The Annals of Statistics
2014, Vol. 42, No. 5, 2026–2057
DOI: 10.1214/14-AOS1250
© Institute of Mathematical Statistics, 2014

ESTIMATING TIME-CHANGES IN NOISY LÉVY MODELS

BY ADAM D. BULL[1]

University of Cambridge

In quantitative finance, we often model asset prices as a noisy Itô semimartingale. As this model is not identifiable, approximating by a time-changed Lévy process can be useful for generative modelling. We give a new estimate of the normalised volatility or time change in this model, which obtains minimax convergence rates, and is unaffected by infinite-variation jumps. In the semimartingale model, our estimate remains accurate for the normalised volatility, obtaining convergence rates as good as any previously implied in the literature.

1. Introduction. In quantitative finance, we often wish to predict the distribution of future asset prices using historical data; this problem is of interest when pricing options or evaluating investment strategies. From economic considerations, we know that log-prices must be given by a noisy semimartingale; however, this model cannot in general be identified from price data.

We will therefore consider modelling log-prices as a noisy time-changed Lévy process. We note that this model is general enough to describe the salient features of price data—stochastic volatility, jumps and noise—while still being simple enough to identify its parameters from data. It thus serves as a useful approximation to the semimartingale model for generative modelling.

Our goal will be to estimate the normalised volatility or time-change process in this model. Previous estimates have failed to achieve minimax convergence rates when the jumps are of infinite variation, as is suggested by empirical evidence. We will therefore describe a new estimate, which obtains minimax rates, and is unaffected by arbitrary jump activity.

We will further show that in the semimartingale model, our estimate remains accurate for the normalised volatility, obtaining convergence rates as good as any previously implied in the literature. Our estimate thus achieves the best of both worlds: good convergence when the time-changed approximation is accurate, and no penalty when it is not.

We begin by describing the statistical models we will consider. We will suppose we have a single asset whose efficient log-price is given by an *Itô semimartin-*

Received December 2013; revised June 2014.

[1]Supported by EPSRC Grant EP/K000993/1.

MSC2010 subject classifications. Primary 62P20; secondary 62G08, 62G20, 62G35.

Key words and phrases. Itô semimartingale, Lévy process, microstructure noise, volatility, time-change.

gale,

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sqrt{c_s}\,dB_s + \int_0^t \int_{\mathbb{R}} x\big(\mu(dx, ds) - 1_{|x|<1}\nu_s(dx)\,ds\big), \tag{1}$$

where $b_t \in \mathbb{R}$ is a drift process, $c_t > 0$ a volatility process, ν_t a jump measure process, $\mu(dx, dt)$ a Poisson random measure with intensity $\nu_t(dx)\,dt$, and the above decomposition holds with respect to a filtration $\mathcal{F}_t$. [We refer to Jacod and Shiryaev (2003), for definitions.]

We note that the assumption (1) is extremely common in quantitative finance, and is motivated by economic no-arbitrage arguments, as in Delbaen and Schachermayer (1994). Model (1) reproduces common features of price data, such as *stochastic volatility*, given by the dependence of the characteristics (b_t, c_t, ν_t) on time, and *jumps*, given by the presence of the jump measure process ν_t.

To fit this model to price data, however, it is widely accepted that we must also account for a third feature, known as *microstructure noise*. The quoted price of assets in general can diverge from the efficient market price, due to economic artefacts such as the bid-ask spread, tick sizes, transaction costs, and communication delays. Indeed, empirical studies confirm that high-frequency price data is too volatile to be explained solely by an efficient price process [Andersen et al. (2000), Hansen and Lunde (2006), Mykland and Zhang (2005)].

A popular model for microstructure noise is to assume that the log-prices are observed under zero-mean errors. We thus consider observations

$$Y_j = X_{Tj/n} + \varepsilon_j, \qquad j = 0, \dots, n-1, \tag{2}$$

over a time interval $[0, T]$, and with errors ε_j satisfying $\mathbb{E}[\varepsilon_j | \mathcal{F}_{Tj/n}] = 0$. We refer to Jacod et al. (2009) for a discussion of this model.

Unfortunately, the observations Y_j are insufficient to identify the parameters of model (1). Even given noiseless observations, letting the time horizon $T \to \infty$, and the step size $T/n \to 0$, we cannot in general identify the drift process b_t, or jump measure process ν_t.

In the following, we will therefore also consider a *time-changed Lévy process* model. Here, we instead suppose the log-price

$$X_t = L_{R_t}, \tag{3}$$

for a Lévy process

$$L_t = L_0 + bt + \sqrt{c}B_t + \int_0^t \int_{\mathbb{R}} x\big(\mu(dx, ds) - 1_{|x|<1}\nu(dx)\,ds\big),$$

with drift $b \in \mathbb{R}$, volatility $c > 0$, jump measure ν, and Poisson random measure $\mu(dx, dt)$ with intensity $\nu(dx)\,dt$, and a time-change process

$$R_t = \int_0^t r_s\,ds,$$

given by a rate process $r_t > 0$.

Model (3) was popularised by Carr and Wu (2004), and its applications also discussed by Cont and Tankov (2004). Intuitively, this model describes prices which move faster or slower according to an activity rate r_t; this rate can be thought of as the cumulative effect of factors such as trading activity and volume, investor liquidity, and general economic uncertainty.

Formally, the time-changed model (3) is the subset of the semimartingale model (1) which satisfies the separability condition

$$b_t = br_t, \qquad c_t = cr_t, \qquad \nu_t = \nu r_t. \tag{4}$$

This condition requires, for example, that the jump measure ν_t be governed by the rate process r_t, and contain no idiosyncratic jump component.

Since these parameters are defined only up to a multiplicative constant, we must also choose a normalisation for r_t. In the following, for simplicity we will set r_t to integrate to one (although we will also discuss alternative normalisations). Equivalently, using (4) we may define

$$r_t = \frac{c_t}{\int_0^t c_s \, ds}; \tag{5}$$

we note that this definition is then also meaningful for the semimartingale model (1).

The separability condition (4) can be thought of as similar to the additivity condition in an additive model. We take a fully nonparametric model, which is difficult to fit, and restrict it to a lower-dimensional one, which is less so. As our smaller model (3) reproduces the salient features of price data—stochastic volatility, jumps and noise—it can potentially offer a good approximation to the full model (1).

This approximation can be useful in a variety of settings. If we wish to predict the distribution of future asset prices, for example, to price options or evaluate investment strategies, we must fit a generative model to the data. We already know we cannot fit the full model (1), as we cannot identify its parameters b_t and ν_t. As we will see below, the parameters of model (3) can all be identified from price data; it may thus be used either directly as a generative model, or as a starting point to identify suitable parametric alternatives [Carr and Wu (2004), Cont and Tankov (2004)].

To fit model (3) to data, we must estimate the parameters b, c, ν and r_t. If the time horizon $T \to \infty$, and the step size $T/n \to 0$, the drift b and volatility c can be estimated using standard techniques. Estimation of the Lévy measure ν, while more involved, has also been considered by several authors [Belomestny (2011), Belomestny and Panov (2013), Figueroa-López (2009, 2011)], and extensions of Figueroa-López's approach to include noise are possible, as in Vetter (2014).

In the following, we will focus specifically on estimation of the rate process r_t. We first note that some of the factors contributing to this process, in particular trading activity and volume, can be observed directly. While such side information may be useful in practice, we can expect that not all such factors are observable, and the relationship between observable factors and efficient prices may be unclear, especially after accounting for microstructure noise.

In the following, we will therefore restrict ourselves to estimating r_t directly from price data. While previous work has provided such estimates in a variety of settings [Figueroa-López (2012), Rosenbaum and Tankov (2011), Winkel (2001), Woerner (2007)], these authors have not considered our setting (2) and (3). Even accounting for microstructure noise, we cannot apply their methods here to obtain minimax rates of convergence.

An alternative route to estimating r_t is to first use identification (4), and then estimate the volatility c_t in the semimartingale model (1). Many authors have described approaches for this problem, under various assumptions on the jump measure process ν_t.

If there are no jumps present, the integrated volatility $\int_0^1 c_s\,ds$ can be recovered using multiscale estimators [Zhang (2006), Zhang, Mykland and Aït-Sahalia (2005)], realised kernels [Barndorff-Nielsen et al. (2008)], or pre-averaging [Jacod et al. (2009), Podolskij and Vetter (2009a)]. The spot volatility c_t can likewise be recovered using kernel estimators [Kristensen (2010), Mancini, Mattiussi and Reno (2014)], Fourier series [Munk and Schmidt-Hieber (2010a), Reiss (2011)], or wavelets [Hoffmann, Munk and Schmidt-Hieber (2012)].

In each case, these methods can achieve minimax convergence rates, equivalent for fixed T to observing c_t under Gaussian white noise of size $n^{-1/4}$. In fact, it can be shown this link is a formal statistical equivalence [Reiss (2011)].

When jumps are present, however, we must account for them before estimating c_t. Methods for doing so include jump thresholding [Mancini (2001, 2009, Fan and Wang (2007), Jing, Kong and Liu (2011)], bipower variation [Barndorff-Nielsen and Shephard (2004), Hautsch and Podolskij (2013), Podolskij and Vetter (2009b, 2009a)], and characteristic functions [Jacod and Reiss (2014), Jacod and Todorov (2014), Todorov and Tauchen (2012)].

Unfortunately, if the jumps are of infinite variation, in general these methods can no longer achieve the same convergence rates. Even given noiseless observations of the efficient prices, it is known that the minimax convergence rate for c_t suffers, unless we assume the infinite-variation part is a scaled β-stable process [Jacod and Reiss (2014), Jacod and Todorov (2014)].

Nonetheless, empirical evidence suggests that price data does indeed contain infinite-variation jumps [Aït-Sahalia and Jacod (2009), Jing, Kong and Liu (2011)]. In the following, we will therefore construct a novel estimate of the rate process r_t. We will show that our estimate achieves good rates of convergence in both models, and in the time-changed model is unaffected by arbitrary jump activity.

Our estimate will be constructed in three stages. We will first obtain pre-averaged estimates of price increments, and estimates of the microstructure noise, as in Jacod et al. (2009) or Podolskij and Vetter (2009a).

We will then construct local estimates of the spot volatility, derived by estimating the characteristic function of the price process. While our approach will be similar to ones considered by previous authors [Jacod and Reiss (2014), Jacod and Todorov (2014), Todorov and Tauchen (2012)], the precise construction necessary to obtain minimax rates will be new.

Finally, we will smooth our local estimates of the volatility, using standard tools from nonparametric regression. While many such approaches are possible, we will use local polynomials, as described, for example, by Tsybakov (2009). We will also discuss how the various parameters required can be chosen automatically from the data.

We will then prove results on the convergence rates of our estimates. We note that our results will apply in two settings: a standard nonparametric setting, where the characteristics of X_t are assumed fixed and smooth; and a setting more natural in quantitative finance, where these characteristics are themselves described by Itô semimartingales with locally bounded characteristics.

For simplicity, our results will focus on the high-frequency case, where the fixed time horizon $T = 1$. We note, however, that similar results can also be proved when $T \to \infty$, provided that the step size $T/n \to 0$.

In the time-changed Lévy model (3), we will then show that our procedure estimates r_t with minimax convergence rates, equal to those in the Gaussian white noise model with noise level $n^{-1/4}$. Our results will hold under arbitrary jump activity, and without knowledge of the Lévy parameters b, c and ν.

In the general semimartingale model (1), we will show that our procedure continues to estimate r_t. While lower bounds for this problem are still unknown, the convergence rates we will obtain are as good as any implied by previous work. Our estimate thus achieves the best of both worlds: good convergence when the time-changed approximation is accurate, and no penalty when it is not.

In Section 2, we will give the construction of our estimates, and in Section 3, describe the specific assumptions we consider. In Section 4, we will then state our results on convergence rates, and in Section 5, give proofs.

2. Local characteristic-function estimates. In this section, we will define our estimates of the volatility and rate processes. As described in the Introduction, our estimates are constructed in three stages: pre-averaging, spot volatility estimation and smoothing.

We begin with the pre-averaging step, and proceed using the construction of Reiss (2011). We must first subdivide the time interval [0, 1] into a number n_0 of equal bins. To define n_0, we choose $n_1, n_2 \in \mathbb{N}$ in terms of n, so that

$$n_m \sim h_m^{-1} n^{(2m-1)/8}, \qquad m = 1, 2,$$

for bandwidths $h_1, h_2 > 0$, and set $n_0 = n_1 n_2$.

We then divide [0, 1] into n_0 bins, and compute on each one a pre-averaged estimate $\widehat{X}_k$ of the increments of X_t. We will compute $\widehat{X}_k$ by integrating the observed increments against a scaling function $\Phi_n(t)$; we define

$$\Phi_n(t) = \sqrt{n_0}\Phi(n_0 t), \qquad \Phi(t) = 2\sin(2\pi t).$$

The specific choice of scaling function Φ_n is motivated by Reiss (2011), who shows that in a Gaussian setting, functions of this form are most efficient at extracting information from noisy data. We note that our choice of Φ includes a full period of the sine wave in each bin, rather than a half period; this choice allows us to ensure that the pre-averaged increments are approximately symmetrically distributed, a property we will require when modelling the behaviour of infinite-variation jumps.

We may now define the pre-averaged increments $\widehat{X}_k$. For $k = 0, \dots, n_0 - 1$, define index sets $J_k = (n/n_0)[k, k+1) \cap \mathbb{Z}$, and let

$$\widehat{X}_k = \sum_{j, j+1 \in J_k} p_j (Y_{j+1} - Y_j), \qquad p_j = \Phi_n(j/n).$$

The estimate $\widehat{X}_k$ thus averages the observed increments of X_t over the time interval $[k, k+1)/n_0$. We can also define an estimate $\widehat{\sigma}_k^2$ of the microstructure noise over the interval. We set

$$\widehat{\sigma}_k^2 = \frac{n_0}{2n} \sum_{j, j+1 \in J_k} (Y_{j+1} - Y_j)^2,$$

proportional to the realised quadratic variation of the observations.

We next describe our spot volatility estimation step. We will subdivide [0, 1] into n_2 larger bins, and on each one, construct an estimate $\widehat{c}_l(u)$ of the volatility c_t. While our approach will be based on local characteristic function estimates, similar to those considered by previous authors [Jacod and Reiss (2014), Jacod and Todorov (2014), Todorov and Tauchen (2012)], the precise construction necessary to obtain minimax rates will be new.

For $l = 0, \dots, n_2 - 1$, we define index sets $K_l = n_1[l, l+1) \cap \mathbb{Z}$, and local estimates $\widehat{\varphi}_l(u)$ of the characteristic function of increments of X_t, given by

$$\widehat{\varphi}_l(u) = \frac{1}{n_1} \sum_{k \in K_l} \cos(u \widehat{X}_k).$$

We note that $\widehat{\varphi}_l(u)$ thus averages the cosines of the $\widehat{X}_k$ over the time interval $[l, l+1)/n_2$. We can also define an estimate $\widehat{\psi}_l(u)$ of the corresponding contribution of the microstructure noise; we set

$$\widehat{\psi}_l(u) = \frac{1}{n_1} \sum_{k \in K_l} \exp(-\kappa u^2 \widehat{\sigma}_k^2), \qquad \kappa = \frac{4\pi^2 n_0^2}{n}.$$

If the log-prices X_t and noises ε_j were Gaussian, then by considering their characteristic functions, we would expect

$$\widehat{\varphi}_l(u) \approx \exp\big(-(c_{l/n_2} + \kappa\sigma^2_{l/n_2})u^2\big), \qquad \widehat{\psi}_l(u) \approx \exp\big(-\kappa\sigma^2_{l/n_2}u^2\big).$$

We could then rearrange these quantities to obtain an estimate

$$-\frac{1}{u^2}\log\left|\frac{\widehat{\varphi}_l(u)}{\widehat{\psi}_l(u)}\right|$$

of the volatility c_{l/n_2}.

In fact, such an estimate would be biased. We can, however, provide bias-corrected estimates $\widehat{c}_l(u)$ of c_{l/n_2}; we define

$$\widehat{c}_l(u) = -\frac{1}{u^2}\left(\log\left|\frac{\widehat{\varphi}_l(u)}{\widehat{\psi}_l(u)}\right| + \frac{\widehat{\tau}^2_l(u)}{2}\right),$$

where the bias-correction term

$$\widehat{\tau}^2_l(u) = \frac{1}{n_1}\left(\frac{1+\widehat{\varphi}_l(2u)}{2\widehat{\varphi}_l(u)^2} - 1\right).$$

We have thus defined an estimate $\widehat{c}_l(u)$ of the spot volatility. The advantage of this procedure over other such estimates is that it naturally accounts for the presence of jump activity: we will show that, for general semimartingales, $\widehat{c}_l(u)$ is an asymptotically-unbiased estimate of the quantity $c_{l/n_2}(u)$, given by the adjusted volatility process

$$c_t(u) = c_t + \frac{1}{n_0u^2}\int_0^1\int_{\mathbb{R}}\big(1 - \cos\big(\sqrt{n_0}\Phi(w)ux\big)\big)\nu_t(dx)\,dw.$$

The process $c_t(u)$ thus includes both the volatility c_t and a term depending on the jump measure ν_t. As $n \to \infty$, the term involving ν_t vanishes; however, when the jump activity β is large, this term will not vanish fast enough to be negligible, and so we must consider it explicitly. Crucially in the time-changed model (3), we have that both terms enter $c_t(u)$ linearly, and so $c_t(u)$ is proportional to the rate process r_t.

In either model, since r_t integrates to one, we may estimate it by normalising our estimates of $c_t(u)$. However, to estimate r_t optimally we will not be able to use the preliminary estimates $\widehat{c}_l(u)$ directly, as their variance is too large. First, we must smooth them, using standard tools from nonparametric regression.

While many such approaches are possible, in the following we will use a local polynomial estimate of $c_t(u)$, as described, for example, by Tsybakov (2009). To define our estimate, fix a nonnegative kernel function $K : \mathbb{R} \to \mathbb{R}$, supported on $[-1, 1]$, and satisfying $\int_{\mathbb{R}} K(t)\,dt = 1$. Also fix an order $N \in \mathbb{N}$, and bandwidth $h > 0$. Then let $\widetilde{c}_t(u)$ denote a local polynomial estimate of $c_t(u)$ of degree $N - 1$, using the observations $\widehat{c}_l(u)$, kernel K, and bandwidth h.

In other words, let

$$\widetilde{c}_t(u) = \sum_{l=0}^{n_2-1} W_{n,l}(t)\widehat{c}_l(u),$$

where the weight functions $W_{n,l}(t)$ are given by

$$W_{n,l}(t) = \frac{1}{n_2 h} K\big(\lambda_{n,l}(t)\big) U(0)^T V_n(t)^{-1} U\big(\lambda_{n,l}(t)\big),$$

for the terms

$$\lambda_{n,l}(t) = \frac{1}{h}\left(t - \frac{l}{n_2}\right),$$

$$U(\lambda) = \left(1, \lambda, \dots, \frac{\lambda^{N-1}}{(N-1)!}\right)^T,$$

$$V_n(t) = \frac{1}{n_2 h} \sum_{l=0}^{n_2-1} K\big(\lambda_{n,l}(t)\big) U\big(\lambda_{n,l}(t)\big) U\big(\lambda_{n,l}(t)\big)^T.$$

The constant $N \in \mathbb{N}$ serves as an upper bound for the smoothness we expect of the volatility process c_t, and other processes related to X_t. We include here the case where N is large, so that our estimate can match-known nonparametric lower bounds for a wide range of smoothness.

In practice, however, we may believe that these processes are Itô semimartingales, as in most common financial models. We will see later that in this case it suffices to take $N = 1$; the above estimate then reduces to the Nadaraya–Watson kernel estimate, with weights $W_{n,l}(t)$ given by

$$W_{n,l}(t) = \frac{K(\lambda_{n,l}(t))}{\sum_{l=0}^{n_2-1} K(\lambda_{n,l}(t))}.$$

In either case, we then have an estimate $\widetilde{c}_t(u)$ of the volatility c_t. To estimate the normalised volatility or rate process r_t, we likewise define the normalised estimate

$$\widetilde{r}_t(u) = \frac{\widetilde{c}_t(u)}{(1/n_2)\sum_{m=0}^{n_2-1} \widehat{c}_m(u)} = \sum_{l=0}^{n_2-1} W_{n,l}(t)\widehat{r}_l(u),$$

where

$$\widehat{r}_l(u) = \frac{\widehat{c}_l(u)}{(1/n_2)\sum_{m=0}^{n_2-1} \widehat{c}_m(u)}.$$

In the following sections, we will prove results on the theoretical performance of our estimates $\widetilde{c}_t(u)$ and $\widetilde{r}_t(u)$. We first, however, briefly discuss their implementation. In particular, we note that the above estimates require the choice of a number of parameters: the kernel K, order N, frequency u, and bandwidths h_1, h_2 and h.

In general, good performance in nonparametric regression can be obtained with a range of kernels K; popular choices include the uniform, Epanechnikov and biweight kernels, given by Beta(k,k) densities for $k = 1, 2$ and 3, respectively. If we believe the volatility c_t and other characteristic processes are given by Itô semimartingales, then as noted above, we may also take $N = 1$.

The remaining parameters u, h_1, h_2 and h are more important. We will show in the following that the variances of our estimates $\widehat{c}_l(u)$ depend crucially on the choice of the frequency u and bandwidths h_1, h_2. The correct choice of the bandwidth h is likewise known to be crucial generally in nonparametric regression.

To select these parameters, we can borrow methods from nonparametric statistics. While many such approaches are available, we will briefly mention the heuristic of generalised cross-validation, a popular method for choosing the bandwidth h in nonparametric regression [Golub, Heath and Wahba (1979)].

The GCV criterion

$$\mathrm{GCV}(u, h_1, h_2, h) = \frac{(1/n_2)\sum_{l=0}^{n_2-1}(\widetilde{r}_{l/n_2}(u) - \widehat{r}_l(u))^2}{((1/n_2)\sum_{l=0}^{n_2-1} W_{n,l}(l/n_2))^2}$$

provides an estimate of the L^2 error in $\widetilde{r}_t(u)$. We can then choose u, h_1, h_2 and h to minimise this criterion, using any standard global optimisation algorithm.

Simple tests on simulated data show that minimising this criterion provides sensible choices of the parameters, for both estimates $\widetilde{c}_t(u)$ and $\widetilde{r}_t(u)$. [We note that it is inadvisable to apply GCV to $\widetilde{c}_t(u)$ directly, as the criterion then favours parameters which shrink the estimate to zero.]

We have thus described new estimates of the volatility c_t, and normalised volatility or rate process r_t; however, we have yet to consider their performance. In the following sections, we will show that, for suitable choices of the parameters, these estimates can obtain good rates of convergence over both the semimartingale model (1) and time-changed Lévy model (3).

3. Semimartingale and Lévy models. In this section, we will describe the assumptions we make on our data. Our assumptions will be satisfied by common models in both nonparametric statistics and quantitative finance. Under these assumptions, we may then proceed to show that our estimates $\widetilde{c}_t(u)$ and $\widetilde{r}_t(u)$ achieve good rates of convergence.

We first assume that the log-prices X_t are generated under the general Itô semimartingale model (1), and our observations Y_j come from the microstructure noise model (2), with fixed time horizon $T = 1$. For simplicity, we do not consider further other choices of T, but we note that similar results can also be proved when $T \to \infty$, $T/n \to 0$.

Our assumptions will then be stated in terms of a filtration $\mathcal{F}_t$, $t \in [0, 1]$, with respect to which the semimartingale decomposition (1) and zero-mean condition of (2) hold. As in Jacod et al. (2009), to allow for the modelling of microstructure noise, we will not assume that the filtration $\mathcal{F}_t$ is right-continuous. Instead, we will require that the semimartingale decomposition (1) is also valid with respect to the filtration $\mathcal{F}_t^+ = \bigcap_{s>t} \mathcal{F}_s$, and that the noises ε_j are $\mathcal{F}_{j/n}^+$-measurable.

We then let $\mathcal{S}$ denote the class of probability measures $\mathbb{P}$ satisfying the above conditions, on some filtered measurable space $(\Omega, \mathcal{F}, \mathcal{F}_t)$. We will also make some further assumptions on the characteristics b_t, c_t and ν_t, and errors ε_j.

We begin by defining a smoothness assumption on $\mathcal{F}_t$-adapted processes. We will require in the following that the volatility c_t, and other characteristic processes of the log-prices X_t and noises ε_j, satisfy this assumption with high probability.

DEFINITION 1. Let $S \in [0, 1]$ be an $\mathcal{F}_t$-stopping time, $\alpha \geq \frac{1}{2}$, $D > 0$, and set $\alpha_0 = 1 \wedge \alpha$. We define $\mathcal{I}^\alpha(D, S)$ to be the class of $\mathcal{F}_t$-adapted complex-valued processes Z_t, for which the stopped process $Z_{t\wedge S}$ satisfies:

(i) $|Z_{t\wedge S}| \leq D, t \in [0, 1]$;
(ii) $\mathbb{E}[|Z_{s\wedge S} - Z_{t\wedge S}|^2 | \mathcal{F}_t^+] \leq D^2(s-t)^{2\alpha_0}, 0 \leq t \leq s \leq 1$;
(iii) if $\alpha > 1$, then letting m denote the largest integer smaller than α, $Z_{t\wedge S}$ has mth real derivative $Z_{t\wedge S}^{(m)}$ satisfying

$$\mathbb{E}\big[\big|Z_{s\wedge S}^{(m)} - Z_{t\wedge S}^{(m)}\big|^2 | \mathcal{F}_t^+\big] \leq D^2(s-t)^{2(\alpha-m)}, \qquad 0 \leq t \leq s \leq 1.$$

The classes $\mathcal{I}^\alpha(D, S)$ thus contain all processes Z_t which, when stopped by S, are bounded and smooth in quadratic mean. We note that these classes describe a variety of processes. Firstly, the classes $\mathcal{I}^\alpha(D, 1)$ contain all processes which are almost-surely α-Hölder, with constant D. More generally, the following lemma shows that the classes $\mathcal{I}^{1/2}(D, S)$ can describe all càglàd Itô semimartingales with locally bounded characteristics.

LEMMA 1. *Let Z_t be a càglàd Itô semimartingale, having decomposition*

$$Z_t = Z_0 + \int_0^{t^-} b_{Z,s}\, ds + \int_0^{t^-} \sqrt{c_{Z,s}}\, dB_{Z,s} + \int_0^{t^-} \int_{\mathbb{R}} x\big(\mu_Z(dx, ds) - 1_{|x|<1} \nu_{Z,s}(dx)\, ds\big)$$

with respect to both filtrations $\mathcal{F}_t$ and $\mathcal{F}_t^+$. Suppose the processes $b_{Z,t}$ and $c_{Z,t}$ are locally bounded, as are the processes $\int_{|x|<R} x^2 \nu_{Z,t}(dx)$ for all $R > 0$. Then for each $D > 0$, there exists an event $\Omega_0 \in \mathcal{F}_0$, and $\mathcal{F}_t$-stopping time S, such that on Ω_0, $Z_t \in \mathcal{I}^{1/2}(D, S)$, with $\mathbb{P}(\Omega_0 \cap \{S = 1\}) \to 1$ as $D \to \infty$.

We now define our assumptions on the observations Y_j. Essentially, our results will be proved in models where with high probability, the drift b_t is bounded, and the stochastic volatility c_t, jump process ν_t and noise variance σ_t^2 are bounded and smooth.

DEFINITION 2. Let $\alpha \geq \frac{1}{2}$, $\beta \in [0, 2]$, $\gamma \in [0, 1]$ and $0 < C \leq D$. We define $\mathcal{S}_\gamma^{\alpha,\beta}(C, D)$ to be the class of probability measures $\mathbb{P} \in \mathcal{S}$ which satisfy the following conditions, on an event $\Omega_0 \in \mathcal{F}_0$, and for a stopping time $S \in [0, 1]$, with $\mathbb{P}(\Omega_0 \cap \{S = 1\}) \geq 1 - \gamma$:

(i) The noises ε_j have variance

$$\mathbb{E}[\varepsilon_j^2 | \mathcal{F}_{j/n}] = \sigma_{j/n}^2,$$

for a latent process $\sigma_t^2 \in \mathcal{I}^\alpha(D, S)$ and have bounded fourth moment

$$\mathbb{E}[\varepsilon_j^4 | \mathcal{F}_{j/n}] \leq D^2, \qquad j/n \leq S.$$

(ii) The drift process b_t is bounded,

$$|b_t| \leq D, \qquad t \in [0, S].$$

(iii) The volatility process $c_t \in \mathcal{I}^\alpha(D, S)$, and is also bounded below

$$c_t \geq C, \qquad t \in [0, S].$$

(iv) The jump activity is of index at most β,

$$\int_{\mathbb{R}} (1 \wedge |x|^\beta) \nu_t(dx) \leq D, \qquad t \in [0, S],$$

and if $\beta > 1$, for any measurable function $f : \mathbb{R} \to \mathbb{C}$ with

$$|f(x)| \leq (1 \wedge x^2),$$

we have

$$\int_{\mathbb{R}} f(x) \nu_t(dx) \in \mathcal{I}^\alpha(D, S).$$

Also define $\mathcal{S}^{\alpha,\beta}(C, D) = \mathcal{S}_0^{\alpha,\beta}(C, D)$, and let $\mathcal{S}^{\alpha,\beta}$ denote the class of probability measures $\mathbb{P}$ which lie in some $\mathcal{S}_\gamma^{\alpha,\beta}(C, D)$ for each $0 < C \leq D$, with $\gamma \to 0$ as $C \to 0$, $D \to \infty$.

In the following, our theoretical results will be first proved for the classes $\mathcal{S}^{\alpha,\beta}(C, D)$, where the characteristics of the log-price X_t and noises ε_j are almost-surely bounded and smooth. These classes will be the most convenient for our analysis and will allow us to draw comparisons with previous nonparametric results in the literature.

We note that these classes impose quite strong conditions on our process X_t; in particular, they require the volatility c_t to be bounded away from zero. However, we will also generalise our results to the larger classes $\mathcal{S}^{\alpha,\beta}$, which only require these conditions to hold locally; in particular, they only impose the weaker bound that $c_t > 0$ almost-surely.

The parameters α and β govern two different smoothness properties of X_t. The parameter α measures the smoothness of the characteristics of X_t and the ε_j over time: if X_t is a Lévy process, and the ε_j have constant variance, the above conditions can hold for any value of α. In contrast, the parameter β governs the jump activity of X_t and thus also the smoothness of its sample paths.

We note that in typical semimartingale models, the parameter $\alpha = \frac{1}{2}$; we include here the case $\alpha > \frac{1}{2}$ to allow for comparison with previous results in nonparametrics. We also note that since the smoothness α is measured in square mean, the case $\alpha = \frac{1}{2}$ allows for jumps in the volatility or in other characteristics; it thus allows the characteristics to depend smoothly on X_t, for any level β of jump activity. More generally, Lemma 1 shows that the classes $\mathcal{S}^{1/2,\beta}$ contain most common models for financial processes.

As some of our results will be specific to the time-changed model (3), we additionally define submodels describing this case. We note that our definition includes a choice of normalisation; as in (5), we assume the rate process r_t integrates to one.

DEFINITION 3. Let $\mathcal{T}$ denote the class of probability measures $\mathbb{P} \in \mathcal{S}$ satisfying (4), for a drift $b \in \mathbb{R}$, volatility $c > 0$, Lévy measure ν, and rate process $r_t > 0$ given by (5). Also define the models $\mathcal{T}^\alpha(C, D) = \mathcal{S}^{\alpha,2}(C, D) \cap \mathcal{T}$ and $\mathcal{T}^\alpha = \mathcal{S}^{\alpha,2} \cap \mathcal{T}$.

This choice of normalisation is most convenient for our results, but we note that others are also possible; for example, we might prefer the deterministic normalisation $\mathbb{E}[\int_0^T r_s\, ds] = T$. If we made an ergodicity assumption on the process r_t, as in Figueroa-López (2009), then for suitable $T \to \infty$, $T/n \to 0$, we would have that $\int_0^T r_s\, ds$ is close to $\mathbb{E}[\int_0^T r_s\, ds]$. The two normalisations would then also be close, and our arguments would apply equally in either setting.

In the following, for simplicity, we will concentrate on Definition 3. We then have that in particular, the class $\mathcal{T}^{1/2}$ covers most common financial models for time-changed Lévy processes. With these definitions, we are now ready to give our results on the performance of our estimates.

4. Convergence results. In this section, we will show that our estimates $\widetilde{c}_t(u)$ and $\widetilde{r}_t(u)$ have good rates of convergence, in both the general semimartingale models $\mathcal{S}^{\alpha,\beta}$ and time-changed Lévy models $\mathcal{T}^\alpha$. In particular, we will establish that in $\mathcal{T}^\alpha$, the time-change r_t can be recovered at minimax rates under arbitrary jump activity.

We first define some additional processes which will be relevant to our results. We set

$$\varphi_t(u) = \exp\bigl(-c_t(u)u^2\bigr)\psi_t(u), \qquad \psi_t(u) = \exp\bigl(-\kappa\sigma_t^2 u^2\bigr),$$

processes we will show describe the means of the estimates $\widehat{\varphi}_l(u)$ and $\widehat{\psi}_l(u)$. We also set

$$\rho_t^2(u) = \tfrac{1}{2}\big(1 + \varphi_t(2u)\big) - \varphi_t^2(u), \qquad \tau_t^2(u) = \rho_t^2(u)/n_1\varphi_t^2(u),$$

processes we will show describe the variances of the estimates $\widehat{\varphi}_l(u)$ and $\widehat{c}_l(u)$.

We now begin with a result on the accuracy of the preliminary estimates $\widehat{c}_l(u)$. At this stage, our results will be proved solely in the bounded semimartingale model $\mathcal{S}^{\alpha,\beta}(C, D)$; we will return later to the consequences for our other models.

We can establish that, on events with high probability, our preliminary estimates $\widehat{c}_l(u)$ have asymptotic mean $c_{l/n_2}(u)$ and variance $\tau^2_{l/n_2}(u)$. We can further show that the errors in these statements are of order $n^{-\alpha_1}$ and $n^{-\alpha_2}$, respectively, where the rates

$$\alpha_1 = \frac{1}{4} \wedge \frac{3\alpha}{8}, \qquad \alpha_2 = \frac{\alpha_1}{2} + \frac{1}{16}.$$

THEOREM 1. *Fix $u, h_1, h_2 > 0$, $\alpha \geq \frac{1}{2}$, $\beta \in [0, 2]$, $0 < C \leq D$, and suppose $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$. Then the local volatility estimates $\widehat{c}_l(u)$ are $\mathcal{F}_{(l+1)/n_2}$-measurable, and we have events $E_l \in \mathcal{F}_{(l+1)/n_2}$, satisfying*

$$\mathbb{P}\big(E_l^c|\mathcal{F}_{l/n_2}\big) \leq \exp\big(-An^{1/8}\big)$$

for a constant $A > 0$, on which

$$\mathbb{E}\big[\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)1(E_l)|\mathcal{F}_{l/n_2}\big] = O\big(n^{-\alpha_1}\big),$$

$$\mathbb{E}\big[\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)^2 1(E_l)|\mathcal{F}_{l/n_2}\big] = \tau^2_{l/n_2}(u) + O\big(n^{-\alpha_2}\big).$$

Furthermore, these results are uniform over $l = 0, \ldots, n_2 - 1$ and $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$.

We thus have that the estimates $\widehat{c}_l(u)$ behave roughly like $n^{3/8}$ observations of the adjusted volatility process $c_t(u)$, under errors with variance $n^{-1/8}$. In other words, we obtain an accuracy like observing the process $c_t(u)$ under $n^{-1/4}$ white noise. While our estimates $\widehat{c}_l(u)$ also include an additional bias term, and are accurate only on a set of high probability, we will nonetheless see that they are good enough to accurately recover the volatility c_t or time-change r_t.

We now establish that our regression estimate $\widetilde{c}_t(u)$ is a good estimate of the adjusted volatility $c_t(u)$. We will measure the accuracy of our estimates both pointwise and in the L^2-norm

$$\|f\|_2^2 = \int_0^1 f(t)^2\,dt.$$

In these metrics, we will show that $c_t(u)$ can be recovered at the rate $n^{-\alpha_3}$, where

$$\alpha_3 = \frac{\alpha}{2(2\alpha + 1)}$$

is the standard minimax rate for recovering a function of smoothness α under $n^{-1/4}$ white noise.

THEOREM 2. *Fix a kernel K as in Section 2, $N \in \mathbb{N}$, $u \in \mathbb{R}$, $h_1, h_2 > 0$, $\alpha \in [\frac{1}{2}, N]$, $\beta \in [0, 2]$, $0 < C \le D$, let $h \sim n^{-1/2(2\alpha+1)}$, and suppose $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$. We then have an event E, satisfying*

$$\mathbb{P}(E^c | \mathcal{F}_0) \le \exp(-An^{1/8})$$

for a constant $A > 0$, on which

$$\mathbb{E}\big[|\widetilde{c}_t(u) - c_t(u)|^2 1(E) | \mathcal{F}_0\big]^{1/2} = O(n^{-\alpha_3}),$$

uniformly in $t \in [0, 1]$, and

$$\mathbb{E}\big[\|\widetilde{c}(u) - c(u)\|_2^2 1(E) | \mathcal{F}_0\big]^{1/2} = O(n^{-\alpha_3}).$$

Furthermore, these results are uniform over $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$.

We thus have that the regression estimates $\widetilde{c}_t(u)$ accurately recover $c_t(u)$ in the model $\mathcal{S}^{\alpha,\beta}(C, D)$. It remains to deduce consequences for the volatility c_t and time-change r_t in the more general models $\mathcal{S}^{\alpha,\beta}$ and $\mathcal{T}^{\alpha}$. In $\mathcal{S}^{\alpha,\beta}$, we will obtain the rate $n^{-\alpha_4}$, where

$$\alpha_4 = \alpha_3 \wedge \frac{2 - \beta}{4}$$

depends also on the jump activity β of the log-price X_t. When estimating r_t in $\mathcal{T}^{\alpha}$, however, we will retain the convergence rate $n^{-\alpha_3}$, even under arbitrary jump activity.

COROLLARY 1. *Let the parameters $K, N, u, h_1, h_2, \alpha, \beta, C, D$ and h be as in Theorem 2.*

(i) *If $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}$, the estimates $\widetilde{c}_t(u)$ and $\widetilde{r}_t(u)$ satisfy*

$$|\widetilde{c}_t(u) - c_t|, |\widetilde{r}_t(u) - r_t| = O_p(n^{-\alpha_4}),$$

uniformly in $t \in [0, 1]$, and

$$\|\widetilde{c}(u) - c\|_2, \|\widetilde{r}(u) - r\|_2 = O_p(n^{-\alpha_4}).$$

Furthermore, these results are uniform over $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$.

(ii) *If also $\mathbb{P} \in \mathcal{T}$, the results for $\widetilde{r}_t(u)$ hold with improved convergence rate $n^{-\alpha_3}$.*

We note that convergence does depend on the choice of parameters K, N, u, h_1, h_2 and h, and in particular requires the bandwidth h to be chosen as in Theorem 2. Adaptive results in this setting are possible, for example, by applying Lepski's method to choose h and using Azuma's inequality to control the deviations in $\widehat{\varphi}_l(u)$ and $\widehat{\psi}_l(u)$ [Lepski, Mammen and Spokoiny (1997)]. For simplicity, however, we will treat these parameters as fixed, noting that they can be chosen heuristically as in Section 2.

For the time-changed Lévy model $\mathcal{T}^\alpha$, as a simple consequence of results in Munk and Schmidt-Hieber (2010b), we can further show that our rates are optimal. We can likewise provide a partially matching lower bound for the general semimartingale model $\mathcal{S}^{\alpha,\beta}$.

THEOREM 3. *Let* $\alpha > \frac{1}{2}$, $\beta \in [0, 2]$, *and* $0 < C < D$.

(i) *No estimate* c_t^* *of* c_t *can satisfy*

$$\|c^* - c\|_2 = o_p(n^{-\alpha_3}),$$

uniformly over $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D) \cap \mathcal{T}$, *or*

$$|c_t^* - c_t| = o_p(n^{-\alpha_3}),$$

uniformly over $t \in [0, 1]$ *and* $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D) \cap \mathcal{T}$.

(ii) *The same results hold for any estimate* r_t^* *of* r_t.

In the general semimartingale model $\mathcal{S}^{\alpha,\beta}$, if β is large, we have $\alpha_3 > \alpha_4$, and matching lower bounds are more difficult to establish. We note, however, that our estimates $\widetilde{c}_t(u)$ and $\widetilde{r}_t(u)$ already obtain rates as good as those implied by previous work under noise. Furthermore, the recent paper of Jacod and Reiss (2014) on the noiseless problem suggests that the rate $n^{-\alpha_4}$ is indeed optimal, up to log factors.

It may at first be surprising that the results for r_t in the time-changed model $\mathcal{T}^\alpha$ are better than in the general semimartingale model $\mathcal{S}^{\alpha,\beta}$, when the jump activity is large. However, we know that the difficulty in estimating the volatility c_t in $\mathcal{S}^{\alpha,\beta}$ comes primarily from distinguishing c_t and ν_t. We obtain improved convergence rates in $\mathcal{T}^\alpha$ because in this model, we can estimate the rate process r_t without having to separate c_t and ν_t.

We have thus shown that our estimate $\widetilde{r}_t(u)$ can recover the time change in a noisy Lévy model at the minimax rate, equivalent to observing r_t under $n^{-1/4}$ white noise. It can do so without knowledge of the distribution of the Lévy process and under arbitrary jump activity.

Furthermore, in the general semimartingale setting, where the Lévy assumption may be violated, $\widetilde{r}_t(u)$ remains a valid estimate of the normalised volatility. In this setting, we again achieve good rates, governed either by the noise level of $n^{-1/4}$ or by a bias due to jump activity, common to all volatility estimates.

5. Proofs. We now give proofs of our results. We prove results on the preliminary estimates $\widehat{c}_l(u)$ in Section 5.1 and results on convergence rates in Section 5.2. Technical proofs, and a list of notations, are given in the supplemental article [Bull (2014)].

5.1. *Proofs on preliminary estimates.* We first prove Theorem 1, our result bounding the error in our preliminary estimates $\widehat{c}_l(u)$. Our proof will use a series of lemmas, controlling the behaviour of the various components of $\widehat{c}_l(u)$. We begin by stating some technical lemmas; proofs are given in the supplemental article [Bull (2014)].

LEMMA 2. *In the setting of Theorem 1, fix $u \in \mathbb{R}$, and let ξ_t denote* (i) $c_t(u)$, (ii) $\varphi_t(u)$, *or* (iii) $\psi_t(u)$. *In each case, for $n \in \mathbb{N}$, $0 \le t \le s \le 1$, we have*

$$\mathbb{E}\big[\big(\xi_s(u) - \xi_t(u)\big)^2 | \mathcal{F}_t^+\big] = O\big((s-t)^{2\alpha_0} + n^{-1/2}\big).$$

Furthermore, we have (iv) $c_t(u) \le 3D$, *almost surely.*

LEMMA 3. *In the setting of Theorem 1, for $k = 0, \dots, n_0 - 1$, and $u \in \mathbb{R}$, we have*

$$\mathbb{E}\big[\cos(u\widehat{X}_k) | \mathcal{F}_{k/n_0}\big] = \varphi_{k/n_0}(u) + O\big(n^{-1/4}\big),$$
$$\mathbb{V}\mathrm{ar}\big[\cos(u\widehat{X}_k) | \mathcal{F}_{k/n_0}\big] = \rho_{k/n_0}^2(u) + O\big(n^{-1/4}\big).$$

LEMMA 4. *In the setting of Theorem 1, for $k = 0, \dots, n_0 - 1$, and $u \in \mathbb{R}$, we have*

$$\mathbb{E}\big[\exp\big(-\kappa\widehat{\sigma}_k^2 u^2\big) | \mathcal{F}_{k/n_0}\big] = \psi_{k/n_0}(u) + O\big(n^{-1/4}\big),$$
$$\mathbb{V}\mathrm{ar}\big[\exp\big(-\kappa\widehat{\sigma}_k^2 u^2\big) | \mathcal{F}_{k/n_0}\big] = O\big(n^{-1/4}\big).$$

We are now in a position to describe the behaviour of the estimates $\widehat{\varphi}_l(u)$ and $\widehat{\psi}_l(u)$. First, we will define the event E_l mentioned in the statement of Theorem 1. We set

$$E_l = \big\{\widehat{\varphi}_l(u) \ge \zeta(u)\big\} \cap \big\{\widehat{\psi}_l(u) \ge \zeta(u)\big\},$$

where the constant $\zeta(u) = \frac{1}{2}\exp(-(\kappa+3)Du^2)$. We then have the following result.

LEMMA 5. *In the setting of Theorem 1, for $l = 0, \dots, n_2 - 1$, we have*:

(i) $\mathbb{E}[\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u) | \mathcal{F}_{l/n_2}] = O(n^{-\alpha_1})$;
(ii) $\mathbb{E}[\widehat{\psi}_l(u) - \psi_{l/n_2}(u) | \mathcal{F}_{l/n_2}] = O(n^{-\alpha_1})$;
(iii) $\mathbb{E}[(\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u))^2 | \mathcal{F}_{l/n_2}] = \rho_{l/n_2}^2(u)/n_1 + O(n^{-\alpha_1})$;
(iv) $\mathbb{E}[(\widehat{\psi}_l(u) - \psi_{l/n_2}(u))^2 | \mathcal{F}_{l/n_2}] = O(n^{-1/4})$;
(v) *for $p = 3, 4$,* $\mathbb{E}[(\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u))^p | \mathcal{F}_{l/n_2}] = O(n^{-\alpha_1})$;
(vi) $\mathbb{P}(E_l^c | \mathcal{F}_{l/n_2}) \le \exp(-An^{1/8})$, *for a constant $A > 0$.*

PROOF. We first note that

$$\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u) = \frac{1}{n_1} \sum_{k \in K_l} Z_{\delta,k},$$

where the random variables

$$Z_{\delta,k} = \cos(u \widehat{X}_k) - \varphi_{l/n_2}(u), \qquad k \in K_l;$$

we will begin by proving some facts about the $Z_{\delta,k}$. We have $|Z_{\delta,k}| \le 2$ and

$$\begin{aligned} &\mathbb{E}\big[\mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]^2|\mathcal{F}_{l/n_2}\big] \\ &\quad = \mathbb{E}\big[\big(\varphi_{k/n_0}(u) - \varphi_{l/n_2}(u) + O\big(n^{-1/4}\big)\big)^2|\mathcal{F}_{l/n_2}\big], \end{aligned}$$

using Lemma 3,

$$\begin{aligned} &\quad = O(1)\mathbb{E}\big[\big(\varphi_{k/n_0}(u) - \varphi_{l/n_2}(u)\big)^2|\mathcal{F}_{l/n_2}\big] + O\big(n^{-1/2}\big) \\ &\quad = O\big(n^{-2\alpha_1}\big), \end{aligned} \tag{6}$$

using Lemma 2(ii).

We also have

$$\begin{aligned} &\mathbb{E}\big[Z_{\delta,k}^2|\mathcal{F}_{k/n_0}\big] \\ &\quad = \mathbb{V}\mathrm{ar}\big[Z_{\delta,k}^2|\mathcal{F}_{k/n_0}\big] + \mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]^2 \\ &\quad = \rho_{k/n_0}^2(u) + \big(\varphi_{k/n_0}(u) - \varphi_{l/n_2}(u)\big)^2 + O\big(n^{-1/4}\big), \end{aligned}$$

using Lemma 3, so

$$\begin{aligned} &\mathbb{E}\big[\big(\mathbb{E}\big[Z_{\delta,k}^2|\mathcal{F}_{k/n_0}\big] - \rho_{l/n_2}^2(u)\big)^2|\mathcal{F}_{l/n_2}\big] \\ &\quad = O(1)\mathbb{E}\big[\big(\rho_{k/n_0}^2(u) - \rho_{l/n_2}^2(u)\big)^2 + \big(\varphi_{k/n_0}(u) - \varphi_{l/n_2}(u)\big)^2|\mathcal{F}_{l/n_2}\big] \\ &\qquad + O\big(n^{-1/2}\big) \\ &\quad = O\big(n^{-2\alpha_1}\big), \end{aligned} \tag{7}$$

using Lemma 2(ii). We may now prove the claims of the theorem:

(i) We have

$$\begin{aligned} \mathbb{E}\big[\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u)|\mathcal{F}_{l/n_2}\big] &= \frac{1}{n_1} \sum_{k \in K_l} \mathbb{E}[Z_{\delta,k}|\mathcal{F}_{l/n_2}] \\ &= \frac{O(1)}{n_1} \sum_{k \in K_l} \mathbb{E}\big[\big|\mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]\big||\mathcal{F}_{l/n_2}\big] \\ &= O\big(n^{-\alpha_1}\big), \end{aligned}$$

using (6).

(ii) The result follows similarly to (i), using Lemmas 2(iii) and 4.

(iii) We have

$$\begin{aligned}\mathbb{E}[Z_{\delta,k}^2|\mathcal{F}_{l/n_2}] &= \rho_{l/n_2}^2(u) + \mathbb{E}[\mathbb{E}[Z_{\delta,k}^2|\mathcal{F}_{k/n_0}] - \rho_{l/n_2}^2(u)|\mathcal{F}_{l/n_2}] \\ &= \rho_{l/n_2}^2(u) + O(n^{-\alpha_1}),\end{aligned}$$

using (7). Likewise, for $k, k_1 \in K_l$, $k > k_1$, we have

$$\begin{aligned}\mathbb{E}[Z_{\delta,k}Z_{\delta,k_1}|\mathcal{F}_{l/n_2}] &= \mathbb{E}[\mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]Z_{\delta,k_1}|\mathcal{F}_{l/n_2}] \\ &= O(1)\mathbb{E}[|\mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]||\mathcal{F}_{l/n_2}] \\ &= O(n^{-\alpha_1}),\end{aligned}$$

using (6). We deduce that

$$\begin{aligned}&\mathbb{E}[(\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u))^2|\mathcal{F}_{l/n_2}] \\ &\quad = \mathbb{E}\biggl[\frac{1}{n_1^2}\sum_{k\in K_l} Z_{\delta,k}^2 + \frac{2}{n_1^2}\sum_{\substack{k,k_1\in K_l,\\ k>k_1}} Z_{\delta,k}Z_{\delta,k_1}\Big|\mathcal{F}_{l/n_2}\biggr] \\ &\quad = \rho_{l/n_2}^2(u)/n_1 + O(n^{-\alpha_1}).\end{aligned}$$

(iv) For $k \in K_l$, by a similar argument, we have

$$\mathbb{E}[(\exp(-\kappa\widehat{\sigma}_k^2u^2) - \psi_{l/n_2}(u))^2|\mathcal{F}_{l/n_2}] = O(n^{-1/4}),$$

using Lemma 4. The result follows.

(v) We first consider the case $p = 3$. For $k \in K_l$, we have

$$\mathbb{E}[Z_{\delta,k}^3|\mathcal{F}_{l/n_2}] = O(1),$$

and for $k, k_1 \in K_l$, $k > k_1$,

$$\begin{aligned}&\mathbb{E}[Z_{\delta,k}^2 Z_{\delta,k_1}|\mathcal{F}_{l/n_2}] \\ &\quad = \mathbb{E}[\mathbb{E}[Z_{\delta,k}^2|\mathcal{F}_{k/n_0}]Z_{\delta,k_1}|\mathcal{F}_{l/n_2}] \\ &\quad = \rho_{l/n_2}^2(u)\mathbb{E}[Z_{\delta,k_1}|\mathcal{F}_{l/n_2}] \\ &\qquad + O(1)\mathbb{E}[|\mathbb{E}[Z_{\delta,k}^2|\mathcal{F}_{k/n_0}] - \rho_{l/n_2}^2(u)||\mathcal{F}_{l/n_2}] \\ &\quad = O(n^{-\alpha_1}),\end{aligned}$$

using (6) and (7).

Similarly, for $k, k_1, k_2 \in K_l$, $k > k_1, k_2$, we have

$$\begin{aligned}\mathbb{E}[Z_{\delta,k}Z_{\delta,k_1}Z_{\delta,k_2}|\mathcal{F}_{l/n_2}] &= \mathbb{E}[\mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]Z_{\delta,k_1}Z_{\delta,k_2}|\mathcal{F}_{l/n_2}] \\ &= O(1)\mathbb{E}[|\mathbb{E}[Z_{\delta,k}|\mathcal{F}_{k/n_0}]||\mathcal{F}_{l/n_2}] \\ &= O(n^{-\alpha_1}),\end{aligned}$$

using (6). We deduce that

$$
\begin{aligned}
&\mathbb{E}\big[\big(\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u)\big)^3 | \mathcal{F}_{l/n_2}\big] \\
&\quad = \mathbb{E}\bigg[\bigg(\frac{1}{n_1}\sum_{k\in K_l} Z_{\delta,k}\bigg)^3 \bigg| \mathcal{F}_{l/n_2}\bigg] \\
&\quad = \frac{O(1)}{n_1^3}\mathbb{E}\bigg[\sum_{k\in K_l} Z_{\delta,k}^3 + \sum_{\substack{k,k_1\in K_l,\\ k>k_1}} Z_{\delta,k}^2 Z_{\delta,k_1} \\
&\qquad\qquad + \sum_{\substack{k,k_1,k_2\in K_l,\\ k>k_1,k_2}} Z_{\delta,k} Z_{\delta,k_1} Z_{\delta,k_2} \bigg| \mathcal{F}_{l/n_2}\bigg] \\
&\quad = O\big(n^{-\alpha_1}\big).
\end{aligned}
$$

For $p = 4$, by a similar argument, we have that for $k, k_1, k_2, k_3 \in K_l$, $k > k_1, k_2, k_3$,

$$
\begin{aligned}
\mathbb{E}\big[Z_{\delta,k}^4 | \mathcal{F}_{l/n_2}\big], \mathbb{E}\big[Z_{\delta,k}^3 Z_{\delta,k_1} | \mathcal{F}_{l/n_2}\big], \mathbb{E}\big[Z_{\delta,k}^2 Z_{\delta,k_1}^2 | \mathcal{F}_{l/n_2}\big] &= O(1), \\
\mathbb{E}[Z_{\delta,k} Z_{\delta,k_1} Z_{\delta,k_2} Z_{\delta,k_3} | \mathcal{F}_{l/n_2}] &= O\big(n^{-\alpha_1}\big),
\end{aligned}
$$

and if $k_1 > k_2$,

$$
\mathbb{E}\big[Z_{\delta,k}^2 Z_{\delta,k_1} Z_{\delta,k_2} | \mathcal{F}_{l/n_2}\big] = O\big(n^{-\alpha_1}\big).
$$

We thus obtain that

$$
\begin{aligned}
&\mathbb{E}\big[\big(\widehat{\varphi}_l(u) - \varphi_{l/n_2}(u)\big)^4 | \mathcal{F}_{l/n_2}\big] \\
&\quad = \mathbb{E}\bigg[\bigg(\frac{1}{n_1}\sum_{k\in K_l} Z_{\delta,k}\bigg)^4 \bigg| \mathcal{F}_{l/n_2}\bigg] \\
&\quad = \frac{O(1)}{n_1^4}\mathbb{E}\bigg[\sum_{k\in K_l} Z_{\delta,k}^4 + \sum_{\substack{k,k_1\in K_l,\\ k>k_1}} Z_{\delta,k}^3 Z_{\delta,k_1} \\
&\qquad\qquad + \sum_{\substack{k,k_1\in K_l,\\ k>k_1}} Z_{\delta,k}^2 Z_{\delta,k_1}^2 + \sum_{\substack{k,k_1,k_2\in K_l,\\ k>k_1>k_2}} Z_{\delta,k}^2 Z_{\delta,k_1} Z_{\delta,k_2} \\
&\qquad\qquad + \sum_{\substack{k,k_1,k_2,k_3\in K_l,\\ k>k_1,k_2,k_3}} Z_{\delta,k} Z_{\delta,k_1} Z_{\delta,k_2} Z_{\delta,k_3} \bigg| \mathcal{F}_{l/n_2}\bigg] \\
&\quad = O\big(n^{-\alpha_1}\big).
\end{aligned}
$$

(vi) We first note that the quantity

$$\begin{aligned}\overline{\varphi}_l(u) &= \frac{1}{n_1}\sum_{k\in K_l}\mathbb{E}\big[\cos(u\widehat{X}_k)|\mathcal{F}_{k/n_0}\big]\\ &= \frac{1}{n_1}\sum_{k\in K_l}\varphi_{k/n_0}(u) + O\big(n^{-1/4}\big),\end{aligned}$$

using Lemma 3,

$$\geq 2\zeta(u) + O\big(n^{-1/4}\big),$$

using Lemma 2(iv). Then using Azuma's inequality, we have

$$\begin{aligned}&\mathbb{P}\big(\widehat{\varphi}_l(u) \leq \zeta(u)|\mathcal{F}_{l/n_2}\big)\\ &\quad\leq \mathbb{P}\big(\widehat{\varphi}_l(u) - \overline{\varphi}_l(u) \leq -\zeta(u) + O\big(n^{-1/4}\big)|\mathcal{F}_{l/n_2}\big)\\ &\quad\leq \exp\big(-A'n^{1/8}\big),\end{aligned}$$

for a constant $A' > 0$. By a similar argument, we also have

$$\mathbb{P}\big(\widehat{\psi}_l(u) \leq \zeta(u)|\mathcal{F}_{l/n_2}\big) \leq \exp\big(-A''n^{1/8}\big),$$

for a constant $A'' > 0$. The result follows. □

Finally, we may prove Theorem 1.

PROOF OF THEOREM 1. From the definitions, we have that the estimates $\widehat{c}_l(u)$ are $\mathcal{F}_{(l+1)/n_2}$-measurable, and the events $E_l \in \mathcal{F}_{(l+1)/n_2}$. The bound on the probability of E_l^c likewise follows directly from Lemma 5(vi).

It thus remains to prove the bounds on the mean and variance of $\widehat{c}_l(u)$. We will decompose the error in $\widehat{c}_l(u)$ into three terms, controlling the error in each of $\log(\widehat{\varphi}_l(u))$, $\log(\widehat{\psi}_l(u))$ and $\widehat{\tau}_l^2(u)$.

We first consider $\log(\widehat{\varphi}_l(u))$ and define the random variable

$$Z_{\varphi,l} = \frac{\widehat{\varphi}_l(u)}{\varphi_{l/n_2}(u)} - 1.$$

We then have that

$$\begin{aligned}&\big(\log\big(\widehat{\varphi}_l(u)\big) - \log\big(\varphi_{l/n_2}(u)\big)\big)1(E_l)\\ &\quad= \log(1 + Z_{\varphi,l})1(E_l)\\ &\quad= \big(Z_{\varphi,l} - \tfrac{1}{2}Z_{\varphi,l}^2 + \tfrac{1}{3}Z_{\varphi,l}^3 + O\big(Z_{\varphi,l}^4\big)\big)1(E_l),\end{aligned}$$

using Taylor's theorem, since on E_l,

$$1 + Z_{\varphi,l} \geq \frac{\zeta(u)}{\varphi_{l/n_2}(u)} \geq \zeta(u) > 0. \tag{8}$$

To bound the error in $\log(\widehat{\varphi}_l(u))$, we will now take expectations of the $Z_{\varphi,l}$ terms. We have that

$$\mathbb{E}[Z_{\varphi,l}1(E_l)|\mathcal{F}_{l/n_2}] = \mathbb{E}\left[\frac{\widehat{\varphi}_l(u)}{\varphi_{l/n_2}(u)} - 1 \Big| \mathcal{F}_{l/n_2}\right] + O(\mathbb{P}(E_l^c|\mathcal{F}_{l/n_2})),$$

since $\widehat{\varphi}_l(u)$ is bounded, and $\varphi_{l/n_2}(u) \geq 2\zeta(u) > 0$,

$$= O(n^{-\alpha_1}),$$

using Lemma 5(i) and (vi). Similarly, we also have

$$\begin{aligned}
\mathbb{E}[Z_{\varphi,l}^2 1(E_l)|\mathcal{F}_{l/n_2}] &= \tau_{l/n_2}^2(u) + O(n^{-\alpha_1}),\\
\mathbb{E}[Z_{\varphi,l}^3 1(E_l)|\mathcal{F}_{l/n_2}] &= O(n^{-\alpha_1}),\\
\mathbb{E}[Z_{\varphi,l}^4 1(E_l)|\mathcal{F}_{l/n_2}] &= O(n^{-\alpha_1}),
\end{aligned}$$

using Lemma 5(i), (iii), (v) and (vi); as a consequence, we deduce

$$\begin{aligned}
\mathbb{E}[|Z_{\varphi,l}|^3 1(E_l)|\mathcal{F}_{l/n_2}] &\leq \mathbb{E}[Z_{\varphi,l}^2 1(E_l)|\mathcal{F}_{l/n_2}]^{1/2}\mathbb{E}[Z_{\varphi,l}^4 1(E_l)|\mathcal{F}_{l/n_2}]^{1/2}\\
&= O(n^{-\alpha_2}),
\end{aligned}$$

using Cauchy–Schwarz.

We can now bound the error in $\log(\widehat{\varphi}_l(u))$. We conclude that

$$\begin{aligned}
&\mathbb{E}[(\log(\widehat{\varphi}_l(u)) - \log(\varphi_{l/n_2}(u)))1(E_l)|\mathcal{F}_{l/n_2}]\\
&\quad = \mathbb{E}[(Z_{\varphi,l} - \tfrac{1}{2}Z_{\varphi,l}^2 + \tfrac{1}{3}Z_{\varphi,l}^3 + O(Z_{\varphi,l}^4))1(E_l)|\mathcal{F}_{l/n_2}]\\
&\quad = -\tfrac{1}{2}\tau_{l/n_2}^2(u) + O(n^{-\alpha_1}),
\end{aligned}$$

and similarly,

$$\begin{aligned}
&\mathbb{E}[(\log(\widehat{\varphi}_l(u)) - \log(\varphi_{l/n_2}(u)))^2 1(E_l)|\mathcal{F}_{l/n_2}]\\
&\quad = \mathbb{E}[(Z_{\varphi,l} + O(Z_{\varphi,l}^2))^2 1(E_l)|\mathcal{F}_{l/n_2}]\\
&\quad = \mathbb{E}[(Z_{\varphi,l}^2 + O(|Z_{\varphi,l}|^3 + Z_{\varphi,l}^4))1(E_l)|\mathcal{F}_{l/n_2}]\\
&\quad = \tau_{l/n_2}^2(u) + O(n^{-\alpha_2}).
\end{aligned}$$

We next consider the error in $\log(\widehat{\psi}_l(u))$. By a similar argument, we can obtain that

$$\begin{aligned}
\mathbb{E}[(\log(\widehat{\psi}_l(u)) - \log(\psi_{l/n_2}(u)))1(E_l)|\mathcal{F}_{l/n_2}] &= O(n^{-\alpha_1}),\\
\mathbb{E}[(\log(\widehat{\psi}_l(u)) - \log(\psi_{l/n_2}(u)))^2 1(E_l)|\mathcal{F}_{l/n_2}] &= O(n^{-1/4}),
\end{aligned}$$

using Lemma 5(ii), (iv) and (vi).

Finally, we prove bounds on $\widehat{\tau}_l^2(u)$, defining the random variable

$$Z_{\tau,l} = \widehat{\varphi}_l(2u) - \varphi_{l/n_2}(2u).$$

We then have

$$\begin{aligned}
&\big(\widehat{\tau}_l^2(u) - \tau_{l/n_2}^2(u)\big)1(E_l) \\
&\quad = \frac{1}{n_1}\left(\frac{1+\varphi_{l/n_2}(2u)+Z_{\tau,l}}{2\varphi_{l/n_2}^2(u)(1+Z_{\varphi,l})^2} - \frac{1+\varphi_{l/n_2}(2u)}{2\varphi_{l/n_2}^2(u)}\right)1(E_l) \\
&\quad = \frac{1}{n_1}\left(\frac{-2(1+\varphi_{l/n_2}(2u))Z_{\varphi,l}+Z_{\tau,l}}{2\varphi_{l/n_2}^2(u)} + O\big(Z_{\varphi,l}^2 + Z_{\varphi,l}Z_{\tau,l}\big)\right)1(E_l),
\end{aligned}$$

using (8), and that $\varphi_t(u)$ is bounded below.

Using Lemma 5(i), (iii) and (vi) as above, we also have that

$$\begin{aligned}
\mathbb{E}\big[Z_{\tau,l}1(E_l)|\mathcal{F}_{l/n_2}\big] &= O\big(n^{-\alpha_1}\big), \\
\mathbb{E}\big[Z_{\tau,l}^2 1(E_l)|\mathcal{F}_{l/n_2}\big] &= O\big(n^{-1/8}\big).
\end{aligned}$$

We therefore conclude that

$$\begin{aligned}
&\mathbb{E}\big[\big(\widehat{\tau}_l^2(u) - \tau_{l/n_2}^2(u)\big)1(E_l)|\mathcal{F}_{l/n_2}\big] \\
&\quad = O\big(n^{-1/8}\big)\mathbb{E}\big[Z_{\varphi,l} + Z_{\tau,l} + Z_{\varphi,l}^2 + |Z_{\varphi,l}||Z_{\tau,l}||\mathcal{F}_{l/n_2}\big],
\end{aligned}$$

since $\varphi_t(u)$ is bounded below,

$$= O\big(n^{-1/4}\big),$$

using Cauchy–Schwarz. Likewise,

$$\begin{aligned}
\mathbb{E}\big[\big(\widehat{\tau}_l^2(u) - \tau_{l/n_2}^2(u)\big)^2 1(E_l)|\mathcal{F}_{l/n_2}\big] &= O\big(n^{-1/4}\big)\mathbb{E}\big[Z_{\varphi,l}^2 + Z_{\tau,l}^2|\mathcal{F}_{l/n_2}\big] \\
&= O\big(n^{-3/8}\big).
\end{aligned}$$

We have thus bounded the error in each of $\log(\widehat{\varphi}_l(u))$, $\log(\widehat{\psi}_l(u))$ and $\widehat{\tau}_l^2(u)$. Combining these results, we deduce that

$$\begin{aligned}
&\mathbb{E}\big[\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)1(E_l)|\mathcal{F}_{l/n_2}\big] \\
&\quad = O(1)\mathbb{E}\big[\big(\big(\log\big(\widehat{\varphi}_l(u)\big) - \log\big(\varphi_{l/n_2}(u)\big)\big) + \tfrac{1}{2}\tau_{l/n_2}^2(u) \\
&\qquad\qquad - \big(\log\big(\widehat{\psi}_l(u)\big) - \log\big(\psi_{l/n_2}(u)\big)\big) \\
&\qquad\qquad\qquad + \tfrac{1}{2}\big(\widehat{\tau}_l^2(u) - \tau_{l/n_2}^2(u)\big)\big)1(E_l)|\mathcal{F}_{l/n_2}\big] \\
&\quad = O\big(n^{-\alpha_1}\big)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\big[\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)^2 1(E_l)|\mathcal{F}_{l/n_2}\big] \\
&\quad = O(1)\mathbb{E}\big[\big(\big(\log(\widehat{\varphi}_l(u)) - \log(\varphi_{l/n_2}(u))\big)^2 \\
&\qquad\qquad + \big(\log(\widehat{\psi}_l(u)) - \log(\psi_{l/n_2}(u))\big)^2 \\
&\qquad\qquad + \big(\widehat{\tau}_l^2(u) - \tau_{l/n_2}^2(u)\big)^2 + O\big(n^{-1/4}\big)\big)1(E_l)|\mathcal{F}_{l/n_2}\big] \\
&\quad = \tau_{l/n_2}^2(u) + O\big(n^{-\alpha_2}\big).
\end{aligned}$$

Finally, it can be checked that these results are uniform over $l = 0, \dots, n_2 - 1$, and $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$. □

5.2. *Proofs of convergence rates.* We next prove Theorem 2, our result on the performance of our regression estimate $\widetilde{c}_t(u)$. Our argument follows from Tsybakov (2009), taking care to account for the extra error terms in the statement of Theorem 1 and the stochastic nature of the target $c_t(u)$.

PROOF OF THEOREM 2. To begin, we will state some facts about local polynomial regression, as given in the proof of Theorem 1.7 in Tsybakov (2009). Since the design points l/n_2 are uniform, we have that for large n, the matrices $V_n(t)$ are invertible and the weight functions $W_{n,l}(t)$ well defined. Furthermore, the weights $W_{n,l}(t)$ satisfy:

$$|W_{n,l}(t)| = O\left(\frac{1}{n_2 h}\right) 1\left(\left|t - \frac{l}{n_2}\right| \le h\right), \tag{9}$$

uniformly in $l = 0, \dots, n_2 - 1$;

$$\sum_{l=0}^{n_2-1} |W_{n,l}(t)| = O(1); \tag{10}$$

$$\sum_{l=0}^{n_2-1} \left(t - \frac{l}{n_2}\right)^p W_{n,l}(t) = \begin{cases} 1, & p = 0, \\ 0, & p = 1, \dots, N-1. \end{cases} \tag{11}$$

We now prove the results on our estimate $\widetilde{c}_t(u)$. We must first define the high-probability event E given in the statement of the theorem. We let $E_{a,b} = \bigcap_{l=a}^{b-1} E_l$, and set $E = E_{0,n_2}$. We then note that from Theorem 1, we have

$$\begin{aligned}
\mathbb{P}(E^c|\mathcal{F}_0) &\le \sum_{l=0}^{n_2-1} \mathbb{E}\big[\mathbb{P}(E_l^c|\mathcal{F}_{l/n_2})|\mathcal{F}_0\big] \\
&= O\big(n^{3/8}\big)\exp\big(-An^{1/8}\big) \\
&\le \exp\big(-A'n^{1/8}\big),
\end{aligned}$$

for constants $A, A' > 0$. Similarly, for $l = 0, \dots, n_2 - 1$, $k \geq l$, we have

$$\mathbb{P}\big(E^c_{k,n_2} | \mathcal{F}_{l/n_2}\big) \leq \exp\big(-A' n^{1/8}\big). \tag{12}$$

We next split the estimates $\widehat{c}_l(u)$ into bias and variance parts. Let

$$\widehat{c}_l(u) = c_{l/n_2}(u) + \widehat{c}_{1,l}(u) + \widehat{c}_{2,l}(u), \tag{13}$$

where the bias term

$$\widehat{c}_{1,l}(u) = \frac{\mathbb{E}[(\widehat{c}_l(u) - c_{l/n_2}(u))1(E)|\mathcal{F}_{l/n_2}]}{\mathbb{P}(E|\mathcal{F}_{l/n_2})},$$

setting $\widehat{c}_{1,l}(u) = 0$ when $\mathbb{P}(E|\mathcal{F}_{l/n_2}) = 0$, and the variance term $\widehat{c}_{2,l}(u)$ is then defined by (13).

We can similarly split the regression estimates $\widetilde{c}_t(u)$ into bias and variance parts. Let

$$\widetilde{c}_t(u) = c_t(u) + \widetilde{c}_{1,t}(u) + \widetilde{c}_{2,t}(u) + \widetilde{c}_{3,t}(u), \tag{14}$$

where the estimator bias and variance, $\widetilde{c}_{1,t}(u)$ and $\widetilde{c}_{2,t}(u)$, are given by

$$\widetilde{c}_{k,t}(u) = \sum_{l=0}^{n_2-1} W_{n,l}(t)\widehat{c}_{k,l}(u), \qquad k = 1, 2,$$

and the regression bias

$$\widetilde{c}_{3,t}(u) = \sum_{l=0}^{n_2-1} W_{n,l}(t) c_{l/n_2}(u) - c_t(u).$$

To bound the error in $\widetilde{c}_t(u)$, we must show that all three terms $\widetilde{c}_{k,t}(u)$ are small. We begin with the estimator bias $\widetilde{c}_{1,t}(u)$ and note that for large n,

$$\begin{aligned}|\widehat{c}_{1,l}(u)| &= 1(E_{0,l}) \frac{\mathbb{E}[(\widehat{c}_l(u) - c_{l/n_2}(u))1(E_{l,n_2})|\mathcal{F}_{l/n_2}]}{\mathbb{P}(E_{l,n_2}|\mathcal{F}_{l/n_2})} \\ &= O(1)\mathbb{E}\big[\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)1(E_{l,n_2})|\mathcal{F}_{l/n_2}\big],\end{aligned}$$

using (12),

$$= O(1)\big(\mathbb{E}\big[\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)1(E_l)|\mathcal{F}_{l/n_2}\big] + \mathbb{P}\big(E^c_{l+1,n_2}|\mathcal{F}_{l/n_2}\big)\big),$$

since $(\widehat{c}_l(u) - c_{l/n_2}(u))1(E_l)$ is bounded,

$$= O\big(n^{-\alpha_1}\big), \tag{15}$$

using (12) and Theorem 1. We thus have

$$|\widetilde{c}_{1,t}(u)| \leq \sum_{l=0}^{n_2-1} |W_{n,l}(t)||\widehat{c}_{1,l}(u)| = O\big(n^{-\alpha_1}\big),$$

using (10) and (15).

We next consider the estimator variance $\widetilde{c}_{2,t}(u)$. We first note that

$$\begin{aligned}\mathbb{E}[\widehat{c}_{2,l}(u)1(E)|\mathcal{F}_{l/n_2}] &= \mathbb{E}[(\widehat{c}_l(u) - c_{l/n_2}(u) - \widehat{c}_{1,l}(u))1(E)|\mathcal{F}_{l/n_2}] \\ &= 0\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[\widehat{c}^2_{2,l}(u)1(E)|\mathcal{F}_{l/n_2}] &= O(1)\big(\mathbb{E}[(\widehat{c}_l(u) - c_{l/n_2}(u))^2 1(E)|\mathcal{F}_{l/n_2}] + \widehat{c}^2_{1,l}(u)\big) \\ &= O(n^{-1/8}),\end{aligned}$$

using (15) and Theorem 1. We thus have

$$\begin{aligned}\mathbb{E}[\widetilde{c}^2_{2,t}(u)1(E)|\mathcal{F}_0] &= \mathbb{E}\left[\left(\sum_{l=0}^{n_2-1} W_{n,l}(t)\widehat{c}_{2,l}(u)1(E)\right)^2 \middle| \mathcal{F}_0\right] \\ &= \sum_{l=0}^{n_2-1} W^2_{n,l}(t)\mathbb{E}[\widehat{c}^2_{2,l}(u)1(E)|\mathcal{F}_0] \\ &= O(n^{-1/8})\Big(\max_{l=0}^{n_2-1}|W_{n,l}(t)|\Big)\left(\sum_{l=0}^{n_2-1}|W_{n,l}(t)|\right) \\ &= O(n^{-2\alpha_3}),\end{aligned}$$

using (9) and (10).

Finally, we bound the regression bias $\widetilde{c}_{3,t}(u)$. Let m denote the largest integer smaller than α. Using Taylor's theorem, for $t \in [0, 1]$, and $l = 0, \ldots, n_2 - 1$, we then have that

$$c_{l/n_2}(u) = c_t(u) + \sum_{r=1}^{m-1} \frac{(t - l/n_2)^r}{r!} c_t^{(r)}(u) + \frac{(t - l/n_2)^m}{m!} c_{t_l}^{(m)}(u),$$

for some $t_l \in [0, 1]$ lying between t and l/n_2. We deduce that

$$\begin{aligned}&\mathbb{E}[\widetilde{c}^2_{3,t}(u)|\mathcal{F}_0] \\ &\quad = \mathbb{E}\left[\left(\sum_{l=0}^{n_2-1} W_{n,l}(t)\big(c_{l/n_2}(u) - c_t(u)\big)\right)^2 \middle| \mathcal{F}_0\right],\end{aligned}$$

using (11),

$$= \mathbb{E}\left[\left(\sum_{l=0}^{n_2-1} W_{n,l}(t)\frac{(t - l/n_2)^m}{m!}\big(c_{t_l}^{(m)}(u) - c_t^{(m)}(u)\big)\right)^2 \middle| \mathcal{F}_0\right],$$

again using (11),

$$= O(h^{2m}) \sum_{k,l=0}^{n_2-1} |W_{n,k}(t)||W_{n,l}(t)| 1\left(\left|t - \frac{k}{n_2}\right|, \left|t - \frac{l}{n_2}\right| \leq h\right)$$
$$\times \mathbb{E}\big[|c_{t_k}^{(m)}(u) - c_t^{(m)}(u)||c_{t_l}^{(m)}(u) - c_t^{(m)}(u)||\mathcal{F}_0\big],$$

using (9),

$$= O(h^{2\alpha}) \left(\sum_{l=0}^{n_2-1} |W_{n,l}(t)|\right)^2,$$

using Cauchy–Schwarz,

$$= O(n^{-2\alpha_3}),$$

using (10).

Combining these results, we obtain that

$$\begin{aligned}
&\mathbb{E}\big[(\widetilde{c}_t(u) - c_t(u))^2 1(E)|\mathcal{F}_0\big] \\
&\quad = O(1)\mathbb{E}\big[(\widetilde{c}_{1,t}^2(u) + \widetilde{c}_{2,t}^2(u) + \widetilde{c}_{3,t}^2(u)) 1(E)|\mathcal{F}_0\big] \\
&\quad = O(n^{-2\alpha_3}),
\end{aligned}$$

as required. For the L^2 risk, we likewise obtain

$$\begin{aligned}
\mathbb{E}\big[\|\widetilde{c}_t(u) - c_t(u)\|_2^2 1(E)|\mathcal{F}_0\big] &= \mathbb{E}\left[\int_0^1 (\widetilde{c}_t(u) - c_t(u))^2 1(E)\, dt \middle| \mathcal{F}_0\right] \\
&= \int_0^1 \mathbb{E}\big[(\widetilde{c}_t(u) - c_t(u))^2 1(E)|\mathcal{F}_0\big]\, dt \\
&= O(n^{-2\alpha_3}).
\end{aligned}$$

Finally, we can check that these rates are uniform over $t \in [0, 1]$, and $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$. □

We may now deduce our corollary describing the performance of $\widetilde{r}_t(u)$ and $\widetilde{c}_t(u)$.

PROOF OF COROLLARY 1. We first fix $0 < C \leq D$ and prove bounds on the error of $\widetilde{c}_t(u)$ under the assumption that $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$. For $t \in [0, 1]$, we have

$$|\widetilde{c}_t(u) - c_t| \leq |\widetilde{c}_t(u) - c_t(u)| + |c_t(u) - c_t|,$$

and from Theorem 2,

$$|\widetilde{c}_t(u) - c_t(u)| = O_p(n^{-\alpha_3}).$$

It thus remains to bound $|c_t(u) - c_t|$. We have that

$$\begin{aligned}
|c_t(u) - c_t| &= \frac{1}{n_0 u^2} \int_{\mathbb{R}} \int_0^1 \left(1 - \cos\left(\sqrt{n_0}\Phi(w)ux\right)\right) dw\, \nu_t(dx) \\
&= O\left(n_0^{-1}\right) \int_{\mathbb{R}} \left(1 \wedge n_0 x^2\right) \nu_t(dx) \\
&= O\left(n_0^{-1}\right) \int_{\mathbb{R}} \left(1 \wedge \left(n_0 x^2\right)^{\beta/2}\right) \nu_t(dx) \\
&= O\left(n_0^{-(2-\beta)/2}\right) \int_{\mathbb{R}} \left(1 \wedge |x|^\beta\right) \nu_t(dx) \\
&= O\left(n_0^{-(2-\beta)/2}\right) = O\left(n^{-(2-\beta)/4}\right).
\end{aligned}$$

We thus conclude that

$$|\widetilde{c}_t(u) - c_t| = O_p\left(n^{-\alpha_4}\right); \tag{16}$$

by a similar argument, the same holds also for the L^2 error, $\|\widetilde{c}(u) - c\|_2$. We can further check that these limits hold conditionally on $\mathcal{F}_0$ and uniformly over all $t \in [0, 1]$, $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$.

We next consider the case that $\mathbb{P} \in \mathcal{S}_\gamma^{\alpha,\beta}(C, D)$, for some $\gamma \in [0, 1]$. Create, on an extended probability space, a process X_t^S, $t \in [0, 1]$, which almost-surely agrees with X_t at times $t \in [0, S]$. For times $t \in [S, 1]$, we require that X_t is a Lévy process with respect to both $\mathcal{F}_t$ and $\mathcal{F}_t^+$, with characteristic triplet (b_S, c_S, ν_S).

Also create observations

$$Y_j^S = X_{j/n}^S + \varepsilon_j^S, \qquad j = 0, \dots, n-1,$$

where for $j/n \le S$, the errors $\varepsilon_j^S = \varepsilon_j$. When $j/n > S$, we require that the errors ε_j^S are $\mathcal{F}_{j/n}^+$ -measurable, and equal to $\pm\sigma_S$ each with probability $\frac{1}{2}$ given $\mathcal{F}_{j/n}$.

Then let $\widetilde{c}_t^S(u)$ denote the estimate of c_t defined similarly to $\widetilde{c}_t(u)$, but using the observations Y_j^S. Conditionally on Ω_0, the law of the X_t^S and Y_j^S lies in $\mathcal{S}^{\alpha,\beta}(C, D)$, so we can apply (16) to $\widetilde{c}_t^S(u)$. We obtain that

$$|\widetilde{c}_t^S(u) - c_{t\wedge S}| 1(\Omega_0) = O_p\left(f(C, D) n^{-\alpha_4}\right), \tag{17}$$

uniformly in γ, C and D, for a function $f(C, D) > 0$.

We now consider the case $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}$, and suppose we are given an arbitrary sequence $\delta_n > 0$, $\delta_n \to \infty$. If we choose $C_n \to 0$, $D_n \to \infty$ slowly enough as $n \to \infty$, we will obtain that $f(C_n, D_n) = O(\delta_n)$. Since $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}$, we also have that $\mathbb{P} \in \mathcal{S}_{\gamma_n}^{\alpha,\beta}(C_n, D_n)$ for some $\gamma_n \to 0$; let $\Omega_{0,n} \in \mathcal{F}_0$ and $S_n \in [0, 1]$ denote the associated events and stopping times.

Applying (17), we deduce that

$$|\widetilde{c}_t^{S_n}(u) - c_{t\wedge S_n}| 1(\Omega_{0,n}) = O_p\left(\delta_n n^{-\alpha_4}\right).$$

Since

$$\mathbb{P}(\Omega_{0,n} \cap \{S_n = 1\}) \geq 1 - \gamma_n \to 1,$$

this implies that

$$\left|\widetilde{c}_t(u) - c_t\right| = O_p(\delta_n n^{-\alpha_4}).$$

Since this result holds for any diverging sequence δ_n, we conclude that

$$\left|\widetilde{c}_t(u) - c_t\right| = O_p(n^{-\alpha_4}).$$

Again, the result for the L^2 error follows similarly.

Next, we suppose that $\mathbb{P} \in \mathcal{T}^\alpha(C, D)$, and bound the accuracy of the estimate $\widetilde{r}_t(u)$. We begin by bounding its normalising constant,

$$\frac{1}{n_2}\sum_{l=0}^{n_2-1} \widehat{c}_l(u) = \sum_{l=0}^{n_2-1} \widetilde{W}_{n,l}(t)\widehat{c}_l(u),$$

where the weights $\widetilde{W}_{n,l}(t) = 1/n_2$. Since these weights satisfy (9) and (10) for the bandwidth $h = 1$, we have that

$$\begin{aligned}\left|\frac{1}{n_2}\sum_{l=0}^{n_2-1}\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)\right| &= \left|\sum_{l=0}^{n_2-1}\widetilde{W}_{n,l}(t)\big(\widehat{c}_l(u) - c_{l/n_2}(u)\big)\right| \\ &= O_p(n^{-\alpha_1}),\end{aligned}$$

arguing as in Theorem 2.

We also have

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{1}{n_2}\sum_{l=0}^{n_2-1} c_{l/n_2}(u) - \int_0^1 c_t(u)\,dt\right)^2\right] \\ &\quad= \mathbb{E}\left[\left(\sum_{l=0}^{n_2-1}\int_{l/n_2}^{(l+1)/n_2}\big(c_{l/n_2}(u) - c_t(u)\big)\,dt\right)^2\right] \\ &\quad\leq \mathbb{E}\left[\sum_{l=0}^{n_2-1}\int_{l/n_2}^{(l+1)/n_2}\big(c_{l/n_2}(u) - c_t(u)\big)^2\,dt\right],\end{aligned}$$

by Jensen's inequality,

$$\begin{aligned}&\quad= \sum_{l=0}^{n_2-1}\int_{l/n_2}^{(l+1)/n_2}\mathbb{E}\big[\big(c_{l/n_2}(u) - c_t(u)\big)^2\big]\,dt \\ &\quad= O(n^{-2\alpha_1}).\end{aligned}$$

We thus deduce that

$$\left|\frac{1}{n_2}\sum_{l=0}^{n_2-1}\widehat{c}_l(u) - \int_0^1 c_t(u)\,dt\right| = O_p(n^{-\alpha_1}). \tag{18}$$

From Theorem 2, we also have that

$$\left|\widetilde{c}_t(u) - c_t(u)\right| = O_p(n^{-\alpha_3}). \tag{19}$$

Combining these results, we obtain that

$$\left|\widetilde{r}_t(u) - r_t\right| = \left|\frac{\widetilde{c}_t(u)}{(1/n_2)\sum_{l=0}^{n_2-1}\widehat{c}_l(u)} - \frac{c_t(u)}{\int_0^1 c_t(u)\,dt}\right| = O_p(n^{-\alpha_3}),$$

since $\int_0^1 c_t(u)\,dt \geq C > 0$.

We can again check that this limit holds conditionally on $\mathcal{F}_0$ and uniformly over all $t \in [0, 1]$, $\mathbb{P} \in \mathcal{T}^\alpha(C, D)$. Arguing as above, we then conclude that for $\mathbb{P} \in \mathcal{T}^\alpha$,

$$\left|\widetilde{r}_t(u) - r_t\right| = O_p(n^{-\alpha_3}).$$

As above, we can also conclude that these results likewise hold for the L^2 error $\|\widetilde{r}(u) - r\|_2$.

Finally, we bound the performance of $\widetilde{r}_t(u)$ for $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}(C, D)$. Combining (16), (18) and (19), we have

$$\left|\frac{1}{n_2}\sum_{l=0}^{n_2-1}\widehat{c}_l(u) - \int_0^1 c_t\,dt\right|, \left|\widetilde{c}_t(u) - c_t\right| = O_p(n^{-\alpha_4}).$$

Arguing as above, we obtain that

$$\left|\widetilde{r}_t(u) - r_t\right| = O_p(n^{-\alpha_4}),$$

and that this can be extended to $\mathbb{P} \in \mathcal{S}^{\alpha,\beta}$, and the L^2 error $\|\widetilde{r}(u) - r\|_2$. □

Finally, we can also prove our lower bound on the rate of estimation, which is a simple corollary of results in Munk and Schmidt-Hieber (2010b).

PROOF OF THEOREM 3. We begin with part (i) and appeal to the proof of Theorem 2.1 in Munk and Schmidt-Hieber (2010b). The authors give a lower bound on the L^2 estimation rate of c_t, in a setting similar to our $\mathcal{S}^{\alpha,\beta}(C, D)$.

Munk and Schmidt-Hieber consider a setting where $\sigma_t^2 = \sigma^2 > 0$ is a deterministic constant, c_t is deterministic, and $b_t = \nu_t = 0$. They then construct a large number of choices $c_{\omega,t}$ for the volatility, separated from each other in L^2 norm at a rate at least $n^{-\alpha_3}$. They further establish that, given observations Y_j under one such volatility function $c_{\omega,t}$, we cannot consistently estimate ω. They thus show that no estimate c_t^* of c_t can satisfy $\|c^* - c\|_2 = o_p(n^{-\alpha_3})$.

It can be checked that, when $C < 1 < D$, their models lie in $\mathcal{S}^{\alpha,\beta}(C, D) \cap \mathcal{T}$ for large n, so their lower bound holds also in that setting. By rescaling their volatility functions $c_{\omega,t}$, we can obtain the same results also for general $0 < C < D$.

A pointwise lower bound can be proved by a similar argument; we sketch a proof below. Define two choices for the volatility,

$$c_{0,t} = 1, \qquad c_{1,t} = 1 + h^{\alpha} K\big((1-t)/h\big),$$

where $h = n^{-1/2(2\alpha+1)}$, and $K : \mathbb{R} \to \mathbb{R}$ is a smooth nonincreasing nonnegative function, satisfying $K(0) = 1$, $K(1) = 0$.

We note that when $C < 1 < D$, these models lie within $\mathcal{S}^{\alpha,\beta}(C, D)$ for large n; as above, by rescaling we can work with general $0 < C < D$. We also have that $c_{0,1}$ and $c_{1,1}$ are separated at a rate $n^{-\alpha_3}$. It thus suffices to show that we cannot consistently distinguish c_0 from c_1 given the Y_j.

We begin by moving to a more informative model, where we additionally observe one efficient price X_t, at a time $t = \lfloor (1-h)n \rfloor / n$. Given X_t, the observations Y_j, $j \le nt$ are independent of the Y_j, $j > nt$; furthermore, the former are identically distributed under c_0 and c_1.

We therefore need consider only the observations X_t and Y_j, $j > nt$. Arguing similarly to Munk and Schmidt-Hieber, it can be shown that these observations are insufficient to distinguish c_0 and c_1, thereby establishing our lower bound.

For part (ii), it can be checked that the rate functions $r_{\omega,t} = c_{\omega,t} / \int_0^1 c_{\omega,s}\,ds$ are again separated, in L^2 norm or pointwise, at a rate at least $n^{-\alpha_3}$. We thus conclude that our lower bounds hold also for r_t. □

SUPPLEMENTARY MATERIAL

Supplement to "Estimating time-changes in noisy Lévy models" (DOI: 10.1214/14-AOS1250SUPP; .pdf). Proofs of some technical results.

REFERENCES

AÏT-SAHALIA, Y. and JACOD, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* **37** 2202–2244. MR2543690

ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. and LABYS, P. (2000). Great realizations. *Risk* **13** 105–108.

BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2004). Power and bipower variation with stochastic volatility and jumps. *J. Financ. Econom.* **2** 1–37.

BARNDORFF-NIELSEN, O. E., HANSEN, P. R., LUNDE, A. and SHEPHARD, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* **76** 1481–1536. MR2468558

BELOMESTNY, D. (2011). Statistical inference for time-changed Lévy processes via composite characteristic function estimation. *Ann. Statist.* **39** 2205–2242. MR2893866

BELOMESTNY, D. and PANOV, V. (2013). Estimation of the activity of jumps in time-changed Lévy models. *Electron. J. Stat.* **7** 2970–3003. MR3151759

BULL, A. D. (2014). Supplement to "Estimating time-changes in noisy Lévy models." DOI:10.1214/14-AOS1250SUPP.

CARR, P. and WU, L. (2004). Time-changed Lévy processes and option pricing. *J. Financ. Econom.* **71** 113–141.

CONT, R. and TANKOV, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton, FL. MR2042661

DELBAEN, F. and SCHACHERMAYER, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300** 463–520. MR1304434

FAN, J. and WANG, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *J. Amer. Statist. Assoc.* **102** 1349–1362. MR2372538

FIGUEROA-LÓPEZ, J. E. (2009). Nonparametric estimation of time-changed Lévy models under high-frequency data. *Adv. in Appl. Probab.* **41** 1161–1188. MR2663241

FIGUEROA-LÓPEZ, J. E. (2011). Central limit theorems for the nonparametric estimation of time-changed Lévy models. *Scand. J. Stat.* **38** 748–765. MR2859748

FIGUEROA-LÓPEZ, J. E. (2012). Statistical estimation of Lévy-type stochastic volatility models. *Ann. Finance* **8** 309–335. MR2922800

GOLUB, G. H., HEATH, M. and WAHBA, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* **21** 215–223. MR0533250

HANSEN, P. R. and LUNDE, A. (2006). Realized variance and market microstructure noise. *J. Bus. Econom. Statist.* **24** 127–218. MR2234447

HAUTSCH, N. and PODOLSKIJ, M. (2013). Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence. *J. Bus. Econom. Statist.* **31** 165–183. MR3055330

HOFFMANN, M., MUNK, A. and SCHMIDT-HIEBER, J. (2012). Adaptive wavelet estimation of the diffusion coefficient under additive error measurements. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 1186–1216. MR3052408

JACOD, J. and REISS, M. (2014). A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. *Ann. Statist.* **42** 1131–1144. MR3224283

JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877

JACOD, J. and TODOROV, V. (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *Ann. Statist.* **42** 1029–1069. MR3210995

JACOD, J., LI, Y., MYKLAND, P. A., PODOLSKIJ, M. and VETTER, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.* **119** 2249–2276. MR2531091

JING, B.-Y., KONG, X.-B. and LIU, Z. (2011). Estimating the jump activity index under noisy observations using high-frequency data. *J. Amer. Statist. Assoc.* **106** 558–568. MR2847970

KRISTENSEN, D. (2010). Nonparametric filtering of the realized spot volatility: A kernel-based approach. *Econometric Theory* **26** 60–93. MR2587103

LEPSKI, O. V., MAMMEN, E. and SPOKOINY, V. G. (1997). Optimal spatial adaptation to inhomogeneous smoothness: An approach based on kernel estimates with variable bandwidth selectors. *Ann. Statist.* **25** 929–947. MR1447734

MANCINI, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *G. Ist. Ital. Attuari* **64** 19–47.

MANCINI, C. (2009). Nonparametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* **36** 270–296. MR2528985

MANCINI, C., MATTIUSSI, V. and RENO, R. (2014). Spot volatility estimation using delta sequences. *Finance Stoch.* To appear.

MUNK, A. and SCHMIDT-HIEBER, J. (2010a). Nonparametric estimation of the volatility function in a high-frequency model corrupted by noise. *Electron. J. Stat.* **4** 781–821. MR2684388

MUNK, A. and SCHMIDT-HIEBER, J. (2010b). Lower bounds for volatility estimation in microstructure noise models. In *Borrowing Strength*: *Theory Powering Applications—a Festschrift for Lawrence D*. *Brown*. *Inst*. *Math*. *Stat*. *Collect*. **6** 43–55. IMS, Beachwood, OH. MR2798510

MYKLAND, P. A. and ZHANG, L. (2005). Comment: A selective overview of nonparametric methods in financial econometrics. *Statist*. *Sci*. **20** 347–350. MR2210224

PODOLSKIJ, M. and VETTER, M. (2009a). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli* **15** 634–658. MR2555193

PODOLSKIJ, M. and VETTER, M. (2009b). Bipower-type estimation in a noisy diffusion setting. *Stochastic Process*. *Appl*. **119** 2803–2831. MR2554029

REISS, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Ann*. *Statist*. **39** 772–802. MR2816338

ROSENBAUM, M. and TANKOV, P. (2011). Asymptotic results for time-changed Lévy processes sampled at hitting times. *Stochastic Process*. *Appl*. **121** 1607–1632. MR2802468

TODOROV, V. and TAUCHEN, G. (2012). The realized Laplace transform of volatility. *Econometrica* **80** 1105–1127. MR2963883

TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer, New York. MR2724359

VETTER, M. (2014). Inference on the Lévy measure in case of noisy observations. *Statist*. *Probab*. *Lett*. **87** 125–133. MR3168946

WINKEL, M. (2001). The recovery problem for time-changed Lévy processes. Unpublished manuscript.

WOERNER, J. H. C. (2007). Inference in Lévy-type stochastic volatility models. *Adv*. *in Appl*. *Probab*. **39** 531–549. MR2343676

ZHANG, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multiscale approach. *Bernoulli* **12** 1019–1043. MR2274854

ZHANG, L., MYKLAND, P. A. and AÏT-SAHALIA, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J*. *Amer*. *Statist*. *Assoc*. **100** 1394–1411. MR2236450

STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
WILBERFORCE ROAD
CAMBRIDGE CB3 0WB
UNITED KINGDOM
E-MAIL: a.bull@statslab.cam.ac.uk