Auxiliary Results for Activity Signature Functions

Viktor Todorov * and George Tauchen †

June 9, 2009

Abstract

This document contains additional results pertaining to the paper "Activity Signature Functions for High-Frequency Data Analysis," Todorov and Tauchen (2009). The notation is the same as in that paper.

*Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@kellogg.northwestern.edu

†Department of Economics, Duke University, Durham, NC 27708; e-mail: george.tauchen@duke.edu.

Contents

1 The Effects of Measurement Error

We analyze the effect of measurement error on our activity signature function. We will suppose that instead of the true process we observe the process with some error. For example, it is well known that at very high-frequency, financial data can be contaminated by microstructure noise; the literature on microstructure noise is vast, see Barndorff-Nielsen and Shephard (2007) for a timely and extensive review. We will focus only on the i.i.d. case. The following theorem gives the behavior of the realized power variation in this situation.

Theorem 1 *Let $(\epsilon_i)_i$ be an i.i.d. sequence of random variables with finite moments and set $\widetilde{\Upsilon}_t = \Upsilon_t + \epsilon_t$ where Υ is one of the processes X, Y or Z defined in the paper. Then, if assumption A1 holds, we have*

$$\Delta_n V(p, \widetilde{\Upsilon}, \Delta_n)_t \stackrel{u.c.p.}{\longrightarrow} K_p t,$$

where K_p is some constant depending on p.

As we can see from the theorem, the asymptotic behavior of the realized power variation is completely determined by the noise in the price process. Therefore, it is easy to see that for any p and $\Upsilon = Y, X$ or Z, $\widehat{\beta}_{(0,T]}(\Upsilon, p)$ will diverge to infinity as we start sampling more frequently. Intuitively, the measurement error becomes the "most active" part of the observed process, since its magnitude does not decrease as we start sampling more frequently.

Proof of Theorem 1: First, using a localization argument we can (and will) assume the stronger assumption A1′. Second, by a standard LLN we have

$$\Delta_n \sum_{i=1}^{[t/\Delta_n]} |\epsilon_{i\Delta_n} - \epsilon_{(i-1)\Delta_n}|^p \stackrel{\mathbb{P}}{\longrightarrow} t\mathbb{E}\left|\epsilon_u - \epsilon_s\right|^p,$$

for $u \neq s$. Thus, to prove Theorem 1 it suffices to show

$$\Delta_n \left| V(p, \widetilde{\Upsilon}, \Delta_n)_t - \sum_{i=1}^{[t/\Delta_n]} |\epsilon_{i\Delta_n} - \epsilon_{(i-1)\Delta_n}|^p \right| \stackrel{u.c.p.}{\longrightarrow} 0.$$

This follows from

1. the elementary inequality

   $$\left| |x+y|^p - |x|^p \right| \leq K_p \left(|x|^{p-1\vee 0} + |y|^{p-1} \right) |y|,$$

   for arbitrary x and y and K_p being some constant that depends on p.
2. the following well-known result (see Jacod (2008))

   $$\mathbb{E}_{i-1}^n |\Delta_i^n Z|^p \leq K \Delta_n^{p/2\wedge 1}.$$

3. application of Hölder inequality.

□

2 Scaled log-Power Variation

An alternative to using differences across time scales to determine the activity level is, from Woerner (2006), to compute a direct scaled log-power variation measure

$$\beta^*_{(0,T]}(X, p) = p \ln(\Delta_n) / \ln(\Delta_n V(p, X, k\Delta_n)_T).$$

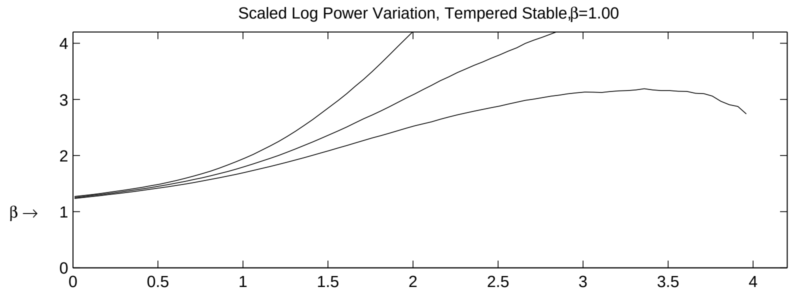

Figure 1: Scaled log power variation for the symmetric tempered stable process with $\beta = 1.00$ and other parameters given by **Case TS1.0** of Table 1; the figures show $\mathcal{B}^*(\Upsilon, p, \alpha)$ as a function of p for the quantiles $\alpha = 0.25$, $\alpha = 0.50$, and $\alpha = 0.75$, reading bottom up.

This measure asymptotically behaves as $\widehat{\beta}_{(0,T]}(X,p)$ in Section 4 of the text; however, Theorem 1 suggests a bias which is apparent in Figure 1. The figure shows three quantile plots of the function

$$\mathcal{B}^*(\Upsilon, p, \alpha) := \beta^*_{[\alpha N]}(p),$$

which is defined analogously to the quantile function in Section 4 of the paper. For the figure, Υ is simulated from case **TS1.0**. Evidentally, $\mathcal{B}^*(\Upsilon, p, \alpha)$ is off center and the interquantile range is quite large. The bias evident in Figure 1 is canceled out using the the two-scale approach.

3 Monte Carlo

3.1 Monte Carlo Scenarios

As noted in the paper, some readers, particularly in econometrics, prefer point estimates and formal test statistics to graphically-based interval estimates. Therefore, this appendix examines in detail via Monte Carlo the properties of the point estimator of the activity index presented in Section 4.3 of the paper. The estimator is an equi-weighted average of the activity signature function over a data-dependent interval. Clearly there are many other weighting schemes that could be used to develop estimators of activity, and exploring all of their properties is far beyond the scope of the paper and this appendix. Rather, our aim is to check if this particular easily-computed estimator performs reasonably well over a wide variety of estimation contexts. We also examine the properties of the formal test of a jump diffusion against the alternative of a pure jump model developed in Subsection 3.3 of the paper.

The various scenarios governing the dynamics of the process Z are shown in Table 1. The table shows the parameter settings and the implied population value of the activity index, denoted β_0, which is always determined by the dominant component of the Z process. Cases **A**–**I** and **N** are Lévy processes. In **A**–**C** the Lévy process is pure jump while in **D** it is a continuous/Brownian martingale. In cases **E**–**G** rare jumps, i.e. compound Poisson jumps, are added to a continuous martingale. Case **H** represents a particular challenge where the activity index is zero, as does **I** where the estimator faces the superposition of two Lévy processes with indices 1.00 and 1.50, respectively, with dominant component 1.50. Cases **J**–**M** are Lévy-driven OU or CARMA$(1,0)$, with strong persistence and very little mean reversion, where the background driving Lévy processes are Cases **A**–**D**, respectively. Case **N** is a continuous/Brownian martingale plus an infinitely active pure jump process. The values of the tuning parameters are set to be realistic for high-frequency financial data and were determined by knowledge gleamed from earlier studies, so the findings below need to be interpreted and qualified accordingly.

A very intuitive way to set up the sampling scheme is to think of the basic time unit, or time period, as one 24-hour day with observations taken at sampling intervals of 1-min, 5-min, 10-min, and 30-min intervals. The first two are typical of many financial applications while the latter two are relatively coarse. These values correspond to $M = 1440, 250, 144, 72$ observations per time period, respectively, or sampling intervals of width $\Delta = 0.0007, 0.0035, 0.0069, 0.02089$, which ranges over three orders of magnitude. Each Monte Carlo pseudo data set is comprised of multiple days worth of data, which we take as $N = 250, 1500, 3000$, or 1-year, 5-years, and 10-years worth of data. Thus M determines the sampling frequency and N the span of the data set, and the values here are typical of high frequency financial data sets. Of course for a given scenario in Table 1 the outcome only depends upon the width Δ of the sampling interval and the notions of "days" and "years" are simple labels to guide intuition.

Table 1: Monte Carlo Scenarios: Dynamics for $Z = Y + X$

Case	β_0	ρ	σ_1^2	σ_2^2	Jump Specification (X)
A	0.50	0.00	0.0	1.0	tempered stable with $A = 1$, $\beta = 0.5$ and $\lambda = 0.05$
B	1.00	0.00	0.0	1.0	tempered stable with $A = 1$, $\beta = 1.0$ and $\lambda = 0.05$
C	1.50	0.00	0.0	1.0	tempered stable with $A = 1$, $\beta = 1.5$ and $\lambda = 0.05$
D	2.00	0.00	0.8	0.0	none
E	2.00	0.00	0.8	1.0	rare-jump with $\lambda_J = 0.0400$, $\tau = 2.2361$
F	2.00	0.00	0.8	1.0	rare-jump with $\lambda_J = 0.3333$, $\tau = 0.7746$
G	2.00	0.00	0.8	1.0	rare-jump with $\lambda_J = 2.0000$, $\tau = 0.3162$
H	0.00	0.00	0.0	1.0	variance gamma: tempered stable with $A = 1$, $\beta = 0.0$ and $\lambda = 0.05$
I	1.50	0.00	0.0	1.0	superposition of two independent tempered stables: one with $A = 1$, $\beta = 1.5$ and $\lambda = 0.05$ and the other with $A = 1$, $\beta = 1.0$ and $\lambda = 0.05$
J	1.00	-0.0693	0.0	1.0	tempered stable with $A = 1$, $\beta = 0.5$ and $\lambda = 0.05$
K	1.00	-0.0693	0.0	1.0	tempered stable with $A = 1$, $\beta = 1.0$ and $\lambda = 0.05$
L	1.50	-0.0693	0.0	1.0	tempered stable with $A = 1$, $\beta = 1.5$ and $\lambda = 0.05$
M	1.50	-0.0693	0.8	0.0	none
N	2.00	0.00	0.8	0.2	tempered stable with $A = 1$, $\beta = 1.5$ and $\lambda = 0.05$

Note: With $Z = Y + X$ the continuous component Y is nontrivial if $\sigma_1^2 > 0$, and the jump component X is nontrivial if $\sigma_2^2 > 0$.

3.2 Precision of the Point Estimator of Activity

For each scenerio and (M, N) pair we generated 1000 Monte Carlo replicates, computed the point estimate of β_0 on a period-by-period basis, i.e., day-by-day, and took the median of the N values as the overall estimate for that replicate. This entails some loss in efficiency because it does not impose equality of β_0 across time time periods, but it mimics actual practice where things are done day-by-day to provide some robustness. Across the 1000 Monte Carlo replications we compute the usual robust summary statistics given by the median, inter-quartile range, and median-absolute-deviation about the the population value β_0.

In order to assess the practical accuracy of the estimator, we also computed error rates as the proportion of the 1000 replicates that the absolute deviation of the overall estimator from the population value β_0 is no more than the pre-specified value $\alpha = 0.15, 0.10, 0.05$. For the activity parameter $\beta_0 \in [0, 2]$, estimating it to within 0.10 of the true value seems sufficiently accurate for many applications, while to within 0.15 is a little loose and 0.05 very precise.

Tables 2–8 displays the summary measures from this Monte Carlo experiment. One immediately sees from the inter-quartile ranges and median absolute deviations that the estimator is very tightly concentrated around a value that is generally, but not always close to the true population variable. In other words, the sampling variance is nearly always very small and the bias generally small, or modest, but the bias can be quite large in certain extreme circumstances. The sampling distribution concentrates around a value that is usually close to the correct value β_0, but in some circumstances it concentrates around a value somewhat distant from the correct value.

The conclusions from Tables 2–8 are readily summarized. The key control parameter is the sampling frequency, M, or equivalently the sampling interval $\Delta = 1/M$. For sampling corresponding to 1-min and 5-min, the point estimator is generally quite accurate across essentially all scenarios and data spans N. On the other hand, for sampling at 10-min and coarser the estimator can sometimes be quite off the mark as in Cases **E**, **F**, **G**, and **H**, for example. The fact that reducing Δ sharply decreases the bias while increasing the span N only moderately reduces variance is not surprising because the asymptotic theory is always of the fill-in type where $\Delta \to 0$. A reasonable level of accuracy requires a large, dense, data set of the type that is commonly used in financial econometrics and presumably other applications as well.

Another point of note is that for values of the index β_0 at or above 0.50 the estimator is generally quite accurate at the higher sampling frequencies, while smaller values could require other approaches. Note that in case **H** $\beta_0 = 0$ the estimator concentrates at the value $\tau = 0.10$, which is the lower bound of the of interval used to average the activity signature function. Keep in mind, however, that if $\beta_0 = 0$, or is close to zero, then the plot of the activity signature function will reveal just that.

Table 2: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth $\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
A	250	1440	0.50	0.54	0.002	0.038	0.00	0.00	0.00
A	250	288	0.50	0.57	0.005	0.071	0.00	0.00	1.00
A	250	144	0.50	0.60	0.006	0.095	0.00	0.17	1.00
A	250	48	0.50	0.66	0.012	0.158	0.81	1.00	1.00
A	1500	1440	0.50	0.54	0.001	0.038	0.00	0.00	0.00
A	1500	288	0.50	0.57	0.002	0.071	0.00	0.00	1.00
A	1500	144	0.50	0.60	0.003	0.095	0.00	0.01	1.00
A	1500	48	0.50	0.66	0.004	0.158	0.98	1.00	1.00
A	3000	1440	0.50	0.54	0.001	0.038	0.00	0.00	0.00
A	3000	288	0.50	0.57	0.001	0.071	0.00	0.00	1.00
A	3000	144	0.50	0.60	0.001	0.095	0.00	0.00	1.00
A	3000	48	0.50	0.66	0.003	0.158	1.00	1.00	1.00
B	250	1440	1.00	1.04	0.005	0.040	0.00	0.00	0.00
B	250	288	1.00	1.07	0.012	0.071	0.00	0.00	0.99
B	250	144	1.00	1.09	0.018	0.092	0.00	0.27	1.00
B	250	48	1.00	1.13	0.027	0.131	0.18	0.95	1.00
B	1500	1440	1.00	1.04	0.001	0.040	0.00	0.00	0.00
B	1500	288	1.00	1.07	0.004	0.071	0.00	0.00	1.00
B	1500	144	1.00	1.09	0.006	0.092	0.00	0.04	1.00
B	1500	48	1.00	1.13	0.011	0.131	0.02	1.00	1.00
B	3000	1440	1.00	1.04	0.001	0.040	0.00	0.00	0.00
B	3000	288	1.00	1.07	0.003	0.071	0.00	0.00	1.00
B	3000	144	1.00	1.09	0.004	0.092	0.00	0.00	1.00
B	3000	48	1.00	1.13	0.008	0.131	0.00	1.00	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\hat{\beta} - \beta_0| > \alpha$.

Table 3: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth		
							$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
C	250	1440	1.50	1.49	0.009	0.009	0.00	0.00	0.00
C	250	288	1.50	1.53	0.021	0.031	0.00	0.00	0.10
C	250	144	1.50	1.56	0.034	0.057	0.00	0.02	0.60
C	250	48	1.50	1.56	0.046	0.055	0.01	0.11	0.57
C	1500	1440	1.50	1.49	0.003	0.009	0.00	0.00	0.00
C	1500	288	1.50	1.53	0.006	0.031	0.00	0.00	0.00
C	1500	144	1.50	1.55	0.015	0.054	0.00	0.00	0.65
C	1500	48	1.50	1.55	0.019	0.052	0.00	0.00	0.57
C	3000	1440	1.50	1.49	0.002	0.009	0.00	0.00	0.00
C	3000	288	1.50	1.53	0.005	0.031	0.00	0.00	0.00
C	3000	144	1.50	1.55	0.011	0.054	0.00	0.00	0.71
C	3000	48	1.50	1.55	0.015	0.052	0.00	0.00	0.57
D	250	1440	2.00	2.00	0.014	0.007	0.00	0.00	0.00
D	250	288	2.00	2.00	0.032	0.016	0.00	0.00	0.03
D	250	144	2.00	1.99	0.045	0.025	0.00	0.01	0.16
D	250	48	2.00	1.88	0.074	0.123	0.32	0.64	0.90
D	1500	1440	2.00	2.00	0.006	0.003	0.00	0.00	0.00
D	1500	288	2.00	2.00	0.012	0.007	0.00	0.00	0.00
D	1500	144	2.00	1.99	0.018	0.014	0.00	0.00	0.00
D	1500	48	2.00	1.87	0.029	0.125	0.12	0.89	1.00
D	3000	1440	2.00	2.00	0.004	0.002	0.00	0.00	0.00
D	3000	288	2.00	2.00	0.009	0.006	0.00	0.00	0.00
D	3000	144	2.00	1.99	0.012	0.013	0.00	0.00	0.00
D	3000	48	2.00	1.87	0.020	0.125	0.05	0.95	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\tilde{\beta} - \beta_0| > \alpha$.

Table 4: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth		
							$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
E	250	1440	2.00	1.99	0.013	0.008	0.00	0.00	0.00
E	250	288	2.00	1.99	0.031	0.019	0.00	0.00	0.07
E	250	144	2.00	1.97	0.043	0.032	0.00	0.01	0.24
E	250	48	2.00	1.85	0.067	0.145	0.46	0.81	0.96
E	1500	1440	2.00	1.99	0.005	0.006	0.00	0.00	0.00
E	1500	288	2.00	1.99	0.013	0.015	0.00	0.00	0.00
E	1500	144	2.00	1.97	0.018	0.026	0.00	0.00	0.03
E	1500	48	2.00	1.86	0.030	0.143	0.40	0.98	1.00
E	3000	1440	2.00	1.99	0.004	0.006	0.00	0.00	0.00
E	3000	288	2.00	1.98	0.009	0.015	0.00	0.00	0.00
E	3000	144	2.00	1.97	0.013	0.026	0.00	0.00	0.01
E	3000	48	2.00	1.86	0.020	0.142	0.31	1.00	1.00
F	250	1440	2.00	1.98	0.014	0.018	0.00	0.00	0.00
F	250	288	2.00	1.96	0.030	0.038	0.00	0.00	0.29
F	250	144	2.00	1.95	0.045	0.053	0.00	0.07	0.54
F	250	48	2.00	1.83	0.071	0.173	0.66	0.92	0.99
F	1500	1440	2.00	1.98	0.005	0.019	0.00	0.00	0.00
F	1500	288	2.00	1.96	0.012	0.038	0.00	0.00	0.08
F	1500	144	2.00	1.95	0.018	0.054	0.00	0.00	0.64
F	1500	48	2.00	1.83	0.027	0.175	0.89	1.00	1.00
F	3000	1440	2.00	1.98	0.003	0.019	0.00	0.00	0.00
F	3000	288	2.00	1.96	0.009	0.038	0.00	0.00	0.02
F	3000	144	2.00	1.94	0.012	0.055	0.00	0.00	0.72
F	3000	48	2.00	1.83	0.020	0.174	0.95	1.00	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\hat{\beta} - \beta_0| > \alpha$.

Table 5: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth		
							$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
G	250	1440	2.00	1.96	0.014	0.037	0.00	0.00	0.11
G	250	288	2.00	1.93	0.029	0.069	0.00	0.08	0.81
G	250	144	2.00	1.91	0.042	0.089	0.01	0.34	0.89
G	250	48	2.00	1.82	0.068	0.175	0.68	0.92	1.00
G	1500	1440	2.00	1.96	0.005	0.038	0.00	0.00	0.00
G	1500	288	2.00	1.93	0.012	0.070	0.00	0.00	0.98
G	1500	144	2.00	1.91	0.018	0.089	0.00	0.19	1.00
G	1500	48	2.00	1.83	0.026	0.173	0.89	1.00	1.00
G	3000	1440	2.00	1.96	0.004	0.038	0.00	0.00	0.00
G	3000	288	2.00	1.93	0.009	0.070	0.00	0.00	1.00
G	3000	144	2.00	1.91	0.012	0.089	0.00	0.11	1.00
G	3000	48	2.00	1.83	0.018	0.174	0.95	1.00	1.00
H	250	1440	0.00	0.10	0.000	0.101	0.00	1.00	1.00
H	250	288	0.00	0.11	0.001	0.109	0.00	1.00	1.00
H	250	144	0.00	0.12	0.001	0.124	0.00	1.00	1.00
H	250	48	0.00	0.18	0.003	0.183	1.00	1.00	1.00
H	1500	1440	0.00	0.10	0.000	0.101	0.00	1.00	1.00
H	1500	288	0.00	0.11	0.000	0.109	0.00	1.00	1.00
H	1500	144	0.00	0.12	0.000	0.124	0.00	1.00	1.00
H	1500	48	0.00	0.18	0.001	0.183	1.00	1.00	1.00
H	3000	1440	0.00	0.10	0.000	0.101	0.00	1.00	1.00
H	3000	288	0.00	0.11	0.000	0.109	0.00	1.00	1.00
H	3000	144	0.00	0.12	0.000	0.124	0.00	1.00	1.00
H	3000	48	0.00	0.18	0.001	0.183	1.00	1.00	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\hat{\beta} - \beta_0| > \alpha$.

Table 6: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth		
							$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
I	250	1440	1.50	1.43	0.008	0.068	0.00	0.00	1.00
I	250	288	1.50	1.45	0.017	0.047	0.00	0.00	0.42
I	250	144	1.50	1.47	0.024	0.033	0.00	0.00	0.16
I	250	48	1.50	1.47	0.047	0.034	0.00	0.01	0.28
I	1500	1440	1.50	1.43	0.002	0.069	0.00	0.00	1.00
I	1500	288	1.50	1.45	0.006	0.048	0.00	0.00	0.35
I	1500	144	1.50	1.47	0.010	0.034	0.00	0.00	0.02
I	1500	48	1.50	1.47	0.016	0.031	0.00	0.00	0.07
I	3000	1440	1.50	1.43	0.002	0.069	0.00	0.00	1.00
I	3000	288	1.50	1.45	0.005	0.048	0.00	0.00	0.33
I	3000	144	1.50	1.47	0.006	0.034	0.00	0.00	0.01
I	3000	48	1.50	1.47	0.011	0.032	0.00	0.00	0.01
J	250	1440	1.00	0.86	0.014	0.139	0.14	1.00	1.00
J	250	288	1.00	0.83	0.018	0.171	0.91	1.00	1.00
J	250	144	1.00	0.82	0.021	0.177	0.95	1.00	1.00
J	250	48	1.00	0.83	0.021	0.166	0.84	1.00	1.00
J	1500	1440	1.00	0.86	0.003	0.137	0.00	1.00	1.00
J	1500	288	1.00	0.83	0.005	0.169	1.00	1.00	1.00
J	1500	144	1.00	0.82	0.006	0.176	1.00	1.00	1.00
J	1500	48	1.00	0.84	0.006	0.165	1.00	1.00	1.00
J	3000	1440	1.00	0.86	0.002	0.137	0.00	1.00	1.00
J	3000	288	1.00	0.83	0.003	0.169	1.00	1.00	1.00
J	3000	144	1.00	0.82	0.003	0.176	1.00	1.00	1.00
J	3000	48	1.00	0.84	0.004	0.164	1.00	1.00	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\hat{\beta} - \beta_0| > \alpha$.

Table 7: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth		
							$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
K	250	1440	1.00	1.04	0.005	0.044	0.00	0.00	0.06
K	250	288	1.00	1.08	0.012	0.078	0.00	0.01	1.00
K	250	144	1.00	1.10	0.018	0.101	0.00	0.52	1.00
K	250	48	1.00	1.15	0.028	0.149	0.49	0.99	1.00
K	1500	1440	1.00	1.04	0.002	0.044	0.00	0.00	0.00
K	1500	288	1.00	1.08	0.004	0.078	0.00	0.00	1.00
K	1500	144	1.00	1.10	0.006	0.100	0.00	0.52	1.00
K	1500	48	1.00	1.15	0.011	0.147	0.35	1.00	1.00
K	3000	1440	1.00	1.04	0.001	0.044	0.00	0.00	0.00
K	3000	288	1.00	1.08	0.003	0.078	0.00	0.00	1.00
K	3000	144	1.00	1.10	0.004	0.100	0.00	0.57	1.00
K	3000	48	1.00	1.15	0.007	0.147	0.28	1.00	1.00
L	250	1440	1.50	1.50	0.008	0.004	0.00	0.00	0.00
L	250	288	1.50	1.55	0.021	0.048	0.00	0.00	0.44
L	250	144	1.50	1.57	0.031	0.069	0.00	0.10	0.80
L	250	48	1.50	1.58	0.047	0.075	0.02	0.25	0.77
L	1500	1440	1.50	1.50	0.003	0.001	0.00	0.00	0.00
L	1500	288	1.50	1.55	0.007	0.047	0.00	0.00	0.27
L	1500	144	1.50	1.57	0.011	0.068	0.00	0.00	0.98
L	1500	48	1.50	1.57	0.019	0.074	0.00	0.02	0.97
L	3000	1440	1.50	1.50	0.002	0.001	0.00	0.00	0.00
L	3000	288	1.50	1.55	0.005	0.047	0.00	0.00	0.20
L	3000	144	1.50	1.57	0.007	0.068	0.00	0.00	1.00
L	3000	48	1.50	1.57	0.013	0.074	0.00	0.00	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\hat{\beta} - \beta_0| > \alpha$.

Table 8: Summary Statistics for 1000 Replicates of $\tilde{\beta}$ Defined in Subsection 4.3

Case	N	M	β_0	MED	IQR	MAD	Error rates: α-away from the truth		
							$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$
M	250	1440	2.00	2.00	0.014	0.007	0.00	0.00	0.00
M	250	288	2.00	2.00	0.032	0.016	0.00	0.00	0.04
M	250	144	2.00	1.99	0.044	0.025	0.00	0.00	0.14
M	250	48	2.00	1.88	0.074	0.122	0.30	0.65	0.89
M	1500	1440	2.00	2.00	0.005	0.003	0.00	0.00	0.00
M	1500	288	2.00	2.00	0.013	0.007	0.00	0.00	0.00
M	1500	144	2.00	1.99	0.017	0.014	0.00	0.00	0.00
M	1500	48	2.00	1.88	0.029	0.124	0.11	0.86	1.00
M	3000	1440	2.00	2.00	0.004	0.002	0.00	0.00	0.00
M	3000	288	2.00	2.00	0.009	0.006	0.00	0.00	0.00
M	3000	144	2.00	1.99	0.012	0.013	0.00	0.00	0.00
M	3000	48	2.00	1.88	0.020	0.125	0.03	0.95	1.00
N	250	1440	2.00	1.97	0.013	0.025	0.00	0.00	0.01
N	250	288	2.00	1.97	0.031	0.029	0.00	0.00	0.19
N	250	144	2.00	1.96	0.045	0.039	0.00	0.02	0.35
N	250	48	2.00	1.86	0.072	0.139	0.42	0.77	0.95
N	1500	1440	2.00	1.97	0.005	0.026	0.00	0.00	0.00
N	1500	288	2.00	1.97	0.016	0.030	0.00	0.00	0.02
N	1500	144	2.00	1.96	0.015	0.038	0.00	0.00	0.16
N	1500	48	2.00	1.86	0.027	0.142	0.35	0.98	1.00
N	3000	1440	2.00	1.97	0.003	0.026	0.00	0.00	0.00
N	3000	288	2.00	1.97	0.012	0.031	0.00	0.00	0.00
N	3000	144	2.00	1.96	0.012	0.039	0.00	0.00	0.09
N	3000	48	2.00	1.86	0.019	0.142	0.30	1.00	1.00

Note: N is the number of units of time in a simulation; M is the number of high-frequency observations per unit of time. β_0 is the population value. MED is the median, IQR the inter-quartile range, MAD the median absolute deviation, each computed over the 1000 replications. Error rates are the proportion of replicates where $|\hat{\beta} - \beta_0| > \alpha$.

3.3 Size and Power of the Test for Continuous Martingale

Tables 9–11 summarize the size and power properties of the test for a continuous martingale proposed in Section 3.3 of the paper. The scenarios are as in Table 1 above and the sampling rate M as per Tables 2–8. For each scenario–M pair, the test is conducted on each of 10000 replicated periods — i.e., "days" — mimicking how such tests are conducted in actual practice.

The conclusions from these three tables accord well with those of the preceding Section 3.2. For the highest sampling frequency (1-min) the size and power properties are excellent, and with 5-min sampling the test is only slightly undersized but with still quite good power properties. On the other hand, at the courser sampling intervals there can be large size distortions and the test thereby unreliable. The size and power appear appropriate, at least for sampling frequencies commonly used in finance. These findings very likely carry over to other disciplines where large dense data sets are also available.

Table 9: Size and Power of Test for Presence of Continuous Martingale

Sampling Frequency	Percentage of Rejection of Null Hypothesis		
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
		Case A (H_0 **false**)	
$M = 1440$ (1 min)	0.9998	0.9998	0.9998
$M = 288$ (5 min)	0.9942	0.9981	0.9990
$M = 144$ (10 min)	0.9735	0.9914	0.9942
$M = 48$ (30 min)	0.5143	0.8721	0.9464
		Case B (H_0 **false**)	
$M = 1440$ (1 min)	0.9999	0.9999	0.9999
$M = 288$ (5 min)	0.9823	0.9964	0.9977
$M = 144$ (10 min)	0.7121	0.9321	0.9700
$M = 48$ (30 min)	0.0274	0.3490	0.6229
		Case C (H_0 **false**)	
$M = 1440$ (1 min)	0.9967	0.9995	0.9998
$M = 288$ (5 min)	0.2634	0.6207	0.7736
$M = 144$ (10 min)	0.0428	0.2942	0.4920
$M = 48$ (30 min)	0.0021	0.0374	0.1682
		Case D (H_0 **true**)	
$M = 1440$ (1 min)	0.0052	0.0407	0.0913
$M = 288$ (5 min)	0.0012	0.0248	0.0743
$M = 144$ (10 min)	0.0002	0.0185	0.0623
$M = 48$ (30 min)	0.0086	0.0134	0.0480

Note: *The test is one-sided and is based on asymptotic distribution under the null of* $\log(\widehat{\beta}_t(\Upsilon, p))$ *for* $p = 0.9$*, given in Theorem 3 in the paper. The number of Monte Carlo replications is* $10,000$ *units of time and* α *denotes the (asymptotic) size of the test.*

Table 10: Size and Power of Test for Presence of Continuous Martingale

Sampling Frequency	Percentage of Rejection of Null Hypothesis		
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
		Case E (H_0 **true**)	
$M = 1440$ (1 min)	0.0052	0.0407	0.0913
$M = 288$ (5 min)	0.0012	0.0248	0.0743
$M = 144$ (10 min)	0.0002	0.0185	0.0623
$M = 48$ (30 min)	0.0086	0.0134	0.0480
		Case F (H_0 **true**)	
$M = 1440$ (1 min)	0.0052	0.0407	0.0913
$M = 288$ (5 min)	0.0012	0.0248	0.0743
$M = 144$ (10 min)	0.0002	0.0185	0.0623
$M = 48$ (30 min)	0.0086	0.0134	0.0480
		Case G (H_0 **true**)	
$M = 1440$ (1 min)	0.0052	0.0407	0.0913
$M = 288$ (5 min)	0.0012	0.0248	0.0743
$M = 144$ (10 min)	0.0002	0.0185	0.0623
$M = 48$ (30 min)	0.0086	0.0134	0.0480
		Case H (H_0 **false**)	
$M = 1440$ (1 min)	0.9992	0.9993	0.9994
$M = 288$ (5 min)	0.9909	0.9940	0.9957
$M = 144$ (10 min)	0.9780	0.9860	0.9898
$M = 48$ (30 min)	0.8831	0.9307	0.9583
		Case I (H_0 **false**)	
$M = 1440$ (1 min)	0.9997	0.9999	1.0000
$M = 288$ (5 min)	0.4653	0.7842	0.8887
$M = 144$ (10 min)	0.0999	0.4401	0.6379
$M = 48$ (30 min)	0.0012	0.062	0.2392

Note: *The test is one-sided and is based on asymptotic distribution under the null of* $\log(\widehat{\beta}_t(\Upsilon, p))$ *for* $p = 0.9$*, given in Theorem 3 in the paper. The number of Monte Carlo replications is* $10,000$ *units of time and* α *denotes the (asymptotic) size of the test.*

Table 11: Size and Power of Test for Presence of Continuous Martingale

Sampling Frequency	Percentage of Rejection of Null Hypothesis		
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
		Case J (H_0 **false**)	
$M = 1440$ (1 min)	0.9997	0.9997	0.9997
$M = 288$ (5 min)	0.9967	0.9981	0.9987
$M = 144$ (10 min)	0.9815	0.9937	0.9959
$M = 48$ (30 min)	0.2714	0.8325	0.9326
		Case K (H_0 **false**)	
$M = 1440$ (1 min)	1.0000	1.0000	1.0000
$M = 288$ (5 min)	0.9824	0.9960	0.9977
$M = 144$ (10 min)	0.6940	0.9318	0.9706
$M = 48$ (30 min)	0.0115	0.3094	0.5969
		Case L (H_0 **false**)	
$M = 1440$ (1 min)	0.9953	0.9994	0.9998
$M = 288$ (5 min)	0.2216	0.5802	0.7462
$M = 144$ (10 min)	0.0259	0.2608	0.4640
$M = 48$ (30 min)	0.0021	0.0251	0.1429
		Case M (H_0 **true**)	
$M = 1440$ (1 min)	0.0050	0.0406	0.0912
$M = 288$ (5 min)	0.0020	0.0258	0.0735
$M = 144$ (10 min)	0.0006	0.0193	0.0640
$M = 48$ (30 min)	0.0103	0.0148	0.0486
		Case N (H_0 **true**)	
$M = 1440$ (1 min)	0.0083	0.0600	0.1308
$M = 288$ (5 min)	0.0021	0.0343	0.0913
$M = 144$ (10 min)	0.0006	0.0219	0.0758
$M = 48$ (30 min)	0.0084	0.0131	0.0471

Note: *The test is one-sided and is based on asymptotic distribution under the null of* $\log(\widehat{\beta}_t(\Upsilon, p))$ *for* $p = 0.9$, *given in Theorem 3 in the paper. The number of Monte Carlo replications is* $10,000$ *units of time and* α *denotes the (asymptotic) size of the test.*

4 Proofs of Theoretical Results

4.1 Preliminary Results

First, it is easy to see that the law of the pure-jump model X_t, given with equation (2.2), can be generated with the following process defined on some different probability space (we will still call it X_t)

$$\begin{aligned} X_t &= \int_0^t b_{2s}ds + \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}_+} \sigma_{2s-}\kappa(x)1_{\{y<a_s\}}\tilde{\underline{\mu}}(ds,dx,dy) \\ &+ \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}_+} \sigma_{2s-}\kappa'(x)1_{\{y<a_s\}}\underline{\mu}(ds,dx,dy), \end{aligned} \tag{4.1}$$

where now $\underline{\mu}$ is a Poisson measure but on $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$ and with compensator $ds \otimes \nu(x)dx \otimes dy$; $\tilde{\underline{\mu}}(ds,dx,dy)$ is the compensated version of $\underline{\mu}$. The process X_t in (4.1) is written as a (stochastic) integral with respect to a time-homogenous Poisson random measure, but we note that $\underline{\mu}$ is defined on a three dimensional space. Intuitively, the first dimension is the time, the second one is the size of the jumps (but note that the jumps are multiplied by σ_{2t-}). The role of the third dimension is to generate thinning of the jumps according to the process a_t and thus create time-varying intensity of the jumps in the process X_t.

In all of the proofs in this section we will assume the representation in (4.1) for X_t, and we will work with the corresponding probability space and filtration which support it. This is convenient as when Δ_n changes we do not need to make a change of the probability space in our proofs. We start with showing some preliminary results. Denote with $\xrightarrow{\mathcal{L}}$ convergence in law on the space $\mathbb{D}(\mathbb{R})$ of one-dimensional càdlàg functions: $\mathbb{R}_+ \to \mathbb{R}$, which is equipped with the Skorokhod topology. We have the following two Lemmas.

Lemma 1 *Let Ψ_t denote a Lévy process with characteristic triplet $(b,0,F)$ with respect to some truncation function $\kappa(x)$, where F is a Lévy measure with $F(dx) = \nu(x)1_{\{|x|\le\epsilon\}}dx$ and $\nu(x)$ is the density defined in (2.4), while ϵ is an arbitrary positive number. Further, let $b = \int_{\mathbb{R}} \kappa(x)1_{\{|x|\le\epsilon\}}\nu(x)dx$ if $\beta < 1$ and $b = 0$ if $\beta \ge 1$. Finally, assume that ν is symmetric when $\beta = 1$. Denote with L_t another Lévy process with characteristic triplet $(b_1, 0, F_1)$ (with respect to the same truncation function). $b_1 = -\int_{\mathbb{R}} \kappa'(x)\nu_1(x)dx$ if $\beta \ge 1$ and $b_1 = \int_{\mathbb{R}} \kappa(x)\nu_1(x)dx$ for $\beta < 1$. F_1 is a Lévy measure with $F_1(dx) = \nu_1(x)dx$ where $\nu_1(x)$ is defined in (2.5). Finally, assume $\kappa(x)$ is F_1-a.s. continuous and in addition that it is symmetric when $\beta = 1$. Then as $\Delta_n \to 0$ we have:*

(a)

$$\Delta_n^{-1/\beta}\Psi_{\Delta_n t} \xrightarrow{\mathcal{L}} L_t, \tag{4.2}$$

(b)

$$\Delta_n^{-p/\beta}\mathbb{E}\left(|\Psi_{\Delta_n t}|^p\right) \longrightarrow \mathbb{E}\left(|L_t|^p\right), \quad \textit{locally uniformly in } t \textit{ for some } p < \beta. \tag{4.3}$$

Proof:
Part (a). Since $\Delta_n^{-1/\beta}\Psi_{\Delta_n t}$ is a Lévy process to prove the convergence of the sequence we need to show the convergence of its characteristics, see e.g. Jacod and Shiryaev (2003), Corollary III.3.6. The case $\beta < 1$ is straightforward, we show the case $\beta \ge 1$. We need to establish the following for $\Delta_n \to 0$

$$\Delta_n \int_{\mathbb{R}} \left(\kappa(\Delta_n^{-1/\beta}x) - \Delta_n^{-1/\beta}\kappa(x)\right)\nu(x)1_{\{|x|\le\epsilon\}}dx \longrightarrow -\int_{\mathbb{R}} \kappa'(x)\nu_1(x)dx, \tag{4.4}$$

$$\Delta_n \int_{\mathbb{R}} \kappa^2(\Delta_n^{-1/\beta}x)\nu(x)1_{\{|x|\le\epsilon\}}dx \longrightarrow \int_{\mathbb{R}} \kappa^2(x)\nu_1(x)dx, \tag{4.5}$$

$$\Delta_n \int_{\mathbb{R}} g(\Delta_n^{-1/\beta}x)\nu(x)1_{\{|x|\leq\epsilon\}}dx \longrightarrow \int_{\mathbb{R}} g(x)\nu_1(x)dx, \tag{4.6}$$

where g is an arbitrary continuous and bounded function on $\mathbb{R}$, which is 0 around 0.

First, the convergence in (4.5) and (4.6) trivially follow from a change of variable in the integration and the fact that the function ϕ in (2.5) is bounded around the origin.

Therefore, we are left with showing (4.4). For $\beta = 1$ this follows from the symmetry of the truncation function and ν. We show (4.4) therefore only in the case $\beta > 1$. First, we have

$$\begin{aligned}
&\Delta_n \int_{\mathbb{R}} \left(\kappa(\Delta_n^{-1/\beta}x) - \Delta_n^{-1/\beta}\kappa(x)\right)\nu_1(x)1_{\{|x|\leq\epsilon\}}dx \\
&= \Delta_n \int_{\mathbb{R}} \left(\kappa(\Delta_n^{-1/\beta}x) - \Delta_n^{-1/\beta}x\right)\nu_1(x)1_{\{|x|\leq\epsilon\}}dx + \Delta_n^{1-1/\beta}\int_{\mathbb{R}}(x-\kappa(x))\,\nu_1(x)1_{\{|x|\leq\epsilon\}}dx \\
&\longrightarrow -\int_{\mathbb{R}} \kappa'(x)\nu_1(x)dx.
\end{aligned} \tag{4.7}$$

Thus we will prove (4.4) if we can show

$$\Delta_n \int_{\mathbb{R}} \left(\kappa(\Delta_n^{-1/\beta}x) - \Delta_n^{-1/\beta}\kappa(x)\right)\nu_2(x)dx \longrightarrow 0. \tag{4.8}$$

It suffices to show (4.4) for $\kappa(x) = x1\{|x|\leq 1\}$. For Δ_n small enough we have

$$\begin{aligned}
&\Delta_n \int_{\mathbb{R}} \left(\Delta_n^{-1/\beta}\kappa(x) - \kappa(\Delta_n^{-1/\beta}x)\right)\nu_2(x)dx \\
&= \Delta_n^{1-1/\beta}\int_{\Delta_n^{1/\beta}}^{1} x\nu_2(x)dx + \Delta_n^{1-1/\beta}\int_{-1}^{\Delta_n^{1/\beta}} x\nu_2(x)dx \\
&\leq \Delta_n^{1-1/\beta}\int_{\Delta_n^{1/\beta}}^{1} x^{-\beta'}\phi(x)dx + \Delta_n^{1-1/\beta}\int_{-1}^{-\Delta_n^{1/\beta}} |x|^{-\beta'}\phi(x)dx.
\end{aligned} \tag{4.9}$$

For sufficiently small Δ_n, using the fact that $\phi(x)$ slowly varies around zero, has a finite limit at 0 and is integrable around 0, we have

$$\int_{\Delta_n^{1/\beta}}^{1} x^{-\beta'}\phi(x)dx \leq K + K\phi(\Delta_n^{1/\beta})\Delta_n^{1/\beta-\beta'/\beta}, \tag{4.10}$$

for some constant K. The same result obviously holds for $\int_{-1}^{-\Delta_n^{1/\beta}} |x|^{-\beta'}\phi(x)dx$, and from here (recall we are looking at the case $\beta > 1$) we have the result in (4.8). Thus, combining (4.7) and (4.8), we have the result in (4.4) and therefore (4.2) is established.

Part (b). First we establish the pointwise convergence for arbitrary $t > 0$. Given the result in Part (a), we need only show that for some p' such that $p < p' < \beta$ we have

$$\sup_{\Delta_n} \mathbb{E}\left(\Delta_n^{-p'/\beta}|\Psi_{\Delta_n t}|^{p'}\right) < \infty. \tag{4.11}$$

In order to consider the cases $\beta \geq 1$ and $\beta < 1$ together we make a slight abuse of notation. In what follows we assume that $\kappa(x) \equiv 0$ and "turn off" the drift b when $\beta < 1$, while when $\beta \geq 1$ we make no change and $\kappa(x)$ continues to be a true truncation function (i.e. it has bounded support and coincides with the identity around the origin). With this assumption on $\kappa(x)$, denoting with $\underline{\underline{\mu}}$ the jump measure associated with the process Ψ_t and with $\tilde{\underline{\underline{\mu}}}$ its compensated version, we can write $\Delta_n^{-1/\beta}\Psi_{\Delta_n t}$ (in all cases of β) as

$$\Delta_n^{-1/\beta}\Psi_{\Delta_n t} = \Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{\mathbb{R}} \kappa(x)\tilde{\underline{\underline{\mu}}}(ds,dx) + \Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{\mathbb{R}} \kappa'(x)\underline{\underline{\mu}}(ds,dx). \tag{4.12}$$

Then, we can proceed with the following decomposition

$$\begin{aligned}\left|\Delta_n^{-1/\beta}\Psi_{\Delta_n t}\right|^{p'} \quad \le \quad & K\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{|x|>s^{1/\beta}}\kappa(x)\underline{\underline{\tilde{\mu}}}(ds,dx)\right|^{p'} \\ & +K\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{|x|\le s^{1/\beta}}\kappa(x)\underline{\underline{\tilde{\mu}}}(ds,dx)\right|^{p'} \\ & +K\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{\mathbb{R}}\kappa'(x)\underline{\underline{\mu}}(ds,dx)\right|^{p'}, \end{aligned} \tag{4.13}$$

where K is some constant. Obviously, given the fact that $\kappa(x) \equiv 0$ if $\beta < 1$, for the first two terms on the right hand side of (4.13) we look only at the case $\beta \ge 1$. For the second integral on the right-hand side of (4.11) for some $1 \le q \le 2$ such that $q > \beta$ we have

$$\begin{aligned}&\mathbb{E}\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{|x|\le s^{1/\beta}}\kappa(x)\underline{\underline{\tilde{\mu}}}(ds,dx)\right|^{p'} \le \left(\mathbb{E}\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{|x|\le s^{1/\beta}}\kappa(x)\underline{\underline{\tilde{\mu}}}(ds,dx)\right|^{q}\right)^{p'/q} \\ &\le K\left(\Delta_n^{-q/\beta}\mathbb{E}\left(\int_0^{\Delta_n t}\int_{|x|\le s^{1/\beta}}|\kappa(x)|^q ds\nu(x)dx\right)\right)^{p'/q} \le K, \end{aligned} \tag{4.14}$$

where the first inequality follows from the fact that $p' < q$; the second one is an application of Burkholder-Davis-Gundy inequality (see e.g. Protter (2004)); and the third one uses the fact that $q \le 2$, the properties of the function $\phi(x)$ and a similar argument as in (4.10).

For the first integral on the right-hand side of (4.13) if $\beta > 1$ we can pick some q such that $\max\{p', 1\} \le q < \beta$ and make exactly the same steps as in (4.14) to get

$$\mathbb{E}\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{|x|>s^{1/\beta}}\kappa(x)\underline{\underline{\tilde{\mu}}}(ds,dx)\right|^{p'} \le K. \tag{4.15}$$

When $\beta = 1$, ν is symmetric so the integration with respect to $\underline{\underline{\mu}}$ is the same as the integration with respect to the compensated measure $\underline{\underline{\tilde{\mu}}}$ around the origin. Thus, when $\beta = 1$, for the first integral on the right-hand side of (4.13) we can write

$$\begin{aligned}&\mathbb{E}\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{|x|>s^{1/\beta}}\kappa(x)\underline{\underline{\tilde{\mu}}}(ds,dx)\right|^{p'} \\ &= \Delta_n^{-p'/\beta}\mathbb{E}\left|\int_0^{\Delta_n t}\int_{|x|>s^{1/\beta}}\kappa(x)\underline{\underline{\mu}}(ds,dx) - \int_0^{\Delta_n t}\int_{|x|>s^{1/\beta}}\kappa(x)ds\nu(dx)\right|^{p'} \\ &\le \Delta_n^{-p'/\beta}\mathbb{E}\left(\int_0^{\Delta_n t}\int_{|x|>s^{1/\beta}}|\kappa(x)|^{p'}\nu(ds,dx)\right) + K \le K, \end{aligned} \tag{4.16}$$

where for the first inequality we made use of the fact that $p' \le 1$. Finally, for the last integral on the right-hand side of (4.13), since the jump measure has a bounded support (i.e. there are no jumps higher than ϵ in absolute value) the expectation exists and for the case $\beta \ge 1$ we trivially have

$$\mathbb{E}\left|\Delta_n^{-1/\beta}\int_0^{\Delta_n t}\int_{\mathbb{R}}\kappa'(x)\underline{\underline{\mu}}(ds,dx)\right|^{p'} \le K\Delta_n^{1-p'/\beta}. \tag{4.17}$$

When $\beta < 1$ (recall $\kappa(x) = 0$ in this case) we split

$$\int_0^{\Delta_n t} \int_{\mathbb{R}} \kappa'(x) \underline{\underline{\mu}}(ds, dx) = \int_0^{\Delta_n t} \int_{|x| \le s^{1/\beta}} x \underline{\underline{\mu}}(ds, dx) + \int_0^{\Delta_n t} \int_{|x| > s^{1/\beta}} x \underline{\underline{\mu}}(ds, dx),$$

and use the fact that for arbitrary $0 < m \le 1$ and some a and b we have $|a+b|^m \le |a|^m + |b|^m$, to prove

$$\mathbb{E} \left| \Delta_n^{-1/\beta} \int_0^{\Delta_n t} \int_{\mathbb{R}} \kappa'(x) \underline{\underline{\mu}}(ds, dx) \right|^{p'} \le K. \tag{4.18}$$

Combining (4.14)-(4.18) we get (4.13), and thus the pointwise convergence.

To prove the convergence of $\Delta_n^{-p/\beta} \mathbb{E}\left(|\Psi_{\Delta_n t}|^p\right)$ for the local uniform topology on t we need only show that this sequence is relatively compact. For this we use the Ascoli-Arzela's Theorem (see e.g. Jacod and Shiryaev (2003), Theorem VI.1.5). The relative compactness of $\Delta_n^{-p/\beta} \mathbb{E}\left(|\Psi_{\Delta_n t}|^p\right)$ follows from the fact that the bounds in (4.14)-(4.17) are continuous in t. This proves part (b) of the Lemma. □

Lemma 2 *Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the process X_t is defined via (4.1). Let u_s, l_s and h_s be some processes on this probability space which are adapted, bounded, and have càdlàg paths. Assume that A2 holds for $\nu(x)$ and that $\nu(x)$ is symmetric when $\beta = 1$. Finally, if $\beta < 1$ set $\kappa(x) \equiv 0$ and when $\beta \ge 1$ let $\kappa(x)$ be a truncation function, which is symmetric in the case $\beta = 1$. Then, for arbitrary $\epsilon > 0$ and some $q_1 > \beta$ and $p \le q_2 < \beta$ we have*

$$\begin{aligned} &\mathbb{E} \left| \int_0^{\Delta_n} \int_{|x| \le \epsilon} \int_{\mathbb{R}_+} u_{s-} \kappa(x) 1_{\{l_{s-} < y < h_{s-}\}} \underline{\tilde{\mu}}(ds, dx, dy) + \int_0^{\Delta_n} \int_{|x| \le \epsilon} \int_{\mathbb{R}_+} u_{s-} \kappa'(x) 1_{\{l_{s-} < y < h_{s-}\}} \underline{\mu}(ds, dx, dy) \right|^p \\ &\le K \Delta_n^{p/\beta} \left(\mathbb{E} \left(\sup_{0 \le s \le \Delta_n} \{|u_s|^{q_1} |h_s - l_s|\} \right) \right)^{p/q_1} + K \Delta_n^{p/\beta} \left(\mathbb{E} \left(\sup_{0 \le s \le \Delta_n} \{|u_s|^{q_2} |h_s - l_s|\} \right) \right)^{p/q_2}, \end{aligned} \tag{4.19}$$

where the constant K does not depend on Δ_n.

Proof: The proof of the Lemma is very similar to the proof of Lemma 1, part (b). First we make the decomposition

$$\begin{aligned} &\int_0^{\Delta_n} \int_{|x| \le \epsilon} \int_{\mathbb{R}_+} u_{s-} 1_{\{l_{s-} < y < h_{s-}\}} \kappa(x) \underline{\tilde{\mu}}(ds, dx, dy) \\ &= \int_0^{\Delta_n} \int_{s^{1/\beta} < |x| \le \epsilon} \int_{\mathbb{R}_+} u_{s-} 1_{\{l_{s-} < y < h_{s-}\}} \kappa(x) \underline{\tilde{\mu}}(ds, dx, dy) \\ &+ \int_0^{\Delta_n} \int_{|x| \le s^{1/\beta} \wedge \epsilon} \int_{\mathbb{R}_+} u_{s-} 1_{\{l_{s-} < y < h_{s-}\}} \kappa(x) \underline{\tilde{\mu}}(ds, dx, dy). \end{aligned} \tag{4.20}$$

We proceed with bounding the p-th absolute moment of the two integrals on the right hand side of the above equation for the case $\beta \ge 1$. For the last one we choose some $1 \le q \le 2$, such that

$q > \beta$ and apply the Burkholder-Davis-Gundy inequality to get

$$\mathbb{E}\left|\int_0^{\Delta_n}\int_{|x|\le s^{1/\beta}\wedge\epsilon}\int_{\mathbb{R}_+} u_{s-}1_{\{l_{s-}<y<h_{s-}\}}\kappa(x)\underline{\tilde{\mu}}(ds,dx,dy)\right|^p$$

$$\le \left(\mathbb{E}\left|\int_0^{\Delta_n}\int_{|x|\le s^{1/\beta}\wedge\epsilon}\int_{\mathbb{R}_+} u_{s-}1_{\{l_{s-}<y<h_{s-}\}}\kappa(x)\underline{\tilde{\mu}}(ds,dx,dy)\right|^q\right)^{p/q}$$

$$\le K\left(\mathbb{E}\left(\int_0^{\Delta_n}\int_{|x|\le s^{1/\beta}}\int_{\mathbb{R}_+} |u_{s-}|^q|\kappa(x)|^q 1_{\{l_{s-}<y<h_{s-}\}}\nu(x)dsdxdy\right)\right)^{p/q}$$

$$\le K\Delta_n^{p/\beta}\left(\mathbb{E}\sup_{0\le s\le\Delta_n}\{|u_s|^q|h_s-l_s|\}\right)^{p/q}, \tag{4.21}$$

where the constant K is changing from line to line. For the second integral in (4.20) we proceed similar to (4.15) and (4.16) to get

$$\mathbb{E}\left|\int_0^{\Delta_n}\int_{s^{1/\beta}<|x|\le\epsilon}\int_{\mathbb{R}_+} u_{s-}1_{\{l_{s-}<y<h_{s-}\}}\kappa(x)\underline{\tilde{\mu}}(ds,dx,dy)\right|^p$$

$$\le K\Delta_n^{p/\beta}\left(\mathbb{E}\left(\sup_{0\le s\le\Delta_n}\{|u_s|^q|h_s-l_s|\}\right)\right)^{p/q}, \tag{4.22}$$

for some q such that $\max\{p,1\}\le q<\beta$ if $\beta>1$ and $q=p$ if $\beta=1$.

Similar transformations as in (4.17)-(4.18) and using the fact that we integrate over the bounded region $|x|\le\epsilon$ give

$$\mathbb{E}\left|\int_0^{\Delta_n}\int_{|x|\le\epsilon}\int_{\mathbb{R}_+} u_{s-}1_{\{l_{s-}<y<h_{s-}\}}\kappa'(x)\underline{\mu}(ds,dx,dy)\right|^p$$

$$\le K\Delta_n^{p/\beta}\mathbb{E}\left(\sup_{0\le s\le\Delta_n}\{|u_s|^p|h_s-l_s|\}\right)+K\Delta_n^{p/\beta}\left(\mathbb{E}\left(\sup_{0\le s\le\Delta_n}\{|u_s|^q|h_s-l_s|\}\right)\right)^{p/q}, \tag{4.23}$$

where $\beta<q<1$. Combining inequalities (4.21)-(4.23) we prove the Lemma. □

4.2 Limit theorems for realized power variation in pure-jump case

Using the preliminary results we now prove the following theorem.

Theorem 2 *For the process X_t defined in (2.2) assume that A1(b), A2 and A3 hold. Denote with L_s a pure-jump Lévy process (defined on some probability space) which has Lévy measure $ds\otimes\nu_1(x)dx$ and drift*

$$b=\begin{cases} -\int_{\mathbb{R}}\kappa'(x)\nu_1(x)dx & \text{if } \beta\ge 1\\ \int_{\mathbb{R}}\kappa(x)\nu_1(x)dx & \text{if } \beta<1,\end{cases}$$

with respect to the truncation function κ used in the definition of X_t. If $g_p(s)=\mathbb{E}(|L_s|^p)$, then for $p<\beta$ we have

$$\Delta_n^{1-p/\beta}V(p,X,\Delta_n)_t \xrightarrow{u.c.p.} \int_0^t g_p(a_s)|\sigma_{2s}|^p ds. \tag{4.1}$$

Proof: We prove the theorem under the following stronger assumption A1′:

Assumption A1′. *In addition to assumption A1 assume that the processes a_s, b_{2s} and σ_{2s} are bounded.*

To prove that if the claim in the theorem is true under the stronger assumption A1$'$, it holds also under assumption A1, we can use exactly the same localization argument as in the proof of Lemma 4.6 in Jacod (2008). Therefore we omit this part of the proof and proceed with proving Theorem 1 under the stronger assumption A1$'$. Also as for the lemmas in the previous subsection we will use the equivalent (in probability) representation of X given in (4.1).

First, we need some notation. Recall that for X_t we work with its representation in (4.1). For arbitrary $\epsilon > 0$ we set

$$X(\epsilon)_t = X_t - \int_0^t \int_{|x|>\epsilon} \int_{\mathbb{R}_+} \sigma_{2s-} x 1_{\{y<a_{s-}\}} \underline{\mu}(ds, dx, dy), \quad X'(\epsilon)_t = X_t - X(\epsilon)_t. \tag{4.2}$$

Further, denote with $S_1, S_2, ..., S_m, ...$ the sequence of successive jump times of $X'(\epsilon)$ (i.e. the jumps of the process X_t which after scaling by σ_{2S_m-} are higher than ϵ in absolute value). Furthermore, if $\Omega_n(T)$ denotes the set of ω-s for which each interval $[0,T] \cap ((i-1)\Delta_n, i\Delta_n]$ contains at most one S_m and in addition $|\Delta_i^n X(\epsilon)| \leq K\epsilon$ for some constant K (recall the process σ_{2t} is bounded pathwise), then $\Omega_n(T) \to \Omega$ as $\Delta_n \to 0$. Finally, for each S_m, we set $R_m = \Delta_i^n X(\epsilon)$, where $((i-1)\Delta_n, i\Delta_n]$ is the interval containing S_m. With this notation on the set $\Omega_n(T)$ we have

$$\begin{aligned} V(p, X, \Delta_n)_t - V(p, X(\epsilon), \Delta_n)_t &= \sum_{0<S_m\leq t} \{|R_m + \Delta X_{S_m}|^p - |R_m|^p\} \\ &\leq K \sum_{0\leq S_m \leq t} |\Delta X_{S_m}|^p + K, \quad t \leq T, \end{aligned} \tag{4.3}$$

where K is some constant and the sum in (4.3) is of course well defined and almost surely finite (the inequality is trivial for $p \leq 1$ and for $p > 1$ it from a Taylor expansion). Therefore, since $\Omega_n(T) \to \Omega$ we have

$$\Delta_n^{1-p/\beta} \{V(p, X, \Delta_n)_t - V(p, X(\epsilon), \Delta_n)_t\} \stackrel{u.c.p.}{\longrightarrow} 0.$$

Thus, to finish the proof we need only show

$$\Delta_n^{1-p/\beta} V(p, X(\epsilon), \Delta_n)_t \stackrel{u.c.p.}{\longrightarrow} \int_0^t g_p(a_s)|\sigma_{2s}|^p ds. \tag{4.4}$$

We make a slight abuse of notation for $\kappa(x)$, as in the proof of Lemma 2 part (b), in order to analyze the different cases for β together. In particular, we set $\kappa(x) \equiv 0$ when $\beta < 1$, which corresponds to no compensation of the small jumps, and set $b_{2t} \equiv 0$ (see assumption A3). We start the proof of (4.4) by making the following decomposition.

$$\Delta_n^{1-p/\beta} V(p, X(\epsilon), \Delta_n)_t - \int_0^t g_p(a_s)|\sigma_{2s}|^p ds = A_1 + A_2 + A_3, \tag{4.5}$$

$$A_1 = \Delta_n^{1-p/\beta} \sum_{i=1}^{[t/\Delta_n]} \xi_i^n,$$

$$\xi_i^n = \{|\xi_{i1}^n|^p - |\xi_{i2}^n|^p\}, \quad \xi_{i1}^n = \Delta_i^n X(\epsilon),$$

$$\begin{aligned} \xi_{i2}^n = & \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x|\leq\epsilon} \int_{\mathbb{R}_+} \sigma_{2(i-1)\Delta_n-} 1_{\{y<a_{(i-1)\Delta_n-}\}} \kappa(x) \underline{\tilde{\mu}}(ds, dx, dy) \\ & + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x|\leq\epsilon} \int_{\mathbb{R}_+} \sigma_{2(i-1)\Delta_n-} 1_{\{y<a_{(i-1)\Delta_n-}\}} \kappa'(x) \underline{\mu}(ds, dx, dy) \end{aligned}$$

$$A_2 = \Delta_n \sum_{i=1}^{[t/\Delta_n]} g_p(a_{(i-1)\Delta_n-})|\sigma_{2(i-1)\Delta_n-}|^p - \int_0^t g_p(a_s)|\sigma_{2s}|^p ds,$$

$$\begin{aligned} A_3 &= \Delta_n \sum_{i=1}^{[t/\Delta_n]} |\sigma_{2(i-1)\Delta_n-}|^p \Big\{ \Delta_n^{-p/\beta} \Big| \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x|\leq\epsilon} \int_{\mathbb{R}_+} 1_{\{y<a_{(i-1)\Delta_n-}\}} \kappa(x) \underline{\tilde{\mu}}(ds,dx,dy) \\ &\quad + \int_0^{\Delta_n} \int_{|x|\leq\epsilon} \int_{\mathbb{R}_+} 1_{\{y<a_{(i-1)\Delta_n-}\}} \kappa'(x) \underline{\mu}(ds,dx,dy) \Big|^p - g_p(a_{(i-1)\Delta_n-}) \Big\}. \end{aligned}$$

We will show that each of the three terms on the right-hand side of (4.5) converges in probability, uniformly in time, to 0. We start with A_1. In what follows $\mathbb{E}_{i-1}^n(\cdot)$ is shorthand for $\mathbb{E}\left(\cdot|\mathscr{F}_{(i-1)\Delta_n}\right)$. We want to derive bound for $\mathbb{E}_{i-1}^n|\xi_i^n|$. If $p \leq 1$ we can make use of the basic inequality $|a+b|^p \leq |a|^p + |b|^p$ for arbitrary a and b. If $p > 1$, we can use first-order Taylor expansion and then apply the Hölder's inequality. Both cases lead to

$$\mathbb{E}_{i-1}^n|\xi_i| \leq K\left(\mathbb{E}_{i-1}^n|\zeta_i^n|^p\right)^{\frac{1}{p}\wedge 1}\left(\mathbb{E}_{i-1}^n|\xi_{i1}^n|^p\right)^{\frac{p-1}{p}\vee 0} + K\left(\mathbb{E}_{i-1}^n|\zeta_i^n|^p\right)^{\frac{1}{p}\wedge 1}\left(\mathbb{E}_{i-1}^n|\xi_{i2}^n|^p\right)^{\frac{p-1}{p}\vee 0},$$

where

$$\begin{aligned} \zeta_i^n &= \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x|\leq\epsilon} \int_{\mathbb{R}_+} (\sigma_{2s-} 1_{\{y<a_{s-}\}} - \sigma_{2(i-1)\Delta_n-} 1_{\{y<a_{(i-1)\Delta_n-}\}}) \kappa(x) \underline{\tilde{\mu}}(ds,dx,dy) \\ &\quad + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x|\leq\epsilon} \int_{\mathbb{R}_+} (\sigma_{2s-} 1_{\{y<a_{s-}\}} - \sigma_{2(i-1)\Delta_n-} 1_{\{y<a_{(i-1)\Delta_n-}\}}) \kappa'(x) \underline{\mu}(ds,dx,dy). \end{aligned}$$

Using Lemma 2 and the fact that the processes σ_{2s} and a_s are bounded, we have

$$\mathbb{E}_{i-1}^n|\xi_i| \leq K\Delta_n^{\frac{p-1}{\beta}\vee 0}\left(\mathbb{E}_{i-1}^n|\zeta_i^n|^p\right)^{\frac{1}{p}\wedge 1}. \tag{4.6}$$

To bound $\mathbb{E}_{i-1}^n|\zeta_i^n|^p$ we first use the following decomposition

$$\sigma_{2s} 1_{\{y<a_s\}} - \sigma_{2(i-1)\Delta_n} 1_{\{y<a_{(i-1)\Delta_n}\}} = \sigma_{2s}(1_{\{y<a_s\}} - 1_{\{y<a_{(i-1)\Delta_n}\}}) + (\sigma_{2s} - \sigma_{2(i-1)\Delta_n}) 1_{\{y<a_{(i-1)\Delta_n}\}},$$

and then apply Lemma 2 to get

$$\begin{aligned} \mathbb{E}_{i-1}^n|\zeta_i|^p &\leq K\Delta_n^{p/\beta}\mathbb{E}_{i-1}^n \sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} |\sigma_{2s-} - \sigma_{2(i-1)\Delta_n-}|^p \\ &\quad + K\Delta_n^{p/\beta}\mathbb{E}_{i-1}^n \sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} |a_{s-} - a_{(i-1)\Delta_n-}| \\ &\quad + K\Delta_n^{p/\beta}\left(\mathbb{E}_{i-1}^n \sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} |\sigma_{2s-} - \sigma_{2(i-1)\Delta_n-}|^q\right)^{p/q} \\ &\quad + K\Delta_n^{p/\beta}\left(\mathbb{E}_{i-1}^n \sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} |a_{s-} - a_{(i-1)\Delta_n-}|\right)^{p/q}, \end{aligned}$$

where q is some number higher or at most equal to p. But

$$\begin{aligned} &\Delta_n \mathbb{E}\left(\sum_{i=1}^{[t/\Delta_n]}\left(\sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} |\sigma_{2s} - \sigma_{2(i-1)\Delta_n-}|^r\right)\right) \\ &= \int_0^t \mathbb{E}\left(\sup_{[u/\Delta_n]\Delta_n \leq u \leq [u/\Delta_n]\Delta_n + \Delta_n} |\sigma_{2u} - \sigma_{2[u/\Delta_n]\Delta_n}|^r\right) du, \end{aligned}$$

for some arbitrary r. Then for arbitrary $u > 0$ we have

$$0 \leq \lim_{\Delta_n \to 0} \mathbb{E} \left(\sup_{[u/\Delta_n]\Delta_n \leq u \leq [u/\Delta_n]\Delta_n + \Delta_n} |\sigma_{2u} - \sigma_{2[u/\Delta_n]\Delta_n}|^r \right) \leq C\mathbb{E} \left(|\Delta \sigma_u|^r \right).$$

Therefore, using the Lebesque's convergence theorem and the fact that σ_{2s} is a bounded càdlàg process (under A1′), we have

$$\Delta_n \mathbb{E} \left(\sum_{i=1}^{[t/\Delta_n]} \left(\sup_{(i-1)\Delta_n \leq s \leq i\Delta_n} |\sigma_{2s} - \sigma_{2(i-1)\Delta_n -}|^r \right) \right) \to 0.$$

Similar result holds for a_s. Therefore we have

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}_{i-1}^n |\xi_i^n| \longrightarrow 0,$$

and thus $A_1 \overset{u.c.p.}{\longrightarrow} 0$.

For A_2 using Riemann integrability we have that A_2 converges pointwise in ω and locally uniformly in time to 0. Finally, for A_3 first note that

$$\int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x| \leq \epsilon} \int_{\mathbb{R}_+} 1_{\{y < a_{(i-1)\Delta_n -}\}} \kappa(x) \underline{\tilde{\mu}}(ds, dx, dy) + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x| \leq \epsilon} \int_{\mathbb{R}_+} 1_{\{y < a_{(i-1)\Delta_n -}\}} \kappa'(x) \underline{\mu}(ds, dx, dy),$$

is equal in distribution to $L_{\Delta_n a_{(i-1)\Delta_n -}}$, where L is the Lévy process defined in Lemma 1 (it is defined on some (possibly) different probability space from the original one on which X_t is defined, but the choice of the constant ϵ is of course the same). Then we can use the convergence result in (4.3), which recall is locally uniform in time. Therefore, since a_s is bounded, we have

$$\begin{aligned} &\mathbb{E}_{i-1}^n \Bigg| \Delta_n^{-p/\beta} \Bigg| \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x| \leq \epsilon} \int_{\mathbb{R}_+} 1_{\{y < a_{(i-1)\Delta_n -}\}} \kappa(x) \underline{\tilde{\mu}}(ds, dx, dy) \\ &\qquad + \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x| \leq \epsilon} \int_{\mathbb{R}_+} 1_{\{y < a_{(i-1)\Delta_n -}\}} \kappa'(x) \underline{\mu}(ds, dx, dy) \Bigg|^p - g_p(a_{(i-1)\Delta_n -}) \Bigg| \\ &\leq f(\Delta_n), \end{aligned}$$

where $f(\cdot)$ is some function such that $\lim_{\Delta_n \to 0} f(\Delta_n) = 0$ and therefore $A_3 \overset{u.c.p.}{\longrightarrow} 0$. □

4.3 Proof of Theorem 1 in the paper

We prove all three parts together. On a set $\Omega_n \uparrow \Omega$, $\widehat{\beta}_{(0,T]}(\Upsilon, p)$ is a continuous transformation of $V(p, \Upsilon, k\Delta_n)_T$ and $V(p, \Upsilon, \Delta_n)_T$. Therefore we will be done if we can show that for a *fixed* $T > 0$ and some $p_l < p_u$ (which for part (b) and (c) of the Theorem are such that β is outside $[p_l, p_u]$) we have

$$\Pi_n(p) := \begin{pmatrix} \Delta_n^{1-p/\beta_{\Upsilon,T}} V(p, \Upsilon, k\Delta_n)_T \\ \Delta_n^{1-p/\beta_{\Upsilon,T}} V(p, \Upsilon, \Delta_n)_T \end{pmatrix} - \begin{pmatrix} k^{p/\beta_{\Upsilon,T}-1} C_T(p) \\ C_T(p) \end{pmatrix} \overset{\mathbb{P}}{\longrightarrow} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{uniformly on } [p_l, p_u], \tag{4.1}$$

where $C_T(p)$ is one of the limits in (3.3) and (3.5) depending on whether $\Upsilon \equiv Y, Z$ or $\Upsilon \equiv X$ respectively and the convergence in (4.1) is in $\mathcal{C}([p_l, p_u], \mathbb{R}_+^2)$ (the space of $\mathbb{R}_+^2$-valued continuous functions on $[p_l, p_u]$) equipped with the uniform metric. The proof of (4.1) consists of establishing finite-dimensional convergence and tightness of the sequence on the left-hand side of (4.1). The finite-dimensional convergence is trivial. It follows from the pointwise (in p) convergence results of Section 3.1 in the paper. Therefore we need only prove tightness. To this end, denote the modulus of continuity on $\mathcal{C}([p_l,p_u],\mathbb{R}_+^2)$ (see e.g. Jacod and Shiryaev (2003), VI.1)

$$w(\Pi_n,\theta) := \sup\{\sup_{u,v\in[p,p+\theta]} |\Pi_n(u)-\Pi_n(v)| : p_l \le p \le p+\theta \le p_u\}.$$

Set

$$U(\Upsilon,p,q,\Delta_n)_T = \Delta_n \sum_{i=1}^{[T/\Delta_n]} ||\Delta_n^{-1/\beta_{\Upsilon,T}}\Delta_i^n\Upsilon|^p - |\Delta_n^{-1/\beta_{\Upsilon,T}}\Delta_i^n\Upsilon|^q|. \tag{4.2}$$

With this notation for sufficiently small θ we have

$$\begin{aligned} w(\Pi_n,\theta) \quad \le \quad & U(\Upsilon,p_l+\theta,p_l,\Delta_n)_T + U(\Upsilon,p_u,p_u-\theta,\Delta_n)_T \\ & +U(\Upsilon,p_l+\theta,p_l,k\Delta_n)_T + U(\Upsilon,p_u,p_u-\theta,k\Delta_n)_T. \end{aligned} \tag{4.3}$$

Then we can use the inequality

$$U(\Upsilon,p,q,\Delta_n)_T \le K|p-q|\Delta_n \sum_{i=1}^{[T/\Delta_n]} \left(|\Delta_n^{-1/\beta_{\Upsilon,T}}\Delta_i^n\Upsilon|^{p\wedge q-\epsilon} + |\Delta_n^{-1/\beta_{\Upsilon,T}}\Delta_i^n\Upsilon|^{p\vee q+\epsilon}\right), \tag{4.4}$$

for arbitrary small ϵ and a constant K. Therefore the Ascoli-Arzela's criteria for tightness is satisfied (see Jacod and Shiryaev (2003), VI.3.26(ii)) and together with the finite-dimensional convergence, this implies (4.1) and hence the result of the Theorem.

4.4 Proof of Theorem 2 in the paper

Part (b) of the theorem follows from Barndorff-Nielsen and Shephard (2004, 2006) and Barndorff-Nielsen et al. (2006), hence we show only part (a). We proof the result in (3.13) under the stronger assumption A1′, the extension to the weaker assumption A1 follows by a localisation argument. Using the fact that $p \le 1$ we have trivially

$$|V(p,Z,\Delta_n)_t - V(p,Y,\Delta_n)_t| \le V(p,X,\Delta_n)_t. \tag{4.1}$$

At the same time a straightforward application of Jensen's inequality yields

$$\mathbb{E}_{i-1}^n|\Delta_i^n X|^p \le K\Delta_n^{\left(\frac{p}{BG_T(X)}-\varepsilon\right)\wedge 1} \quad \text{for any } \varepsilon>0. \tag{4.2}$$

Combining (4.1) and (4.2) we have

$$\Delta_n^{1/2-p/2}|V(p,Z,\Delta_n)_t - V(p,Y,\Delta_n)_t| \overset{u.c.p.}{\longrightarrow} 0, \tag{4.3}$$

under the conditions of the theorem. Therefore it is enough to prove (3.13) for $\Upsilon \equiv Y$. But this trivially follows from the following application of Theorem 7.3 in Jacod (2007)

$$\frac{1}{\sqrt{\Delta_n}}\left(\begin{array}{c} (k\Delta_n)^{1-p/2}V(p,Y,k\Delta_n)_t - \mu_p\int_0^t|\sigma_{1u}|^p du \\ \Delta_n^{1-p/2}V(p,Y,\Delta_n)_t - \mu_p\int_0^t|\sigma_{1u}|^p du \end{array}\right) \overset{\mathcal{L}-s}{\longrightarrow} \Xi_t, \tag{4.4}$$

where the process Ξ_t is defined on an extension of the original probability space, is continuous, and conditionally on the σ-field $\mathcal{F}$ is centered Gaussian with variance-covariance matrix given by

$$\int_0^t |\sigma_{1u}|^2 p du \left(\begin{array}{cc} k(\mu_{2p}-\mu_p^2) & k^{1-p/2}\mu_p(k) - k\mu_p^2 \\ k^{1-p/2}\mu(k,p) - k\mu_p^2 & \mu_{2p}-\mu_p^2 \end{array}\right).$$

Then the convergence in (3.13) follows trivially from the result in (4.4) and a Delta method (note that the convergence in (4.3) is stable).

References

Barndorff-Nielsen, O. and N. Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics 2*, 1–37.

Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics 4*, 1–30.

Barndorff-Nielsen, O. and N. Shephard (2007). Variation, jumps, Market Frictions and High Frequency Data in Fnancial Econometrics. In R. Blundell, T. Persson, and W. Newey (Eds.), *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress.* Cambridge University Press.

Barndorff-Nielsen, O., N. Shephard, and M. Winkel (2006). Limit Theorems for Multipower Variation in the Presence of Jumps in Financial Econometrics. *Stochastic Processes and Their Applications 116*, 796–806.

Jacod, J. (2007). Statistics and High-Frequency Data. Semstat Course in La Manga.

Jacod, J. (2008). Asymptotic Properties of Power Variations and Associated Functionals of Semimartingales. *Stochastic Processes and their Applications 118*, 517–559.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems For Stochastic Processes* (2nd ed.). Berlin: Springer-Verlag.

Protter, P. (2004). *Stochastic Integration and Differential Equations* (2nd ed.). Berlin: Springer-Verlag.

Todorov, V. and G. Tauchen (2009). Activity Signature Functions for High-Frequency Data Analysis. Working paper, Duke University and Northwestern University.

Woerner, J. (2006). Analyzing the Fine Structure of Continuous-Time Stochastic Processes. Working paper.