3/2-stochastic volatility process. See for example [9] and the references therein. Let B and W be two correlated Brownian motions. The model prescribes the following dynamics for the stock price and its volatility

$$dS_t = S_t \sqrt{v_t} dB_t$$
$$dv_t = v_t (p + qv_t) dt + \epsilon v_t^{\frac{3}{2}} dW_t$$

where p, q and ϵ are constants such that $q < \frac{\epsilon^2}{2}$ and $\epsilon > 0$. The reason for the upper bound on q is to avoid explosion of v in finite time. To see this, consider the process $R_t = \frac{1}{v_t}$, the reciprocal of v, which satisfies the SDE

$$dR_t = (\epsilon^2 - q - pR_t)dt - \epsilon \sqrt{R_t}dW_t.$$

The process R is a square-root process, and it is well-known that this process avoids zero when $\epsilon^2 - q > \frac{\epsilon^2}{q}$, which is exactly the condition $q < \frac{\epsilon^2}{2}$. Let again M be the stochastic logarithm of S, i.e. $M_t = \int_0^t \sqrt{v_s} dB_s$, so that $\langle M, M \rangle_t = \int_0^t v_s ds$. Carr and Sun [9] provide the Laplace transform of the integrated variance $\int_0^T v_s ds$ in closed form.

Proposition 4 In the 3/2-model, the Laplace transform of the realized variance $\int_0^T v_s ds$ is given by

$$E(e^{-\lambda \int_0^T v_s ds}) = \frac{\Gamma(\gamma - \alpha)}{\Gamma(\gamma)} \left(\frac{2}{\epsilon^2 y_0}\right)^{\alpha} M(\alpha, \gamma, \frac{-2}{\epsilon^2 y_0})$$

where $y_0 = v_0 \frac{e^{pT} - 1}{p}$, $\alpha = -(\frac{1}{2} - \frac{q}{\epsilon^2}) + \sqrt{(\frac{1}{2} - \frac{q}{\epsilon^2})^2 + 2\frac{\lambda}{\epsilon^2}}$, $\gamma = 2(\alpha + 1 - \frac{q}{\epsilon^2})$, Γ is the Gamma function, and M is the confluent hypergeometric function

$$M(\alpha, \gamma, z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n}{(\gamma)_n} \frac{z^n}{n!}$$

with the notation $(x)_n = \prod_{i=0}^{n-1} (x+i)$.

Proof. We refer the reader to Carr and Sun [9]. \blacksquare

Since the Laplace transform of the realized variance exists in a neighborhood of zero, all moments of $\int_0^T v_s ds$ are finite. This implies in particular that $E(\langle M, M \rangle_T) \in L^2$. From Proposition 1, both the true variance swap payoff and its approximation have finite expectation.

Recall now the following result proved by Dufresne [15] on the finiteness of moments of the square-root process.

Proposition 5 Let $\overline{v} = \frac{2(\epsilon^2 - q)}{\epsilon^2}$. Then

$$\begin{array}{ll} \forall \ p < \overline{v}, \qquad E(R_t^{-p}) < \infty \\ \forall \ p \geq \overline{v}, \qquad E(R_t^{-p}) = \infty \end{array}$$

and for all $p \geq -\overline{v}$,

$$E(R_t^p) = \mu_t^p e^{-\lambda_t} \frac{\Gamma(\overline{v} + p)}{\Gamma(\overline{v})} M(\overline{v} + p, p, \lambda_t)$$

where $\mu_t = \frac{\epsilon^2}{2} \frac{1 - e^{-pt}}{p}$, $\lambda_t = \frac{2pv_0}{\epsilon^2(e^{-pt} - 1)}$, Γ is the Gamma function, and M is the congruent hypergeometric function defined in Proposition 4.

If q < 0, define $\kappa = -q > 0$ and $\theta = \frac{p}{\kappa}$. Then the SDE satisfied by v can be re-written as

$$dv_t = \kappa v_t (\theta - v_t) dt + \epsilon v_t^{\frac{3}{2}} dW_t$$

So under the condition q < 0, the process v is mean-reverting with a rate of mean-reversion proportional to v. Also $\overline{v} > 2$ when q < 0, so using Proposition 5, it can be seen that $E(v_t^2) = E(R_t^{-2})$ is finite and integrable on [0, T] as a continuus function on this compact time interval. Hence the condition of Theorem 4 is satisfied and the expectation of $P^n(T)$ converges to the expectation of P(T) as $n \to \infty$.

Under the condition $0 \le q < \frac{\epsilon^2}{2}$, it follows that $1 < \overline{v} \le 2$ and Proposition 5 implies that $E(v_t^2) = \infty$. By Fubini's theorem, $E(\int_0^T v_s^2 ds) = \infty$ so that the condition of Theorem 4 fails and the convergence of the $E(P^n(T))$ to E(P(T)) is not guaranteed anymore.

We now summarize the above findings.

Example 4 Suppose the stock price follows the 3/2-stochastic volatility model. The above discussion shows that

(i) Both the true payoff $P^n(T)$ and the approximation P(T) have finite expectation.

- (ii) If q < 0, i.e. when the squared volatility process is mean reverting, $P^n(T)$ converges to P(T) as $n \to \infty$.
- (iii) If $q \ge 0$, our sufficient condition fails and we can no longer guarantee that $P^n(T)$ converges to P(T). It is an open problem to establish whether or not this convergence actually takes place.

References

- L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. Finance and Stochastics, 11:29–50, 2007.
- [2] J. Ansel and C. Stricker. Couverture des actifs contingents et prix maximum. Annales de l'IHP, 30(2):303–315, 1994.
- [3] K. Back, personal communication, 2010.
- [4] M. Broadie and A. Jain. Pricing and hedging volatility derivatives. Journal of Derivatives, 15(3):7–24, 2008.
- [5] M. Broadie and A. Jain. The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps, *International Journal of Theoretical and Applied Finance*, 11 (8), 761 - 797, 2008.
- [6] P. Carr, A. Cherny, and M. Urusov. On the martingale property of time-homogeneous diffusions. *Preprint*, 2007.
- [7] P. Carr and R. Lee. Volatility derivatives. Annual Review of Financial Economics, 1:319–339, 2009.
- [8] P. Carr and R. Lee. Variance swaps on time-changed Levy processes, working paper, University of Chicago, 2009.
- [9] P. Carr and J. Sun. A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, 10:87–150, 2007.
- [10] L. Chan and E. Platen. Exact pricing and hedging formulas of long dated variance swaps under a 3/2 volatility model, working paper, University of Technology, Sydney, Austrialia, 2009.
- [11] K.L Chung and R.J. Williams, An Introduction to Stochastic Integration, Second Edition, Birkhäuser, Boston, 1990.
- [12] A. Cox and D. Hobson. Local martingales, bubbles and option prices. Finance and Stochastics, 9:477–492, 2005.
- [13] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathemathische Annalen*, 300:463–520, 1994.

- [14] F. Delbaen and W. Schachermayer. Fundamental theorem of asset pricing for unbounded stochastic processes. *Mathemathische Annalen*, 312:215–250, 1998.
- [15] D. Dufresne. The integrated square-root process. *Research paper no. 90, University* of Melbourne, 2001.
- [16] E. Ekström and J. Tysk. Bubbles, convexity and the black-scholes equation. Annals of Applied Probability, 19:1369–1384, 2009.
- [17] D. Hobson. Comparison results for stochastic volatility models via coupling. *Finance and Stochastics*, 14:129–152, 2010.
- [18] J. Jacod and A. Shiryaev. Limit Theorems for Stochastic Processes. Springer-Verlag, second edition, 2003.
- [19] R. Jarrow, P. Protter, and K. Shimbo. Asset price bubbles in a complete market. Advances in Mathematical Finance, In Honor of Dilip B. Madan:105–130, 2006.
- [20] R. Jarrow, P. Protter, and K. Shimbo. Asset price bubbles in incomplete markets. Mathematical Finance, 20(2):145–185, 2010.
- [21] L. Rogers and D. Williams. Diffusions, Markov Processes and Martingales, Volume 2: Itô Calculus. Cambridge University Press, 1994.
- [22] C. A. Sin. Complications with stochastic volatility models. Advances in Applied Probability, 30(1):256–268, 1998.