Finance Research Letters 59 (2024) 104814

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

A note on the Gumbel convergence for the Lee and Mykland jump tests

João Pedro Vidal Nunes [*], João Pedro Ruas

Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
Business Research Unit (BRU-IUL), Lisboa, Portugal

A R T I C L E I N F O

JEL classification:
C12
C14
C58

Keywords:
Extreme-value theory
Gumbel law
Folded normal distribution
Jump detection

A B S T R A C T

The Lee and Mykland (2008, 2012) nonparametric jump tests have been widely used in the literature but its critical region is stated with reference to the asymptotic distribution of the maximum of a set of standard normal variates. However, such reference would imply a typo (of a non-negligible order) for the norming constants adopted. By using the asymptotic distribution of the maximum of a set of *folded* normal random variables instead, this paper shows that there is no typo at all, thus preserving the validity of all the empirical findings based on these tests.

1. Introduction

Extreme (and rare) events take place in financial markets, and the identification of these jumps in financial time series is a relevant topic for risk management purposes. For this purpose, several nonparametric jump tests have been proposed in the literature—see, for instance, Andersen et al. (2007) or Corsi et al. (2010), among many others—at the same pace as new robust (to jumps) measures of (integrated) variance were also derived—as, for example, the bipower variation of Barndorff-Nielsen and Shephard (2004) or the threshold bipower variation of Corsi et al. (2010). The Lee and Mykland (2008) test also belongs to such class of nonparametric jump tests, with three advantages: it identifies both jump arrival times and realized jump sizes, and does not suffer from the *multiple testing* issue identified by Bajgrowicz et al. (2016).

The more recent widespread availability of tick-by-tick data prompted the application of intraday jump tests at sampling frequencies as high as a few seconds. However, those tests require even more robust measures of volatility—as, for instance, the ones provided by the pre-averaging approach followed by Podolskij and Vetter (2009, Theorem 2) or Christensen et al. (2014, Proposition 1)—, because at frequencies higher than 5 min the observed market prices are often distorted by microstructure noise (arising, for example, from bid–ask spreads, price discreteness or data bugs). One of the most well known noise- and jump-robust test is the one proposed by Lee and Mykland (2012), which extends Lee and Mykland (2008) to a setup with microstructure noise.

Both Lee and Mykland (2008, 2012) tests are widely used in the literature, since its finite sample properties (in terms of size and power) have been successfully tested by different authors—as, for instance, Dumitru and Urga (2012). For example, Schneider et al. (2010) resorted to the Lee and Mykland (2008) test to find strong evidence of jumps (and cojumps) in the time series of credit default swap spreads of different maturities. Bradley et al. (2014) identified jumps (at a 15-min frequency and in stocks listed on the NYSE between 2002 and 2007), also through the Lee and Mykland (2008) test, to evaluate the relative importance of analyst recommendations, earnings announcements and management guidance. Cremers et al. (2015) identified daily jumps in the S&P 500

* Correspondence to: ISCTE-IUL, Edifício II, Gabinete D531, Av. Prof. Aníbal Bettencourt, 1600-189 Lisboa, Portugal.
E-mail addresses: joao.nunes@iscte.pt (J.P.V. Nunes), joao.ruas@iscte.pt (J.P. Ruas).

https://doi.org/10.1016/j.frl.2023.104814
Received 28 September 2023; Received in revised form 8 November 2023; Accepted 27 November 2023
Available online 1 December 2023

index (from January 1988 to December 2011) through the Lee and Mykland (2008) test to evaluate their "jump factor" returns. Zhao (2017) applied the Lee and Mykland (2008) test to daily returns of a large sample of US stocks (between 1994 and 2009) to assess whether Securities and Exchange Commission current reports are likely to cause extreme price movements. Brogaard et al. (2018) used the Lee and Mykland (2012) test to detect jumps on high frequency trading data from Nasdaq (in 2008 and 2009), and found little evidence of high frequency traders causing extreme price movements.

However, Tsai and Shackleton (2016), Ferriani and Zoi (2020) and Bibinger (2021) all argued that both the original papers as well as all the subsequent implementations of the Lee and Mykland jump tests replicate a typo in Lee and Mykland (2008, Equation (13)) and Lee and Mykland (2012, Equation (13)); and, as shown in Section 3, such typo would produce a material impact on the definition of the critical region. Therefore, and given the non-negligible order of such typo, its correction along the lines of Ferriani and Zoi (2020, Footnote 2) or of Bibinger (2021, Equation (4b)) might jeopardize many of the empirical findings contained in the extensive literature that adopts these jump tests.

Fortunately, and as explained in Section 4, the original definition of the tests critical region is correct: even though both Lee and Mykland (2008, 2012) test statistics are normally distributed, the critical region of each test should be defined in terms of the maximum not of those (intraday) statistics but rather of their absolute value (that follows a *folded* normal distribution); i.e., Lee and Mykland (2008, 2012) should have referred to the asymptotic distribution of the maximum of a set of standard normal but *folded* random variables to define the critical region. Both Tsai and Shackleton (2016) and Ferriani and Zoi (2020) missed this point, while Bibinger (2021) failed to realize that the original norming constants proposed by Lee and Mykland (2008, 2012) are asymptotically equivalent to the ones he derived in Bibinger (2021, Equation (4b)). And since the *folded* normal distribution is still in the same *domain of maximal attraction* as the one considered by Lee and Mykland (2008, 2012)—the Gumbel law—, the main contribution of this paper to the literature follows: adopting the correct limiting distribution for both tests, the critical region defined through Lee and Mykland (2008, Equation (13)) and Lee and Mykland (2012, Equation (13)) is exactly recovered, and all the empirical findings based on these tests are preserved. In addition, Section 4 also recovers the set of equivalent norming constants derived by Bibinger (2021, Equation (4b)) and shows that the correction suggested by Tsai and Shackleton (2016, Equation (6.16)) or Ferriani and Zoi (2020, Footnote 2) would not yield the right asymptotic distribution for both tests (under the null).

2. The Lee and Mykland (2008, 2012) tests: a recap

In what follows, and denoting by 0 the current time, the time-interval $[0, T]$, with fixed $T \geq 0$, is divided into n equally spaced discrete dates $0 = t_0 < t_1 < \cdots < t_n = T$, such that $t_i - t_{i-1} = \Delta t$, for all $i \in \{1, 2, \ldots, n\}$, and $r_i := Y_i - Y_{i-1}$ represents the continuously compounded rate of return on some financial asset (with time-t_i log price Y_i) in the subinterval $\left(t_{i-1}, t_i\right]$. Since a too high absolute return can be due not to a jump but simply to a high volatility state, to detect jumps, returns must be first standardized by the prevailing instantaneous volatility. For this purpose, Barndorff-Nielsen and Shephard (2004, Page 9) propose a consistent estimator of the (integrated) variance—the bipower (quadratic) variation—that—unlike the usual quadratic variance measure—is not affected by the presence of (finite activity) jumps in the rate of return process. Hence, and similarly, for instance, to Andersen et al. (2007), the test statistic proposed by Lee and Mykland (2008, Definition 1.1) that tests, at time t_i, whether there was a jump in the time-interval $\left(t_{i-1}, t_i\right]$ is defined as

$$L_{\hat{m}_i}\left(t_i\right) := \frac{r_i - \hat{m}_i}{\hat{\sigma}_i}, \tag{1}$$

where $\hat{m}_i$ is the average rate of return realized over some suitable time window (preceding time t_i)—as given, for instance, in Lee and Mykland (2008, Page 2556)—, and $\hat{\sigma}_i$ is some (robust to jumps) measure of local volatility realized over the same estimation window—as given, for instance, by the standardized realized bipower variation in Andersen et al. (2007, Equation (12)).

Under the null hypothesis ($\mathcal{H}_0$) that there is no jump at any time in $\left(t_{i-1}, t_i\right]$, Andersen et al. (2007, Page 134) as well as Lee and Mykland (2008, Theorem 1.1) proved that the test statistic (1) is asymptotically standard normally distributed: Andersen et al. (2007, Equation (11)) only require returns to be sampled over intervals of identical quadratic variation whereas Lee and Mykland (2008, Assumption 1) assume that the underlying asset price is driven by an Itô process whose drift and diffusion coefficients do not change "dramatically over a short time interval".[1] Moreover, Lee and Mykland (2008, Theorem 2) have also shown that $L_{\hat{m}_i}\left(t_i\right) \to \infty$ as $\Delta t \to 0$, under the alternative hypothesis ($\mathcal{H}_a$). Therefore, the (clever) intuition behind the Lee and Mykland (2008) jump test is that a jump exists if the absolute value of the test statistic (1) is "too high". To define such "too high" threshold, Lee and Mykland (2008, Lemma 1) argued that the asymptotic distribution of the maximum of the test statistic (1) absolute values is, under $\mathcal{H}_0$, a Gumbel law, i.e., for all $x \in \mathbb{R}$ and as $\Delta t \to 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{n}\left(\left|L_{\hat{m}_i}\left(t_i\right)\right|\right) < \alpha_n^{LM} x + \beta_n^{LM}\right) = \Lambda(x), \tag{2}$$

where $\mathbb{P}$ denotes the real world (or physical) probability measure, n is total number of observations,

$$\beta_n^{LM} := (2\log n)^{\frac{1}{2}} - \frac{\log\log n + \log \pi}{2\,(2\log n)^{\frac{1}{2}}}, \tag{3}$$

[1] More specifically, Lee and Mykland (2008) assumed that both the drift and diffusion terms are α-Hölder continuous for every $\alpha < \frac{1}{2}$ (and that the additive jumps are of finite activity). Later, Palmes and Woerner (2013) relaxed these assumptions by only requiring a general pathwise Hölder-continuity for the volatility process, and Palmes and Woerner (2016) even accommodate (finite and not too large) volatility jumps.

and

$$\alpha_n^{LM} := (2\log n)^{-\frac{1}{2}} \tag{4}$$

are norming constants that the authors claim to borrow from Aldous (1989) and Galambos (1978), and

$$\Lambda(x) := \exp(-e^{-x}) \tag{5}$$

is the distribution function of a standard Gumbel random variable. Lee and Mykland (2008, Subsection 1.4) argued that if the observed value of $\left|L_{\hat{m}_i}(t_i)\right|$ is above the usual region of maximums, it is unlikely that the realized return arises from a diffusion model with no jumps (and, hence, $\mathcal{H}_0$ is rejected). More specifically, for a significance level of α^*, Lee and Mykland (2008) reject $\mathcal{H}_0$ if

$$\left|L_{\hat{m}_i}(t_i)\right| > \alpha_n^{LM}\beta^* + \beta_n^{LM}, \tag{6}$$

where β^* is such that $\Lambda(\beta^*) = 1 - \alpha^*$, i.e., $\beta^* = -\log(-\log(1-\alpha^*))$.

For sampling frequencies higher than 5 min, log prices might be contaminated by microstructure noise. Therefore, to average away (most of) the noise, Lee and Mykland (2012) used pre-averaging to obtain denoised log prices $\bar{Y}_i = \frac{1}{M}\sum_{j=i}^{i+M-1} Y_j$, for $i = 0, M, 2M, \ldots, \left(\left\lfloor \frac{n+1}{M} \right\rfloor - 1\right) M$, where the block length $M = \left\lceil \theta\sqrt{n} \right\rceil$ is computed using the tuning parameter θ prescribed in Lee and Mykland (2012, Table 5). Hence, the test statistic is now computed not on raw returns r_i but rather on the non-overlapping denoised returns $\bar{Y}_{i+M} - \bar{Y}_i$, and is equal to $\bar{L}_0(t_i) := \frac{\sqrt{M}(\bar{Y}_{i+M}-\bar{Y}_i)}{\sqrt{\hat{V}_i}}$, where $\hat{V}_i$ is a (jump- and noise-robust) estimate of the variance of $\sqrt{M}\left(\bar{Y}_{i+M} - \bar{Y}_i\right)$, for $i = 0, 2M, 4M, \ldots, (u-1)M$, with $u = \left(\left\lfloor \frac{n+1}{M} \right\rfloor - 1\right)$, if odd, or $u = \left(\left\lfloor \frac{n+1}{M} \right\rfloor - 2\right)$, otherwise. Under the null hypothesis that there is no jump at any time in $\left(t_{i-1}, t_i\right]$, and as $\Delta t \to 0$, Lee and Mykland (2012, Lemma 1) have shown that the test statistic $\bar{L}_0(t_i)$ is standard normally distributed, and, hence,[2]

$$\lim_{u\to\infty} \mathbb{P}\left(\max_{l=0}^{\frac{u-1}{2}}\left(\left|\bar{L}_0\left(t_{2lM}\right)\right|\right) < \alpha_{\frac{u+1}{2}}^{LM} x + \beta_{\frac{u+1}{2}}^{LM}\right) = \Lambda(x),$$

yielding a critical region similar to the one defined in Eq. (6)—after changing the number of raw observations n for the number of non-overlapping denoised returns $\frac{u+1}{2}$. Therefore, the subsequence analysis will be focused in only one of the two test statistics, namely $L_{\hat{m}_i}(t_i)$ instead of $\bar{L}_0(t_i)$.

3. Problems with the test statistics

The proof of Lee and Mykland (2008, Lemma 1) is not explicitly given by the authors; instead, it is simply stated that the "proof of Lemma 1 follows from Aldous (1989) and the proof in Galambos (1978) ". Likewise, the proof of Lee and Mykland (2012, Lemma 2) is replaced by references to Berman (1964) and Ljung (1993). However, all these references deal with the asymptotic distribution of the maximum of a set of standard (but not *folded*) random variables that will be now summarized.

It is well known—see, for instance, Resnick (1987, Page 42)—that the standard normal distribution is a *Von Mises function*, and, hence, is in the domain of maximal attraction of the Gumbel law. So, if $U_1, U_2, \ldots, U_n$ denote n independent standard normal random variables (and under some probability measure $\mathbb{P}$), then

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{n}\left(U_i\right) < a_n x + b_n\right) = \Lambda(x), \tag{7}$$

for all $x \in \mathbb{R}$, and for suitably chosen centering and scaling constants $b_n \in \mathbb{R}$ and $a_n > 0$, respectively. More specifically, and following, for instance, David and Nagaraja (2003, Part (c) of Theorem 10.5.2), b_n can be obtained as the solution to the nonlinear equation

$$1 - \Phi\left(b_n\right) = \frac{1}{n}, \tag{8}$$

while

$$a_n = \left[n\phi\left(b_n\right)\right]^{-1}, \tag{9}$$

where

$$\phi(u) := \frac{\exp\left(-\frac{u^2}{2}\right)}{\sqrt{2\pi}} \tag{10}$$

[2] This result is even extended by Bibinger et al. (2019, Proposition 3.1) from a jump–diffusion setup to a more general semimartingale model that encompasses volatility jumps.

and $\Phi(u)$ represent, for $u \in \mathbb{R}$, the density and the cumulative distribution function of a standard normal probability law, respectively. Moreover, and following, for instance, Hall (1979, Equation (1b)), the asymptotic approximation offered by Feller (1968, Page 175) for the Mills ratio,

$$\lim_{u\to\infty} \frac{1-\Phi(u)}{\phi(u)} = \frac{1}{u}, \tag{11}$$

and Eq. (8) allow Eq. (9) to be rewritten as[3]

$$a_n \sim \left\{ n b_n \left[1-\Phi\left(b_n\right)\right]\right\}^{-1} = \left(b_n\right)^{-1}. \tag{12}$$

Instead of solving the nonlinear Eq. (8), since

$$H(u) := 1 - \frac{\phi(u)}{u} \tag{13}$$

is a right tail equivalent distribution function, in the sense that $\lim_{u\to\infty} \frac{1-\Phi(u)}{1-H(u)} = 1$, and following again Hall (1979, Equation (1a)), the centering constant b_n can be approximated by the (easier) solution b_n^H of

$$1 - H\left(b_n^H\right) = \frac{1}{n}, \tag{14}$$

while the scaling constant a_n in Eq. (12) is approximated by

$$a_n^H := \left(b_n^H\right)^{-1}. \tag{15}$$

Finally, and even though an explicit solution to Eq. (14) is not available, Galambos (1978, Page 65)—referenced in the proof of Lee and Mykland (2008, Lemma 1)—has shown that the norming constants b_n^H and a_n^H can be approximated, up to an error of order $o\left((\log n)^{-\frac{1}{2}}\right)$, by

$$\beta_n^H := (2\log n)^{\frac{1}{2}} - \frac{\log\log n + \log 4\pi}{2\,(2\log n)^{\frac{1}{2}}}, \tag{16}$$

and

$$\alpha_n^H := \left(\beta_n^H\right)^{-1}, \tag{17}$$

respectively—see also, for example, Cramér (1946, Page 374), Berman (1964, Equation (3.3)), Resnick (1987, Page 71), Aldous (1989, Page 46), Ljung (1993, Equation (4.2)), David and Nagaraja (2003, Equation (10.5.20)), or Haan and Ferreira (2006, Example 1.1.7).

As shown, for instance, by Palmes and Woerner (2013, Proposition 4.1), the test statistics $\left\{L_{\hat{m}_i}\left(t_i\right)\right\}_{i=1}^n$ are asymptotically independent and normally distributed with zero mean and variance equal to 1, and, therefore, Eqs. (7) to (17) imply that

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{n}\left(L_{\hat{m}_i}\left(t_i\right)\right) < \alpha_n^H x + \beta_n^H\right) = \Lambda(x), \tag{18}$$

for all $x \in \mathbb{R}$. Comparing Eqs. (2) to (4) with Eqs. (16) to (18), two differences emerge: first, Eq. (18) offers the limit distribution of the maximum of the test statistic (1) whereas Eq. (2)—used by Lee and Mykland (2008)—describes the asymptotic distribution of the maximum of such test statistic absolute values; second, $\alpha_n^H \sim (2\log n)^{-\frac{1}{2}} = \alpha_n^{LM}$ but $\beta_n^H \neq \beta_n^{LM}$, and such difference should be qualitatively relevant because it induces an error of order strictly higher than $o\left((\log n)^{-\frac{1}{2}}\right)$: $\beta_n^{LM} - \beta_n^H = \frac{\log 4}{2(2\log n)^{\frac{1}{2}}} = O\left((\log n)^{-\frac{1}{2}}\right)$. This is why Tsai and Shackleton (2016, Equation (6.16)) as well as Ferriani and Zoi (2020, Footnote 2) claimed that Lee and Mykland (2008, Equation (13)) contains a typo: they argued that $\log \pi$ should be replaced by $\log 4\pi$.[4]

4. Corrected tests

The extreme value distribution (2) might be ill-defined because, under $\mathcal{H}_0$, $\left(\left|L_{\hat{m}_i}\left(t_i\right)\right|\right)_{i=1}^n$ is a sequence not of independent standard normal random variables but rather of independent *folded* (standard) normal random variables, which are not considered in Aldous (1989), Galambos (1978) and Berman (1964) or Ljung (1993). Following, for instance, Johnson and Kotz (1970, Page 136), it is well known that for any standard normal random variable U its absolute value $|U|$ possesses a folded normal density equal to

$$f(u) := 2\phi(u) \tag{19}$$

[3] For any real-valued functions $\varphi, \gamma : \mathbb{R} \to \mathbb{R}$, the statement $\varphi(x) \sim \gamma(x)$, as $x \to \infty$, is intended to mean $\lim_{x\to\infty} \frac{\varphi(x)}{\gamma(x)} = 1$.

[4] Note, however, that this problem does not affect the Gumbel tests proposed by Palmes and Woerner (2013, 2016) because these authors consider positive and negative jumps separately—i.e., use a two-tailed critical region—but at the expense of a lower statistical power (for the same significance level).

and a cumulative distribution function given by

$$F(u) := 2\Phi(u) - 1, \tag{20}$$

for $u \geq 0$. Fortunately, and as shown in the next proposition, the folded normal distribution is still in the domain of maximal attraction of the Gumbel law, but the corresponding norming constants can no longer be given by Eqs. (16) and (17).

Proposition 1. *If $U_1, U_2, \ldots, U_n$ denote n independent standard normal random variables (under some probability measure $\mathbb{P}$), then*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{n}\left(|U_i|\right) < A_n x + B_n\right) = \Lambda(x), \tag{21}$$

for all $x \in \mathbb{R}$, where the norming constants $B_n \in \mathbb{R}$ and $A_n > 0$ are equal to

$$B_n := b_{2n} \tag{22}$$

and

$$A_n := a_{2n}, \tag{23}$$

while b_{2n} and a_{2n} are defined by Eqs. (8) and (9), respectively.

Proof. Proposition 1 follows easily from Bibinger (2021, Lemma 1) because the standard normal density is symmetric around its mean. Nevertheless, and for the sake of completeness, an alternative proof is provided in Appendix A. ■

Since no exact solution exists for Eq. (8), an approximate solution will be adopted along the lines of Hall (1979). As described in Section 3, and for the standard normal probability law, Hall (1979, Equations (1a) and (1b)) found the centering constant b_n^H as the solution of Eq. (14)—using the right tail equivalent distribution (13)—and obtained the scale constant a_n^H as the inverse of b_n^H—see Eq. (15). Similarly, it is easy to show that the right tail distribution function equivalent to $F(u)$ is not $H(u)$—as defined in Eq. (13)—but rather $\bar{H}(u) := 1 - 2\frac{\phi(u)}{u}$, because $\lim_{u\to\infty} \frac{1-F(u)}{1-\bar{H}(u)} = 1$. Therefore, the centering constant B_n can be approximated by the solution $B_n^{\bar{H}}$ of $1 - \bar{H}\left(B_n^{\bar{H}}\right) = \frac{1}{n}$ that is equivalent to

$$1 - H\left(B_n^{\bar{H}}\right) = \frac{1}{2n}. \tag{24}$$

Comparing Eqs. (14) and (24), it follows that

$$B_n^{\bar{H}} = b_{2n}^H, \tag{25}$$

and, hence, Eqs. (15) and (23) imply that the scale constant A_n can be approximated by

$$A_n^{\bar{H}} = \left(b_{2n}^H\right)^{-1}. \tag{26}$$

However, and even though b_n^H can be approximated by β_n^H but not by β_n^{LM}, b_{2n}^H can be approximated by both β_{2n}^H or β_n^{LM}. This new result is stated in the next proposition and completely justifies the use of the original Lee and Mykland (2008, 2012) jump tests by the previous literature.

Proposition 2. *The norming constants $B_n^{\bar{H}}$ and $A_n^{\bar{H}}$—and, therefore, B_n and A_n as well—can be approximated, up to an error of order $o\left((\log n)^{-\frac{1}{2}}\right)$, by*

$$\beta_n^{\bar{H}} = \beta_n^{LM}, \tag{27}$$

as given in Eq. (3), and

$$\alpha_n^{\bar{H}} := \left(\beta_n^{LM}\right)^{-1}, \tag{28}$$

respectively.

Proof. This proof follows exactly the same steps as in Resnick (1987, Page 72). Combining Eqs. (10), (13) and (24), $B_n^{\bar{H}}$ must be such that

$$\frac{(2\pi)^{-\frac{1}{2}} \exp\left(-\frac{\left(B_n^{\bar{H}}\right)^2}{2}\right)}{B_n^{\bar{H}}} = \frac{1}{2n}, \tag{29}$$

i.e.,

$$\left(\frac{\pi}{2}\right)^{-\frac{1}{2}} \left(B_n^{\bar{H}}\right)^{-1} \exp\left(-\frac{\left(B_n^{\bar{H}}\right)^2}{2}\right) = n^{-1}. \tag{30}$$

Following Resnick (1987, Equation (1.33)), i.e., taking -log of both sides of Eq. (30), gives

$$\frac{1}{2}\left(B_n^{\bar{H}}\right)^2 + \log B_n^{\bar{H}} + \frac{1}{2}\log\frac{\pi}{2} = \log n, \tag{31}$$

and dividing both sides of Eq. (31) by $\left(B_n^{\bar{H}}\right)^2$ it follows that

$$B_n^{\bar{H}} \sim (2\log n)^{\frac{1}{2}}, \tag{32}$$

since $\lim_{n\to\infty} B_n^{\bar{H}} = \infty$, and, hence,

$$A_n^{\bar{H}} \sim (2\log n)^{-\frac{1}{2}} \tag{33}$$

from Eqs. (25) and (26). Therefore, and using, for instance, Galambos (1978, Lemma 2.2.2), we seek an expansion of the form

$$B_n^{\bar{H}} = \beta_n^{\bar{H}} + o\left((\log n)^{-\frac{1}{2}}\right), \tag{34}$$

for some sequence $\left\{\beta_n^{\bar{H}}, n > 1\right\}$ of real numbers and of order $O\left((\log n)^{\frac{1}{2}}\right)$ to be determined, since $\lim_{n\to\infty} \frac{B_n^{\bar{H}} - \beta_n^{\bar{H}}}{A_n^{\bar{H}}} = \frac{o\left((\log n)^{-\frac{1}{2}}\right)}{O\left((\log n)^{-\frac{1}{2}}\right)} \to 0.$

Eq. (32) implies that

$$B_n^{\bar{H}} = (2\log n)^{\frac{1}{2}} + r_n, \tag{35}$$

where the reminder is $r_n = o\left((\log n)^{\frac{1}{2}}\right)$, and combining Eqs. (31) and (35), it follows that

$$\frac{1}{2}r_n^2 + (2\log n)^{\frac{1}{2}} r_n + \frac{1}{2}(\log\log n + \log\pi) + \log\left(1 + (2\log n)^{-\frac{1}{2}} r_n\right) = 0, \tag{36}$$

which is exactly the same as Resnick (1987, Equation (1.36)) when the term $\log\pi$ is replaced by $\log 4\pi$. Therefore, Resnick (1987, Equations (1.36) and (1.38)) imply that

$$r_n = -\frac{1}{2}\frac{\log\log n + \log\pi}{(2\log n)^{\frac{1}{2}}} + o\left((\log n)^{-\frac{1}{2}}\right), \tag{37}$$

i.e., Eq. (27) follows from Eqs. (3), (34), (35) and (37).

Finally, and using Eq. (33) as well as Galambos (1978, Lemma 2.2.2), $A_n^{\bar{H}}$ can be approximated through Eq. (28) because $\lim_{n\to\infty}\frac{A_n^{\bar{H}}}{\alpha_n^{\bar{H}}} = 1$ follows from $\beta_n^{LM} = \beta_n^{\bar{H}} \sim (2\log n)^{\frac{1}{2}}$. ■

Remark 1. Since $\left(\beta_n^{LM}\right)^{-1} \sim (2\log n)^{-\frac{1}{2}}$, then Proposition 2 yields exactly the norming constants (3) and (4) proposed by Lee and Mykland (2008). Note that, to the authors knowledge, these norming constants were first derived (only in 2012) by Mutangi and Matarise (2012), using the asymptotic equivalence relation between $\lim_{n\to\infty}\mathbb{P}\left(\max_{i=1}^n\left(U_i\right) < \alpha_n^H x + \beta_n^H\right) = \Lambda(x)$ and $\lim_{n\to\infty} n\left[1 - \Phi\left(\alpha_n^H x + \beta_n^H\right)\right] = -\log\Lambda(x)$.

Remark 2. Applying -log not to both sides of Eq. (30) but rather to Eq. (29), and following exactly the same steps as in the proof of Proposition 2, would lead to the equivalent asymptotic expansion

$$B_n^{\bar{H}} = \beta_{2n}^H + o\left((\log 2n)^{-\frac{1}{2}}\right), \tag{38}$$

where β_{2n}^H is given by Eq. (16). Such equivalent asymptotic expansion cannot be obtained through the simpler Mutangi and Matarise (2012) approach. Instead, it was independently derived by Bibinger (2021, Proposition 2.1), using a generalization of our Proposition 1—Bibinger (2021, Lemma 1). Note, however, that Bibinger (2021, Lemma 1) cannot yield the asymptotic solution (27), which leads Bibinger (2021, Page 2) to erroneously assume that the original Lee and Mykland (2008) norming constant (3) involves "a small but relevant typo".

Proposition 2 shows that the norming constants of Eqs. (3) and (4)—proposed by Lee and Mykland (2008) and extensively used in the literature—are correct, and, therefore, all the favorable evidence provided by the literature on the finite sample performance of the Lee and Mykland (2008, 2012) jump tests is preserved. In the case of both Tsai and Shackleton (2016) and Ferriani and Zoi (2020), the alleged typo in Lee and Mykland (2008, Equation (13)) was based on the wrong reference to a standard (instead of a folded) normal distribution. Moreover, Proposition 2 also shows that the correction adopted by Tsai and Shackleton (2016, Equation (16)) or Ferriani and Zoi (2020, Footnote 2) shall not be used, because $B_n^{\bar{H}}$ can be approximated by both β_n^{LM} or β_{2n}^H but never by β_n^H. Finally, Bibinger (2021) use the correct asymptotic distribution of the maxima of a set of standard folded normal random variables but misses the asymptotic equivalence between β_{2n}^H and β_n^{LM}. Of course, in practice, the econometrician deals with finite samples, and the sets $\left(\beta_{2n}^H, \alpha_{2n}^H\right)$ and $\left(\beta_n^{LM}, \alpha_n^{LM}\right)$ of norming constants can yield different results. The finite-sample performance of these norming constants is outside the scope of this paper but a Monte Carlo study is presented in the online supplementary file.

5. Conclusion

This paper proves that the norming constants proposed by Lee and Mykland (2008, 2012) yield the correct specification of the critical region for their nonparametric jump tests. Therefore, the typo mentioned by Tsai and Shackleton (2016) and Ferriani and Zoi (2020) or Bibinger (2021) does not exist, and the validity of all the empirical findings based on these tests is entirely preserved. Additionally, the norming constants later proposed by Bibinger (2021) are also shown to be asymptotically equivalent to the original ones. In other words, the original Lee and Mykland (2008, 2012) tests are still *alive and kicking*.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

Financial support from Fundação para a Ciência e Tecnologia (FCT), through the grant number UIDB/00315/2020, as well as helpful comments from Claudio Tejerina are gratefully acknowledged by the authors.

Appendix A. Proof of Proposition 1

The sufficient condition stated in David and Nagaraja (2003, Part (c) of Theorem 10.5.2) yields Eq. (21) with B_n such that

$$F\left(B_n\right) = 1 - \frac{1}{n}, \tag{A.1}$$

and

$$A_n = \left[n f\left(B_n\right)\right]^{-1}, \tag{A.2}$$

as long as $f(u) > 0$ and

$$\lim_{u\to\infty} \frac{d}{du}\left[\frac{1-F(u)}{f(u)}\right] = 0, \tag{A.3}$$

for all $u > 0$. Using relations (19) and (20), it follows that $\frac{dF(u)}{du} = 2\phi(u)$ and $\frac{df(u)}{du} = -2u\phi(u)$, and, therefore,

$$\begin{aligned}\frac{d}{du}\left[\frac{1-F(u)}{f(u)}\right] &= \frac{-[2\phi(u)]^2 + 2u\phi(u)[2-2\Phi(u)]}{[2\phi(u)]^2} \\ &= -1 + u\frac{1-\Phi(u)}{\phi(u)}.\end{aligned} \tag{A.4}$$

Applying the asymptotic approximation (11) to Eq. (A.4), Eq. (A.3) arises immediately, and Eq. (21) follows, for suitably chosen norming constants B_n and A_n.

Combining Eqs. (20) and (A.1), the centering constant B_n must be such that

$$1 - \Phi\left(B_n\right) = \frac{1}{2n}, \tag{A.5}$$

and since the norming constant b_n solves the nonlinear Eq. (8), then Eq. (22) follows. Finally, Eqs. (19), (A.2), (11) and (A.5) imply that

$$A_n = \frac{1}{2n\phi\left(B_n\right)} \sim \frac{1}{2nB_n\left[1-\Phi\left(B_n\right)\right]} = \frac{1}{B_n}, \tag{A.6}$$

and Eq. (23) arises from Eqs. (12), (22) and (A.6). ■

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.frl.2023.104814.

References

Aldous, D., 1989. Probability Approximations Via the Poisson Clumping Heuristic. Springer-Verlag, New York.
Andersen, T., Bollerslev, T., Dobrev, D., 2007. No-arbitrage semi-martingale restrictions for continuous time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. J. Econometrics 138, 125–180.
Bajgrowicz, P., Scaillet, O., Treccani, A., 2016. Jumps in high-frequency data: Spurious detections, dynamics, and news. Manage. Sci. 62, 2198–2217.
Barndorff-Nielsen, O., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. J. Financ. Econom. 2, 1–37.
Berman, S., 1964. Limit theorems for the maximum term in stationary sequences. Ann. Math. Stat. 35, 502–516.
Bibinger, M., 2021. Gumbel Convergence of the Maximum of Convoluted Half-Normally Distributed Random Variables. Working paper, Julius-Maximilians-Universität Würzburg.
Bibinger, M., Neely, C., Winkelmann, L., 2019. Estimation of the discontinuous leverage effect: Evidence from the NASDAQ order book. J. Econometrics 209, 158–184.
Bradley, D., Clarke, J., Lee, S., Ornthanalai, C., 2014. Are analysts' recommendations informative? Intraday evidence on the impact of time stamp delays. J. Finance 69, 645–673.
Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A., Sokolov, K., 2018. High frequency trading and extreme price movements. J. Financ. Econ. 128, 253–265.
Christensen, K., Oomen, R., Podolskij, M., 2014. Fact or friction: Jumps at ultra high frequency. J. Financ. Econ. 114, 576–599.
Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. J. Econometrics 159, 276–288.
Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton.
Cremers, M., Halling, M., Weinbaum, D., 2015. Aggregate jump and volatility risk in the cross-section of stock returns. J. Finance 70, 577–614.
David, H., Nagaraja, H., 2003. Order Statistics, third ed. Wiley, New Jersey.
Dumitru, A., Urga, G., 2012. Identifying jumps in financial assets: A comparison between nonparametric jump tests. J. Bus. Econom. Statist. 30, 242–255.
Feller, W., 1968. An Introduction to Probability Theory and Its Applications. Volume I, third ed. Wiley, New York.
Ferriani, F., Zoi, P., 2020. The dynamics of price jumps in the stock market: An empirical study on Europe and U.S.. Euro. J. Finance 28, 718–742.
Galambos, J., 1978. The Asymptotic Theory of Extreme Order Statistics. John Wiley & Sons, New York.
Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. Springer.
Hall, P., 1979. On the rate of convergence of normal extremes. J. Appl. Probab. 16, 433–439.
Johnson, N., Kotz, S., 1970. Distributions in Statistics: Continuous Univariate Distributions. Volume 2. Wiley, New York.
Lee, S., Mykland, P., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Rev. Financ. Stud. 21, 2535–2563.
Lee, S., Mykland, P., 2012. Jumps in equilibrium prices and market microstructure noise. J. Econometrics 168, 396–406.
Ljung, G., 1993. On outlier detection in time series. J. R. Stat. Soc. Ser. B Stat. Methodol. 55, 559–567.
Mutangi, K., Matarise, F., 2012. On how to find the norming constants for the maxima of a folded normally distributed variable. J. Stat. Res. 46, 31–35.
Palmes, C., Woerner, J., 2013. The Gumbel Test for Jumps in Stochastic Volatility Models. Working paper, TU Dortmund.
Palmes, C., Woerner, J., 2016. The Gumbel test and jumps in the volatility process. Stat. Inference Stoch. Process. 19, 235–258.
Podolskij, M., Vetter, M., 2009. Bipower-type estimation in a noisy diffusion setting. Stochastic Process. Appl. 119, 2803–2831.
Resnick, S., 1987. Extreme Values, Regular Variation, and Point Processes. Springer-Verlag, New York.
Schneider, P., Sögner, L., Veža, T., 2010. The economic role of jumps and recovery rates in the market for corporate default risk. J. Financ. Quant. Anal. 45, 1517–1547.
Tsai, P., Shackleton, M., 2016. Detecting jumps in high-frequency prices under stochastic volatility: A review and a data-driven approach. In: Florescu, I., Mariani, M., S, H., Viens, F. (Eds.), Handbook of High-Frequency Trading and Modeling in Finance. Chapter 6. John Wiley & Sons, pp. 137–181.
Zhao, X., 2017. Does information intensity matter for stock returns? Evidence from form 8-K filings. Manage. Sci. 63, 1382–1404.