CRDE

CAHIER 30-2001

A THEORETICAL COMPARISON BETWEEN INTEGRATED AND REALIZED VOLATILITIES

Nour MEDDAHI

**Centre de recherche
et développement en économique**

C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7

Téléphone : (514) 343-6557
Télécopieur : (514) 343-5831
crde@crde.umontreal.ca
http://www.crde.umontreal.ca/

Université de Montréal

CAHIER 30-2001

A THEORETICAL COMPARISON BETWEEN INTEGRATED AND REALIZED VOLATILITIES

Nour MEDDAHI[1]

[1] Centre de recherche et développement en économique (C.R.D.E.), CIRANO and Département de sciences économiques (Université de Montréal), and CEPR

December 2001

The author would like to thank Bryan Campbell, Jean-Marie Dufour, René Garcia, Benoit Perron, Éric Renault, and Neil Shephard for helpful comments on an earlier draft. The author acknowledges financial support from Fonds FCAR, IFM2, and MITACS, and is solely responsible for any remaining errors.

RÉSUMÉ

Dans cet article, nous quantifions qualitativement et quantitativement la précision de la mesure de la volatilité intégrée par la volatilité réalisée quand la fréquence d'observations est fixée. Nous commençons par caractériser pour une diffusion générale la différence entre les volatilités réalisée et intégrée pour une fréquence d'observations donnée. Ensuite, nous calculons l'espérance et la variance de ce bruit ainsi que sa corrélation avec la volatilité intégrée en supposant que la diffusion est un modèle à volatilité stochastique par fonctions propres de Meddahi (2001a). Ce modèle contient, comme exemples particuliers, les modèles de diffusion log-normale, affine et GARCH. En utilisant certains résultats empiriques, nous montrons que l'écart-type du bruit n'est pas négligeable par rapport à la moyenne et à l'écart-type de la volatilité intégrée même si on considère des rendements à cinq minutes. Nous proposons aussi une approche simple pour extraire l'information sur la volatilité intégrée contenue dans les rendements via l'effet de levier.

Mots clés : volatilité intégrée, volatilité réalisée, générateur infinitésimal, modèles à volatilité stochastique par fonctions propres, effet de levier, moments exacts

ABSTRACT

In this paper, we provide both qualitative and quantitative measures of the cost of measuring the integrated volatility by the realized volatility when the frequency of observation is fixed. We start by characterizing for a general diffusion the difference between the realized and the integrated volatilities for a given frequency of observations. Then, we compute the mean and variance of this noise and the correlation between the noise and the integrated volatility in the Eigenfunction Stochastic Volatility model of Meddahi (2001a). This model has, as special examples, log-normal, affine, and GARCH diffusion models. Using some previous empirical works, we show that the standard deviation of the noise is not negligible with respect to the mean and the standard deviation of the integrated volatility, even if one considers returns at five minutes. We also propose a simple approach to capture the information about the integrated volatility contained in the returns through the leverage effect.

Key words : integrated volatility, realized volatility, infinitesimal generator, eigenfunction stochastic volatility models, leverage effect, exact moments

1 Introduction

Several recent works highlighted the importance of using the high frequency data to measure the volatility. These include Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Ebens (2001), Andersen, Bollerslev, Diebold and Labys (2001a, ABDL hereafter), Barndorff-Nielsen and Shephard (2001a-b), Taylor and Xu (1997) and Zhou (1996); for a survey on this literature, see Andersen, Bollerslev and Diebold (2001) and Dacorogna et al. (2001). Typically, when one is interested on the volatility over, say, a day, then these papers propose to measure this volatility by the sum of the intra-daily squared returns, like returns over five or thirty minutes. This measure of volatility is called the realized volatility. The theoretical justification is that when the length of the intra-daily returns tends to zero, this sum tends in probability to the quadratic variation of the underlying diffusion process (ABDL, 2001a). This quadratic variation plays a central role in the option pricing literature. In particular, when there are no jumps, the quadratic variation corresponds to the integrated volatility highlighted by Hull and White (1987).

An important characteristic of the high frequency data is the presence of microstructure effects. More precisely, the observed prices (or quotes) are in general the bid and/or the ask ones; they are ticked, i.e. the difference between two consecutive prices (quotes) is in $\{\Delta i, i \in \mathbf{Z}\}$, where Δ is a fraction of, say, a Dollar; zero returns are observed frequently. Therefore, using data at the highest available frequency to measure the volatility is not necessarily the best solution since such measures may be contaminated by these microstructure effects. The solution adopted by the literature is to consider intra-daily returns over an intermediate frequency. For instance, when ABDL (2001b) address the issue of forecasting the volatility by using realized volatilities, the latter were based on intra-daily returns over thirty minutes.

The main objective of the paper is to provide both qualitative and quantitative measures of the cost of measuring the integrated volatility by the realized volatility for a given frequency. In particular, we characterize the quality of the measures when one moves from a frequency to another one.

In all the paper, we will neither consider microstructure effects (see Bai, Russell and Tiao, 2001) nor parameters uncertainty. We will assume that the underlying data generating process is a continuous time model. We will derive the properties of the difference between the integrated volatility and the realized volatility computed with intra-daily returns for a given frequency. Thus, our study may be viewed as a benchmark when one considers the difference between the integrated and realized volatilities. In the sequel, we call the noise the random variable defined as the realized volatility minus the integrated volatility. Hence, we adopt the traditional terminology of the econometric literature when a variable of interest, here the integrated volatility, is observed with a measurement error, the observation being the realized volatility.

We start by characterizing this noise term in a general setting. The form of the noise allows us to give three of its qualitative characteristics. First, the unconditional mean of the noise is nonzero if and only if the drift of the diffusion characterizing the asset returns is nonzero. Second, the noise is heteroskedastic. In particular, its conditional variance is correlated with the integrated and realized volatilities. Third, the noise is correlated with the integrated volatility if and only if there is leverage effect or the drift depends on the instantaneous volatility.

In order to quantify these three characteristics, we specify the continuous time model. We assume that the underlying continuous time process is an Eigenfunctions Stochastic Volatility model (ESV) of Meddahi (2001). This class contains most of the popular SV models, in particular the log-normal model of Hull and White (1987) and Wiggins (1987), the square-root and affine models of Heston (1993) and Duffie, Pan and Singleton (2000) respectively, and the GARCH diffusion model of Nelson (1990). In this setting, we derive explicitly the mean and the variance of the noise and its correlation with the integrated volatility.

These theoretical results complement those of Barndorff-Nielsen and Shephard (2001b). These authors provide two important theoretical results. In the first one, they give in a general setting a Central Limit Theorem of the convergence of the realized volatility to the integrated when the length of the intra-daily returns tends to zero. Thus, they provide the speed of convergence and the asymptotic variance of the noise term. In the second result, they characterize the mean and variance of the noise when the underlying instantaneous variance process is a linear combination of independent positive Lévy processes of Barndorff-Nielsen and Shephard (2001a).[1] In both cases, the authors ruled out the leverage effect while they assumed a driftless model in the second case. Thus, our results are the extension of the second results of Barndorff-Nielsen and Shephard (2001b) to the case where the underlying diffusion process governing the volatility is general and where there is leverage effect and drift. Moreover, we provide also the first order limit of the mean and variance of the noise term. Therefore, while it is not a Central Limit Theorem, our results complement those of Barndorff-Nielsen and Shephard (2001b). In particular, we show that this first order limit does not depend on the leverage effect. This may suggest that the asymptotic result of Barndorff-Nielsen and Shephard (2001b) holds also when there is leverage effect.

After deriving the theoretical formulae of the mean and variance of the noise and its correlation with the integrated volatility, we quantify these values by taking explicit examples from the literature. These examples are: i) the GARCH diffusion models without drift and leverage effect used by Andersen and Bollerslev (1998) and Andreou and Ghysels (2001); ii) the affine models with drift and leverage effect estimated by Andersen, Benzoni and Lund (2001) on

[1] As advocated by these authors, their results hold also when the variance process is a marginalization of a vector of factors, where this vector admits a Vector Autoregressive representation of order one, VAR(1). Andersen (1994) firstly introduced such models in discrete time and called them the Square-Root Stochastic Autoregressive Volatility (SR-SARV) models while Meddahi and Renault (1996, 2000) introduced them in continuous time and showed their robustness against temporal and cross-sectional aggregations.

the S&P500; iii) the log-normal model with drift and leverage effect estimated by Andersen, Benzoni and Lund (2001) on the S&P500.

The main findings of this empirical illustration are the following. First, the mean of the noise is very small relatively to the mean of the integrated volatility when one uses intra-daily observations. In particular, it is smaller (in absolute value) than .2%. Second, the standard deviation of the noise is relatively important with respect to the mean and the standard deviation of the integrated volatility. In particular, when one uses realized volatility based on returns at five (resp thirty) minutes, the ratio of the standard deviation of the noise over the mean of the integrated volatility is around 10% (resp 25%). At the same frequencies, the ratio of the variance of the noise over the variance of the integrated volatility is around 5% (resp 10%) and some times much more. These two ratios suggest that the noise is important even when one considers five minutes returns. Third, under leverage effect, the autocorrelation between the noise and the integrated volatility is very small. Besides, the results based on square-root and log-normal models are almost the same. Finally, we found that by using the first order asymptotic approximation, one gets results that are very close to ones obtained by using exact formulae. This is very interesting because of the simplicity of the first order results with respect to the exact ones.

We also suggest an approach to extract the information about the integrated volatility contained in the returns through the leverage effect. It turns out that, in practice, this additional information is negligible.

Finally, note that we also characterize the difference between the integrated and realized co-volatilities. These co-volatilities are important in a multivariate setting. Studying the properties of this difference is left for future research.

The paper is organized as follows. In Section 2, we characterize the noise in the univariate and bivariate cases and discuss their qualitative properties. In Section 3, we recap the main properties of the ESV models of Meddahi (2001a). In the fourth Section, we compute explicitly the mean and variance of the noise and the correlation between the noise and the integrated volatility. At each step, we give an empirical illustration of the importance of these terms. The last section concludes while all the proofs are provided in the Appendix.

2 Preliminary Results

In this section, we characterize the difference between the realized and integrated volatilities in a general framework. This will be useful in Section 3. We also characterize this difference in the multivariate case.

2.1 Relationship between the integrated and realized volatilities

Consider S_t a continuous time process representing the price of an asset or the exchange rate between two currencies. Assume that it is characterized by the following stochastic differential equation:

$$d\log(S_t) = m_t dt + \sigma_t dW_t \tag{2.1}$$

where W_t is a standard Brownian process. We assume that m_t is general and may depend, for instance, on σ_t and $log(S_t)$. The process σ_t is also general and we allow for leverage effect, i.e., if one assumes that σ_t^2 is characterized by

$$d\sigma_t^2 = \tilde{m}_t dt + \tilde{\sigma}_t d\tilde{W}_t,$$

then we allow $d\tilde{W}_t$ to be correlated with dW_t. We assume here (without loss of generality) that the time t corresponds to a day. Consider a real h such that $1/h$ is an integer and define the realized volatility $RV_t(h)$ by

$$RV_t(h) \equiv \sum_{i=1}^{1/h} r_{t-1+ih}^{(h)2} \tag{2.2}$$

where $r_{t-1+ih}^{(h)}$ is the return over the period $[t-1+(i-1)h; t-1+ih]$, i.e.

$$r_{t-1+ih}^{(h)} \equiv \log\left(\frac{S_{t-1+ih}}{S_{t-1+(i-1)h}}\right). \tag{2.3}$$

When h goes to zero, the realized volatility converges in L^2. The limit is called the quadratic variation or the integrated volatility in the financial literature. It is denoted by IV_t and defined by[2]

$$IV_t \equiv \int_{t-1}^{t} \sigma_u^2 du. \tag{2.4}$$

Barndorff-Nielsen and Shephard (2001b) provide an asymptotic theory of the convergence of the realized volatility to the integrated volatility. In particular, they show that given the information $\sigma(\sigma_u, t-1 \leq u \leq t)$, we have

$$\sqrt{h^{-1}}(RV_t(h) - IV_t) \longrightarrow \mathcal{N}(0, 2\int_{t-1}^{t} \sigma_u^4 du). \tag{2.5}$$

While $RV_t(h)$ converges to IV_t when $h \to +\infty$, the difference may be not negligible for a given h. In order to study the difference, define $\mu_{t-1+ih}^{(h)}$ and $\varepsilon_{t-1+ih}^{(h)}$ by

$$\mu_{t-1+ih}^{(h)} \equiv \int_{t-1+(i-1)h}^{t-1+ih} m_u du \quad \text{and} \quad \varepsilon_{t-1+ih}^{(h)} \equiv \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u. \tag{2.6}$$

It is clear that

$$r_{t-1+ih}^{(h)} = \mu_{t-1+ih}^{(h)} + \varepsilon_{t-1+ih}^{(h)}. \tag{2.7}$$

[2]If one incorporates jumps in (2.1), then the quadratic variation will be equal to the integrated volatility plus an additional term due to the jumps.

Therefore,

$$(r^{(h)}_{t-1+ih})^2 = (\mu^{(h)}_{t-1+ih})^2 \ + \ 2\mu^{(h)}_{t-1+ih}\varepsilon^{(h)}_{t-1+ih} \ + \ (\varepsilon^{(h)}_{t-1+ih})^2.$$

Hence,

$$(r^{(h)}_{t-1+ih})^2 = \int_{t-1+(i-1)h}^{t-1+ih} \sigma^2_u du + (\mu^{(h)}_{t-1+ih})^2 + 2\mu^{(h)}_{t-1+ih}\varepsilon^{(h)}_{t-1+ih} + \left((\varepsilon^{(h)}_{t-1+ih})^2 - \int_{t-1+(i-1)h}^{t-1+ih} \sigma^2_u du\right).$$

To understand the properties of the third term, it is useful to rewrite it in terms of a stochastic integral. This is the purpose of the following Proposition, where we use Ito's Lemma to characterize the noise defined as the difference between the realized and integrated volatilities:

Proposition 2.1 Characterizing the noise. *Let h be a positive real such that $1/h$ is an integer, i an integer and consider the processes S_t, $RV_t(h)$, $r^{(h)}_{t-1+ih}$, IV_t, $\mu^{(h)}_{t-1+ih}$ and $\varepsilon^{(h)}_{t-1+ih}$ defined respectively in (2.1), (2.2), (2.3), (2.4), the left part of (2.6) and right part of (2.6). Then:*

$$(r^{(h)}_{t-1+ih})^2 = \int_{t-1+(i-1)h}^{t-1+ih} \sigma^2_u du + u^{(h)}_{t-1+ih}, \text{ where} \tag{2.8}$$

$$u^{(h)}_{t-1+ih} = (\mu^{(h)}_{t-1+ih})^2 \ + \ 2\mu^{(h)}_{t-1+ih}\varepsilon^{(h)}_{t-1+ih} + 2\int_{t-1+(i-1)h}^{t-1+ih} (\int_{t-1+(i-1)h}^{u} \sigma_s dW_s)\sigma_u dW_u. \tag{2.9}$$

Hence,

$$RV_t(h) = IV_t + u_t(h) \tag{2.10}$$

where

$$u_t(h) = \sum_{i=1}^{1/h} u^{(h)}_{t-1+ih}. \tag{2.11}$$

From this proposition, we can make some general remarks about the characteristics of $u^{(h)}_{t-1+ih}$ and, hence, $u_t(h)$. The first term in the right part of (2.9) implies that when the drift m_u is not zero, the mean of $u^{(h)}_{t-1+ih}$ and $u_t(h)$ are in general nonzero. The drift m_u is obviously nonzero in the stocks cases. This is also the case for exchange rates when one considers intra-daily data. For example, Andersen and Bollerslev (1997) showed that intra-daily returns of exchange rates $r^{(h)}_{t-1+ih}$ are correlated when one considers five minutes sample data. This nonzero mean is due to microstructure effects. This lead, for instance, Bai, Russel and Tiao (2001) to study the microstructure effects on measuring the integrated volatility by the realized volatility.

Note that this nonzero mean of $u_t(h)$ is not in contradiction with the asymptotic result (2.5) of Barndorff-Nielsen and Shephard (2001b). This asymptotic result implies that

$$lim_{h\to 0}\frac{E[u_t(h)]}{\sqrt{h}} = 0.$$

Consider now the third term in the right part of (2.9). It implies that $u^{(h)}_{t-1+ih}$ and $u_t(h)$ are in general heteroskedastic. This is problematic since the difference between the integrated and realized volatilities will have a higher variance when the instantaneous variance σ^2_t and integrated volatility IV_t are high. Note that this heteroskedasticity is not a surprising result since it is implicit in (2.5).

This third term implies also that under leverage effect, $u_{t-1+ih}^{(h)}$ and $u_t(h)$ are in general correlated with the integrated volatility IV_t. Note as well that under leverage effect, if the drift m_u depends on the volatility, the mean of the second term in (2.9) is nonzero.

Most of the previous remarks are well known. For instance, Barndorff-Nielsen and Shephard (2001b) pointed out that the mean of the noise is nonzero and that the noise is heteroskedastic. However, the impact of the leverage effect is not considered in the literature. We will consider explicit examples in the fourth section to quantify the characteristics of $u_t(h)$.

Finally it is important to observe that we consider a different approach than Barndorff-Nielsen and Shephard (2001a-b) to study the finite sample properties of $u_t(h)$. Their proofs are done given the sample path of the volatility. However, they exclude leverage effects. Moreover, they assume that the drift is an affine function of the variance. Our proofs are based on Ito calculus in order to take into account the leverage effect, in particular if it is time-varying, and general drifts.

2.2 Relationship between the integrated and realized covolatilities

We consider now the multivariate case. Without loss of generality, we consider the bivariate case only. As in the previous section, assume that two processes s_{1t} and S_{2t} are given by

$$d\log(S_{jt}) = m_{jt}dt + \sigma_{jt}dW_{jt}, \quad j = 1, 2, \tag{2.12}$$

where W_{1t} and W_{2t} are two standard Brownian processes that may be correlated. This correlation, denoted ρ_t, may be constant or time-varying (for instance if it depends on the volatilities σ_{1t} and σ_{2t}). Define the realized covolatility $RV_{1,2,t}(h)$ by

$$RCoV_{1,2,t}(h) \equiv \sum_{i=1}^{1/h} r_{1,t-1+ih}^{(h)} r_{2,t-1+ih}^{(h)} \tag{2.13}$$

where $r_{j,t-1+ih}^{(h)}$ for $j = 1, 2$, is defined by

$$r_{j,t-1+ih}^{(h)} \equiv \log\left(\frac{S_{j,t-1+ih}}{S_{j,t-1+(i-1)h}}\right), \ j = 1, 2. \tag{2.14}$$

When h goes to zero, the realized covolatility converges in L^2. The limit, denoted $ICoV_{1,2,t}$, is called the quadratic covariation or integrated covolatility and is defined by

$$ICoV_{1,2,t} \equiv \int_{t-1}^{t} \rho_u \ \sigma_{1,u}\sigma_{2,u}du. \tag{2.15}$$

We will now characterize the difference between the realized and integrated covolatilities. For this purpose we also define, for $j = 1, 2$, $\mu_{j,t-1+ih}^{(h)}$ and $\varepsilon_{j,t-1+ih}^{(h)}$ by

$$\mu_{j,t-1+ih}^{(h)} \equiv \int_{t-1+(i-1)h}^{t-1+ih} m_{j,u}du \quad \text{and} \quad \varepsilon_{j,t-1+ih}^{(h)} \equiv \int_{t-1+(i-1)h}^{t-1+ih} \sigma_{j,u}dW_{j,u}, \ j = 1, 2. \tag{2.16}$$

Proposition 2.2 Characterization of the noise between the co-volatilities. *Let h be a positive real such that $1/h$ is an integer, i an integer and for $j = 1, 2$, consider the processes S_{jt}, $RCoV_{1,2,t}(h)$, $r^{(h)}_{j,t-1+ih}$, $ICoV_t$, $\mu^{(h)}_{j,t-1+ih}$ and $\varepsilon^{(h)}_{j,t-1+ih}$ defined respectively in (2.12), (2.13), (2.14), (2.15), the left part of (2.16) and right part of (2.16). Then:*

$$r^{(h)}_{1,t-1+ih} r^{(h)}_{2,t-1+ih} = \int_{t-1+(i-1)h}^{t-1+ih} \rho_u \ \sigma_{1,u}\sigma_{2,u} du \ + \ u^{(h)}_{1,2,t-1+ih} \tag{2.17}$$

where

$$u^{(h)}_{1,2,t-1+ih} = \mu^{(h)}_{1,t-1+ih}\mu^{(h)}_{2,t-1+ih} + \mu^{(h)}_{1,t-1+ih}\varepsilon^{(h)}_{2,t-1+ih} + \mu^{(h)}_{2,t-1+ih}\varepsilon^{(h)}_{1,t-1+ih}$$
$$+ \int_{t-1+(i-1)h}^{t-1+ih} (\int_{t-1+(i-1)h}^{u} \sigma_{2,s} dW_{2,s})\sigma_{1,u} dW_{1,u} + \int_{t-1+(i-1)h}^{t-1+ih} (\int_{t-1+(i-1)h}^{u} \sigma_{1,s} dW_{1,s})\sigma_{2,u} dW_{2,u}. \tag{2.18}$$

Hence,

$$RCoV_{1,2,t}(h) = ICoV_{1,2,t} + u_{1,2,t}(h) \tag{2.19}$$

where

$$u_{1,2,t}(h) = \sum_{i=1}^{1/h} u^{(h)}_{1,2,t-1+ih}. \tag{2.20}$$

As in the univariate case, if both drifts are nonzero, then the mean of $u_{1,2,t}(h)$ is nonzero. Moreover, $u_{1,2,t}(h)$ is heteroskedastic. Finally, observe that these results hold even if the integrated covolatility $ICoV_t$ is zero, which holds when $\rho_u = 0$. Little has been done in the literature about the realized covolatility; see however ABDL (2001b). We will not study more the covolatilities and leave this for future research.

3 Eigenfunction Stochastic Volatility Models

In this section, we recap the main important properties of the Eigenfunction Stochastic Volatility (ESV) models introduced in Meddahi (2001a).

3.1 The General Theory

Consider a univariate Markov stationary process f_t characterized by

$$df_t = \mu(f_t) + \sigma(f_t) dW^{(2)}_t \tag{3.1}$$

where $W^{(2)}_t$ is a standard Brownian process. Let $\mathcal{A}$ be the infinitesimal generator operator associated to f_t (see, e.g., Hansen and Scheinkman, 1995):

$$\mathcal{A}\phi(f_t) \equiv \mu(f_t)\phi'(f_t) + \frac{\sigma^2(f_t)}{2}\phi''(f_t) \tag{3.2}$$

where $\phi(f_t)$ is a square-integrable function and twice differentiable. A function ϕ is called an eigenfunction of the infinitesimal generator $\mathcal{A}$ with a corresponding eigenvalue $-\delta$ if

$$\mathcal{A}\phi(f_t) = -\delta\phi(f_t). \tag{3.3}$$

For a review on operator methods for continuous time Markov models, see Ait-Sahalia, Hansen and Scheinkman (2001). An obvious eigenvalue is zero associated to any nonzero constant. When $\{f_t\}$ is time-reversible, i.e. the conditional distributions of f_t given f_{t-1} and of f_t given f_{t+1} are the same,[3] the eigenvalues are reals. Hansen, Scheinkman and Touzi (1998) show that under appropriate boundary protocol, stationary scalar diffusions are time-reversible. So we make the time-reversibility assumption:

Assumption A1. The stationary process $\{f_t\}$ is time reversible.

The set of eigenvalues is called the spectrum of the operator $\mathcal{A}$. In the following, we assume:

Assumption A2. The spectrum of the infinitesimal generator operator $\mathcal{A}$ of $\{f_t\}$ is discrete and denoted $\{-\delta_i, i \in \mathbf{N}\}$ with $\delta_0 = 0$ and $\delta_0 < \delta_1 < \delta_2 < ... < \delta_i < \delta_{i+1}...$; the corresponding eigenfunctions are denoted $E_i(f_t)$, $i \in \mathbf{N}$.

This assumption is true for most of the important examples considered by the literature, excluding the GARCH diffusion model that we will consider at the end of this section. A sufficient assumption is that the operator $\mathcal{A}$ is compact; see Hansen, Scheinkman and Touzi (1998) and Ait-Sahalia, Hansen and Scheinkman (2001) for more details.

The eigenfunctions have some interesting properties. In particular:

i) two eigenfunctions $E_i(f_t)$ and $E_j(f_t)$ associated to two different eigenvalues are orthogonal:

$$E[E_i(f_t)E_j(f_t)] = 0; \tag{3.4}$$

ii) as a consequence, any nonconstant eigenfunction is centered:

$$E[E_i(f_t)] = 0; \tag{3.5}$$

iii) any eigenfunction is an autoregressive process of order one, in general heteroscedastic:

$$\forall h > 0,\ E[E_i(f_{t+h}) \mid f_\tau, \tau \leq t] = \exp(-\delta_i h)E_i(f_t); \tag{3.6}$$

iv) any square-integrable function g, i.e. $E[g(f_t)^2] < \infty$, may be written as a linear combination of the eigenfunctions, i.e.

$$g(f_t) = \sum_{i=0}^{\infty} a_i E_i(f_t) \quad \text{where} \quad a_i = E[g(f_t)E_i(f_t)] \quad \text{and} \quad \sum_{i=0}^{\infty} a_i^2 = E[g(f_t)^2] < \infty. \tag{3.7}$$

Therefore, $g(f_t)$ is the limit in mean-square of $\sum_{i=0}^{p} a_i E_i(f_t)$ when p goes to $+\infty$.

[3]Recall that f_t is assumed to be Markovian. Therefore, the conditional distribution of f_t given f_{t-1} (resp f_{t+1}) is also the conditional distribution of f_t given $\{f_\tau, \tau \leq t-1\}$ (resp $\{f_\tau, \tau \geq t+1\}$).

Meddahi (2001a) defines a a continuous time process $\{\log(S_t)\}$ as an ESV model of order p, ESV(p), with $\{f_t\}$ the underlying diffusion process if:

$$d\log(S_t) = m_t dt + \sigma_t \left[\sqrt{1-\rho^2} dW_t^{(1)} + \rho dW_t^{(2)}\right], \quad \text{with} \tag{3.8}$$

$$\sigma_t^2 = \sum_{i=0}^{p} a_i E_i(f_t), \quad \text{where} \quad \sum_{i=0}^{p} a_i^2 < \infty \tag{3.9}$$

and $W_t^{(1)}$ and $W_t^{(2)}$ are two independent standard Brownian processes.

The ESV class contains as special cases all the popular SV models, including the square-root model of Heston (1993), the affine one of Duffie, Pan and Singleton (2000), the log-normal model of Hull and White (1987) and Wiggins (1987) and the GARCH diffusion model of Nelson (1990). The main reason of this result is due to (3.7). Most volatilty models define the variance process as a function of a particular state-variable f_t. It turns out that this function is always square-integrable. Therefore, this function may be expanded onto the eigenfunctions associated to the state variable. In the following section, we consider these examples in detail.

3.2 Examples

3.2.1 The square-root case

A popular SV model in the continuous time literature is the Heston (1993) model where the variance process σ_t^2 is square-root, i.e.

$$d\sigma_t^2 = k(\theta - \sigma_t^2)dt + \eta\sigma_t dW_t^{(2)}, \qquad k > 0.$$

Define the real α and the process f_t by

$$\alpha = \frac{2k\theta}{\eta^2} - 1, \quad f_t = \frac{2k}{\eta^2}\sigma_t^2. \tag{3.10}$$

Then, by Ito's Lemma, we have

$$df_t = k(\alpha + 1 - f_t) + \sqrt{2k}\sqrt{f_t} dW_t^{(2)}. \tag{3.11}$$

It turns out that the diffusion (3.11) admits as eigenfunctions the Laguerre polynomials $L_i^{(\alpha)}(f_t)$ associated to the eigenvalues $\delta_i = ki$. The Laguerre polynomials $L_i^{(\alpha)}$ are characterized by

$$\binom{i+\alpha}{i}^{1/2} i\, L_i^{(\alpha)}(x) = \binom{i-1+\alpha}{i-1}^{1/2} (-x+2i+\alpha-1) L_{i-1}^{(\alpha)}(x) - \binom{i-2+\alpha}{i-2}^{1/2} (i+\alpha-1) L_{i-2}^{(\alpha)}(x), \tag{3.12}$$

$$\text{where} \quad L_0^{(\alpha)}(x) = 1, \quad L_1^{(\alpha)}(x) = \frac{1+\alpha-x}{\sqrt{1+\alpha}}.$$

Thus, the variance process σ_t^2 is a linear combination of the constant and the first eigenfunction:

$$\sigma_t^2 = \theta - \frac{\theta}{\sqrt{\alpha+1}} L_1^{(\alpha)}(f_t), \quad \text{or}$$

$$\sigma_t^2 = a_0 L_0^{(\alpha)}(f_t) + a_1 L_1^{(\alpha)}(f_t) \quad \text{where} \quad a_0 = \theta \quad \text{and} \quad a_1 = -\frac{\sqrt{\theta}\eta}{\sqrt{2k}}. \tag{3.13}$$

It is easy to show that the affine model of Duffie, Pan and Singleton (2000) has also a variance process which is a linear function of the state variable (see Meddahi, 2001a).

3.2.2 The log-normal example

Another popular SV model is the log-normal model of Hull and White (1987) and Wiggins (1987) where σ_t is defined by

$$d\log(\sigma_t^2) = k[\theta - \log(\sigma_t^2)]dt + \sigma dW_t^{(2)}. \tag{3.14}$$

Define the state variable f_t by

$$f_t \equiv \frac{\sqrt{2k}}{\sigma}(\log \sigma_t^2 - \theta). \tag{3.15}$$

By using Ito's Lemma, we get

$$df_t = -kf_t dt + \sqrt{2k}\, dW_t^{(2)}. \tag{3.16}$$

The eigenfunction associated to the Ornstein-Uhlenbeck process (3.16) are the Hermite polynomials H_i associated to the eigenvalues $\delta_i = ki$ and characterized by

$$H_0(x) = 1,\; H_1(x) = x \text{ and } \forall i > 1, H_i(x) = \frac{1}{\sqrt{i}}\{xH_{i-1}(x) - \sqrt{i-1}H_{i-2}(x)\}. \tag{3.17}$$

Meddahi (2001a) shows that

$$\sigma_t^2 = \sum_{i=0}^{\infty} a_i H_i(f_t), \quad \text{where} \quad a_i = \exp(\theta + \frac{\sigma^2}{4k})\frac{(\sigma/\sqrt{2k})^i}{\sqrt{i!}}. \tag{3.18}$$

3.2.3 The GARCH diffusion example

The third popular SV model is the GARCH diffusion one considered by Nelson (1990):

$$d\sigma_t^2 = k(\theta - \sigma_t^2)dt + \sigma\sigma_t^2 dW_t^{(2)}. \tag{3.19}$$

This process was first introduced by Wong (1964). The polynomial solutions of

$$k(\theta - x)\phi'(x) + \frac{\sigma^2 x^2}{2}\phi''(x) = -\delta\phi(x)$$

are known as the Bessel polynomials. However, there exists an integer i_0 such that for all $i > i_0$, the Bessel polynomials are not square-integrable with respect to the stationary marginal density of σ_t^2. As a consequence, the Bessel polynomials of order higher than i_0 are not in the domain of the infinitesimal generator and, hence, are not eigenfunctions. Fortunately, Wong (1964) derives the eigenfunctions of the diffusion infinitesimal generator by using results on

Sturm-Liouville equations. In particular, he shows that the spectrum is mixed or continuous and that the eigenfunctions are hypergeometric functions.

In the general theory of ESV models, we exclude such example since we assume that the spectrum is discrete. However, we can easily incorporate this example (see Meddahi, 2001a). The main difference is in the expansion result (3.7). When the spectrum is continuous, one has to consider an integral instead of a sum. We will not consider such expansion in this paper. However, we will assume that the variance is a GARCH diffusion model and that the second moment of the variance σ_t^2 is finite. This means that the first eigenfunction is an affine function and that the variance depends only on this eigenfunction and the constant. Observe that Andersen and Bollerslev (1998) and Andreou and Ghysels (2001) who consider this example also assume the existence of the second moment of σ_t^2, in order to use the weak GARCH results of Drost and Werker (1996).

By using Wong (1964) results, we know that the stationary density of σ_t^2 is given by[4]

$$f(x) = \frac{(\sigma^2/2k\theta)^{[-(2k/\sigma^2)-1]}}{\Gamma((2k/\sigma^2)+1)} x^{[-(2k/\sigma^2)-2]} \exp(-\frac{2k\theta}{\sigma^2} x^{-1}) \tag{3.20}$$

where $\Gamma(.)$ is the Gamma function defined by

$$\Gamma(a) = \int_0^{+\infty} u^{-a-1} \exp(-u) du. \tag{3.21}$$

A moment of order r of σ_t^2 exists if and only if $r + [-(2k/\sigma^2) - 2] < -1$. Consider an alternative parameterization considered by Andersen and Bollerslev (1998) and Andreou and Ghysels (2001):

$$\sigma = \sqrt{2k\lambda} \quad \text{with} \quad \lambda > 0.$$

Then a moment of order r exists if and only if $r < 1 + \frac{1}{\lambda}$. In their Monte Carlo simulations, Andersen and Bollerslev (1998) and Andreou and Ghysels (2001) consider $\lambda = .296$ and $\lambda = .480$. Thus, they assume the following bounds for moments order: 4.378 and 3.084 respectively. Moreover, the second moment of σ_t^2 exists if and only if $\lambda < 1$. While it is well known that the mean of σ_t^2 is θ, it is easy to show by using the marginal distribution of σ_t^2 that its variance is $Var[\sigma_t^2] = (\theta^2\lambda)/(1-\lambda)$. Therefore, the first orthonormal eigenfunction is given by

$$E_1(\sigma_t^2) = \frac{\sqrt{1-\lambda}}{\theta\sqrt{\lambda}}(\sigma_t^2 - \theta). \tag{3.22}$$

Moreover,

$$\sigma_t^2 = a_0 E_0(\sigma_t^2) + a_1 E_1(\sigma_t^2) \quad \text{where} \quad a_0 = \theta \;\text{ and }\; a_1 = \frac{\theta\sqrt{\lambda}}{\sqrt{1-\lambda}}. \tag{3.23}$$

[4]For a stationary scalar diffusion process x_t solution of $dx_t = \mu(x_t)dt + \sigma(x_t)dW_t$, the stationary density function of x_t is, up to a scale, given by $\sigma(x)^{-2}\exp(\int_z^x \frac{\mu(u)}{\sigma^2(u)})du$. The scale parameter is chosen such that the density integral is equal to one.

3.3 The multifactor case

Meddahi (2001a) considers also the case where the variance is a function of several factors. Without loss of generality, we consider the two factors case. Let $f_{1,t}$ and $f_{2,t}$ be two independent stochastic processes characterized by

$$df_{j,t} = \mu_j(f_{j,t})dt + \sigma_j(f_{j,t})dW_{j,t}, \; j = 1, 2, \tag{3.24}$$

where the eigenfunctions (resp eigenvalues) of the corresponding infinitesimal generator are denoted $E_{1,i}(f_{1,t})$ and $E_{2,i}(f_{2,t})$ (resp $\delta_{1,i}$ and $\delta_{2,i}$). Then, the variance process σ_t^2 is defined by

$$\sigma_t^2 = \sum_{0 \leq i,j \leq p} a_{i,j} E_{1,i}(f_{1,t}) E_{2,j}(f_{2,t}) \quad \text{where} \quad \sum_{0 \leq i,j \leq p} a_{i,j}^2 < \infty.$$

It turns out that the four properties of the eigenfunctions defined in (3.4), (3.5), (3.6) and (3.7) also hold for the functions $E_{i,j}(f_t)$ defined by

$$E_{i,j}(f_t) \equiv E_{1,i}(f_{1,t}) E_{2,j}(f_{2,t}) \quad \text{where} \quad f_t \equiv (f_{1,t}, f_{2,t})'. \tag{3.25}$$

In other words, $E_{i,j}(f_t)$ are the eigenfunctions associated to the bivariate state variable $(f_{1,t}, f_{2,t})$.[5] Note that all the results that we will show later hold also when one considers a multifactor model without leverage effect.

4 Characteristics of the noise

In this section, we quantify the importance of the noise term. We start by computing its mean and variance. At each step, we illustrate this importance by considering examples from the literature.

4.1 Mean of the noise

We assume that the processes $\log(S_t)$, σ_t and f_t are defined by (3.8), (3.9) and (3.1). Besides, the drift m_u is assumed to be

$$m_u = \sum_{i=0}^{p} b_i E_i(f_u) \quad \text{with} \quad \sum_{i=0}^{p} | b_i | < +\infty. \tag{4.1}$$

Observe that the condition (4.1) implies that $\sum_{i=0}^{p} b_i^2 < +\infty$. Thus, we include any example where the drift is assumed to be a square-integrable function of f_t. In particular, if the drift is assumed to be an affine function of the variance, i.e.

$$m_u = c + d\sigma_u^2, \tag{4.2}$$

[5] See Chen, Hansen and Scheinkman (2000) for a general approach of eigenfunctions modeling in the multivariate case.

then the coefficients b_i are given by

$$b_0 = c + da_0, \quad b_i = da_i, \quad i \geq 1. \tag{4.3}$$

In the following Proposition, we use, for a give i, the reals $\{e_{i,j}\}$ and p_i defined in the (L^2) expansion of $\sigma_t\sigma(f_t)E_i'(f_t)$ onto the eigenfunctions, i.e.

$$\sigma_t\sigma(f_t)E_i'(f_t) = \sum_{j=0}^{p_i} e_{i,j}E_j(f_t), \tag{4.4}$$

where $E_i'(.)$ is the first derivative of $E_i(.)$. Thus, we assume that $\sigma_t\sigma(f_t)E_i'(f_t)$ is square-integrable. For more details, see Meddahi (2001b).

Proposition 4.1 Mean of the noise. *Let h be a positive real such that $1/h$ is an integer, and consider the processes $\log(S_t)$, σ_t, f_t, m_t, $RV_t(h)$, IV_t and $u_t(h)$ defined respectively in (3.8), (3.9), (3.1), (4.1), (2.2), (2.4) and (2.10). Then:*

$$E[u_t(h)] = hb_0^2 + \frac{2}{h}\left(\sum_{i=1}^{p} \frac{b_i(b_i + \rho e_{i,0})}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]\right). \tag{4.5}$$

As a consequence, when $h \to 0$, we get

$$E[u_t(h)] \;\sim\; h[b_0^2 + \sum_{i=1}^{p} b_i(b_i + \rho e_{i,0})]. \tag{4.6}$$

As pointed out in Section 2, the mean of $u_t(h)$ in nonzero when the drift is nonzero. Besides, (4.5) involves the leverage effect parameter ρ. This is due to the mean of the second term in (2.9) which is nonzero. The equation (4.6) gives additional information than the asymptotic result (2.5) of Barndorff-Nielsen and Shephard (2001b); (4.6) means that when $h \to 0$:

$$\frac{E[u_t(h)]}{\sqrt{h}} \;\sim\; \sqrt{h}[b_0^2 + \sum_{i=1}^{p} b_i(b_i + \rho e_{i,0})].$$

To assess the importance of this mean, we consider the empirical results of Andersen, Benzoni and Lund (2001). These authors estimated several models on daily returns of the S&P500.[6] In particular they estimated the square-root and log-normal models without and with leverage effect. They consider an affine drift as in (4.3). They rejected all these models. However we consider their empirical results in order to get a first order approximation of the importance of (4.5) and (4.6) since their estimated models are the best ones in these popular SV classes.[7] Note also that models without leverage effect were strongly rejected by these authors.

[6] The sample period is 01/02/1953-12/31/1996.

[7] Models including jumps in (3.8) that we exclude in our study were not rejected by Andersen, Benzoni and Lund (2001). Note however that ESV models of Meddahi (2001a) can have jumps.

To measure the importance of the mean of the noise, we consider the following criterion:

$$Ratio = 100\ \frac{E[u_t(h)]}{E[IV_t]}. \tag{4.7}$$

In other words, we present the ratio, in percentage, of the mean of the noise term over the mean of the integrated volatility. The results are provided in Table 1. We give the value of the ratio defined in (4.7) by using both the exact formula (4.5) and the asymptotic one (4.6).

Table 1

Model		Affine		Log-normal	
1/h	freq	Ratio-Ex	Ratio-Asym	Ratio-Ex	Ratio-Asym
1	day	.168	.168	.179	.179
24	1 hour	.00701	.00701	.00747	.00745
48	30 mn	.00351	.00351	.00373	.00373
96	15 mn	.00175	.00175	.00187	.00186
144	10 mn	.00117	.00117	.00124	.00124
288	5 mn	.000584	.000584	.000622	.000621

From Table 1,[8] it is clear that the results based on both exact and asymptotic formulae are the same. Moreover, the mean of the noise is almost the same in both affine and log-normal models. Finally and more importantly, the mean of the noise is relatively negligible with intra-daily data, for instance when one uses returns based on hourly data or higher frequency.[9]

4.2 Variance of the noise

In the sequel, we will assume that the drift is constant, i.e.

$$m_u = b_0. \tag{4.8}$$

Proposition 4.2 Variance of the noise term. *Let h be a positive real such that $1/h$ is an integer, and consider the processes $\log(S_t)$, σ_t, f_t, m_t, $RV_t(h)$, IV_t and $u_t(h)$ defined respectively in (3.8), (3.9), (3.1), (4.1), (2.2), (2.4) and (2.10). Assume that the drift m_u is given by (4.8). Then:*

$$Var[u^{(h)}_{t+ih}] = 4a_0 b_0^2 h^3 + 8b_0 h\rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h] + 4\sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$

$$+8\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j}\left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)}\right]\right] \tag{4.9}$$

where $e_{i,j}$ and p_i are defined in (4.4) and under the convention

$$\frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i-\delta_j)} - \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} = -\frac{\exp(-\delta_i h)(1+\delta_i h)}{\delta_i^2} \quad \text{when} \quad \delta_i = \delta_j.$$

[8]The results of the log-normal model are based on the expansion (4.5) by taking the first 100 terms.
[9]As we mentioned in the introduction, we do not consider impact the of microstructure effects.

Moreover, the random variables $u^{(h)}_{t+ih}$ are uncorrelated. Hence,

$$Var[u_t(h)] = \frac{Var[u^{(h)}_{t+ih}]}{h}. \tag{4.10}$$

Finally, when $h \to 0$, we have:

$$Var[u_t(h)] \; \sim \; h \; 2 \sum_{i=0}^{p} a_i^2. \tag{4.11}$$

Observe that if there is neither drift nor leverage effect, $Var[u_t(h)]$ becomes

$$Var[u_t(h)] = \frac{1}{h}\left(\frac{a_0^2 h^2}{2} + \sum_{i=1}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]\right).$$

This formula is exactly the same one that Barndorff-Nielsen and Shephard (2001b) derive. However, these authors assume that the variance process is a linear combination of p independent factors. Thus, one needs only orthogonality between the factors (as for the eigenfunctions) to get this formula.

The asymptotic result (4.11) implies that $lim_{h\to 0} Var[u_t(h)] = 0$, i.e. the realized volatility tends to the integrated volatility in mean-square[10] and, hence, in probability. This is the well known result mentioned in Section 2. Moreover, if one considers the asymptotic result equation (2.5), it is easy to show that

$$E[\int_{t-1}^{t} \sigma_u^4 du] = \sum_{i=0}^{p} a_i^2. \tag{4.12}$$

Therefore, we show here that the asymptotic variance of the noise $u_t(h)$, i.e. $Var[\sqrt{h^{-1}} u_t(h)]$, does not depend on the leverage effect. This suggests that the asymptotic result of Barndorff-Nielsen and Shephard (2001b) holds also when there is leverage effect, at least unconditionally.

In order to quantify the importance of the variance of the noise for a given frequency, we consider several examples. We start by considering models without drift and leverage effect. The first examples are the square-root models estimated by Bollerslev and Zhou (2001) on daily exchange rate data by using the realized volatilities.[11] They estimated a model with one factor and another with two factors. They rejected only the first one. The second examples we consider are the GARCH diffusions models by Andersen and Bollerslev (1998) and Andreou and Ghysels (2001). After that, we consider two other examples with leverage effect where the drift is constant. The first examples we consider are the same models than the previous subsection, that is the square-root and log-normal models estimated by Andersen, Benzoni and Lund (2001).[12]

[10] Recall that the mean of the noise tends to zero when h does.

[11] Galbraith and Zinde-Walsh (2001) and Maheu and McCurdy (2001) consider also estimation of GARCH models by using realized volatilities.

[12] Notice that we do not take into account the affine term in the drift, that is we assume that the drift is constant. Before estimating their models, Andersen, Benzoni and Lund (2001) filtered the data in order to remove the dependence in the returns.

We use two criteria in order to measure the importance of the variance of the noise:

$$Ratio1 = 100\,\frac{\sqrt{Var[u_t(h)]}}{E[IV_t]} \quad \text{and} \quad Ratio2 = 100\,\frac{Var[u_t(h)]}{Var[IV_t]}. \tag{4.13}$$

The first criterion is clearly related to the length of the confidence interval of the integrated volatility. The second criterion is appealing because of the randomness of the integrated volatility. Typically, when the noise is uncorrelated with the integrated volatility,[13] the variance of the realized volatility is the sum of the variances of the noise and the integrated volatility. This ratio is crucial when one considers filtering the integrated volatility from the realized volatilities.

Table 2-a Affine without leverage

		1 Fac.		2 Fac.		1 Fac.	2 Fac.	1 Fac.	2 Fac.
1/h	freq	Std-Ex	Std-As	Std-Ex	Std-As	Ratio1	Ratio1	Ratio2	Ratio2
1	day	1.29	1.31	.749	.753	249	149	295	2137
24	1 hour	.267	.268	.154	.154	51.7	30.5	12.7	89.9
48	30 mn	.189	.189	.109	.109	36.6	21.5	6.34	45.0
96	15 mn	.134	.134	.0768	.0768	25.9	15.2	3.17	22.5
144	10 mn	.109	.109	.0627	.0627	21.1	12.4	2.12	15.0
288	5 mn	.0773	.0773	.0444	.0444	14.9	8.80	1.06	7.50

In Table 2-a, we report the results based on the models estimated by Bollerslev and Zhou (2001). The first interesting result is that computing the standard deviation of the noise by using the exact formula or by using the asymptotic first order approximation is almost the same when one uses intra-daily data. Besides, the standard deviation of the noise is almost divided by two when one goes from the one factor model to the two factors one. Therefore, since the unconditional mean of the integrated volatility is almost the same for both models (.517 and .504 respectively), the first criterion is also divided by two when one goes from the one factor model to the two factors model. Consider the two factors model that was not rejected by Bollerslev and Zhou (2001). The first criterion is 8.8% and 21.5% when one considers realized volatilities computed with five and thirty minutes returns respectively. This is clearly not negligible since it means that the length of the confidence interval of the integrated volatility is relatively large with respect to the integrated volatility. Of course, in this criterion we clearly did not take into account the dependence between the conditional standard deviation of the noise and the integrated volatility. Therefore, one has to be cautious with our results. However, by using the asymptotic theory developed in Barndorff-Nielsen and Shephard (2001b), Barndorff-Nielsen and Shephard (2001c) estimated empirically at each day the confidence interval of the integrated volatility and showed that its length is large and positively correlated with the integrated volatility.

[13] This holds when there is neither drift nor leverage effect; see the next subsection.

Consider now the second criterion. For the two factors model, this criterion is 7.5% and 45% when one considers realized volatilities computed with five and thirty minutes returns respectively. Again, this is not negligible, especially when one uses thirty minutes returns, and suggests that one has to filter the integrated volatility by using all the history of the realized volatility. Notice that for the one factor model, this criterion is relatively small. The main reason is that in this case, the integrated volatility is more volatile than for the two factors model (the standard deviations are .751 and .162 respectively). Finally, as we will see in the following examples, the second criterion is not so bad, in particular when one uses thirty minutes returns. The main reason is that in their inference procedure, Bollerslev and Zhou (2001) did not take into account the difference between the integrated and realized volatilities. More precisely, they derived theoretical moment conditions for the integrated volatilities while they used the realized volatilities in the estimation procedure. By doing this, these authors incorporated the noise term in the variance process. Therefore, they obtained a high variance of the variance which is crucial when one uses realized volatility instead of integrated volatility. In particular, the variance of the variance clearly appears in (4.12) and (2.5).

Table 2-b GARCH

Model		DM-US$		Yen-US$		DM	Yen	DM	Yen
1/h	freq	Std-Ex	Std-Asym	Std-Ex	Std-Asym	Ratio1	Ratio1	Ratio2	Ratio2
1	day	1.07	1.07	.930	.934	168.	195.	681	421
24	1 hour	.219	.219	.191	.191	34.4	40.0	28.5	17.7
48	30 mn	.155	.155	.135	.135	24.3	28.3	14.2	8.84
96	15 mn	.109	.109	.0953	.0953	17.2	20.0	7.12	4.42
144	10 mn	.0893	.0893	.0778	.0778	14.0	16.3	4.75	2.95
288	5 mn	.0632	.0632	.0550	.0550	9.93	11.6	2.37	1.47

Consider now the two GARCH diffusions models considered by Andersen and Bollerslev (1998) and Andreou and Ghysels (2001). They correspond to daily returns of DM-US$ and Yen-US$. The results are presented in Table 2-b.[14] Note again the small difference between using exact and first order approximations formulae for the standard deviation of the noise. Moreover, the results are almost the same for both DM-US$ and Yen-US$ returns. Thus, we consider only the results on DM-US$. The first criterion is still not negligible (around 10% with five minutes returns). Thus, the length of the confidence intervals will be relatively important. However, the second criterion is negligible at five minutes (2.37%) but not at thirty minutes (14.2%). This means that filtering the integrated volatility when one uses realized volatility computed with thirty (resp five) minutes returns will have a large (resp small) impact on the quality of the measure of the integrated volatility.

[14] Note that the variance of the noise corresponds to the MSE computed by simulation in Andersen and Bollerslev (1998). The exact results are very close to their ones.

Table 2-c Leverage

Model		Affine		Log-normal	Affine		Log-normal	
1/h	freq	Std-Ex	Std-Asym	Std-Asym	Ratio1	Ratio2	Ratio1	Ratio2
1	day	2.16	.892	.993	402.9	4176	180.2	523.8
24	1 hour	.200	.182	.203	37.3	35.8	36.8	21.8
48	30 mn	.135	.129	.143	25.3	16.4	26.0	10.9
96	15 mn	.0933	.0910	.101	17.4	7.83	18.4	5.46
144	10 mn	.0756	.0743	.0827	14.1	5.13	15.0	3.64
288	5 mn	.0530	.0526	.0585	9.91	2.53	10.6	1.82

Consider now the results on the affine and log-normal models with leverage effect estimated by Andersen, Benzoni and Lund (2001). The results are reported in Table 2-c. Consider the affine case. Again, the difference between the exact and asymptotic results is very small when one uses intra-daily observations. The first criterion is still not negligible (around 10% with five minutes returns) while the second one is at five minutes (around 2.5%) but not at thirty minutes (around 15%). Thus, filtering the integrated volatility when one uses realized volatility computed with thirty (resp five) minutes returns will have a large (resp small) impact on the quality of the measure of the integrated volatility. The same results hold for the log-normal model. Note that for this case, we use the asymptotic results.[15]

4.3 Covariance between the noise and the integrated volatility

Given that the variable of interest IV_t is observed with errors, it is of interest to characterize the covariance between IV_t and the noise term. This is important for estimation, filtering and forecasting purposes. This covariance is characterized in the following proposition:

Proposition 4.3 Covariance between the noise and the integrated volatility *Let h be a positive real such that $1/h$ is an integer, and consider the processes $\log(S_t)$, σ_t, f_t, m_t, $RV_t(h)$, IV_t and $u_t(h)$ defined respectively in (3.8), (3.9), (3.1), (4.1), (2.2), (2.4) and (2.10). Assume that the drift m_u is given by (4.8). Then:*

$$Cov(u^{(h)}_{t+ih}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma^2_u du) = 2b_0 h\rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$

$$+2\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j}\frac{e_{j,0}}{\delta_j}\left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i-\delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)}\right]\right]. \quad (4.14)$$

Besides,

$$Cov(u_t(h), IV_t) = \frac{1}{h} Cov(u^{(h)}_{t+ih}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma^2_u du). \quad (4.15)$$

Finally, when $h \to 0$, the correlation between the noise $u_t(h)$ and the integrated volatility IV_t is

$$Corr(u_t(h), IV_t) = O(h^{3/2}). \quad (4.16)$$

[15] We need to compute the coefficients $e_{i,j}$ that appear (4.4). The exact results will be included in the next version of the paper.

This proposition implies that when there is no leverage effect, we have

$$Cov(u_t(h), IV_t) = 0.$$

Moreover, (4.16) implies that the correlation between the noise and the integrated volatility tends to zero very fastly when one increases the frequency of intra-daily observations.

Table 3 Correlation

1/h	freq	Correlation
1	day	-0.0005
24	1 hour	-9.400e-06
48	30 mn	-3.472e-06
96	15 mn	-1.257e-06
144	10 mn	-6.904e-07
2880	5 mn	-2.440e-07

In order to assess the importance of this correlation, we consider models with leverage effect. In Table 3, we report this correlation for the affine model estimated by Andersen, Benzoni and Lund (2001).[16] These results clearly mean that the correlation between the noise and the integrated volatility is very small, for instance .0056 when one considers returns at thirty minutes. Hence, one can ignore this correlation.

4.4 Combining the realized volatility with the constant

We show previously that the realized volatility $RV_t(h)$ is a noisy version of the integrated volatility IV_t, i.e.

$$RV_t(h) = IV_t + u_t(h).$$

However, we are interested in the integrated volatility IV_t. Therefore, the best proxy of the integrated volatility is not the realized volatility but a combination of the latter with the constant. In other words, we have to consider the best linear predictor of IV_t given $RV_t(h)$ and the constant. This comes from the following linear regression

$$IV_t = a(h) + b(h)RV_t(h) + \eta_t(h). \tag{4.17}$$

It is obvious that

$$b(h) = \frac{Cov(IV_t, RV_t(h))}{Var[RV_t(h)]}, \quad a(h) = E[IV_t] - b(h)E[RV_t(h)] \tag{4.18}$$

and that the R2 of the regression (4.17) is given by

$$R2(h) = \frac{Cov(IV_t, RV_t(h))^2}{Var[IV_t]Var[RV_t(h)]}. \tag{4.19}$$

[16]The results of the Log-normal model will be included in the next version of the paper.

Observe that when IV_t and $u_t(h)$ are not correlated, we have:

$$b(h) = \frac{1}{1 + Var[u_t(h)]/Var[IV_t]}, \quad R2(h) = b(h) \quad \text{and} \quad \frac{Var[\psi_t(h)]}{Var[u_t(h)]} = R2(h). \tag{4.20}$$

Table 4

Model		Affine 2 Fac.	GARCH DM-US$	Affine	with	leverage
1/h	freq	b(h),R2(h),Ratio	b(h),R2(h),Ratio	b(h)	R2(h)	Ratio
1	day	.0447	.128	.0233	.0232	.0234
24	1 hour	.527	.778	.736	.736	.736
48	30 mn	.690	.875	.859	.859	.859
96	15 mn	.816	.934	.927	.927	.927
144	10 mn	.870	.955	.951	.951	.951
288	5 mn	.930	.977	.975	.975	.975

We report in Table 4 the different values of b(h), R2(h) and the ratio $Var[\psi_t(h)]/Var[u_t(h)]$. Note that $Var[\psi_t(h)]$ is the MSE of the regression (4.17) while $Var[u_t(h)]$ is the MSE that we considered previously section, that is when one considers the realized volatility as a measure for the integrated volatility.

We report the result of the affine model with two factors, the GARCH diffusion for the DM-US$ and the affine model with leverage effect. The results of the other models are similar. Note that for models without leverage effect, we report only one column for b(h), R2(h) and the ratio $Var[\psi_t(h)]/Var[u_t(h)]$ since they are the same (see (4.20)). From Table 4, it is clear that it is better to combine the constant and the realized volatility when one considers intra-daily returns at fifteen minutes or more. The Ratio of the MSEs suggests that this improvement is important even if one considers five minutes returns.

However, one has to be careful with this criterion when both measures are very good. The reason is the following. Consider three random variables y, x_1 and x_2. Let m_1 (resp m_2) be the best linear regression of y given x_1 (resp x_2), $R2_1$ and MSE_1 (resp $R2_2$ and MSE_2) the corresponding R2 and MSE. Then, it is easy to show that

$$\frac{MSE_1}{MSE_2} = \frac{1 - R2_1}{1 - R2_2}.$$

Thus, the ratio of the MSEs may be high (or small) when $R2_1$ and $R2_2$ are close to one, that is the explanatory variables x_1 and x_2 explain well the variable y. For instance, if $R_1 = .98$ and $R2_2 = .99$, then the ratio of the MSEs is two. This is exactly what happens in Table 4 when h is very small.

4.5 Extracting the information implied by the leverage effect

Assume now that there is leverage effect. This implies that the daily return r_t defined by

$$r_t \equiv \log(S_t/S_{t-1}) \tag{4.21}$$

is correlated with the integrated volatility IV_t. Therefore, a natural question is: How can we extract the information about the integrated volatility contained in the daily return through the leverage effect?

A simple appraoch Hence, a simple is to add this return r_t in the regression (4.17), i.e. by considering the regression

$$IV_t = a_1(h) + b_1(h)RV_t(h) + c_1(h)r_t + \psi_t(h). \tag{4.22}$$

Again, we have the following relationships

$$\begin{bmatrix} b_1(h) \\ c_1(h) \end{bmatrix} = \left(Var \begin{bmatrix} RV_t(h) \\ r_t \end{bmatrix} \right)^{-1} \begin{bmatrix} Cov(RV_t(h), IV_t) \\ Cov(r_t, IV_t) \end{bmatrix}. \tag{4.23}$$

The theoretical formulae of $b_1(h)$ and $c_1(h)$ are provided in the Appendix B. We give in Table 5 the different values of $b(h)$, $b_1(h)$ and $c_1(h)$ for the affine model with leverage effect estimated by Andersen, Benzoni and Lund (2001). These results mean that the contribution of the daily return r_t to explain the integrated volatility is very small when one uses realized volatility computed with intra-daily data. This is not very surprising since the convergence of the realized volatility to the integrated volatility when the length of intra-daily returns tends to zero. However it is surprising for the daily frequency case.

Table 5

h	freq	$b(h)$	$b_1(h)$	$c_1(h)$
1	day	.0233	.0232	-.0205
24	1 hour	.736	.736	-.00501
48	30 mn	.859	.859	-.00265
96	15 mn	.927	.927	-.00135
144	10 mn	.951	.951	-.000908
288	5 mn	.975	.975	-.000457

5 Conclusion

In this paper, we characterize the noise defined as the difference between the realized and integrated volatilities for a given frequency of observations. Then, we provide qualitative and quantitative results about the characteristics of this difference termed the noise term. The main findings are threefold. First, under leverage effect or time varying drift, the mean of the noise is nonzero but negligible compared to the mean of the integrated volatility. Second, the noise is heteroskedastic and its standard deviation is not negligible with respect to the mean and the standard deviation of the integrated volatility even if one considers returns at five minutes. Third, under leverage effect, the correlation of the noise with the integrated volatility is nonzero but very small when one considers intra-daily data.

We also show that combining the realized volatility with the constant or some other variables reduces the noise. In particular, it is better to consider the linear regression of the integrated volatility on the constant and the realized volatility. Moreover, under leverage effect, we can add the daily return to extract the information that it contains about the integrated volatility through the leverage effect. It turns out that the improvement is small.

Our work can be extended in different directions. The first one is to take into account in our analysis the parameter uncertainty, since in practice these parameters have to be estimated. Moreover, we ignore microstructures effects. A simple approach for incorporating them is by assuming that one of the factors is a continuous time Markov chain. It turns out that such processes also admit an eigenfunctions decomposition.

Two other major extensions are currently considered. The first one is to incorporate jumps in the stock or its volatility. Assuming the characteristics of the jumps, i.e. their intensity and sizes, are functions of the same state variable we consider will be very useful. This is exactly what happens in the affine models with jumps of Duffie, Pan and Singleton (2000). The second extension is related to the quadratic power variations considered by Barndorff-Nielsen and Shephard (2001d). As advocated by these authors, the difference between

$$\int_{t-1}^{t} \sigma_u^{\gamma} du \quad \text{and} \quad \sum_{i=1}^{1/h} \mid r_{t-1+ih}^{(h)} \mid^{\gamma}$$

is smaller when one considers γ equal to one (for instance) instead of two. Thus, finding the optimal γ that reduces the importance of the microstructure effects is of interest. Interestingly, the eigenfunctions expansion is still valid in this case and very easy indeed. For instance, if one considers the log-normal model, then we have

$$\sigma_t^{\gamma} = \sum_{i=0}^{\infty} a_i(\gamma) H_i(f_t), \quad \text{where} \quad a_i(\gamma) = \exp\left(\frac{\theta\gamma}{2} + \frac{\sigma^2\gamma^2}{16k}\right) \frac{(\sigma\gamma/\sqrt{8k})^i}{\sqrt{i!}}.$$

Thus, using our approach will be useful in this case.

References

Ait-Sahalia, Y., L.P. Hansen and J. Scheinkman (2001), "Operator Methods for Continuous-Time Markov Models," chapter prepared for the Handbook of Financial Econometrics.

Andersen, T.G. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling," *Mathematical Finance*, 4, 75-102.

Andersen, T.G., L. Benzoni and J. Lund (2001), " An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, forthcoming.

Andersen, T.G. and T. Bollerslev (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115-158.

Andersen, T.G. and T. Bollerslev (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885-905.

Andersen, T.G., T. Bollerslev and F.X. Diebold (2001), "Parametric and Nonparametric Measurements of Volatility," chapter prepared for the Handbook of Financial Econometrics.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001), "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, forthcoming.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001a), "The Distribution of Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42-55.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001b), "Modeeling and Forecasting Realized Volatility," unpublished manuscript, Northwester University.

Andreou, E. and E. Ghysels (2001), "Rolling-Sampling Volatility Estimators: Some New Theoretical, Simulation and Empirical results," *Journal of Business and Economic Statistics,* forthcoming.

Bai, X., J. Russell and G.C. Tiao, "Beyon Merton's Utopia: Effects of Non-Normality and Dependence on the Precision of Variance Estimates Using High-Frequency Financial Data," unpublished manuscript, University of Chicago.

Barndorff-Nielsen, O.E. and N. Shephard (2001a), "Non-Gaussian OU based Models and some of their uses in Financial Economics," *Journal of the Royal Statistical Society, B*, 63, 167-241.

Barndorff-Nielsen, O.E. and N. Shephard (2001b), "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, B*, forthcoming.

Barndorff-Nielsen, O.E. and N. Shephard (2001c), " Estimating Quadratic Variation Using Realised Volatility," unpublished manuscript, Nuffield College, Oxford.

Barndorff-Nielsen, O.E. and N. Shephard (2001d), " Realised Power Variation and Stochastic Volatility," unpublished manuscript, Nuffield College, Oxford.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. and H. Zhou (2001), "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility," *Journal of Econometrics*, forthcoming.

Chen, X., L.P. Hansen and M. Carrasco (1999), "Nonlinearity and Temporal Dependence," unpublished manuscript, University of Chicago.

Chen, X., L.P. Hansen and J. Scheinkman (2000), "Principal Components and the Long Run," unpublished manuscript, University of Chicago.

Dacorogna, M.M., R. Gençay, U. Muller, R.B. Olsen and O.V. Pictet (2001), *An Introduction to High-Frequency Finance,* Academic Press.

Drost, F.C. and TH.E. Nijman (1993), "Temporal Aggregation of GARCH processes," *Econometrica*, 61, 909-927.

Drost, F.C. and B.J.M. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling," *Journal of Econometrics*, 74, 31-58.

Duffie, D., J. Pan and K. Singleton (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1342-1376.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

Engle, R.F. and G.G.J. Lee (1999), "A Permanent and Transitory Component Model for Stock Return Volatility," in *Cointegration, Causality and Forecasting: a festschrift in Honour of Clive W.J. Granger*, R.F. Engle and H. White eds., Oxford University Press.

Galbraith, J.W. and V. Zinde-Walsh (2001), "Properties of Estimates of Daily GARCH Parameters Based on Intra-Day Observations," CIRANO DP 2001s-15.

Hansen, L.P. and J. Scheinkman (1995), "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes," *Econometrica*, 63, 767-804.

Hansen, L.P., J. Scheinkman and N. Touzi (1998), "Spectral Methods for Identifying Scalar Diffusions," *Journal of Econometrics*, 86, 1-32.

Heston, S.L. (1993), "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327-344.

Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, Vol XLII, 281-300.

Maheu, J. and T.H. McCurdy (2001), "Nonlinear Features of Realised FX Volatility," unpublished manuscript, University of Toronto.

Meddahi, N. (2001a), "An Eigenfunction Approach for Volatility Modeling," unpublished manuscript, Université de Montréal, http://www.crde.umontreal.ca/online/nm-30-10-01.pdf.

Meddahi, N. (2001b), "Moments of Continuous Time Stochastic Volatility Models," in preparation.

Meddahi, N. and E. Renault (1996), "Aggregation and Marginalization of GARCH and Stochastic Volatility Models," GREMAQ DP 96.30.433, Univesité de Toulouse and CRDE DP 3597, Université de Montréal.

Meddahi, N. and E. Renault (2000), "Temporal Aggregation of Volatility Models," CIRANO DP 2000s-22.

Nelson, D.B. (1990), "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7-39.

Taylor, S.J. and X. Xu (1997), "The Incremental Volatility Information in One Million Foreign Exchange Qoutations," *Journal of Empirical Finance*, 4, 317-340.

Wiggins, J.B. (1987), "Options Values under Stochastic Volatility: Theory and Empirical Estimates", *Journal of Financial Economics*, 19, 351-372.

Wong, E. (1964), "The Construction of a Class of Stationary Markoff Processes," in Belleman, R., (Ed.), Sixteenth Symposium in Applied Mathematics, Stochastic Processes in Mathematical Physics and Engineering, American Mathematical Society, Providence, 264-276.

Zhou, B. (1996), "High-Frequency Data and Volatility in Foreign Exchange Rates," *Journal of Business and Economic Statistics*, 14, 45-52.

Appendix A

We start the Appendix by giving some lemmas.

Lemma A1 *Let f_t defined by (3.1). Then:*

$$dE_i(f_u) = -\delta_i E_i(f_u)du + E_i'(f_u)\sigma(f_u)dW_u^{(2)}, \tag{A.1}$$

$$\forall s < u, \quad E_i(f_u) = \exp(-\delta_i(u-s))E_i(f_s) + \exp(-\delta_i(u-s))\int_s^u \exp(\delta_i(w-s))\sigma(f_w)E_i'(f_w)dW_w^{(2)}, \tag{A.2}$$

$$E[\int_0^u E_i(f_u)\sigma_s dW_s] = \rho e_{i,0}\frac{1-\exp(-\delta_i u)}{\delta_i}, \tag{A.3}$$

where $e_{i,0}$ is given by $e_{i,0} = E[\sigma_s\sigma(f_s)E_i'(f_s)]$ (and defined in (4.4)). Thus,

$$E[\int_0^h(\int_0^u E_i(f_u)\sigma_s dW_s)du] = \rho e_{i,0}\frac{\exp(-\delta_i h)+\delta_i h - 1}{\delta_i^2}. \tag{A.4}$$

Proof: By Ito's Lemma, we get

$$dE_i(f_t) = \mathcal{A}E_i(f_t) + E_i'(f_t)\sigma(f_t)dW_t^{(2)}.$$

By defintion, $\mathcal{A}E_i(f_t) = -\delta_i E_i(f_t)$. Thus, we get (A.1).

Define z_u by $z_u = \exp(\delta_i u)E_i(f_u)$. By using Ito's Lemma we get $dz_u = \exp(\delta_i u)E_i'(f_u)\sigma(f_u)dW_u^{(2)}$. Hence, $z_u = z_s + \int_s^u \exp(\delta_i w)E_i'(f_w)\sigma(f_w)dW_w^{(2)}$. We then get (A.2). We have:

$$\begin{aligned}
E\left[\int_0^u E_i(f_u)\sigma_s dW_s\right] &= E\left[\int_0^u \exp(-\delta_i(u-s))E_i(f_s)\sigma_s dW_s\right] \\
&+E\left[\int_0^u \exp(-\delta_i(u-s))\left(\int_s^u \exp(\delta_i(w-s))\sigma(f_w)E_i'(f_w)dW_w^{(2)}\right)\sigma_s dW_s\right] \\
&= 0 + \int_0^u \exp(-\delta_i(u-s))[\exp(\delta_i(s-s))\sigma(f_s)E_i'(f_s)\sigma_s\rho ds] \\
&= \rho e_{i,0}\int_0^u \exp(-\delta_i(u-s))ds = \rho e_{i,0}\frac{1-\exp(-\delta_i u)}{\delta_i},
\end{aligned}$$

i.e. (A.3). From (A.3), one gets (A.4).□

Proof of Proposition 2.1. We have

$$r_{t-1+ih}^{(h)} = \int_{t-1+(i-1)h}^{t-1+ih} m_u du \; + \; \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u.$$

Therefore,

$$r_{t-1+ih}^{(h)2} = \left(\int_{t-1+(i-1)h}^{t-1+ih} m_u du\right)^2 + 2\left(\int_{t-1+(i-1)h}^{t-1+ih} m_u du\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u\right) + \left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u\right)^2.$$

Let us consider $\int_0^h \sigma_u dW_u$ and compute its square by using Ito's Lemma. We have

$$(\int_0^h \sigma_u dW_u)^2 = 2\int_0^h(\int_0^u \sigma_s dW_s)\sigma_u dW_u + \int_0^h \sigma_u^2 du.$$

Hence,

$$\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u\right)^2 = 2\int_{t-1+(i-1)h}^{t-1+ih}(\int_{t-1+(i-1)h}^{u} \sigma_s dW_s)\sigma_u dW_u + \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du.$$

As a consequence, we get (2.8) and, hence, (2.7).□

Proof of Proposition 2.2. We have

$$r^{(h)}_{1,t-1+ih} = \int_{t-1+(i-1)h}^{t-1+ih} m_{1,u}du \ + \ \int_{t-1+(i-1)h}^{t-1+ih} \sigma_{1,u}dW_{1,u}.$$

$$r^{(h)}_{2,t-1+ih} = \int_{t-1+(i-1)h}^{t-1+ih} m_{2,u}du \ + \ \int_{t-1+(i-1)h}^{t-1+ih} \sigma_{2,u}dW_{2,u}.$$

Hence,

$$\begin{aligned}
r^{(h)}_{1,t-1+ih}r^{(h)}_{2,t-1+ih} &= \left(\int_{t-1+(i-1)h}^{t-1+ih} m_{1,u}du\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} m_{2,u}du\right)\\
&+ \left(\int_{t-1+(i-1)h}^{t-1+ih} m_{1,u}du\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_{2,u}dW_{2,u}\right) + \left(\int_{t-1+(i-1)h}^{t-1+ih} m_{2,u}du\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_{1,u}dW_{1,u}\right)\\
&+ \left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_{1,u}dW_{1,u}\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_{2,u}dW_{2,u}\right).
\end{aligned}$$

By using Ito's Lemma, we get:

$$\begin{aligned}
&\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_{1,u}dW_{1,u}\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_{2,u}dW_{2,u}\right)\\
&= \int_{t-1+(i-1)h}^{t-1+ih}\left(\int_{t-1+(i-1)h}^{u} \sigma_{2,s}dW_{2,s}\right)\sigma_{1,u}dW_{1,u} + \int_{t-1+(i-1)h}^{t-1+ih}\left(\int_{t-1+(i-1)h}^{u} \sigma_{1,s}dW_{1,s}\right)\sigma_{2,u}dW_{2,u}\\
&+ \int_{t-1+(i-1)h}^{t-1+ih} \rho_u\sigma_{1,u}\sigma_{2,u}du.
\end{aligned}$$

As a consequence, we get (2.17) and, hence, (2.19).□

Proof of Proposition 4.1. Let μ_h and ε_h defined respectively by

$$\mu_h \equiv \int_0^h m_u du \quad \text{and} \quad \varepsilon_h \equiv \int_0^h \sigma_u dW. \tag{A.5}$$

By Ito's Lemma, we have:

$\mu_h^2 = 2\int_0^h \mu_u d\mu_u + \int_0^h d[\mu,\mu]_u = 2\int_0^h \mu_u m_u du = 2\int_0^h (\int_0^u m_s m_u ds)du$. Hence,

$E[\mu_h^2] = 2\int_0^h (\int_0^u E[m_s m_u]ds)du$. But, for $u \geq s$:

$E[m_s m_u] = \sum_{0\leq i,j\leq p} b_i b_j E[E_i(f_s)E_j(f_u)] = \sum_{0\leq i,j\leq p} b_i b_j \exp(-\delta_j(u-s))E[E_i(f_s)E_j(f_s)]$

$= \sum_{i=0}^p b_i^2 \exp(-\delta_i(u-s))$. Hence,

$\int_0^u E[m_s m_u]ds = \sum_{i=0}^p \frac{b_i^2}{\delta_i}[1-\exp(-\delta_i u)]$. As a consequence,

$E[\mu_h^2] = 2\sum_{i=0}^p \frac{b_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$.

Let $\tilde{\mu}_h$ defined by

$$\tilde{\mu}_h \equiv 2\int_0^h m_u du \int_0^h \sigma_u dW_u. \tag{A.6}$$

By Ito's Lemma, we have:
$\tilde{\mu}_h = 2\int_0^h (\int_0^u \sigma_s dW_s) m_u du + 2\int_0^h (\int_0^u m_s ds)\sigma_u dW_u$. Hence,
$E[\tilde{\mu}_h] = 2E[\int_0^h (\int_0^u \sigma_s dW_s) m_u du] = 2\sum_{i=0}^{p} b_i \int_0^h E(\int_0^u E_i(f_u)\sigma_s dW_s))du.$
$= 2\rho \sum_{i=0}^{p} \frac{b_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) + \delta_i h - 1]$ by (A.4).
The expectation of the third term in $u_{t+ih}^{(h)}$ is zero. Hence
$E[u_{t+ih}^{(h)}] = E[\mu_h^2] + E[\tilde{\mu}_h] = 2\sum_{i=0}^{p} \frac{b_i^2 + \rho b_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$
$= hb_0^2 + 2\sum_{i=1}^{p} \frac{b_i^2 + \rho b_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$ since $e_{0,0} = 0$.
Therefore we get (4.5) since $E[u_t(h)] = \frac{1}{h}E[u_{t+ih}^{(h)}]$.
For a small h, since $\delta_i \neq 0$, we have: $[\exp(-\delta_i h) - 1 + \delta_i h] \sim \delta_i^2 h^2/2$. Thus, (4.6) is deduced.□

Proof of Proposition 4.2. We have $Var[u_{t+ih}] = Var[b_0^2 h^2 + 2b_0 h\varepsilon_h + 2\tilde{Z}_h]$, where $\tilde{Z}_h \equiv \int_0^h (\int_0^u \sigma_s dW_s)\sigma_u dW_u$. Thus,
$Var[u_{t+ih}] = 4b_0^2 h^2 Var[\varepsilon_h] + 4Var[\tilde{Z}_h] + 8b_0 h Cov[\varepsilon_h, \tilde{Z}_h] = 4b_0^2 h^2 E[\varepsilon_h^2] + 4E[\tilde{Z}_h^2] + 8b_0 h E[\varepsilon_h \tilde{Z}_h]$.
We will compute the three terms:
i) We have: $E[\varepsilon_h^2] = E[\int_0^h \sigma_u^2 du] = a_0 h$.
ii) By Ito's Lemma, we have:
$\tilde{Z}_h^2 = 2\int_0^h \tilde{Z}_u \, d\tilde{Z}_u + \int_0^h d[\tilde{Z}, \tilde{Z}]_u = 2\int_0^h \tilde{Z}_u \, (\int_0^u \sigma_s dW_s)\sigma_u dW_u + \int_0^h (\int_0^u \sigma_s dW_s)^2 \sigma_u^2 du.$
Therefore,
$E[\tilde{Z}_h^2] = E\left[\int_0^h (\int_0^u \sigma_s dW_s)^2 \sigma_u^2 du\right] = \int_0^h E\left[\left(2\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s + \int_0^u \sigma_s^2 ds\right)\sigma_u^2 du\right].$
We have to compute the two terms. Consider the second one. We have:
$\int_0^h E\left[\left(\int_0^u \sigma_s^2 ds\right)\sigma_u^2\right] du = \sum_{i=0}^{p} a_i \int_0^h \left(\int_0^u E[E_i(f_u)\sigma_s^2]ds\right) du$
$= \sum_{i=0}^{p} a_i \int_0^h \left(\int_0^u \exp(-\delta_i(u-s))E[E_i(f_s)\sigma_s^2]ds\right) du = \sum_{i=0}^{p} a_i^2 \int_0^h \left(\int_0^u \exp(-\delta_i(u-s))ds\right) du$
$= \sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h].$
Consider now the first term. Let us compute at a first step $E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right) E_i(f_u)\right]$.
This term is zero when $i = 0$ since $E_0(.) = 1$. For $i \neq 0$, we have:
$E[\tilde{Z}_u E_i(f_u)] = E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right) E_i(f_u)\right] = E\left[\int_0^u \left(\int_0^s \sigma_w dW_w\right) E_i(f_u)\sigma_s dW_s\right]$ by (A.2),
$= E\left[\int_0^u \left(\int_0^s \sigma_w dW_w\right) \rho \exp(-\delta_i(u-s))E_i(f_u)'\sigma(f_s)\sigma_s ds\right]$ by (4.4),
$= \rho E\left[\int_0^u \left(\int_0^s \sigma_w dW_w\right) \exp(-\delta_i(u-s))(\sum_{j=0}^{p_i} e_{i,j} E_j(f_s))ds\right]$

$$= \rho \sum_{j=0}^{p_i} e_{i,j} \int_0^u \left(\int_0^s E[E_j(f_s)\sigma_w dW_w] \right) \exp(-\delta_i(u-s))ds$$
$$= \rho \sum_{j=0}^{p_i} e_{i,j} \int_0^u \left(\int_0^s E[\rho \exp(-\delta_j(s-w))E_j(f_w)'\sigma(f_w)\sigma_w dw] \right) \exp(-\delta_i(u-s))ds \text{ by (A.2),}$$
$$= \rho^2 \sum_{j=0}^{p_i} e_{i,j} \int_0^u \left(\int_0^s \exp(-\delta_j(s-w))e_{j,0}dw \right) \exp(-\delta_i(u-s))ds$$
$$= \rho^2 \sum_{j=1}^{p_i} e_{i,j} \int_0^u \left(\int_0^s \exp(-\delta_j(s-w))e_{j,0}dw \right) \exp(-\delta_i(u-s))ds \quad \text{since } e_{0,0}=0$$
$$= \rho^2 \sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{1-\exp(-\delta_i u)}{\delta_i} - \frac{\exp(-\delta_j u) - \exp(-\delta_i u)}{\delta_i - \delta_j} \right]$$

Hence,

$$E[\int_0^h \tilde{Z}_u \sigma_u^2 du] = \int_0^h E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w \right) \sigma_s dW_s \right) \sigma_u^2 du \right]$$
$$= \sum_{i=1}^{p} a_i \int_0^h E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w \right) \sigma_s dW_s \right) E_i(f_u) du \right]$$
$$= \rho^2 \sum_{i=1}^{p} a_i \int_0^h \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{1-\exp(-\delta_i u)}{\delta_i} - \frac{\exp(-\delta_j u) - \exp(-\delta_i u)}{\delta_i - \delta_j} \right] \right] du$$
$$= \rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)} \right] \right]. \qquad \text{(A.7)}$$

As a summary,

$$E[\tilde{Z}_h^2] = \sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2} [\exp(-\delta_i h) - 1 + \delta_i h]$$
$$+ 2\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)} \right] \right] \qquad \text{(A.8)}$$

iii) $\varepsilon_h \tilde{Z}_h = \int_0^h \tilde{Z}_u d\varepsilon_u + \int_0^h d\tilde{Z}_u \varepsilon_u + \int_0^h d[\tilde{Z}, \varepsilon]_u = \int_0^h \tilde{Z}_u \sigma_u dW_u + \int_0^h \varepsilon_u^2 \sigma_u dW_u + \int_0^h \sigma_u^2 \varepsilon_u du.$
Thus,

$$E[\varepsilon_h \tilde{Z}_h] = E[\int_0^h \sigma_u^2 \varepsilon_u du] = \sum_{i=0}^{p} a_i \int_0^h \int_0^u E[E_i(f_u)\sigma_s dW_s]$$
$$= \sum_{i=0}^{p} a_i \int_0^h \rho \frac{e_{i,0}}{\delta_i}(1 - \exp(-\delta_i u))du = \rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2} [\exp(-\delta_i h) - 1 + \delta_i h] \text{ since } e_{0,0} = 0.$$

Hence,

$$Var[u_{t+ih}] = 4a_0 b_0^2 h^3 + 8b_0 h \rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2} [\exp(-\delta_i h) - 1 + \delta_i h] + 4 \sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2} [\exp(-\delta_i h) - 1 + \delta_i h]$$
$$+ 8\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)} \right] \right], \text{ i.e. (4.9).}$$

The random variables $u_{t-1+(i-1)h}^{(h)}$ are uncorrelated since $\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u$ and $\int_{t-1+(i-1)h}^{t-1+ih} (\int_{t-1+(i-1)h}^{u} \sigma_s dW_s) \sigma_u dW_u$ are martingale difference sequences.

Thus, $Var[u_t(h)] = Var[u^{(h)}_{t-1+ih}]/h$.

For a small h, we have:

i) $[\exp(-\delta_i h) - 1 + \delta_i h] \sim \delta_i^2 h^2/2$;

ii) $\dfrac{h}{\delta_i} - \dfrac{1-\exp(-\delta_i h)}{\delta_i^2} - \dfrac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \dfrac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)} = \dfrac{h^3\delta_j}{6} + o(h^3)$.

Hence, for a small h, the dominant term in (4.9) is the third one. Thus, we get (4.11).□

Proof of Proposition 4.3. We have:

$$Cov(u^{(h)}_{t+ih}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du) = E[u^{(h)}_{t+ih} \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du] - E[u^{(h)}_{t+ih}]E[\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du]$$

$$= E[u^{(h)}_{t+ih} \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du] - h^3 a_0 b_0^2 = E[(b_0^2 h^2 + 2b_0 h\varepsilon_h + 2\tilde{Z}_h) \int_0^h \sigma_u^2 du] - h^3 a_0 b_0^2$$

$$= 2E[(b_0 h\varepsilon_h + \tilde{Z}_h) \int_0^h \sigma_u^2 du].$$

We have: $E[\varepsilon_h \int_0^h \sigma_u^2 du] = E[\int_0^h \varepsilon_u \sigma_u^2 du] = \sum_{i=0}^{p} a_i \int_0^h E[E_i(f_u) \int_0^u \sigma_s dW_s]$

$$= \rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h] \text{ since } e_{0,0} = 0 \text{ and (A.4).}$$

Moreover: $E[\tilde{Z}_h \int_0^h \sigma_u^2 du] = E[\int_0^h (\int_0^u \sigma_u^2 du) d\tilde{Z}_u] + E[\int_0^h \tilde{Z}_u \sigma_u^2 du] = E[\int_0^h \tilde{Z}_u \sigma_u^2 du]$

$$= \rho^2 \sum_{i=0}^{p} a_i \left[\sum_{j=0}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)}\right]\right] \text{ by (A.7).}$$

Hence,

$$Cov(u^{(h)}_{t+ih}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du) = 2b_0 h\rho \sum_{i=0}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$

$$+2\rho^2 \sum_{i=0}^{p} a_i \left[\sum_{j=0}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)}\right]\right].□$$

Appendix B

In this Appendix, we compute some variables used in the text or the Tables.

1- Mean and variance of the integrated volatility: Meddahi (2001b) shows that:

$$E[IV_t] = a_0 \text{ and} Var[IV_t] = 2\sum_{i=1}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i) - 1 + \delta_i]. \tag{B.1}$$

2- Coefficienst $e_{i,j}$: Meddahi (2001b) shows that:

For an affine model: $e_{1,0} = \sqrt{2k\theta}$ and $e_{1,1} = \sigma$.

For a log-normal model: $e_{i,0} = \sqrt{2k}\sqrt{j} \exp\left(\dfrac{\theta}{2} + \dfrac{\sigma^2}{16k}\right) \dfrac{(\sigma/\sqrt{8k})^{j-1}}{\sqrt{(j-1)!}}$.

3- Coefficients $b_1(h)$ $c_1(h)$ We have to compute the coefficients in the right part of (4.23). The coefficients $Cov(RV_t, IV_t)$ and $Var[RV_t(h)]$ are already computed.

i) $Cov(r_t, IV_t) = Cov(\int_{t-1}^t \sigma_u dW_u, \int_{t-1}^t \sigma_u^2 du) = E(\int_{t-1}^t \sigma_u dW_u \int_{t-1}^t \sigma_u^2 du)$

$= E(\int_{t-1}^{t}(\int_{t-1}^{u}\sigma_s dW_s)\sigma_u^2 du + \int_{t-1}^{t}(\int_{t-1}^{u}\sigma_s^2 ds)\sigma_u dW_u)$ by Ito's Lemma

$= E(\int_{t-1}^{t}(\int_{t-1}^{u}\sigma_s dW_s)\sigma_u^2 du) = \rho\sum_{i=1}^{p}\frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i) - 1 + \delta_i]$ by (A.4).

ii) $Cov(r_t, RV_t(h)) = Cov(r_t, IV_t) + Cov(r_t, u_t(h))$. But:

$Cov(r_t, u_t(h)) = Cov(\sum_{i=1}^{1/h} r_{t-1+ih}^{(h)}, \sum_{i=1}^{1/h} u_{t-1+ih}^{(h)}) = \frac{1}{h}Cov(r_{t-1+h}^{(h)}, u_{t-1+h}^{(h)}) = \frac{1}{h}E(r_{t-1+h}^{(h)}u_{t-1+h}^{(h)})$.

By using Ito's Lemma we get:

$E[r_{t-1+h}^{(h)}u_{t-1+h}^{(h)}] = E\left(\int_{t-1}^{t-1+h}\sigma_u dW_u\right)\left(\int_{t-1}^{t-1+h}(\int_{t-1}^{u}\sigma_s dW_s)\sigma_u dW_u\right)$

$= E\int_{t-1}^{t-1+h}(\int_{t-1+h}^{u}\sigma_s dW_s)\sigma_u^2 du = \rho\sum_{i=1}^{p}\frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$ by (A.4).

Hence, $Cov(r_t, u_t(h)) = \frac{\rho}{h}\sum_{i=1}^{p}\frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$.

iii) $Var(r_t) = E[\int_{t-1}^{t}\sigma_u^2 du] = a_0$.