Estimation of integrated volatility of volatility with applications to goodness-of-fit testing

Mathias Vetter [*]

December 2, 2011

Abstract

In this paper we are concerned with non-parametric inference on the volatility of volatility process τ^2 in stochastic volatility models. We construct an estimator for its integrated version $\int_0^t \tau_s^2 ds$ in a high frequency setting which is based on increments of spot volatility estimators, and we are able to prove both feasible and infeasible central limit theorems at the optimal rate $n^{-1/4}$. Such CLTs can be widely used in practice, as they are the key to essentially all tools in model validation for stochastic volatility models. As an illustration we apply our results to goodness-of-fit testing, providing the first consistent test for a certain parametric form of the volatility of volatility.

Keywords: central limit theorem, goodness-of-fit testing, high frequency observations, model validation, semimartingale, stable convergence, stochastic volatility model.

AMS 2000 subject classifications: primary, 60F05, 60G44; secondary, 62G20, 62M99.

[*]Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany. Tel: +49-234-3223283. Fax: +49-234-3214559. Email: mathias.vetter@rub.de. The author is thankful for financial support through the collaborative research center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG).

1 Introduction

Nowadays, stochastic volatility models are standard tools in the continuous-time modelling of financial time series. Typically, the underlying (log) price process is assumed to follow a diffusion process of the form

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \tag{1.1}$$

where μ and σ can be quite general stochastic processes themselves. A classical case is where the volatility $\sigma_s^2 = \sigma^2(s, X_s)$ is a function of time and state–a situation referred to as the one of a local volatility model. It has turned out in empirical finance that such models do not fit the data very well, as some stylised facts such as the leverage effect or volatility clustering cannot be explained using local volatility only. Stochastic volatility models, however, are able to reproduce such features, as they bear an additional source of randomness. In these models the volatility process is a diffusion process itself, having the representation

$$\sigma_t^2 = \sigma_0^2 + \int_0^t \nu_s ds + \int_0^t \tau_s dV_s, \tag{1.2}$$

where ν and τ again are suitable stochastic processes and V is another Brownian motion, correlated with W.

Standard stochastic volatility models are parametric ones, and probably the prime example among those is the Heston model of [17], given by

$$X_t = X_0 + \int_0^t \Big(\beta - \frac{\sigma_s^2}{2}\Big) ds + \int_0^t \sigma_s dW_s, \quad \sigma_t^2 = \sigma_0^2 + \kappa \int_0^t (\alpha - \sigma_s^2) ds + \xi \int_0^t \sigma_s dV_s,$$

for some parameters β, κ, α and ξ, and with $\mathrm{Corr}(W, V) = \rho$. Here, the volatility process follows a Cox-Ingersoll-Ross model, that means it is mean-reverting with mean α and speed κ, and both diffusion coefficients are proportional with parameter ξ. Such a behaviour appears to be rather typical for stochastic volatility models, and in this sense the Heston model can be regarded as prototypic. Popular alternatives are for example coming from the more general (but again parametric) class of CEV models, where the diffusion coefficient τ becomes a power function of σ, whereas the drift part of the volatility remains in principle the same. See [25] for a survey.

For this reason, statistical inference for stochastic volatility models has focused on parametric methods for most times, and usually the authors provide tools for a specific class of models. However, one is faced with two severe problems: First, it is in most cases impossible to assess the distribution of X (or its increments), which makes standard maximum likelihood theory unavailable. Second, the volatility process σ^2 is not observable, and many statistical concepts have in common that they propose to reproduce the unknown volatility process from observed option prices, typically by using proxies based on implied volatility. A survey on early estimation methods in this context can be found in [11]. One remarkable exception where stock price data only is used, is the paper of [10] who construct a GMM estimator for the parameters of the Heston model from increments of realised variance. But also in a general setting with no specific model in mind, the focus has been on parametric approaches. An early approach on parameter estimation when σ^2 is ergodic is the work of [15], optimal rates are discussed in [18] and [16], and a maximum likelihood approach based on proxies for the volatility can be found in [4]. Even non-parametric concepts have been used to identify parameters of a stochastic volatility model, see for example [9] or [30].

Genuine non-parametric inference for stochastic volatility models has typically focused on function estimation. Both [29] and [12] discuss techniques for the estimation of f and g, when the volatility process satisfies $d\sigma_t^2 = f(\sigma_t^2)dt + g(\sigma_t^2)dV_t$. In the more general model-free context of (1.2) only [8] have discussed estimation of functionals of the process τ by providing a consistent estimator for the integrated volatility of volatility $\int_0^t \tau_s^2 ds$. Their approach is inspired by the asymptotic behaviour of realised variance, which states that the sum of squared increments of σ^2 converges in probability to the quantity of interest. Since σ^2 is not observable, the authors use spot volatility estimators instead.

We will pursue their approach and define a slightly different estimator for integrated volatility of volatility which attains the optimal rate of convergence in this context, also using observations of X only. Furthermore, a stable central limit theorem is provided, and by defining appropriate estimators for the asymptotic (conditional) variance we obtain a feasible version as well. The latter result is of theoretical interest on one hand, but is extremely important from an applied point of view as well, as it makes model validation for stochastic volatility models possible. Given the tremendous number of such models with entirely different qualitative behaviours, there is a lack of techniques that help deciding whether a certain model fits the data appropriately or not.

As a first result on model validation in this whole framework we give an example on goodness-of-fit testing, but our method is by no means limited to it. Related procedures can be used to test e.g. whether a Brownian component or jumps are present in the volatility process and what in general the structure of the jump part is. Such problems have been solved for the price process X in recent years (see [22] for an overview), and in principle the methods are all based on the estimation of plain integrated volatility $\int_0^t \sigma_s^2 ds$ and further quantities. Using our main results, they can be translated to the stochastic volatility case by using integrated volatility of volatility instead. See the conclusions in Section 5 for some hints on further research.

Finally, the paper is organised as follows: In Section 2 we introduce our estimator and state the two central limit theorems, whereas Section 3 presents a method for goodness-of-fit testing in stochastic volatility models. Some Monte Carlo results can be found in Section 4, and as noted before we give a conclusion in Section 5. Most proofs can be found in the Appendix, which is Section 6.

2 Main results

Suppose that the process X is given by (1.1), where W is a standard Brownian motion and the drift process μ is left continuous. We assume further that the volatility process σ^2

is a continuous Itô semimartingale itself, having the representation (1.2), where ν is left continuous as well and (W,V) are jointly Brownian with correlation parameter $\rho \in [-1,1]$. Note that $|\rho| = 1$ corresponds to $W = V$ almost surely, in which case we are essentially in the setting of a local volatility model, whereas $|\rho| < 1$ refers to the genuine stochastic volatility case. Our aim is to draw inference on the integrated volatility of volatility, which is $\int_0^t \tau_s^2 ds$. To this end we impose a regularity condition on τ, namely

$$\tau_s^2 = \tau_0^2 + \int_0^t \omega_s ds + \int_0^t \vartheta_s^{(1)} dW_s + \int_0^t \vartheta_s^{(2)} dV_s + \int_0^t \vartheta_s^{(3)} d\overline{W}_s, \tag{2.1}$$

where $\overline{W}$ is another Brownian motion, independent of (W,V), ω is locally bounded and each $\vartheta^{(l)}$ is left continuous, $l = 1,2,3$. Finally, we assume that all processes are defined on the same probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and that all coefficients are specified in such a way that σ^2 and τ^2 are almost surely positive. These assumptions are all rather mild and are covered by a variety of (stochastic) volatility models used.

Any statistical inference in this work will be based on high-frequency observations of X, and we assume that the data is recorded at equidistant times. Thus, without loss of generality let the process be defined on the interval [0,1] and observed at the time points i/n, $i = 0, \dots, n$. Just as standard integrated volatility is estimated using (squared) increments of X, a reasonable estimator for integrated volatility of volatility can be built upon increments of σ^2. These are in general not observable, so a proxy for them is needed. Since we are in a model-free world, a natural estimator for spot volatility $\sigma_{i/n}^2$ is given by

$$\hat{\sigma}_{\frac{i}{n}}^2 = \frac{n}{k_n} \sum_{j=1}^{k_n} (\Delta_i^n X)^2, \quad i = 0, \dots, n - k_n,$$

for some auxiliary (integer-valued) sequence k_n and where we have set $\Delta_i^n Z = Z_{i/n} - Z_{i-1/n}$ for any process Z. See [5] or [30] for details on the asymptotic behaviour of this estimator. Equation (2.9), which is a simple consequence of Itô formula and Lemma 6.1, shows later on that

$$\hat{\sigma}_{\frac{i}{n}}^2 - \sigma_{\frac{i}{n}}^2 = O_p(\sqrt{k_n/n} + \sqrt{1/k_n})$$

holds, which explains the choice $k_n = cn^{1/2} + o(n^{1/4})$ for some $c > 0$ we will use in the following.

An estimator for integrated volatility of volatility will now be defined as a sum over squared increments of $\hat{\sigma}_{(i+l_n)/n}^2 - \hat{\sigma}_{i/n}^2$. What is a reasonable choice for l_n? (2.2) suggests that one should not take l_n of smaller order than k_n, as otherwise the estimation error $\hat{\sigma}_{i/n}^2 - \sigma_{i/n}^2$ (which is of order $\sqrt{k_n/n}$) is dominating the quantity of interest $\sigma_{(i+l_n)/n}^2 - \sigma_{i/n}^2$ (which has order $\sqrt{l_n/n}$). We will see that we can indeed take l_n equal to k_n which guarantees convergence at the optimal rate, but in this case a bias correction becomes necessary. For this reason we define local estimators for the process τ^2 as follows: With a slight abuse of notation we set

$$\hat{\tau}_{\frac{i}{n}}^2 = \frac{3n}{2k_n} (\hat{\sigma}_{\frac{i+k_n}{n}}^2 - \hat{\sigma}_{\frac{i}{n}}^2)^2 - 6\frac{n}{k_n^2} \hat{\sigma}_{\frac{i}{n}}^4, \quad \hat{\sigma}_{\frac{i}{n}}^4 = \frac{n^2}{3k_n} \sum_{j=1}^{k_n} |\Delta_{i+j}^n X|^4, \tag{2.2}$$

where the latter is in general different from $(\hat{\sigma}^2_{i/n})^2$. A global estimator for integrated volatility of volatility then becomes

$$\hat{V}_t = \frac{1}{n} \sum_{i=0}^{\lfloor nt \rfloor - 2k_n} \hat{\tau}^2_{\frac{i}{n}}. \tag{2.3}$$

Its asymptotic behaviour is discussed in the following theorem.

Theorem 2.1 *Under the above assumptions we have the central limit theorem*

$$\sqrt{\frac{n}{k_n}} \Big(\hat{V}_t - \int_0^t \tau_s^2 ds \Big) \stackrel{\mathcal{L}-(s)}{\longrightarrow} U_t \tag{2.4}$$

for all $t > 0$, where the limiting variable has the representation

$$U_t = \int_0^t \alpha_s dW'_s, \quad \alpha_s^2 = \frac{48}{c^4}\sigma_s^8 + \frac{12}{c^2}\sigma_s^4\tau_s^2 + \frac{151}{70}\tau_s^4,$$

W' is a Brownian motion defined on an extension of the original probability space and independent of $\mathcal{F}$ and the convergence in (2.4) is $\mathcal{F}$-stable in law. For details on this mode of convergence see for example [23].

Proof of Theorem 2.1 We will give a sketch of the proof here and relegate some tedious details to the Appendix. In general, $\mathcal{F}$-stable convergence of a sequence Z_n to some limiting variable Z defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original space is equivalent to

$$\mathbb{E}(h(Z_n)Y) \to \widetilde{\mathbb{E}}(h(Z)Y) \tag{2.5}$$

for any bounded Lipschitz function h and any bounded $\mathcal{F}$ measurable Y. For details, see for example [23] and related work. Suppose now that there are additional variables $Z_{n,p}$ and Z_p (the latter defined on the same extension as Z) such that

$$\lim_{p\to\infty} \limsup_{n\to\infty} \mathbb{E}|Z_n - Z_{n,p}| = 0, \tag{2.6}$$

$$Z_{n,p} \stackrel{\mathcal{L}-(s)}{\longrightarrow} Z_p \text{ for all } p, \tag{2.7}$$

$$\lim_{p\to\infty} \widetilde{\mathbb{E}}|Z_p - Z| = 0, \tag{2.8}$$

hold. Then the desired stable convergence $Z_n \stackrel{\mathcal{L}-(s)}{\longrightarrow} Z$ follows. Indeed, let $\epsilon > 0$. Then there exists a $\delta > 0$ such that $|x - y| < \delta$ implies $|h(x) - h(y)| < \epsilon$. Thus we have

$$\begin{aligned} & |\mathbb{E}(h(Z_n)Y) - \mathbb{E}(h(Z_{n,p})Y)| \\ \leq\ & C\Big(\mathbb{E}(|h(Z_n) - h(Z_{n,p})| 1_{\{|Z_n - Z_{n,p}| \geq \delta\}}) + \mathbb{E}(|h(Z_n) - h(Z_{n,p})| 1_{\{|Z_n - Z_{n,p}| < \delta\}})\Big) \\ \leq\ & C\Big(P(|Z_n - Z_{n,p}| \geq \delta) + \epsilon\Big) \end{aligned}$$

for a generic $C > 0$. From Markov inequality, (2.6) and as ϵ was arbitrary, we have

$$\lim_{p\to\infty} \limsup_{n\to\infty} |\mathbb{E}(h(Z_n)Y) - \mathbb{E}(h(Z_{n,p})Y)| = 0.$$

The same argument using (2.8) yields

$$\lim_{p\to\infty} |\widetilde{\mathbb{E}}(h(Z_p)Y) - \widetilde{\mathbb{E}}(h(Z)Y)| = 0,$$

and (2.7) is by definition equivalent to

$$\lim_{n\to\infty} |\mathbb{E}(h(Z_{n,p})Y) - \widetilde{\mathbb{E}}(h(Z_p)Y)| = 0.$$

Putting the latter three claims together [plus the triangle inequality and the fact that all three limiting conditions on p and n are actually the same] gives (2.5).

Our aim in this proof is to employ a certain blocking technique, which allows us to make use of a type of conditional independence between the estimators $\hat{\tau}^2_{i/n}$. To this end we apply the above methodology, so we have to define an appropriate double sequence $U_t^{n,p}$, which will correspond to the sum of approximated versions of $\hat{\tau}^2_{i/n}$ over the big blocks. Some additional notation is necessary. First of all, recall that Itô formula gives

$$\hat{\sigma}^2_{\frac{i}{n}} = \frac{n}{k_n} \sum_{j=1}^{k_n} 2 \int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} (X_s - X_{\frac{i+j-1}{n}}) dX_s + \frac{n}{k_n} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \sigma_s^2 ds =: A_i^n + B_i^n. \tag{2.9}$$

The main part of U_t^n is some functional of increments of A and B, and as noted above we need certain approximations for these in the sequel. Let $p \in \mathbb{N}$ be arbitrary. We set

$$a_\ell(p) = (\ell - 1)(p+2)k_n, \quad b_\ell(p) = a_\ell(p) + pk_n, \quad c(p) = J_n(p)(p+2)k_n + 1,$$

the first two for any $\ell = 1, \ldots, J_n(p)$ with $J_n(p) = \lfloor \lfloor nt - 2k_n \rfloor / (p+2)k_n \rfloor$. These numbers depend on n as well, even though it does not show up in the notation. We define further $H_i^n = \int_{i-1/n}^{i/n} (W_s - W_{(i-1)/n}) dW_s$ and

$$\begin{aligned} \widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}} &= \frac{n}{k_n} \sum_{j=1}^{k_n} 2\sigma^2_{\frac{a_\ell(p)}{n}} (H^n_{i+j+k_n} - H^n_{i+j}) \\ &= \frac{n}{k_n} \sigma^2_{\frac{a_\ell(p)}{n}} \sum_{j=1}^{k_n} \Big((\Delta^n_{i+k_n+j} W)^2 - (\Delta^n_{i+j} W)^2 \Big), \end{aligned} \tag{2.10}$$

where the latter identity is a consequence of Itô formula, and

$$\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}} = \frac{n}{k_n} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \tau_{\frac{a_\ell(p)}{n}} (V_{s+\frac{k_n}{n}} - V_s) ds.$$

These quantities are defined over the big blocks, that is for $i = a_\ell(p), \ldots, b_\ell(p) - 1$. Up to a different standardisation, the role of $Z_{n,p}$ in this proof will be played by $U_t^{n,p} = \sum_{\ell=1}^{J_n(p)} U_\ell^{n,p}$,

$$U_\ell^{n,p} = \sum_{i=a_\ell(p)}^{b_\ell(p)-1} \frac{3}{2k_n} ((\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}}) + (\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}}))^2 - \frac{pk_n}{n} \Big[\frac{6n}{k_n^2} \sigma^4_{\frac{a_\ell(p)}{n}} + \tau^2_{\frac{a_\ell(p)}{n}} \Big], \tag{2.11}$$

which again involves quantities from the big blocks only. The $U_\ell^{n,p}$ can be shown to be martingale differences, and the most involved part in the proof is to obtain

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\Big(\hat{V}_t-\int_0^t\tau_s^2ds\Big)-U_t^{n,p}\Big|=0, \tag{2.12}$$

which is the analogue of (2.6). Let us focus on the remaining two steps as well. We set

$$U_t^p=\int_0^t\alpha(p)_s dW_s',\quad \alpha(p)_s^2=\frac{p}{p+2}\Big(\frac{48p+d_1}{pc^4}\sigma_s^8+\frac{12p+d_2}{pc^2}\sigma_s^4\tau_s^2+\frac{151p+d_3}{70p}\tau_s^4\Big)$$

for certain unspecified constants d_l, $l=1,2,3$. In order to prove the stable convergence

$$\sqrt{\frac{n}{k_n}}U_t^{n,p}\overset{\mathcal{L}-(s)}{\longrightarrow}U_t^p \tag{2.13}$$

we use a well-known result for triangular arrays of martingale differences, which is due to Jacod [19]. In particular, the following three conditions have to be checked, where we call $\mathbb{E}_i^n$ the conditional expectation with respect to $\mathcal{F}_{\frac{i}{n}}$.

$$\frac{n}{k_n}\sum_{\ell=1}^{J_n(p)}\mathbb{E}_{a_\ell(p)}^n[(U_\ell^{n,p})^2]\overset{\mathbb{P}}{\longrightarrow}\int_0^t\alpha(p)_s^2ds, \tag{2.14}$$

$$\frac{n^2}{k_n^2}\sum_{\ell=1}^{J_n(p)}\mathbb{E}_{a_\ell(p)}^n[(U_\ell^{n,p})^4]\overset{\mathbb{P}}{\longrightarrow}0, \tag{2.15}$$

$$\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\mathbb{E}_{a_\ell(p)}^n[U_\ell^{n,p}(N_{\frac{a_{\ell+1}(p)}{n}}-N_{\frac{a_\ell(p)}{n}})]\overset{\mathbb{P}}{\longrightarrow}0, \tag{2.16}$$

where N is any component of (W,V) or a bounded martingale orthogonal to both W and V. The final step $\lim_{p\to\infty}\widetilde{E}|U_t^p-U_t|=0$ is obvious. □

Remark 2.2 It is quite likely that a functional central limit theorem holds as well, but a proof of this result appears to be somewhat more involved. In any case, the claim above is sufficient for most of the statistical applications we have in mind. □

Remark 2.3 The rate of convergence in Theorem 2.1 is $n^{-1/4}$, and it is optimal for this statistical problem. Indeed, a related parametric setting has been discussed in [18] a decade ago, and it was shown therein that this rate is optimal in the special case, where W and V are independent and τ is a function of time and state, known up to a parameter θ. □

Remark 2.4 As noted in the introduction, an alternative estimator has been defined in [8]. Apart from an additional truncation in the spot volatility estimators to make these robust to jumps in the price process, the main difference between both approaches are different orders of k_n and l_n. Their choices grant consistency for the integrated volatility of volatility without a bias correction as above, but as a drawback the optimal rate cannot be attained. □

The limiting distribution in Theorem 2.1 is mixed normal, and in order to obtain a feasible central limit theorem we have to introduce a consistent estimator for the conditional variance $\int_0^t \alpha_s^2 ds$. This term is a sum of three quantities, and regarding the parts involving the process τ, we will rely on the previously introduced local estimators. To be precise, for the integral over τ^4 we will base an estimator on fourth powers of increments of $\hat{\sigma}^2$, and again a suitable bias correction is necessary, whereas the estimator for the mixed part is built directly from $\hat{\tau}^2_{i/n}$ and $\hat{\sigma}^4_{i/n}$. For the term involving powers of σ only, there are several possibilities (involving standard power variations), but for computational reasons we choose one which is based on $\hat{\sigma}^4_{i/n}$ as well. Altogether we obtain the following result, and a proof of this claim (just as for all later ones as well) can be found in the Appendix.

Theorem 2.5 *Set*

$$G^{(1)}_{t,n} = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor - k_n} (\hat{\sigma}^4_{\frac{i}{n}})^2, \quad G^{(2)}_{t,n} = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor - 2k_n} \hat{\tau}^2_{\frac{i}{n}} \hat{\sigma}^4_{\frac{i}{n}}, \quad G^{(3)}_{t,n} = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor - 2k_n} \frac{n^2}{k_n^2} (\hat{\sigma}^2_{\frac{i+k_n}{n}} - \hat{\sigma}^2_{\frac{i}{n}})^4.$$

Then we have the convergence in probability

$$G^{(1)}_{t,n} \xrightarrow{\mathbb{P}} \int_0^t \sigma_s^8 ds, \tag{2.17}$$

$$G^{(2)}_{t,n} \xrightarrow{\mathbb{P}} \int_0^t \sigma_s^4 \tau_s^2 ds, \tag{2.18}$$

$$G^{(3)}_{t,n} \xrightarrow{\mathbb{P}} \int_0^t \Big(\frac{48}{c^4} \sigma_s^8 + \frac{16}{c^2} \sigma_s^4 \tau_s^2 + \frac{4}{3} \tau_s^4 \Big) ds, \tag{2.19}$$

and as a consequence

$$\hat{C}^n_t = \frac{453}{280} G^{(3)}_{t,n} - \frac{n}{k_n^2} \frac{486}{35} G^{(2)}_{t,n} - \frac{n^2}{k_n^4} \frac{1038}{35} G^{(1)}_{t,n} \xrightarrow{\mathbb{P}} \int_0^t \alpha_s^2 ds.$$

Remark 2.6 Theorem 2.5 shows that a consistent estimator for $\int_0^t \tau_s^4 ds$ is given by

$$\hat{T}_t = \frac{3}{4} G^{(3)}_{t,n} - 12 \frac{n}{k_n^2} G^{(2)}_{t,n} - 36 \frac{n^2}{k_n^4} G^{(1)}_{t,n}, \tag{2.20}$$

and its proof suggests that a central limit theorem holds with the same rate of convergence as before. In general, it is quite likely that this methods provides estimates for arbitrary even powers of integrated volatility of volatility. A concise theory is beyond the scope of this paper, however. □

The properties of stable convergence guarantee that dividing by the square root of a consistent estimator for $\int_0^t \alpha_s^2 ds$ gives a feasible central limit theorem for the estimation of integrated volatility of volatility. See for example [28] for details. Therefore we end this section with the following corollary.

Corollary 2.7 *Under the previous assumptions we have weak convergence*

$$\sqrt{\frac{n}{k_n}} \frac{(\hat{V}_t - \int_0^t \tau_s^2 ds)}{\sqrt{\hat{C}_t^n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

for all $t > 0$.

3 Model checks for stochastic volatility models

Suppose that we are given a stochastic volatility model (with continuous paths), that is we have the representation (1.1) for the log price process X and a volatility process satisfying (1.2). There is still a lot of freedom in the modelling of σ^2, and the various proposals in the literature typically differ in the representation of its diffusion part τ. As noted in the introduction, a quite general class of stochastic volatility models is given by the so-called CEV models, in which $\tau_s = \theta(\sigma_s^2)^\gamma$ for some non-negative γ and an unknown parameter θ, and the most popular among these is the Heston model from [17], corresponding to $\gamma = 1/2$.

Given the number of different stochastic volatility models, there is a lack of techniques in goodness-of-fit testing. We will partly fill this gap and employ a technique which was already used in [13] or [31] when dealing with local volatility models. Let us explain the methodology by taking the example of a Heston model, for which $\nu_s = \kappa(\alpha - \sigma_s^2)$ and $\tau_s^2 = \xi^2\sigma_s^2$. Since it is in general impossible to obtain information on the drift part of a semimartingale from high-frequency observations only, we will solely focus on the diffusion process. Thus our aim is to test whether the specific functional relationship of proportionality between τ_s^2 and σ_s^2 is true or not. Let us have a look at the stochastic process

$$N_t = \int_0^t (\tau_s^2 - \theta_{min}\sigma_s^2)ds, \quad \theta_{min} = \text{argmin}_\theta \int_0^1 (\tau_s^2 - \theta\sigma_s^2)^2 ds.$$

Under the null hypothesis of a Heston-like diffusion process σ^2, the process N_t is obviously equal to zero for all t. Therefore a promising approach is to define an estimate $\hat{N}_t$, which will be based on the heuristics from the previous section, and to prove weak convergence of the statistic $\sqrt{n/k_n}(\hat{N}_t - N_t)$ to a certain limiting process A_t, for which we use Theorem 2.1 and related results. Test statistics can then be constructed via functionals of $\sqrt{n/k_n}\hat{N}_t$ which converge weakly as well and to the same functionals of A_t, if the underlying process is indeed coming from a Heston model.

This approach is of course not limited to the Heston model, which is why we return to general case. Suppose that (1.2) holds. We are interested in testing for $\tau_s^2 = \tau^2(s, X_s, \sigma_s^2, \theta)$, where τ^2 is a given function and θ is some unknown (in general multi-dimensional) parameter. For simplicity, we will focus on the one-dimensional linear case only, that is

$$H_0 : \tau_s^2 = \theta\tau^2(s, X_s, \sigma_s^2) \text{ for all } s \in [0,1] \text{ (a.s.)}$$

Extensions to the general case follow along the lines of Section 5 in [31].

A test for the null hypothesis will be based on the observation that H_0 is equivalent to $N_t = 0$ for all $t \in [0,1]$ (a.s.), where the process N_t is given by

$$N_t = \int_0^t \Big(\tau_s^2 - \theta_{min}\tau^2(s, X_s, \sigma_s^2)\Big)ds, \quad \theta_{min} = \text{argmin}_\theta \int_0^1 \Big(\tau_s^2 - \theta\tau^2(s, X_s, \sigma_s^2)\Big)^2 ds.$$

Assume that the function τ^2 is bounded away from zero. Then a standard argument from Hilbert space theory shows that $\theta_{min} = D^{-1}C$ (and therefore $N_t = V_t - B_t D^{-1}C$), where we have set $V_t = \int_0^t \tau_s^2 ds$ and

$$B_t = \int_0^t \tau^2(s, X_s, \sigma_s^2)ds, \quad D = \int_0^1 \tau^4(s, X_s, \sigma_s^2)ds, \quad C = \int_0^1 \tau_s^2\tau^2(s, X_s, \sigma_s^2)ds.$$

To define reasonable estimators for the various quantities above let k_n as before and recall (2.2). We set $\hat{N}_t = \hat{V}_t - \hat{B}_t\hat{D}^{-1}\hat{C}$ with $\hat{V}_t$ from the previous section, whereas we denote

$$\hat{B}_t = \frac{1}{n}\sum_{i=0}^{\lfloor nt\rfloor - k_n} \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}^2_{\frac{i}{n}}), \quad \hat{D} = \frac{1}{n}\sum_{i=0}^{n-k_n} \tau^4(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}^2_{\frac{i}{n}}), \quad \hat{C} = \frac{1}{n}\sum_{i=0}^{n-2k_n} \hat{\tau}^2_{\frac{i}{n}}\tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}^2_{\frac{i}{n}}).$$

In the sequel we will prove weak convergence of $\hat{N}_t - N_t$, up to a suitable normalisation. Theorem 2.1 suggests that $\sqrt{n/k_n}$ is a reasonable choice, and the following claim proves that two of the estimators converge at a faster speed, at least if we impose an additional smoothness condition on the function τ^2.

Lemma 3.1 *Suppose that the function τ^2 has continuous partial derivatives of second order. Then we have*

$$\hat{B}_t - B_t = o_p(n^{-1/4}), \quad \hat{D} - D = o_p(n^{-1/4}),$$

the first result holding uniformly in $t \in [0,1]$.

The above claim indicates that we have to focus on the terms involving $\hat{\tau}^2_{i/n}$ only, which is familiar ground due to the results of Section 2. We start with a proposition on the joint asymptotic behaviour of $\hat{V}_t$ and $\hat{C}$.

Lemma 3.2 *Let d be an integer and $t_1, \ldots, t_d$ be arbitrary in $[0,1]$. Set*

$$\Sigma_{t_1,\ldots,t_d}(s, X_s, \sigma_s^2) = \alpha_s^2 h_{t_1,\ldots,t_d}(s, X_s, \sigma_s^2) h_{t_1,\ldots,t_d}(s, X_s, \sigma_s^2)^T$$

with $h_{t_1,\ldots,t_d}(s, X_s, \sigma_s^2) = (1_{[0,t_1]}, \ldots, 1_{[0,t_d]}, \tau^2(s, X_s, \sigma_s^2))^T$ and α_s^2 as in Theorem 2.1. Under the previous assumptions we have the stable convergence

$$\sqrt{\frac{n}{k_n}}\Big(\hat{V}_{t_1} - V_{t_1}, \ldots, \hat{V}_{t_d} - V_{t_d}, \hat{C} - C\Big)^T \xrightarrow{\mathcal{L}-(s)} \int_0^1 \Sigma^{1/2}_{t_1,\ldots,t_d}(s, X_s, \sigma_s^2)dW'_s,$$

where W' is a $(d+1)$-dimensional standard Brownian motion defined on an extension of the original space and independent of $\mathcal{F}$.

We are intested in the asymptotics of the process $A_n(t) = \sqrt{n/k_n}(\hat{N}_t - N_t)$, and the preceding lemma basically leads to its finite dimensional convergence. The entire result on weak convergence of A_n reads as follows.

Theorem 3.3 *Assume that the previous assumptions hold. Then the process $(A_n(t))_{t\in[0,1]}$ converges weakly to a mean zero process $(A(t))_{t\in[0,1]}$, which is Gaussian conditionally on $\mathcal{F}$ and whose conditional covariance equals the one of the process*

$$\Big\{\alpha_U\Big(1_{\{U\le t\}} - B_t D^{-1}\tau^2(U, X_U, \sigma_U^2)\Big)\Big\}_{t\in[0,1]}$$

where $U \sim \mathcal{U}[0,1]$, independent of $\mathcal{F}$.

As indicated before, convergence of the finite dimensional distributions is a direct consequence of Lemma 3.2, using the Delta method for stable convergence (see e.g. [14]). Tightness follows from Theorem VI. 4.5 in [23] with a minimal amount of work.

Recall that $N_t = 0$ for all t under the null hypothesis. Therefore Theorem 3.3 shows that a consistent test is obtained by rejecting the null hypothesis for large values of a suitable functional of the process $\{\sqrt{n/k_n}\hat{N}_t\}_{t\in[0,1]}$. If we choose the Kolmogorov-Smirnov functional $K_n = \sup_{t\in[0,1]}\sqrt{n/k_n}|\hat{N}_t|$ for example, we have weak convergence under the null to $\sup_{t\in[0,1]}|A_t|$ as a consequence of Theorem 3.3. The distribution of the latter statistic is extremely difficult to assess, as it typically depends on the entire process (X, σ^2). We will thus propose to pursue a different path and to obtain critical values via a bootstrap procedure, which we will discuss in the next section in detail.

Remark 3.4 In practice, one should test beforehand, whether modelling via stochastic volatility is actually appropriate or not. At least two recent procedures should be mentioned here: [27] propose a test which discriminates between local volatility and stochastic volatility models and which is based on the sign of increments of X and of increments of spot volatility, which tend to be equal if both X and its volatility process are driven by the same Brownian motion. [30] discusses more generally semi-parametric techniques for the estimation of the correlation parameter ρ between W and V. □

Remark 3.5 An alternative approach on model validation could be based on an appropriate L^2 distance, instead of working with empirical processes. To be precise, set

$$M^2 = \int_0^1 \Big(\tau_s^2 - \theta_{min}\tau^2(s, X_s, \sigma_s^2)\Big)^2 ds$$

with θ_{min} as above. Then the null hypothesis is equivalent to $M^2 = 0$ almost surely, and a natural estimator for M^2 can be defined in the same way as for N_t. Nevertheless, the asymptotic theory for $\hat{M}^2$ is a bit more involved, since a central limit theorem for an estimator of $\int_0^1 \tau_s^4 ds$ is necessary and a discussion of such a theory is beyond the scope of this paper. See [14] for the asymptotic theory of the analogue of M^2 in the local volatility setting. □

To end this section we define appropriate estimator for the conditional variance of $A(t)$, which is given by

$$s_t^2 = \int_0^t \alpha_s^2 ds - 2B_t D^{-1} \int_0^t \alpha_s^2 \tau^2(s, X_s, \sigma_s^2) ds + B_t^2 D^{-2} \int_0^t \alpha_s^2 \tau^4(s, X_s, \sigma_s^2) ds,$$

due to Theorem 3.3. Empirical counterparts for B_t and D are obviously defined by the statistics $\hat{B}_t$ and $\hat{D}$, whereas Theorem 2.5 suggests that a local estimator for $\alpha_{i/n}^2$ is given by

$$\hat{\alpha}_{\frac{i}{n}}^2 = \frac{n^2}{k_n^2} \Big(\frac{453}{280} (\hat{\sigma}_{\frac{i+k_n}{n}}^2 - \hat{\sigma}_{\frac{i}{n}}^2)^4 - \frac{486}{35} \hat{\tau}_{\frac{i}{n}}^2 \hat{\sigma}_{\frac{i}{n}}^4 \Big) - \frac{n^6}{k_n^5} \frac{346}{1225} \sum_{j=1}^{k_n} |\Delta_{i+j}^n X|^8.$$

We obtain the following result, which can be proven in the same way as Theorem 2.5.

Theorem 3.6 *Let t be arbitrary and set*

$$\begin{aligned} \hat{s}_t^2 &= \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor - 2k_n} \hat{\alpha}_{\frac{i}{n}}^2 - 2\hat{B}_t \hat{D}^{-1} \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor - 2k_n} \hat{\alpha}_{\frac{i}{n}}^2 \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}_{\frac{i}{n}}^2) \\ &+ \hat{B}_t^2 \hat{D}^{-2} \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor - 2k_n} \hat{\alpha}_{\frac{i}{n}}^2 \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}_{\frac{i}{n}}^2). \end{aligned}$$

Then $\hat{s}_t^2$ is consistent for s_t^2.

As a consequence, each statistic $\sqrt{n/k_n}\hat{N}_t/\hat{s}_t$ converges weakly to a normal distribution. This result will be used to construct a feasible bootstrap statistic in the following.

4 Simulation study

Let us start with a simulation study concerning the performance of $\hat{V}_t$ as an estimator for integrated volatility of volatility. Throughout this section we will work with the Heston model only, and the parameters are chosen as follows: $\beta = 0.3$, $\kappa = 5$, $\alpha = 0.2$ and $\xi = 0.5$. Furthermore, we set $X = 0$ and $\sigma_0^2 = \alpha$. Note that the Feller condition $2\kappa\alpha \geq \xi^2$ is satisfied, which ensures that the process σ^2 is almost surely positive as requested. So does τ^2, and it is obvious that (2.1) holds as well. Therefore all conditions from Section 2 are satisfied.

We begin the finite sample properties of the central limit theorem from Theorem 2.1 in its infeasible version, which is

$$\sqrt{\frac{n}{k_n}} \frac{(\hat{V}_t - \int_0^t \tau_s^2 ds)}{\sqrt{\int_0^t \alpha_s^2 ds}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{4.1}$$

That means we use the unobservable (conditional) variance $\int_0^t \alpha_s^2 ds$ to standardise correctly instead of using an estimator for it. We discuss the performance of this result for different choices the correlation parameter ρ and the number of observations n, and in all cases we take n to be a square number and k_n equals $n^{1/2}$, so we have $c = 1$. Finally we set $t = 1$.

[INSERT TABLE 1 ABOUT HERE]

Table 1 shows the finite sample behavior of (4.1) for $\rho = 0$, always based on 10,000 simulations. We see that mean and variance are rather close to the limiting values in most cases. In the remaining columns we show some empirical quantiles in the tails, that is we state both α and the relative number of times where the statistic in (4.1) was below the corresponding α quantile of the standard normal distribution. These values appear to be reproduced in a satisfying way as well.

[INSERT TABLE 2 ABOUT HERE]

The same situation has been analysed for $\rho = -0.2$, which corresponds to a moderate leverage effect of negative correlation between increments in price and volatility, and the results are in general comparable to the previous ones. Note that (4.1) does not depend at all on the choice of ρ, but some smaller order terms do as can be seen from the proof. These might affect the quality of approximation for finite samples, but apparently they do not.

[INSERT TABLE 3 ABOUT HERE]

For the feasible statistic from Corollary 2.7 the situation is somewhat different, as it takes more time for the asymptotics to kick in. Apparent is a slight overestimation of the lower tails of the distribution, which seem to originate from the relation of the estimators $\hat{V}_1$ and $G^{(3)}_{1,n}$. By construction, in cases where $\hat{V}_1$ is underestimating the true quantity, it is typically the case that increments of $\hat{\sigma}^2$ are relatively small. As these increments occur in $G^{(3)}_{1,n}$ as well, most likely the asymptotic variance is underestimated as well, which explains a too large negative standardised statistic. The same effect is visible for the upper quantiles as well (but resulting in an overestimation), and this simple explanation is supported by a detailed look at simulation results not reported here which reveal that the estimation of the asymptotic variance is extremely accurate for moderate sizes of $\hat{V}_1 - \int_0^1 \tau_s^2$, but becomes worse when the deviation is rather large.

[INSERT TABLE 4 ABOUT HERE]

As an example for an application in goodness-of-fit testing, we have constructed a test for a Heston-like volatility structure via a bootstrap procedure as follows: Based on the observation that for each t, $\sqrt{n/k_n}\hat{N}_t/\hat{s}_t$ converges weakly to a standard normal distribution if the null is satisfied, it seems reasonable to reject the hypothesis for large values of the standardised Kolmogorov-Smirnov statistic $Y_n = \sup_{i \leq n-2k_n} |\sqrt{n/k_n}\hat{N}_{i/n}/\hat{s}_{i/n}|$. Since its (asymptotic) distribution is in general hard to assess, we used bootstrap quantiles instead, and precisely we have generated bootstrap data $X^{*(b)}_{i/n}$, $b = 1, \ldots, B$, following the equation

$$X_t^* = \int_0^t \sigma_s^* dW_s^*, \quad (\sigma_t^*)^2 = \hat{\alpha} + \int_0^t \hat{\kappa}(\hat{\alpha} - (\sigma_s^*)^2) ds + \hat{\xi} \int_0^t \sigma_s^* dV_s^*.$$

Here, W^* and V^* are independent Brownian motions, and we have identified $\hat{\alpha}$ with the realised volatility of the original data (which is a measure for the average volatility over [0,1]) and defined $\hat{\xi} = \hat{\theta}^{1/2}$, since both quantities coincide under the null. Finally, we have simply set $\hat{\kappa} = 5\hat{\theta}/\hat{\alpha}$ such that Feller's condition is satisfied. Setting $B = 200$, we have run 500 simulations each.

[INSERT TABLE 5 ABOUT HERE]

Table 5 shows that the simulated levels are rather close to the expected ones, irrespectively of n. We have tested two alternatives from the class of CEV models, namely

$$\sigma_t^2 = \sigma_0^2 + \kappa \int_0^t (\alpha - \sigma_s^2) ds + V_t \quad \text{and} \quad \sigma_t^2 = \sigma_0^2 + \kappa \int_0^t (\alpha - \sigma_s^2) ds + \sqrt{\kappa} \int_0^t \sigma_s^2 dV_s,$$

corresponding to $\gamma = 0$ and $\gamma = 1$, respectively, and using the parameters from above. We see from the simulation results that the rejection probabilities are much larger for the second alternative than for the first, which can partially explained from two observations: First, the Vasicek model does not satisfy the assumptions from the previous sections since the volatility may become negative (in which case it is set to zero); second, our choice of $\hat{\kappa}$ is responsible for a large speed of mean reversion in the bootstrap algorithm which makes it difficult to distinguish between a Heston-like volatility of volatility and a constant one. It is expected that the power improves for an entirely data-driven choice of $\hat{\kappa}$.

[INSERT TABLE 6 ABOUT HERE]

5 Conclusion

In this paper we have discussed a non-parametric method on estimation of the integrated volatility of volatility process $\int_0^t \tau_s^2 ds$ in stochastic volatility models. Our estimator is based on spot volatility estimators, and just as for standard realised volatility we use sums of squares of these to obtain a global estimator $\hat{V}_t$ for $\int_0^t \tau_s^2 ds$, up to a further bias correction. It is shown that $\hat{V}_t$ converges at the optimal rate $n^{-1/4}$, and we provide both an infeasible and a feasible central limit theorem for it.

Given the variety of stochastic volatility models (in continuous time) which are used to describe financial data, there is a severe lack in tools on model validation. Our results somehow fill this gap, as we provide a promising method for goodness-of-fit testing in such models which investigates whether a specific parametric model for volatility of volatility is appropriate given the data or not. But several further applications are possible as well, particularly if we turn to the even more general context of models including jumps in price and volatility. Non-parametric inference in this context is rare as well, but to mention is recent work by Jacod and collaborators on the existence and the form of joint jumps in both processes (see [24] and [21]).

Following our results, much more questions regarding jumps in the volatility process can be tackled now, and to explain possible extensions of our approach let us have a look at the case of jumps in the price process. Several statistical tools have been developed over the past years that help answering e.g. whether there are jumps or not in the process, whether there are finitely or infinitely many, or what in general their degree of activity is (see foremost [1]–[3], but also [7] or [26]). Most of these procedures are based on realised volatility and related quantities, such as truncated versions or bipower variation. Theorem 2.1 indicates that similar methods are likely to work for the volatility process as well, but usually with the slower rate of convergence $n^{-1/4}$. A detailed analysis is beyond the scope of this paper, however.

A different issue to take microstructure issues into account which are likely to be present when data is observed at high frequency. Again it is promising to combine filtering methods for noisy diffusions with the method proposed in this paper to obtain an estimator for integrated volatility of volatility in such models as well, but the rate of convergence is expected to drop further. Again, precise statements on the asymptotics are left for further research.

6 Appendix

Note first that every left-continuous process is locally bounded, thus all processes appearing are. Second, standard localisation procedures as in [6] or [20] allow us to assume that any locally bounded process is actually bounded, and that almost surely positive processes can be regarded as bounded away from zero. Universal constants are denoted by C or C_r, the latter if we want to emphasise depedence on some additional parameter r.

6.1 Proof of (2.12)

We have a couple of somewhat lengthy steps to show. Start with the following observation: If we set

$$U_t^n = \sum_{i=0}^{\lfloor nt \rfloor - 2k_n} \frac{3}{2k_n} (\hat{\sigma}^2_{\frac{i+k_n}{n}} - \hat{\sigma}^2_{\frac{i}{n}})^2 - \frac{6}{c^2} \int_0^t \sigma_s^4 ds - \int_0^t \tau_s^2 ds,$$

then a simple computation proves that

$$\hat{V}_t = \sum_{i=0}^{\lfloor nt \rfloor - 2k_n} \frac{3}{2k_n} (\hat{\sigma}^2_{\frac{i+k_n}{n}} - \hat{\sigma}^2_{\frac{i}{n}})^2 - 2\frac{n^2}{k_n^2} \sum_{i=1}^{\lfloor nt \rfloor} |\Delta_i^n X|^4 + O_p(n^{-1/2}),$$

where the error term is due to boundary effects. Theorem 2.1 in [6] and the definition of k_n give

$$\sqrt{\frac{n}{k_n}} \mathbb{E}\Big|\Big(\hat{V}_t - \int_0^t \tau_s^2 ds\Big) - U_t^n\Big| = o(1),$$

uniformly in t. Therefore we are left to show

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|U_t^n - U_t^{n,p}\Big| = 0. \tag{6.1}$$

To this end, we need an auxiliary result on the increments of A and B and their approximations, and we introduce similar terms over the small blocks. Set

$$\begin{aligned}
\widetilde{C}_{\frac{i+k_n}{n}} - \widetilde{C}_{\frac{i}{n}} &= \frac{n}{k_n}\sigma^2_{\frac{b_\ell(p)}{n}}\sum_{j=1}^{k_n}\Big((\Delta^n_{i+k_n+j}W)^2 - (\Delta^n_{i+j}W)^2\Big),\\
\widetilde{D}_{\frac{i+k_n}{n}} - \widetilde{D}_{\frac{i}{n}} &= \frac{n}{k_n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\tau_{\frac{b_\ell(p)}{n}}(V_{s+\frac{k_n}{n}} - V_s)ds, \quad \text{both for } i = b_\ell(p),\ldots,a_{\ell+1}(p)-1.
\end{aligned}$$

Then the following claim holds.

Lemma 6.1 *We have*

$$\begin{aligned}
\mathbb{E}|A_{\frac{i+k_n}{n}} - A_{\frac{i}{n}} - (\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}})|^r &\le C_r(pn^{-1})^{r/2},\\
\mathbb{E}|B_{\frac{i+k_n}{n}} - B_{\frac{i}{n}} - (\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}})|^r &\le C_r(pn^{-1})^{r/2},
\end{aligned}$$

as well as

$$\mathbb{E}|A_{\frac{i+k_n}{n}} - A_{\frac{i}{n}}|^r \le C_r n^{-r/4} \quad \textit{and} \quad \mathbb{E}|B_{\frac{i+k_n}{n}} - B_{\frac{i}{n}}|^r \le C_r n^{-r/4} \tag{6.2}$$

for every $r > 0$. The latter bounds hold also for the approximated versions, and the same results are true for C and D and their approximations as well.

Proof. Note first that

$$B_{\frac{i+k_n}{n}} - B_{\frac{i}{n}} - (\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}}) = \frac{n}{k_n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\Big[\int_s^{s+\frac{k_n}{n}}\nu_r dr + \int_s^{s+\frac{k_n}{n}}(\tau_r - \tau_{\frac{a_\ell(p)}{n}})dV_r\Big]ds, \tag{6.3}$$

and due to the boundedness of ν the r-th moment of the first summand is bounded by $C_r n^{-r/2}$. Thus we focus on the latter summand only. Since σ^2 and τ^2 are continuous Itô semimartingales and both processes are bounded below by some positive constant, an application of Itô formula shows that σ and τ are continuous Itô semimartingales themselves and with similar representations. Therefore several applications of Hölder and Burkholder inequality yield

$$\begin{aligned}
\mathbb{E}\Big|\frac{n}{k_n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}(\tau_r - \tau_{\frac{a_\ell(p)}{n}})dV_r ds\Big|^r &\le C_r\frac{n}{k_n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\mathbb{E}\Big|\int_s^{s+\frac{k_n}{n}}(\tau_r - \tau_{\frac{a_\ell(p)}{n}})dV_r\Big|^r ds\\
\le \quad C_r\frac{n}{k_n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\mathbb{E}\Big[\Big(\int_s^{s+\frac{k_n}{n}}(\tau_r - \tau_{\frac{a_\ell(p)}{n}})^2 dr\Big)^{r/2}\Big]ds &\le C_r(pn^{-1})^{r/2}.
\end{aligned}$$

To see that $(A_{(i+k_n)/n} - A_{i/n}) - (\widetilde{A}_{(i+k_n)/n} - \widetilde{A}_{i/n})$ has the same property, note that this term can be written as the sum of two quantities, for which the first one is

$$\frac{n}{k_n}\sum_{j=1}^{k_n}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(X_s - X_{\frac{i+j-1}{n}})\mu_s ds$$
$$+\frac{n}{k_n}\sum_{j=1}^{k_n}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}((X_s - X_{\frac{i+j-1}{n}})\sigma_s - (W_s - W_{\frac{i+j-1}{n}})\sigma^2_{\frac{a_\ell(p)}{n}})dW_s. \quad (6.4)$$

The other quantity has a similar representation, but involves integrals within the interval $[(i+k_n)/n, (i+2k_n)/n]$. The r-th moment of the first term in (6.4) is of order $n^{-r/2}$ as before, whereas Burkholder inequality, the martingale property of W plus the semimartingale representation of σ give the desired bound for the approximation error concerning A. The bounds in (6.2) follow in a similar way. □

A simple consequence of Lemma 6.1 is that the remainder terms in U_t^n are negligible, that is

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{i=c(p)}^{\lfloor nt\rfloor-2k_n}\frac{3}{2k_n}(\hat\sigma^2_{\frac{i+k_n}{n}} - \hat\sigma^2_{\frac{i}{n}})^2 - \frac{6}{c^2}\int_{\frac{c(p)}{n}}^t \sigma_s^4 ds - \int_{\frac{c(p)}{n}}^t \tau_s^2 ds\Big| = 0,$$

using also boundedness of the processes on the right hand side and the definition of $c(p)$. A similar claim holds for the approximation of the integrands, and we restrict ourselves to the big blocks and prove

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\int_{\frac{a_\ell(p)}{n}}^{\frac{b_\ell(p)}{n}}(\tau_s^2 - \tau^2_{\frac{a_\ell(p)}{n}})ds\Big| = 0. \quad (6.5)$$

Recall (2.1). The result above follows from

$$\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\int_{\frac{a_\ell(p)}{n}}^{\frac{b_\ell(p)}{n}}\int_{\frac{a_\ell(p)}{n}}^{s}\omega_r dr ds\Big| \le C\frac{n}{pk_n}(\frac{pk_n}{n})^2 \le Cpn^{-1/2}$$

and

$$\mathbb{E}\Big(\sum_{\ell=1}^{J_n(p)}\int_{\frac{a_\ell(p)}{n}}^{\frac{b_\ell(p)}{n}}\int_{\frac{a_\ell(p)}{n}}^{s}\vartheta_r^{(1)}dW_r ds\Big)^2 = \sum_{\ell=1}^{J_n(p)}\mathbb{E}\Big(\int_{\frac{a_\ell(p)}{n}}^{\frac{b_\ell(p)}{n}}\int_{\frac{a_\ell(p)}{n}}^{s}\vartheta_r^{(1)}dW_r ds\Big)^2 \le Cp^2n^{-1},$$

since the terms involving $\vartheta^{(2)}$ and $\vartheta^{(3)}$ can be treated in the same way. Note that the analogue of (6.5) involving σ^4 instead of τ^2 holds for the same reasons. We have further

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\frac{pk_n}{n}\Big(\frac{n}{k_n^2} - \frac{1}{c^2}\Big)\sigma^4_{\frac{a_\ell(p)}{n}}\Big| = 0, \quad (6.6)$$

which by boundedness of σ amounts to prove $n^{-3/4}(k_n^2 - nc^2) = o(1)$, and the latter is satisfied by definition of k_n. Again, (6.6) holds over the small blocks as well.

The most involved part is of course on the error due to the approximation of increments of A and B. Our aim is to prove

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{1}{k_n}\Big(((A_{\frac{i+k_n}{n}}-A_{\frac{i}{n}})+(B_{\frac{i+k_n}{n}}-B_{\frac{i}{n}}))^2 - ((\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})+(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}}))^2\Big)\Big| = 0, \tag{6.7}$$

and from the proof it will become obvious that similar methods give the analogous result for the approximation via $\widetilde{C}_{(i+k_n)/n}-\widetilde{C}_{i/n}$ and $\widetilde{D}_{(i+k_n)/n}-\widetilde{D}_{i/n}$ within the small blocks.

First, the binomial theorem tells us that we can discuss the approximation for B, the one for A and the mixed part separately. Using further $x^2-y^2=2y(x-y)+(x-y)^2$ and $xx'-yy'=(x-y)y'+y(x'-y')+(x-y)(x'-y')$, we see from Lemma 6.1 and the growth conditions that (6.7) follows from $\lim_{p\to\infty}\limsup_{n\to\infty}\sum_{r=1}^{4}\mathbb{E}|L_{n,p}^{(j)}|=0$ with

$$L_{n,p}^{(1)}=\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{1}{k_n}\Big((B_{\frac{i+k_n}{n}}-B_{\frac{i}{n}})-(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}}), \tag{6.8}$$

$$L_{n,p}^{(2)}=\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{1}{k_n}\Big((B_{\frac{i+k_n}{n}}-B_{\frac{i}{n}})-(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big)(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}}), \tag{6.9}$$

$$L_{n,p}^{(3)}=\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{1}{k_n}\Big((A_{\frac{i+k_n}{n}}-A_{\frac{i}{n}})-(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})\Big)(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}}), \tag{6.10}$$

$$L_{n,p}^{(4)}=\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{1}{k_n}\Big((A_{\frac{i+k_n}{n}}-A_{\frac{i}{n}})-(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}}). \tag{6.11}$$

Let us start with the claim for (6.8) and we discuss the part within (6.3) involving ν first. We have $\nu_r=(\nu_r-\nu_{a_\ell(p)/n})+\nu_{a_\ell(p)/n}$. The latter term is treated using Lemma 6.1, as we have

$$\frac{n}{k_n}\mathbb{E}\Big(\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\nu_{\frac{a_\ell(p)}{n}}drds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big)^2 \tag{6.12}$$

$$= \frac{n^3}{k_n^5}\sum_{\ell=1}^{J_n(p)}\mathbb{E}\Big(\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\nu_{\frac{a_\ell(p)}{n}}drds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big)^2\le Cpn^{-1/2},$$

which converges to zero for any fixed p. For the other one we use left continuity of ν. From Fubini theorem and two applications of Cauchy-Schwarz inequality we have

$$\mathbb{E}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}(\nu_r-\nu_{\frac{a_\ell(p)}{n}})drds\Big)^2\le\frac{k_n^3}{n^3}\int_{\frac{i}{n}}^{\frac{i+2k_n}{n}}\mathbb{E}|\nu_r-\nu_{\frac{a_\ell(p)}{n}}|^2dr,$$

therefore Lemma 6.1 shows

$$\begin{aligned}
&\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}(\nu_r-\nu_{\frac{a_\ell(p)}{n}})drds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big| \\
\leq\; & Cn^{-3/4}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\Big(\int_{\frac{i}{n}}^{\frac{i+2k_n}{n}}\mathbb{E}|\nu_r-\nu_{\frac{a_\ell(p)}{n}}|^2dr\Big)^{1/2} \qquad (6.13)\\
\leq\; & C\sqrt{p}n^{-1/4}\sum_{\ell=1}^{J_n(p)}\Big(\int_{\frac{a_\ell(p)}{n}}^{\frac{a_{\ell+1}(p)}{n}}\mathbb{E}|\nu_r-\nu_{\frac{a_\ell(p)}{n}}|^2dr\Big)^{1/2}\leq C\Big(\int_0^1\mathbb{E}|\nu_r-\nu_{[p,n](r)}|^2dr\Big)^{1/2},
\end{aligned}$$

where $[p,n](r)$ denotes the largest $a_\ell(p)$ smaller than r. We call $\gamma(n,p)$ the right hand side above. For fixed r and p, $[p,n](r)$ converges to r from the left, so $|\nu_r-\nu_{p,n(r)}|^2$ converges to zero pointwise as well, since ν is left continuous. By boundedness of ν and Lebesgue theorem, $\gamma(n,p)$ is a zero sequence for all p, and we are done with this part as well. Finally,

$$\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}(\tau_r-\tau_{\frac{a_\ell(p)}{n}})dV_rds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}}) \qquad (6.14)$$

is treated as follows: Call τ_t' the sum of the last three terms in (2.1). Then $\tau_r-\tau_{a_\ell(p)/n}=\int_{a_\ell(p)/n}^r\omega_udu+(\tau_r'-\tau_{a_\ell(p)/n}')$, and using Lemma 6.1 we have

$$\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r\omega_ududV_rds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big|\leq Cp\frac{k_n^{1/2}}{n^{1/2}},$$

which converges to zero for every $p>0$. Have a look at the first Brownian term in $\tau_r'-\tau_{a_\ell(p)/n}'$, for which the decomposition

$$\int_{\frac{a_\ell(p)}{n}}^r\vartheta_u^{(1)}dW_u=\int_{\frac{a_\ell(p)}{n}}^r\vartheta_{\frac{a_\ell(p)}{n}}^{(1)}dW_u+\int_{\frac{a_\ell(p)}{n}}^r(\vartheta_u^{(1)}-\vartheta_{\frac{a_\ell(p)}{n}}^{(1)})dW_u$$

holds. We use the fact that (W,V) is jointly Brownian. Conditioning on $\mathcal{F}_{a_\ell(p)/n}$, properties of the normal distribution show that

$$\begin{aligned}
&\frac{n}{k_n}\mathbb{E}\Big(\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r\vartheta_{\frac{a_\ell(p)}{n}}^{(1)}dW_udV_rds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big)^2 \qquad (6.15)\\
=\;&\frac{n}{k_n}\sum_{\ell=1}^{J_n(p)}\mathbb{E}\Big(\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r\vartheta_{\frac{a_\ell(p)}{n}}^{(1)}dW_udV_rds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big)^2,
\end{aligned}$$

which is of order $p^2n^{-1/2}$. On the other hand, Cauchy-Schwarz and Burkholder inequality give

$$\begin{aligned}
&\mathbb{E}\Big|\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r(\vartheta_u^{(1)}-\vartheta_{\frac{a_\ell(p)}{n}}^{(1)})dW_udV_rds\Big|^2\\
\leq\;&\frac{k_n}{n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r\mathbb{E}(\vartheta_u^{(1)}-\vartheta_{\frac{a_\ell(p)}{n}}^{(1)})^2dudrds,
\end{aligned}$$

thus from Lemma 6.1 and Cauchy-Schwarz inequality again

$$
\begin{aligned}
&\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r(\vartheta_u^{(1)}-\vartheta_{\frac{a_\ell(p)}{n}}^{(1)})dW_u dV_r ds\Big)(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})\Big| \\
\leq\ & Cn^{-1/4}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r \mathbb{E}(\vartheta_u^{(1)}-\vartheta_{\frac{a_\ell(p)}{n}}^{(1)})^2 du dr ds\Big)^{1/2}.
\end{aligned}
$$

Convergence to zero for any fixed p can be deduced in the same way as for (6.13). The remaining two summands in $\tau_r'-\tau_{a_\ell(p)/n}'$ can be treated similarly, thus the claim for (6.8) is entirely shown.

Note that the first two steps go through for (6.9) as well: To obtain the analogue of (6.12) we need the unbiasedness of $\widetilde{A}_{(i+k_n)/n}-\widetilde{A}_{i/n}$ and the bounds of Lemma 6.1 only, and the latter are used for (6.13) as well. In the proof of (6.14) the only difference regards (6.15), as

$$
Y_\ell^n=\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_s^{s+\frac{k_n}{n}}\int_{\frac{a_\ell(p)}{n}}^r\vartheta_{\frac{a_\ell(p)}{n}}^{(1)}dW_u dV_r ds\Big)(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})
$$

is not unbiased. Nevertheless, let us have a look at $\mathbb{E}^n_{a_\ell(p)}(Y_\ell^n)$. Since W and V are jointly Brownian, standard Itô calculus yields

$$
\Big|\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\mathbb{E}\Big[\Big(\int_s^{s+\frac{k_n}{n}}(W_r-W_{\frac{a_\ell(p)}{n}})dV_r\Big)\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(W_u-W_{\frac{i+j-1}{n}})dW_u\Big)\Big]ds\Big|\leq C\frac{k_n}{n^3}
$$

for arbitrary $j=1,\ldots,2k_n$ and regardless of i. Using the representation in (2.10) we obtain $|\mathbb{E}^n_{a_\ell(p)}(Y_\ell^n)|\leq pn^{-1}$. Furthermore, the previously used arguments yield $\mathbb{E}(Y_\ell^n)^2\leq p^3n^{-3/2}$. Therefore

$$
\begin{aligned}
\frac{n}{k_n}\mathbb{E}\Big(\sum_{\ell=1}^{J_n(p)}Y_\ell^n\Big)^2 &= \frac{n}{k_n}\sum_{\ell=1}^{J_n(p)}\mathbb{E}(Y_\ell^n)^2+2\frac{n}{k_n}\sum_{m<\ell}^{J_n(p)}\mathbb{E}(\mathbb{E}^n_{a_m(p)}(Y_\ell^n)Y_m^n) \\
\leq\ & Cp^2n^{-1/2}+2\frac{n}{k_n}\sum_{m<\ell}^{J_n(p)}\mathbb{E}(|\mathbb{E}^n_{a_m(p)}(Y_\ell^n)||Y_m^n|)\leq C(p^2n^{-1/2}+p^{1/2}n^{-1/4}), \quad (6.16)
\end{aligned}
$$

which gives the analogue of (6.15), so the claim for (6.9) is entirely shown.

Let us come to the proof of (6.10). Recall (6.4). As for Lemma 6.1 we will only show the claim for the summands involving integrals over $[i/n,(i+k_n)/n]$, as the entire proof is obtained in the same way. Let us start with the ds-part. Using the bounds from Lemma 6.1 we have

$$
\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\sum_{j=1}^{k_n}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\int_{\frac{i+j-1}{n}}^s\mu_u du\Big)\mu_s ds\Big)(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})\Big|\leq Cn^{-1/2}, \quad (6.17)
$$

and that

$$\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\sum_{j=1}^{k_n}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\int_{\frac{i+j-1}{n}}^{s}\sigma_u dW_u\Big)(\mu_s-\mu_{\frac{a_\ell(p)}{n}})ds\Big)(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})\Big| \quad (6.18)$$

can be bounded by a certain zero sequence $\gamma(n,p)$ in the sense following (6.13). We set further

$$Z_\ell^n=\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n}{k_n^2}\Big(\sum_{j=1}^{k_n}\mu_{\frac{a_\ell(p)}{n}}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\int_{\frac{i+j-1}{n}}^{s}\sigma_r dW_r ds\Big)(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}}).$$

Note for $j_1<j_2$ that

$$\int_{\frac{i+j_1-1}{n}}^{\frac{i+j_1}{n}}\int_{\frac{i+j_2-1}{n}}^{\frac{i+j_2}{n}}\mathbb{E}\Big[\mu^2_{\frac{a_\ell(p)}{n}}\int_{\frac{i+j_1-1}{n}}^{s}\sigma_r dW_r\int_{\frac{i+j_2-1}{n}}^{t}\sigma_u dW_u\Big]dsdt=0$$

by conditioning on $\mathcal{F}_{(i+j_1-1)/n}$ and using the martingale property of W. Therefore and from Lemma 6.1 we have the bound

$$\mathbb{E}|Z_\ell^n|\le Cpk_n\frac{n}{k_n^2}\sqrt{k_n n^{-3}}n^{-1/4}\le Cpn^{-1},\quad \text{thus}\quad \sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\mathbb{E}|Z_\ell^n|\le Cn^{-1/4}. \quad (6.19)$$

The remainder part of (6.4) is decomposed into three terms, namely

$$\begin{aligned}&\frac{n}{k_n}\sum_{j=1}^{k_n}\Big[\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\int_{\frac{i+j-1}{n}}^{s}\sigma_u dW_u\Big)(\sigma_s-\sigma_{\frac{a_\ell(p)}{n}})dW_s \quad (6.20)\\&+\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\int_{\frac{i+j-1}{n}}^{s}(\sigma_u-\sigma_{\frac{a_\ell(p)}{n}})dW_u\Big)\sigma_{\frac{a_\ell(p)}{n}}dW_s+\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\int_{\frac{i+j-1}{n}}^{s}a_u du\Big)\sigma_s dW_s\Big].\end{aligned}$$

We begin with the first term, and to this end recall that σ is a continuous Itô semimartingale (and in particular that its driving Brownian motion is V and we call its volatility process $\tilde{\tau}$ which has a representation similar to (2.1)). It is sufficient to prove that

$$\begin{aligned}&\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n^2}{k_n^3}\sum_{j=1}^{k_n}\sigma^2_{\frac{a_\ell(p)}{n}}\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(W_s-W_{\frac{i+j-1}{n}})dW_s\Big)\\&\qquad\sum_{j=1}^{k_n}\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\int_{\frac{i+j-1}{n}}^{s}\sigma_u dW_u(\sigma_s-\sigma_{\frac{a_\ell(p)}{n}})dW_s\Big) \quad (6.21)\end{aligned}$$

converges to zero in the usual sense, and the same arguments as above show that the error due to replacing $\sigma_s-\sigma_{a_\ell(p)/n}$ by $\int_{a_\ell(p)/n}^{s}\tilde{\tau}_r dV_r$ is bounded by $Cpn^{-1/2}$. Then we successively replace σ_u by the corresponding $\sigma_{a_\ell(p)/n}$ and for $\tilde{\tau}$ as well. This error is of order $Cp^{1/2}n^{-1/4}$ each, and finally we are left with $\sqrt{n/k_n}\sum_{\ell=1}^{J_n(p)}\sigma^3_{a_\ell(p)/n}\tilde{\tau}_{a_\ell(p)/n}R^n_\ell$, setting $R^n_\ell=\sum_{i=a_\ell(p)}^{b_\ell(p)-1}n^2/k_n^3S^{(1)}_{i,n}S^{(2)}_{i,n}$ with

$$S^{(1)}_{i,n}=\sum_{j=1}^{k_n}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(W_s-W_{\frac{i+j-1}{n}})dW_s,\quad S^{(2)}_{i,n}=\sum_{j=1}^{k_n}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\int_{\frac{i+j-1}{n}}^{s}dW_u\int_{\frac{a_\ell(p)}{n}}^{s}dV_r\Big)dW_s.$$

By successive conditioning we obtain $\mathbb{E}(S^{(1)}_{i,n})^4\le Cn^{-3}$, as each summand of the quadruple sum has a non-zero expectation only if it comes from two pairs of equal indices. Proving an upper bound for the fourth moment of $S^{(2)}_{i,n}$ is a bit more involved, as one has to be careful with conditioning. We decompose $S^{(2)}_{i,n}=S^{(2,1)}_{i,n}+S^{(2,2)}_{i,n}$ by splitting the dV_s integral into $(V_s-V_{(i+j-1)/n})+(V_{(i+j-1)/n}-V_{a_\ell(p)/n})$. For the first term we obtain the bound $\mathbb{E}(S^{(2,1)}_{i,n})^4\le Cn^{-5}$ in the same way as for $\mathbb{E}(S^{(1)}_{i,n})^4$, and we focus on

$$S^{(2,2)}_{i,n}=\sum_{j=1}^{k_n}\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(W_s-W_{\frac{i+j-1}{n}})dW_s\Big)(V_{\frac{i+j-1}{n}}-V_{\frac{a_\ell(p)}{n}}).$$

If one index in the corresponding quadruple sum is larger than any other, then conditioning gives a zero expectation. Therefore the only non-trivial case with three indices being different is when the two largest are the same. Nevertheless, a straight-forward computation shows that the expectation is then zero as well, and we obtain $\mathbb{E}(S^{(2,1)}_{i,n})^4\le Cp^2n^{-4}$. Overall, using generalised Hölder inequality,

$$\mathbb{E}(R^n_\ell)^2\le Cp^2k_n^2\frac{n^4}{k_n^6}n^{-3/2}pn^{-2}\le Cp^3n^{-3/2}.$$

Furthermore, properties of the normal distribution prove $\mathbb{E}(R^n_\ell)=0$. Thus

$$\frac{n}{k_n}\mathbb{E}\Big(\sum_{\ell=1}^{J_n(p)}\sigma^3_{\frac{a_\ell(p)}{n}}\tilde{\tau}_{\frac{a_\ell(p)}{n}}R^n_\ell\Big)^2\le C\frac{n}{k_n}\sum_{\ell=1}^{J_n(p)}\mathbb{E}(R^n_\ell)^2\le Cp^2n^{-1/2}. \tag{6.22}$$

The same arguments work for the second term in (6.20) as well. Finally,

$$\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n^2}{k_n^3}\mathbb{E}\Big|\sum_{j=1}^{k_n}\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(W_s-W_{\frac{i+j-1}{n}})dW_s\Big)\sum_{j=1}^{k_n}\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\int_{\frac{i+j-1}{n}}^{s}a_udu\sigma_sdW_s\Big)\Big|$$

can be bounded by $Cn^{-1/4}$.

It remains to discuss (6.11). The analogues of (6.17)–(6.19) follow immediately from Lemma 6.1, and so does the final step above. Thus all we need to prove is negligibility of

$$\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{n^2}{k_n^3}\Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\tau_{\frac{a_\ell(p)}{n}}(V_{s+\frac{k_n}{n}}-V_s)ds\Big)$$
$$\sum_{j=1}^{k_n}\Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\int_{\frac{i+j-1}{n}}^{s}\sigma_udW_u(\sigma_s-\sigma_{\frac{a_\ell(p)}{n}})dW_s\Big),$$

and as before it is sufficient to discuss $\sqrt{\frac{n}{k_n}}\sum_{\ell=1}^{J_n(p)} \tau_{a_\ell(p)/n}\sigma_{a_\ell(p)/n}\tilde{\tau}_{a_\ell(p)/n}G_\ell^n$ only, where

$$G_\ell^n = \sum_{i=a_\ell(p)}^{b_\ell(p)-1} \frac{n^2}{k_n^3} K_{i,n}^{(1)} S_{i,n}^{(2)}, \quad K_{i,n}^{(1)} = \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} (V_{s+\frac{k_n}{n}} - V_s) ds.$$

We have $\mathbb{E}(K_{i,n}^{(1)})^4 \leq Cn^{-3}$ and $\mathbb{E}(S_{i,n}^{(2)})^4 \leq Cp^2 n^{-4}$ as before. However, in contrast to the previous result G_ℓ^n is not unbiased. Therefore we compute a conditional expectation again. We have

$$\Big| \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \sum_{j=1}^{k_n} \mathbb{E}(V_{s+\frac{k_n}{n}} - V_s) \Big(\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} \Big(\int_{\frac{i+j-1}{n}}^{t} dW_u \int_{\frac{a_\ell(p)}{n}}^{t} dV_r \Big) dW_t \Big) ds \Big| \leq Cn^{-2}$$

as before, thus $|\mathbb{E}_{a_\ell(p)}^n (G_\ell^n)| \leq Cpn^{-1}$. Since $|\mathbb{E}_{a_\ell(p)}^n (G_\ell^n)^2| \leq Cp^3 n^{-3/2}$ as well, the result follows as in (6.16).

The final step in the proof of (6.1) is concerned with the contribution of the small blocks, so we have to show that

$$\lim_{p\to\infty} \limsup_{n\to\infty} \sqrt{\frac{n}{k_n}} \mathbb{E}\Big| \sum_{\ell=1}^{J_n(p)} \Big(\sum_{i=b_\ell(p)}^{a_{\ell+1}(p)-1} \frac{3}{2k_n} ((\widetilde{C}_{\frac{i+k_n}{n}} - \widetilde{C}_{\frac{i}{n}}) + (\widetilde{D}_{\frac{i+k_n}{n}} - \widetilde{D}_{\frac{i}{n}}))^2 - \frac{2k_n}{n} \Big[\frac{6n}{k_n^2} \sigma^4_{\frac{b_\ell(p)}{n}} + \tau^2_{\frac{b_\ell(p)}{n}} \Big] \Big) \Big| = 0. \quad (6.23)$$

For this purpose and for later reasons, we compute the conditional expection of the approximated increments, and we will do this for the $\widetilde{A}$ and $\widetilde{B}$ terms only. We have

$$\mathbb{E}_{a_\ell(p)}^n (\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}})^2 = \frac{n^2}{k_n^2} \sigma^4_{\frac{a_\ell(p)}{n}} \sum_{j=1}^{k_n} \mathbb{E}\Big((\Delta_{i+k_n+j}^n W)^2 - (\Delta_{i+k_n}^n W)^2 \Big)^2 = \frac{4}{k_n} \sigma^4_{\frac{a_\ell(p)}{n}} \quad (6.24)$$

as well as

$$\begin{aligned} \mathbb{E}_{a_\ell(p)}^n (\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}})^2 &= 2\frac{n^2}{k_n^2} \tau^2_{\frac{a_\ell(p)}{n}} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{i}{n}}^{s} \mathbb{E}[(V_{s+\frac{k_n}{n}} - V_s)(V_{r+\frac{k_n}{n}} - V_r)] dr ds \\ &= 2\frac{n^2}{k_n^2} \tau^2_{\frac{a_\ell(p)}{n}} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{i}{n}}^{s} \Big(r + \frac{k_n}{n} - s \Big) dr ds \\ &= \frac{n^2}{k_n^2} \tau^2_{\frac{a_\ell(p)}{n}} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \Big(\frac{k_n^2}{n^2} - \Big(\frac{i+k_n}{n} - s \Big)^2 \Big) ds = \frac{2}{3} \frac{k_n}{n} \tau^2_{\frac{a_\ell(p)}{n}}. \end{aligned}$$

The expectation of the mixed part is zero. Therefore the $U_\ell^{n,p}$ are indeed martingale differences, and we have

$$\mathbb{E}_{b_\ell(p)}^n \Big(\sum_{i=a_\ell(p)}^{b_\ell(p)-1} \frac{3}{2k_n} ((\widetilde{C}_{\frac{i+k_n}{n}} - \widetilde{C}_{\frac{i}{n}}) + (\widetilde{D}_{\frac{i+k_n}{n}} - \widetilde{D}_{\frac{i}{n}}))^2 - \frac{2k_n}{n} \Big[\frac{6n}{k_n^2} \sigma^4_{\frac{b_\ell(p)}{n}} + \tau^2_{\frac{b_\ell(p)}{n}} \Big] \Big) = 0$$

as well. (6.23) then follows from the fact that

$$\frac{n}{k_n}\sum_{\ell=1}^{J_n(p)}\mathbb{E}\Big(\sum_{i=b_\ell(p)}^{a_{\ell+1}(p)-1}\frac{3}{2k_n}((\widetilde{C}_{\frac{i+k_n}{n}}-\widetilde{C}_{\frac{i}{n}})+(\widetilde{D}_{\frac{i+k_n}{n}}-\widetilde{D}_{\frac{i}{n}}))^2-\frac{2k_n}{n}\Big[\frac{6n}{k_n^2}\sigma^4_{\frac{b_\ell(p)}{n}}+\tau^2_{\frac{b_\ell(p)}{n}}\Big]\Big)^2$$

is bounded by a constant times p^{-1}, using Lemma 6.1 and $p\to\infty$. □

6.2 Proof of (2.13)

Let us check the conditions for stable convergence in this step, where particularly the proof of (2.14) is tedious. Write $U_\ell^{n,p}=\sum_{s=1}^3 U_\ell^{n,p,s}$ with

$$U_\ell^{n,p,1}=\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{3}{2k_n}\Big((\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})^2-\frac{4}{k_n}\sigma^4_{\frac{a_\ell(p)}{n}}\Big),$$

$$U_\ell^{n,p,2}=\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{3}{2k_n}\Big((\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}})^2-\frac{2k_n}{3n}\tau^2_{\frac{a_\ell(p)}{n}}\Big),$$

$$U_\ell^{n,p,3}=\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{3}{k_n}(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})(\widetilde{B}_{\frac{i+k_n}{n}}-\widetilde{B}_{\frac{i}{n}}).$$

It turns out that only the $(U_\ell^{n,p,s})^2$ terms are responsible for the conditional variance, whereas the remaining mixed ones are of small order each. Let us start with the pure σ part in the conditional variance which is due to

$$\mathbb{E}^n_{a_\ell(p)}(U_\ell^{n,p,1})^2=\sum_{i,m=a_\ell(p)}^{b_\ell(p)-1}\frac{9}{4k_n^2}\Big(\mathbb{E}^n_{a_\ell(p)}\Big[(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})^2(\widetilde{A}_{\frac{m+k_n}{n}}-\widetilde{A}_{\frac{m}{n}})^2\Big]-\frac{16}{k_n^2}\sigma^8_{\frac{a_\ell(p)}{n}}\Big).$$

Due to conditional independence of $\widetilde{A}_{(i+k_n)/n}-\widetilde{A}_{i/n}$ and $\widetilde{A}_{(m+k_n)/n}-\widetilde{A}_{m/n}$ for $|i-m|>2k_n$ and because of (6.24) we have to discuss the remaining cases with $|i-m|\le 2k_n$ only. Let us stay away from the boundary first and compute

$$\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n}\sum_{m=i-2k_n}^{i+2k_n}\frac{9}{4k_n^2}\Big(\mathbb{E}^n_{a_\ell(p)}\Big[(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})^2(\widetilde{A}_{\frac{m+k_n}{n}}-\widetilde{A}_{\frac{m}{n}})^2\Big]-\frac{16}{k_n^2}\sigma^8_{\frac{a_\ell(p)}{n}}\Big)$$

$$=\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n}\sum_{m=i-2k_n}^{i-1}\frac{9}{2k_n^2}\Big(\mathbb{E}^n_{a_\ell(p)}\Big[(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})^2(\widetilde{A}_{\frac{m+k_n}{n}}-\widetilde{A}_{\frac{m}{n}})^2\Big]-\frac{16}{k_n^2}\sigma^8_{\frac{a_\ell(p)}{n}}\Big)+O_p(pn^{-3/2}),$$

using Lemma 6.1. Recall the definition of H_i^n. The task is to simplify

$$\mathbb{E}^n_{a_\ell(p)}\Big[(\widetilde{A}_{\frac{i+k_n}{n}}-\widetilde{A}_{\frac{i}{n}})^2(\widetilde{A}_{\frac{m+k_n}{n}}-\widetilde{A}_{\frac{m}{n}})^2\Big]=16\frac{n^4}{k_n^4}\sigma^8_{\frac{a_\ell(p)}{n}}\sum_{j_1,j_2,j_3,j_4=1}^{k_n}$$
$$\mathbb{E}[(H^n_{i+j_1+k_n}-H^n_{i+j_1})(H^n_{i+j_2+k_n}-H^n_{i+j_2})(H^n_{m+j_3+k_n}-H^n_{m+j_3})(H^n_{m+j_4+k_n}-H^n_{m+j_4})].$$

In order for each expectation on the right hand side to not vanish, at least one index of each paranthesis has to agree with one of another. Note that $j_1 = j_2$ and $j_3 = j_4$ corresponds exactly to the subtracted mean (apart from the small order case of four equal indices), which is why we focus on the few cases left. Suppose that $i + j_1 = m + j_3$. In this case either $i + j_2 = m + j_4$ or $i + j_2 = m + k_n + j_4$. The same options exist for $i+j_1 = m+k_n+j_3$. By symmetry, the cases where indices within the first and the fourth paranthesis agree, can be discussed in the same way, which explains an additional factor 2. Let $m \leq i-k_n$. Then the only possible case is $i+j_1 = m+k_n+j_3$ and $i+j_2 = m+k_n+j_4$, whose contribution to the quadruple sum is

$$2 \sum_{\substack{j_1,j_2=1 \\ 1\leq j_1+i-m-k_n\leq k_n \\ 1\leq j_2+i-m-k_n\leq k_n}}^{k_n} \mathbb{E}[(H^n_{i+j_1})^2(H^n_{m+j_2})^2] = \frac{1}{2n^4}(2k_n-(i-m))^2 + O(n^{-7/2}),$$

using $\mathbb{E}(H^n_i)^2 = 1/(2n^2)$. For $m > i-k_n$ all options are allowed, and a careful observation shows that the quadruple sum becomes

$$\begin{aligned} &\frac{2}{n^4}(k_n-(i-m))^2 - \frac{2}{n^4}(i-m)(k_n-(i-m)) + \frac{1}{2n^4}(i-m)^2 + O(n^{-7/2}) \\ =\ &\frac{4k_n^2 - 12k_n(i-m) + 9(i-m)}{2n^4} + O(n^{-7/2}), \end{aligned}$$

where the first term above is due to $i + j_1 = m + j_3$ and $i + j_2 = m + j_4$, the second one belongs to the mixed parts, and the final one comes from $i + j_1 = m + k_n + j_3$ and $i + j_2 = m + k_n + j_4$ again. An index transformation gives in total

$$\begin{aligned} &\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n} \sum_{m=i-2k_n}^{i-1} \frac{9}{2k_n^2}\Big(\mathbb{E}^n_{a_\ell(p)}\Big[(\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}})^2(\widetilde{A}_{\frac{m+k_n}{n}} - \widetilde{A}_{\frac{m}{n}})^2\Big] - \frac{16}{k_n^2}\sigma^8_{\frac{a_\ell(p)}{n}}\Big) \\ =\ &\sigma^8_{\frac{a_\ell(p)}{n}}\frac{36}{k_n^6}\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n}\sum_{m=1}^{k_n}\Big(m^2 + (4k_n^2 - 12k_n m + 9m^2)\Big) + O_P(n^{-3/2}) \\ =\ &\sigma^8_{\frac{a_\ell(p)}{n}} p\frac{48}{k_n^2} + O_P(n^{-3/2}). \end{aligned}$$

A similar argument for the missing boundary terms reveals that their contribution equals $\sigma^8_{a_\ell(p)/n} d_1/k_n^2$ for some unspecified d_1, independently of p, and again up to some small order terms. Overall,

$$\mathbb{E}^n_{a_\ell(p)}(U^{n,p,1}_\ell)^2 = \sigma^8_{\frac{a_\ell(p)}{n}}\frac{48p+d_1}{k_n^2} + O_P(pn^{-3/2}).$$

Similarly, the main part of $\mathbb{E}^n_{a_\ell(p)}(U^{n,p,2}_\ell)^2$ is due to

$$\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n} \sum_{m=i-2k_n}^{i-1} \frac{9}{2k_n^2}\Big(\mathbb{E}^n_{a_\ell(p)}\Big[(\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}})^2(\widetilde{B}_{\frac{m+k_n}{n}} - \widetilde{B}_{\frac{m}{n}})^2\Big] - \frac{4k_n^2}{9n^2}\tau^4_{\frac{a_\ell(p)}{n}}\Big). \quad (6.25)$$

For a centred normal variable (N_1, N_2, N_3, N_4) we have

$$\mathbb{E}(N_1N_2N_3N_4) = \mathbb{E}(N_1N_2)\mathbb{E}(N_3N_4) + \mathbb{E}(N_1N_3)\mathbb{E}(N_2N_4) + \mathbb{E}(N_1N_4)\mathbb{E}(N_2N_3).$$

Applied to the increments of V in $(\widetilde{B}_{(i+k_n)/n} - \widetilde{B}_{i/n})^2(\widetilde{B}_{(m+k_n)/n} - \widetilde{B}_{m/n})^2$ we see that the first of the three terms above corresponds to the mean, thus by symmetry (6.25) equals

$$\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n} \sum_{m=i-2k_n}^{i-1} \frac{9n^4}{k_n^6} \tau^4_{\frac{a_\ell(p)}{n}} \Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{m}{n}}^{\frac{m+k_n}{n}} \mathbb{E}[(V_{s+\frac{k_n}{n}} - V_s)(V_{r+\frac{k_n}{n}} - V_r)] drds \Big)^2.$$

For $m \le i - k_n$ we have

$$\int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{m}{n}}^{\frac{m+k_n}{n}} \mathbb{E}[(V_{s+\frac{k_n}{n}} - V_s)(V_{r+\frac{k_n}{n}} - V_r)] drds = \int_{\frac{i}{n}}^{\frac{m+2k_n}{n}} \int_{s-\frac{k_n}{n}}^{\frac{m+k_n}{n}} (r + \frac{k_n}{n} - s) drds$$

$$= \frac{1}{2} \int_{\frac{i}{n}}^{\frac{m+2k_n}{n}} (\frac{m+2k_n}{n} - s)^2 ds = \frac{(m+2k_n-i)^3}{6n^3}, \qquad (6.26)$$

and analogously

$$\int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{m}{n}}^{\frac{m+k_n}{n}} \mathbb{E}[(V_{s+\frac{k_n}{n}} - V_s)(V_{r+\frac{k_n}{n}} - V_r)] drds = \int_{\frac{i}{n}}^{\frac{m+k_n}{n}} \int_{\frac{m}{n}}^{s} (r + \frac{k_n}{n} - s) drds$$

$$+ \int_{\frac{i}{n}}^{\frac{m+k_n}{n}} \int_{s}^{\frac{m+k_n}{n}} (s + \frac{k_n}{n} - r) drds + \int_{\frac{m+k_n}{n}}^{\frac{i+k_n}{n}} \int_{s-\frac{k_n}{n}}^{\frac{m+k_n}{n}} (r + \frac{k_n}{n} - s) drds$$

$$= \frac{4k_n^3 - 6(i-m)^2 k_n + 3(i-m)^3}{6n^3}$$

for $m > i - k_n$. Therefore (6.25) becomes

$$\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n} \frac{9n^4}{k_n^6} \tau^4_{\frac{a_\ell(p)}{n}} \sum_{m=1}^{k_n} \frac{1}{36n^6} \Big(m^6 + (4k_n^3 - 6m^2 k_n + 3m^3)^2 \Big) = p \frac{151}{70} \frac{k_n^2}{n^2} \tau^4_{\frac{a_\ell(p)}{n}} + O_P(pn^{-3/2}),$$

so for a certain d_2 we obtain

$$\mathbb{E}^n_{a_\ell(p)} (U_\ell^{n,p,2})^2 = \tau^4_{\frac{a_\ell(p)}{n}} (p \frac{151}{70} + d_2) \frac{k_n^2}{n^2} + O_P(pn^{-3/2}).$$

The term $\mathbb{E}^n_{a_\ell(p)}(U_\ell^{n,p,3})^2$ is responsible for joint part of the conditional variance. Again, we discuss only

$$\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n} \sum_{m=i-2k_n}^{i-1} \frac{18}{k_n^2} \mathbb{E}^n_{a_\ell(p)} \Big[(\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}})(\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}})(\widetilde{A}_{\frac{m+k_n}{n}} - \widetilde{A}_{\frac{m}{n}})(\widetilde{B}_{\frac{m+k_n}{n}} - \widetilde{B}_{\frac{m}{n}}) \Big]$$

in detail. Note that

$$\mathbb{E}^n_{a_\ell(p)} \Big[(\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}})(\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}})(\widetilde{A}_{\frac{m+k_n}{n}} - \widetilde{A}_{\frac{m}{n}})(\widetilde{B}_{\frac{m+k_n}{n}} - \widetilde{B}_{\frac{m}{n}}) \Big]$$

$$= 4 \frac{n^4}{k_n^4} \sigma^4_{\frac{a_\ell(p)}{n}} \tau^2_{\frac{a_\ell(p)}{n}} \sum_{j_1, j_2=1}^{k_n} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{m}{n}}^{\frac{m+k_n}{n}} \qquad (6.27)$$

$$\mathbb{E}[(V_{s+\frac{k_n}{n}} - V_s)(V_{r+\frac{k_n}{n}} - V_r)(H^n_{i+j_1+k_n} - H^n_{i+j_1})(H^n_{m+j_2+k_n} - H^n_{m+j_2})] drds$$

with the previously introduced notation. Thus we have to compute quantities like

$$\mathbb{E}[(V_{s+\frac{k_n}{n}} - V_s)(V_{r+\frac{k_n}{n}} - V_r)H^n_{i+j_1}H^n_{m+j_2}],$$

and one obtains that the expectation above is non-zero only for intervals satisfying the condition $[s, s+k_n/n] \cap [r, r+k_n/n] \neq \emptyset$ and indices with $i + j_1 = m + j_2$. As usual, let $m \leq i - k_n$ first. Then the double sum in (6.27) becomes

$$-\sum_{j=1}^{2k_n-(i-m)} \int_{\frac{i}{n}}^{\frac{m+2k_n}{n}} \int_{s-\frac{k_n}{n}}^{\frac{m+k_n}{n}} \mathbb{E}[(V_{r+\frac{k_n}{n}} - V_s)^2 (H^n_{i+j})^2] dr ds, \tag{6.28}$$

and the expectation factorises for all but a small order amount of choices for s and r. For (6.28) we thus obtain

$$-\frac{1}{2n^2} \sum_{j=1}^{2k_n-(i-m)} \int_{\frac{i}{n}}^{\frac{m+2k_n}{n}} \int_{s-\frac{k_n}{n}}^{\frac{m+k_n}{n}} (r + \frac{k_n}{n} - s) dr ds = -\frac{(2k_n - (i-m))^4}{12n^5}$$

plus a term of small order $O(n^{-7/2})$, where we have used (6.26). Analogously, for $m > i - k_n$ the double sum equals

$$\begin{aligned} &\Big(\sum_{\substack{j=1 \\ 1 \leq j+i-m \leq k_n}}^{k_n} \mathbb{E}(H^n_{i+j+k_n} - H^n_{i+j})^2 - \sum_{\substack{j=1 \\ 1 \leq j+i-m-k_n \leq k_n}}^{k_n} \mathbb{E}(H^n_{i+j})^2 \Big) \\ &\qquad\qquad\qquad\qquad \times \frac{4k_n^3 - 6(i-m)^2 k_n + 3(i-m)^3}{6n^3} \\ &= \frac{(2k_n - 3(i-m))(4k_n^3 - 6(i-m)^2 k_n + 3(i-m)^3)}{12n^5}, \end{aligned}$$

again up to some $O(n^{-7/2})$. Overall, the main part responsible for the mixed terms is

$$\begin{aligned} &\sum_{i=a_\ell(p)+2k_n}^{b_\ell(p)-2k_n} \frac{6}{nk_n^6} \sigma^4_{\frac{a_\ell(p)}{n}} \tau^2_{\frac{a_\ell(p)}{n}} \sum_{m=1}^{k_n} \Big((8k_n^4 - 12k_n^3 m - 12k_n^2 m^2 + 24k_n m^3 - 9m^4) - m^4 \Big) \\ = \; & p\frac{12}{n} \sigma^4_{\frac{a_\ell(p)}{n}} \tau^2_{\frac{a_\ell(p)}{n}} + O_p(pn^{-3/2}), \end{aligned}$$

so finally

$$\mathbb{E}^n_{a_\ell(p)} (U_\ell^{n,p,3})^2 = \sigma^4_{\frac{a_\ell(p)}{n}} \tau^2_{\frac{a_\ell(p)}{n}} \frac{12p + d_3}{n} + O_P(pn^{-3/2}).$$

It remains to prove that the mixed parts of $\mathbb{E}^n_{a_\ell(p)}[(U_\ell^{n,p})^2]$ are $O_p(pn^{-3/2})$ each, which is another simple but tedious task. We will drop these computations for the sake of brevity. Altogether, we have

$$\begin{aligned} &\frac{n}{k_n} \sum_{\ell=1}^{J_n(p)} \mathbb{E}^n_{a_\ell(p)}[(U_\ell^{n,p})^2] \\ = \; & \frac{pk_n}{n} \sum_{\ell=1}^{J_n(p)} \Big(\frac{n^2}{k_n^4}(48 + \frac{d_1}{p}) \sigma^8_{\frac{a_\ell(p)}{n}} + \frac{n}{k_n^2}(12 + \frac{d_2}{p}) \sigma^4_{\frac{a_\ell(p)}{n}} \tau^2_{\frac{a_\ell(p)}{n}} + (\frac{151}{70} + \frac{d_3}{p}) \tau^4_{\frac{a_\ell(p)}{n}} \Big) + O_p(n^{-1/2}), \end{aligned}$$

thus (2.14) holds using $k_n \sim cn^{1/2}$. Simpler to obtain is (2.15), as Lemma 6.1 gives

$$\frac{n^2}{k_n^2} \sum_{\ell=1}^{J_n(p)} \mathbb{E}^n_{a_\ell(p)}[(U_\ell^{n,p})^4] \leq C \frac{n^3}{pk_n^3} p^4 n^{-2},$$

which converges to zero in the usual sense. Finally, one can prove

$$\mathbb{E}^n_{a_\ell(p)}\Big[\sum_{i=a_\ell(p)}^{b_\ell(p)-1} \frac{3}{2k_n}((\widetilde{A}_{\frac{i+k_n}{n}} - \widetilde{A}_{\frac{i}{n}}) + (\widetilde{B}_{\frac{i+k_n}{n}} - \widetilde{B}_{\frac{i}{n}}))^2 (N_{\frac{a_{\ell+1}(p)}{n}} - N_{\frac{a_\ell(p)}{n}})\Big] = 0, \tag{6.29}$$

where N is either W or V or when N is a bounded martingale, orthogonal to (W, V). Focus on the first case and decompose $((\widetilde{A}_{(i+k_n)/n} - \widetilde{A}_{i/n}) + (\widetilde{B}_{(i+k_n)/n} - \widetilde{B}_{i/n}))^2$ via the binomial theorem. For the pure $\widetilde{A}$ and the pure $\widetilde{B}$ term, the claim follows immediately from properties of the normal distribution upon using that $\sigma_{a_\ell(p)/n}$ or $\tau_{a_\ell(p)/n}$ are $\mathcal{F}_{a_\ell(p)/n}$ measurable. For the mixed term, one has to use the special form of $\widetilde{A}_{(i+k_n)/n} - \widetilde{A}_{i/n}$ as a difference of two sums, and a symmetry argument proves (6.29) in this case. For an orthogonal N, we use standard calculus. By Itô formula, both $(\widetilde{A}_{(i+k_n)/n} - \widetilde{A}_{i/n})^2$ and $(\widetilde{B}_{(i+k_n)/n} - \widetilde{B}_{i/n})^2$ are a measurable variable times the sum of a constant and a stochastic integral with respect to W and V, respectively. Thus (6.29) holds. In the mixed case, we use integration by parts formula to reduce $(\widetilde{A}_{(i+k_n)/n} - \widetilde{A}_{i/n})(\widetilde{B}_{(i+k_n)/n} - \widetilde{B}_{i/n})$ to the sum of a constant, a dW and a dV integral. Then the same argument applies. Altogether, this gives (2.16). □

6.3 Proof of Theorem 2.5

Let us begin with the proof of (2.18), for which we write

$$\overline{\tau}^2_{\frac{i}{n}} = \frac{3n}{2k_n}((\overline{A}_{\frac{i+k_n}{n}} - \overline{A}_{\frac{i}{n}}) + (\overline{B}_{\frac{i+k_n}{n}} - \overline{B}_{\frac{i}{n}}))^2 - 6\frac{n}{k_n^2}\overline{\sigma}^4_{\frac{i}{n}}, \quad \overline{\sigma}^4_{\frac{i}{n}} = \frac{n^2}{3k_n}\sum_{j=1}^{k_n} \sigma^4_{\frac{i}{n}} |\Delta^n_{i+j} W|^4, \tag{6.30}$$

with

$$\overline{A}_{\frac{i+k_n}{n}} - \overline{A}_{\frac{i}{n}} = \frac{n}{k_n}\sigma^2_{\frac{i}{n}} \sum_{j=1}^{k_n} \Big(|\Delta^n_{i+k_n+j} W|^2 - |\Delta^n_{i+j} W|^2\Big),$$

$$\overline{B}_{\frac{i+k_n}{n}} - \overline{B}_{\frac{i}{n}} = \frac{n}{k_n} \int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \tau_{\frac{i}{n}} (V_{s+\frac{k_n}{n}} - V_s) ds.$$

First of all, we have

$$G^{(2)}_{t,n} = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor - 2k_n} \overline{\tau}^2_{\frac{i}{n}} \overline{\sigma}^4_{\frac{i}{n}} + O_p(n^{-1/4}),$$

since σ and τ are Itô semimartingales and thus the techniques from the proof of Lemma 6.1 show that both $\hat{\tau}^2_{i/n} - \overline{\tau}^2_{i/n} = O_p(n^{-1/4})$ and $\hat{\sigma}^4_{i/n} - \overline{\sigma}^4_{i/n} = O_p(n^{-1/4})$ hold. Note further by conditional independence that

$$\Big(\frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor - 2k_n} (\overline{\tau}^2_{\frac{i}{n}} \overline{\sigma}^4_{\frac{i}{n}} - \mathbb{E}^n_i [\overline{\tau}^2_{\frac{i}{n}} \overline{\sigma}^4_{\frac{i}{n}}])\Big)^2 = O_p(n^{-1/2}).$$

A simple computation gives $\mathbb{E}^n_i[\overline{\tau}^2_{i/n} \overline{\sigma}^4_{i/n}] = \tau^2_{i/n} \sigma^4_{i/n} + O_p(n^{-1/2})$, so (2.18) follows again from the Itô semimartingale property of both processes. The same techniques prove (2.17), and we also have

$$G^{(3)}_{t,n} = \sum_{i=1}^{\lfloor nt \rfloor - 2k_n} ((\overline{A}_{\frac{i+k_n}{n}} - \overline{A}_{\frac{i}{n}}) + (\overline{B}_{\frac{i+k_n}{n}} - \overline{B}_{\frac{i}{n}}))^4 + O_p(n^{-1/4}),$$

so all we have to do is to compute the conditional expectation of each summand. For the first term, this is simple, as we have

$$\mathbb{E}^n_i[(\overline{A}_{\frac{i+k_n}{n}} - \overline{A}_{\frac{i}{n}})^4] = 3 \frac{n^4}{k_n^4} \sigma^8_{\frac{i}{n}} \Big(\sum_{j=1}^{k_n} \mathbb{E}\Big(|\Delta^n_{i+k_n+j} W|^2 - |\Delta^n_{i+j} W|^2 \Big)^2 \Big)^2 = \frac{48}{k_n^2} \sigma^8_{\frac{i}{n}},$$

up to an error of order $n^{-3/2}$. It is simple to show

$$\mathbb{E}^n_i[(\overline{A}_{\frac{i+k_n}{n}} - \overline{A}_{\frac{i}{n}})^2 (\overline{B}_{\frac{i+k_n}{n}} - \overline{B}_{\frac{i}{n}})^2] = \frac{8}{3n} \sigma^4_{\frac{i}{n}} \tau^2_{\frac{i}{n}} + O_p(n^{-3/2}),$$

since we have seen already that the expectation factorises up to a small term error. Finally, properties of the normal distribution give

$$\mathbb{E}^n_i[(\overline{B}_{\frac{i+k_n}{n}} - \overline{B}_{\frac{i}{n}})^4] = 12 \frac{n^4}{k_n^4} \tau^4_{\frac{i}{n}} \Big(\int_{\frac{i}{n}}^{\frac{i+k_n}{n}} \int_{\frac{i}{n}}^{s} \mathbb{E}[(V_{r+\frac{k}{n}} - V_s)^2] dr ds \Big)^2 = \frac{4k_n^2}{3n^2} \tau^4_{\frac{i}{n}}.$$

The two remaining terms have zero expectation. □

6.4 Proof of Lemma 3.1

We will only prove the first result. Note that

$$B_t - \hat{B}_t = \sum_{i=0}^{\lfloor nt \rfloor - 2k_n} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \Big(\tau^2(s, X_s, \sigma^2_s) - \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}^2_{\frac{i}{n}}) \Big) ds + O_p(n^{-1/2}),$$

the error coming from border terms in B_t, for which we have used boundedness of the function τ^2, due to differentiability and the assumption that any process involved is bounded itself. We have

$$\begin{aligned} & \tau^2(s, X_s, \sigma^2_s) - \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}^2_{\frac{i}{n}}) \qquad (6.31) \\ = \; & (\tau^2(s, X_s, \sigma^2_s) - \tau^2(\frac{i}{n}, X_s, \sigma^2_s)) + (\tau^2(\frac{i}{n}, X_s, \sigma^2_s) - \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \sigma^2_s) \\ & + (\tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \sigma^2_s) - \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \sigma^2_{\frac{i}{n}})) + (\tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \sigma^2_{\frac{i}{n}}) - \tau^2(\frac{i}{n}, X_{\frac{i}{n}}, \hat{\sigma}^2_{\frac{i}{n}}). \end{aligned}$$

The first three terms can be discussed in the same way. From differentiability we may conclude

$$\sum_{i=0}^{\lfloor nt\rfloor-2k_n}\int_{\frac{i-1}{n}}^{\frac{i}{n}}(\tau^2(s,X_s,\sigma_s^2)-\tau^2(\frac{i}{n},X_s,\sigma_s^2))ds=\sum_{i=0}^{\lfloor nt\rfloor-2k_n}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\frac{\partial}{\partial s}\tau^2(\xi,X_s,\sigma_s^2)(s-\frac{i}{n})ds$$

for a suitable ξ, and the term is obviously of order n^{-1}. In the same way, we see that the second and third term in (6.31) are of order $n^{-1/2}$ each. For the last quantity, we use twice differentiability and Lemma 6.1 to obtain

$$\tau^2(\frac{i}{n},X_{\frac{i}{n}},\hat\sigma^2_{\frac{i}{n}})-\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})=\frac{\partial}{\partial\sigma^2}\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})(\hat\sigma^2_{\frac{i}{n}}-\sigma^2_{\frac{i}{n}})+O_p(n^{-1/2}),$$

uniformly in i. Also, $\hat\sigma^2_{i/n}-\sigma^2_{i/n}=M_i^n+O_p(n^{-1/2})$, where the

$$M_i^n=\frac{n}{k_n}\sum_{j=1}^{k_n}2\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(X_s-X_{\frac{i+j-1}{n}})\sigma_s dW_s+\frac{n}{k_n}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}\int_{\frac{i}{n}}^{s}\tau_u dV_u ds$$

are martingale differences of order $n^{-1/4}$. Therefore Lemma 3.1 follows from

$$\mathbb{E}\Big(\frac{1}{n}\sum_{i=0}^{\lfloor nt\rfloor-2k_n}\frac{\partial}{\partial\sigma^2}\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})M_i^n\Big)^2\le Cn^{-1},$$

where we have used Lemma 6.1 again plus $\mathbb{E}[M_i^nM_j^n]=0$ for $|i-j|\ge k_n$. □

6.5 Proof of Lemma 3.2

From (2.1) and by differentiability of the function τ^2 we have

$$\hat C-C=\frac{1}{n}\sum_{i=0}^{n-2k_n}\Big(\hat\tau^2_{\frac{i}{n}}\tau^2(\frac{i}{n},X_{\frac{i}{n}},\hat\sigma^2_{\frac{i}{n}})-\tau^2_{\frac{i}{n}}\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})\Big)+O_p(n^{-1/2}).$$

The first claim is

$$\frac{1}{n}\sum_{i=0}^{n-2k_n}\hat\tau^2_{\frac{i}{n}}\Big(\tau^2(\frac{i}{n},X_{\frac{i}{n}},\hat\sigma^2_{\frac{i}{n}})-\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})\Big)=O_p(n^{-1/2}).\tag{6.32}$$

Recall M_i^n from the previous proof and set

$$\overline{M}_i^n=\frac{n}{k_n}\sum_{j=1}^{k_n}2\sigma^2_{\frac{i}{n}}\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}(W_s-W_{\frac{i+j-1}{n}})dW_s+\frac{n}{k_n}\tau_{\frac{i}{n}}\int_{\frac{i}{n}}^{\frac{i+k_n}{n}}(V_s-V_{\frac{i}{n}})ds.$$

Standard methods give $M_i^n-\overline{M}_i^n=O_p(n^{-1/2})$. Using the mean value theorem, we conclude that the left hand side of (6.32) equals

$$\frac{1}{n}\sum_{i=0}^{n-2k_n}\vartheta_i^n+O_p(n^{-1/2}),\quad \vartheta_i^n=\frac{\partial}{\partial\sigma^2}\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})\overline{\tau}^2_{\frac{i}{n}}\overline{M}_i^n.$$

ϑ_i^n is obviously of order $n^{-1/4}$, so by conditional independence we have

$$\mathbb{E}\Big(\frac{1}{n}\sum_{i=0}^{n-2k_n}(\vartheta_i^n-\mathbb{E}_i^n[\vartheta_i^n])\Big)^2=\frac{1}{n^2}\sum_{\substack{i,j=1\\|i-j|\le 2k_n}}^{n-2k_n}\mathbb{E}[(\vartheta_i^n-\mathbb{E}_i^n[\vartheta_i^n])(\vartheta_j^n-\mathbb{E}_j^n[\vartheta_j^n])]\le Cn^{-1}.$$

We omit to compute $\mathbb{E}_i^n[\vartheta_i^n]$ in details. Standard arguments give $|\mathbb{E}_i^n[\vartheta_i^n]|\le Cn^{-1/2}$, from which (6.32) follows, so we have

$$\hat{C}-C=\frac{1}{n}\sum_{i=0}^{n-2k_n}\Big(\hat{\tau}^2_{\frac{i}{n}}-\tau^2_{\frac{i}{n}}\Big)\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})+O_p(n^{-1/2}).$$

We will use the same blocking technique as in the proof of Theorem 2.1 now. Let $I_n(p)$ be defined as $J_n(p)$ before, but with $t=1$. We proceed in two steps. The first one is

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\frac{1}{n}\sum_{\ell=1}^{I_n(p)}\sum_{i=a_\ell(p)}^{b_\ell(p)-1}(\hat{\tau}^2_{\frac{i}{n}}-\tau^2_{\frac{i}{n}})(\tau^2(\frac{i}{n},X_{\frac{i}{n}},\sigma^2_{\frac{i}{n}})-\tau^2(\frac{a_\ell(p)}{n},X_{\frac{a_\ell(p)}{n}},\sigma^2_{\frac{a_\ell(p)}{n}}))\Big|=0,$$

and there is of course a related result concerning the small blocks. This result is in fact quite simple to show. The assumptions on the function τ^2 and growth conditions of continuous Itô semimartingales reduce the claim to

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\frac{1}{n}\sum_{\ell=1}^{I_n(p)}\frac{\partial}{\partial x}\tau^2(\frac{a_\ell(p)}{n},X_{\frac{a_\ell(p)}{n}},\sigma^2_{\frac{a_\ell(p)}{n}})\sum_{i=a_\ell(p)}^{b_\ell(p)-1}(\hat{\tau}^2_{\frac{i}{n}}-\tau^2_{\frac{i}{n}})(X_{\frac{i}{n}}-X_{\frac{a_\ell(p)}{n}})\Big|=0$$

and an analogous one involving the partial derivative with respect to σ^2 and increments of σ^2, which can be discussed in the same way. Let $\widetilde{\tau}^2_{i/n}$ be defined as $\overline{\tau}^2_{i/n}$ in (6.30), but with $\overline{A}$ and $\overline{B}$ replaced with $\widetilde{A}$ and $\widetilde{B}$, respectively, and $\sigma_{i/n}$ with $\sigma_{a_\ell(p)/n}$. Denote with N_ℓ an unspecified $\mathcal{F}_{a_\ell(p)/n}$-measurable random variable. Then Lemma 6.1 and growth conditions again show that we might as well prove that

$$\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\frac{1}{n}\sum_{\ell=1}^{I_n(p)}N_\ell\sum_{i=a_\ell(p)}^{b_\ell(p)-1}(\widetilde{\tau}^2_{\frac{i}{n}}-\tau^2_{\frac{a_\ell(p)}{n}})\sigma_{\frac{a_\ell(p)}{n}}(W_{\frac{i}{n}}-W_{\frac{a_\ell(p)}{n}})\Big|$$

becomes small, which follows in a similar way as (6.22) by conditional independence. For the second step recall $U_\ell^{n,p}$ from (2.11). We will show that

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{I_n(p)}\tau^2(\frac{a_\ell(p)}{n},X_{\frac{a_\ell(p)}{n}},\sigma^2_{\frac{a_\ell(p)}{n}})\Big(\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{1}{n}(\hat{\tau}^2_{\frac{i}{n}}-\tau^2_{\frac{i}{n}})-U_\ell^{n,p}\Big)\Big|=0 \quad (6.33)$$

holds. For the increments involving A and B within $\hat{\tau}^2_{i/n}-\tau^2_{i/n}$, the proof is identical to the one of (2.12). Let us show

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\mathbb{E}\Big|\sum_{\ell=1}^{I_n(p)}N_\ell\sum_{i=a_\ell(p)}^{b_\ell(p)-1}\frac{2n^2}{k_n^3}\sum_{j=1}^{k_n}\Big((\Delta_{i+j}^nX)^4-\frac{3}{n^2}\sigma^4_{\frac{a_j(p)}{n}}\Big)\Big|=0, \qquad (6.34)$$

for which we use the decomposition

$$\begin{aligned}(\Delta_{i+j}^n X)^4 - 3n^{-2}\sigma^4_{\frac{a_j(p)}{n}} &= ((\Delta_{i+j}^n X)^4 - \sigma^4_{\frac{i+j-1}{n}}(\Delta_{i+j}^n W)^4) \\ &\quad + \sigma^4_{\frac{i+j-1}{n}}((\Delta_{i+j}^n W)^4 - 3n^{-2}) + 3n^{-2}(\sigma^4_{\frac{i+j-1}{n}} - \sigma^4_{\frac{a_j(p)}{n}}).\end{aligned}$$

Plugging in the first term on the right hand side gives a small order in (6.34) due to the growth condition on σ. For the second one, note that

$$\begin{aligned}&\frac{n}{k_n}\mathbb{E}\Big(\sum_{\ell=1}^{I_n(p)} N_\ell \sum_{i=a_\ell(p)}^{b_\ell(p)-1} \frac{2n^2}{k_n^3}\sum_{j=1}^{k_n} \sigma^4_{\frac{i+j-1}{n}}((\Delta_{i+j}^n W)^4 - 3n^{-2})\Big)^2 \\ = \;& \frac{n}{k_n}\sum_{\ell=1}^{I_n(p)} \sum_{i=a_\ell(p)}^{b_\ell(p)-1} \frac{4n^4}{k_n^6}\sum_{j=1}^{k_n} \mathbb{E}\Big(N_\ell^2 \sigma^8_{\frac{i+j-1}{n}}((\Delta_{i+j}^n W)^4 - 3n^{-2})^2\Big) \le Cn^{-1}.\end{aligned}$$

For the third term above we use the standard argument of approximating $\sigma^4_{(i+j-1)/n} - \sigma^4_{a_j(p)/n}$ by an $\mathcal{F}_{a_\ell(p)/n}$-measurable times an increment of V plus conditional independence. For the same reason, the error due to $\tau^2_{i/n} - \tau^2_{a_\ell(p)/n}$ is small, which gives (6.34). Together with (2.12) we have thus shown

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\Big\{(\hat{V}_t - V_t) - \sum_{\ell=1}^{I_n(p)} U_\ell^{n,p} 1_{\{b_l(p)\le \lfloor nt\rfloor - 2k_n\}}\Big\} = 0,$$

$$\lim_{p\to\infty}\limsup_{n\to\infty}\sqrt{\frac{n}{k_n}}\Big\{(\hat{C} - C) - \sum_{\ell=1}^{I_n(p)} \tau^2(\frac{a_\ell(p)}{n}, X_{\frac{a_\ell(p)}{n}}, \sigma^2_{\frac{a_\ell(p)}{n}})U_\ell^{n,p}\Big\} = 0.$$

In order to prove the analogue of (2.7), we use a multivariate version of the result in [23]. The analogues of (2.15) and (2.16) are obtained in exactly the same way as for the one-dimensional result, and it is also quite simple to deduce

$$\lim_{p\to\infty}\limsup_{n\to\infty}\frac{n}{k_n}\sum_{\ell=1}^{I_n(p)} \mathbb{E}[(U_\ell^{n,p})^2 1_{\{b_l(p)\le\lfloor n(t_i\wedge t_j)\rfloor - 2k_n\}}] = \int_0^1 \alpha_s^2 1_{[0,t_i\wedge t_j]}(s)ds,$$

$$\begin{aligned}&\lim_{p\to\infty}\limsup_{n\to\infty}\frac{n}{k_n}\sum_{\ell=1}^{I_n(p)} \mathbb{E}[(U_\ell^{n,p})^2 \tau^2(\frac{a_\ell(p)}{n}, X_{\frac{a_\ell(p)}{n}}, \sigma^2_{\frac{a_\ell(p)}{n}})1_{\{b_l(p)\le\lfloor nt_i\rfloor - 2k_n\}}] \\ &\qquad = \int_0^1 \alpha_s^2\tau^2(s, X_s, \sigma_s^2)1_{[0,t_i]}(s)ds,\end{aligned}$$

$$\lim_{p\to\infty}\limsup_{n\to\infty}\frac{n}{k_n}\sum_{\ell=1}^{I_n(p)} \mathbb{E}[(U_\ell^{n,p})^2 \tau^4(\frac{a_\ell(p)}{n}, X_{\frac{a_\ell(p)}{n}}, \sigma^2_{\frac{a_\ell(p)}{n}})] = \int_0^1 \alpha_s^2\tau^4(s, X_s, \sigma_s^2)ds,$$

for arbitrary t_i, t_j. Proving that we have indeed stable convergence for each fixed p is just another tedious task. □

References

[1] Aït-Sahalia, Y., J. Jacod (2009): Testing for jumps in a discretely observed process. *Ann. Statist.* 37, 184–222.

[2] Aït-Sahalia, Y., J. Jacod (2009): Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* 37, 2202–2244.

[3] Aït-Sahalia, Y., J. Jacod (2011): Testing whether jumps have finite or infinite activity. *Ann. Statist.* 39, 1689–1719.

[4] Aït-Sahalia, Y., R. Kimmel (2007): Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics* 134, 507-551.

[5] Alvares, A., F. Panloup, P. Pontier, N. Savy (2010): Estimation of the instantaneous volatility. *Stat. Inference Stoch. Process.*, to appear.

[6] Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, N. Shephard (2006): A central limit theorem for realised power and bipower variations of continuous semi-martingales. In: Kabanov, Yu., R. Liptser, J. Stoyanov. (Eds.), *From Stochastic Calculus to Mathematical Finance. Festschrift in Honour of A. N. Shiryaev.* Springer-Verlag, Berlin, 33–68.

[7] Barndorff-Nielsen, O., N. Shephard (2004): Power and bipower variation with stochastic volatility and jumps. *Journal of Finanical Econometrics* 2, 1–37.

[8] Barndorff-Nielsen, O., A. Veraart (2009): Stochastic volatility of volatility in continuous time. *Technical report.*

[9] Bandi, F., R. Renò (2008): Nonparametric stochastic volatility. *Technical report.*

[10] Bollerslev, T., H. Zhou (2002): Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *J. Econometrics*, 109, 33–65.

[11] Chernov, M., E. Ghysels (2000): Estimation of stochastic volatility models for the purpose of option pricing. In: Abu-Mostafa, Y., B. LeBaron, A. Lo, A. Weigend (Eds.), *Computational Finance 1999.* MIT Press, Cambridge, 567-581.

[12] Comte, F., V. Genon-Catalot, Y. Rozenholc (2010): Nonparametric estimation for a stochastic volatility model. *Finance Stoch.* 14, 49–80.

[13] Dette, H., M. Podolskij (2008): Testing the parametric form of the volatility in continuous time diffusion models - an empirical process approach. *J. Econometrics* 143, 56–73.

[14] Dette, H., M. Podolskij, M. Vetter (2006): Estimation of integrated volatility in continuous time financial models with applications to goodness-of-fit testing. *Scand. J. Statist.* 33, 259–278.

[15] Genon-Catalot, V., T. Jeantheau, C. Laredo (1999): Parameter estimation for discretely observed stochastic volatility models. *Bernoulli* 5, 855–872.

[16] Gloter, A. (2007): Efficient estimation of drift parameters in stochastic volatility models. *Finance Stoch.* 11, 495-519.

[17] Heston, S. (1993): A closed-form solution for options with stochastic volatility with applications to bonds and currency options. *Review of Financial Studies* 6, 327–343.

[18] Hoffmann, M. (2002): Rate of convergence for parametric estimation in a stochastic volatility model. *Stoch. Proc. Appl.* 97, 147–170.

[19] Jacod, J. (1997): On continuous conditional Gaussian martingales and stable convergence in law. *Seminaire de Probabilites* XXXI, 232–246.

[20] Jacod, J. (2008): Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Proc. Appl.* 118, 517–559.

[21] Jacod, J., C. Klüppelberg, G. Müller (2011): Testing for non-correlation or for a functional relationship between price and volatility jumps. *Working paper.*

[22] Jacod, J., P. Protter (2012): *Discretization of processes.* Springer-Verlag, Berlin.

[23] Jacod, J., A. Shiryaev (2003): *Limit theorems for stochastic processes.* Springer-Verlag, Berlin.

[24] Jacod, J., V. Todorov (2010): Do price and volatility jump together? *Ann. Appl. Probab.* 20, 1425–1469.

[25] Jones, C. (2003): The dynamics of stochastic volatility: evidence from underlying and options markets. *J. Econometrics* 116, 181–224.

[26] Mancini, C. (2009): Non-parametric estimation for models with stochastic diffusion coefficients and jumps. *Scand. J. Statist.* 36, 270–296.

[27] Podolskij, M., M. Rosenbaum (2011): Testing the local volatility assumption: a statistical approach. *Annals of Finance*, to appear.

[28] Podolskij, M., M. Vetter (2010): Understanding limit theorems for semimartingales: a short survey. *Statist. Neerlandica* 64, 329-351.

[29] Renò, R. (2006): Nonparametric estimation of stochastic volatility models. *Econ. Lett.* 90, 390-395.

[30] Vetter, M. (2011): Estimation of correlation for continuous semimartingales. *Technical report.*

[31] Vetter, M., H. Dette (2011): Model checks for the volatility under microstructure noise. *Bernoulli*, to appear.

n	mean	variance	.025	.05	.1	.9	.95	.975
2500	-0.129	0.901	.0098	.0345	.0961	.9215	.9561	.9721
10000	-0.040	1.020	.0152	.0395	.0992	.8974	.9426	.9665
22500	-0.005	0.994	.0180	.0405	.0906	.8993	.9424	.9678
40000	0.024	1.029	.0184	.0428	.0952	.8918	.9446	.9692
52900	0.061	1.033	.0193	.0399	.0911	.8878	.9380	.9672

Table 1: *Mean/variance and simulated quantiles of the infeasible test statistic (4.1) for* $\rho = 0$.

n	mean	variance	.025	.05	.1	.9	.95	.975
2500	-0.132	0.931	.0115	.0358	.0984	.9195	.9527	.9724
10000	-0.048	1.008	.0153	.0400	.0950	.9022	.9457	.9677
22500	-0.126	0.928	.0206	.0463	.1085	.9221	.9579	.9793
40000	0.021	0.995	.0193	.0423	.0945	.8959	.9457	.9717
52900	0.051	1.027	.0187	.0434	.0950	.8907	.9407	.9675

Table 2: *Mean/variance and simulated quantiles of the infeasible test statistic (4.1) for* $\rho = 0.2$.

n	mean	variance	.025	.05	.1	.9	.95	.975
2500	-0.287	0.965	.0526	.0932	.1619	.9572	.9862	.9965
10000	-0.170	1.023	.0449	.0799	.1425	.9325	.9757	.9928
22500	-0.112	1.002	.0404	.0696	.1253	.9271	.9722	.9914
40000	-0.073	1.029	.0401	.0703	.1235	.9203	.9690	.9874
52900	-0.031	1.022	.0368	.0653	.1157	.9154	.9633	.9872

Table 3: *Mean/variance and simulated quantiles of the feasible test statistic from Corollary 2.7 for* $\rho = 0$.

n	mean	variance	.025	.05	.1	.9	.95	.975
2500	-0.295	0.971	.0552	.0963	.1614	.9559	.9864	.9962
10000	-0.176	1.013	.0464	.0808	.1427	.9369	.9770	.9940
22500	-0.226	0.987	.0480	.0840	.1476	.9436	.9776	.9932
40000	-0.075	1.001	.0410	.0673	.1217	.9254	.9713	.9904
52900	-0.040	1.019	.0396	.0677	.1171	.9180	.9663	.9879

Table 4: *Mean/variance and simulated quantiles of the feasible test statistic from Corollary 2.7 for* $\rho = -0.2$.

n	.01	.025	.05	.1	.2
2500	.018	.040	.064	.120	.216
10000	.010	.018	.040	.084	.194
22500	.016	.024	.034	.088	.194
40000	.020	.038	.068	.128	.220
52900	.010	.020	.052	.118	.200

Table 5: *Simulated level of the bootstrap test based on the standardised Kolmogorov-Smirnov statistic* Y_n.

alt	$\gamma = 0$					$\gamma = 1$				
n	.01	.025	.05	.1	.2	.01	.025	.05	.1	.2
2500	.028	.052	.082	.134	.262	.044	.090	.156	.248	.372
10000	.032	.048	.086	.138	.260	.036	.084	.176	.284	.396
22500	.024	.042	.068	.138	.302	.032	.086	.162	.284	.432
40000	.028	.046	.094	.196	.426	.028	.064	.120	.310	.482
52900	.026	.040	.082	.174	.422	.024	.058	.144	.320	.488

Table 6: *Simulated rejection probabilities of the bootstrap test based on the standardised Kolmogorov-Smirnov functional statistic* Y_n *for various alternatives.*