DISEI - Università degli Studi di Firenze

Working Papers – Mathematical Economics

Optimum thresholding using mean and conditional mean square error

José E. Figueroa-López

Cecilia Mancini

Working Paper N. 2017-01

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze, Italia
www.disei.unifi.it

Optimum thresholding using mean and conditional mean square error

José E. Figueroa-López* and Cecilia Mancini†

March 14, 2017

Abstract

We consider a univariate semimartingale model for (the logarithm of) an asset price, containing jumps having possibly infinite activity (IA). The nonparametric threshold estimator $\hat{IV}_n$ of the integrated variance $IV := \int_0^T \sigma_s^2 ds$ proposed in [6] is constructed using observations on a discrete time grid, and precisely it sums up the squared increments of the process when they are under a *threshold*, a deterministic function of the observation step and possibly of the coefficients of X. All the threshold functions satisfying given conditions allow asymptotically consistent estimates of IV, however the finite sample properties of $\hat{IV}_n$ can depend on the specific choice of the threshold. We aim here at optimally selecting the threshold by minimizing either the estimation mean square error (MSE) or the conditional mean square error (cMSE). The last criterion allows to reach a threshold which is optimal not in mean but for the specific path at hand.

A parsimonious characterization of the optimum is established, which turns out to be asymptotically proportional to the Lévy's modulus of continuity of the underlying Brownian motion. Moreover, minimizing the cMSE enables us to propose a novel implementation scheme for the optimal threshold sequence. Monte Carlo simulations illustrate the superior performance of the proposed method.

Keywords: Threshold estimator, integrated variance, Lévy jumps, mean square error, conditional mean square error, modulus of continuity of the Brownian motion paths, numerical scheme

JEL classification codes: C6, C13

1 Introduction

We consider the model

$$dX_t = \sigma_t dW_t + dJ_t, \tag{1}$$

where W is a standard Brownian motion, σ is a cadlag process, and J is a pure jump semimartingale (SM) process. Assume we have at our disposal a record $\{x_0, X_{t_1}, .., X_{t_n}\}$ of discrete observations of X spanned on the fixed time interval $[0, T]$, define $\Delta_i Z$ or $\Delta_i^n Z$ the increment $Z_{t_i} - Z_{t_{i-1}}$ for any process Z, and define *threshold function* $r(\sigma, h)$ any deterministic non-negative function of the observation step h, and possibly of the coefficients of X, such that for any value $\sigma \in \mathbb{R}$

$$\lim_{h\to 0} r(\sigma, h) = 0, \quad \lim_{h\to 0} \frac{r(\sigma, h)}{h \log \frac{1}{h}} = +\infty.$$

We know that then the *Threshold Realized Variance* (or *Truncated Realized Variance*)

$$\hat{IV}_n := \sum_{i=1}^{n} (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r(\sigma_{t_{i-1}}, h_i)\}}, \tag{2}$$

*Department of Mathematics, Washington University in St. Louis, MO, 63130, USA (figueroa@math.wustl.edu)

†Department of Management and Economics, University of Florence, via delle Pandette 9, 50127 (cecilia.mancini@unifi.it)

where $h_i := t_i - t_{i-1}$, gives a consistent estimator of the *Integrated Variance*

$$IV := \int_0^T \sigma_s^2 ds,$$

as $\sup_i h_i \to 0$, as soon as σ is a.s. bounded away from zero on $[0, T]$. In the case where the jump process J has finite activity (FA) and the observations are evenly spaced, the estimator is also asymptotically Gaussian. However the finite sample properties of $\hat{IV}_n$ can depend on the specific choice of the threshold (TH). The estimation error is large when either the threshold is too small or when it is too large. In the first case too many increments are discarded, included the increments bearing only small and negligible jumps, and TRV underestimates IV. In the second case too many increments are kept within TRV, included many increments containing jumps, leading to an overestimation of IV.

In this paper we look for an optimal threshold, by considering the following two optimality criteria: minimization of MSE, the expected quadratic error in the estimation of IV; and minimization of cMSE, the expected quadratic error conditional to the realized paths of the jump process J and of the volatility process $(\sigma_s)_{s\geq 0}$. Assuming evenly spaced observations, the two quantities MSE and cMSE are explicit functions of the TH and under each criterion it turns out that for any semimartingale X, for which the volatility and the jump processes are independent on the underlying Brownian motion, an optimal TH exists, and is a solution of an explicitly given equation, the equation being different under the two criteria. Further, under each criterion the optimal TH is unique, at least for given classes of processes X.

The characterizing equation depends on the observation step h and so does its solution. The optimal TH has to tend to 0 as h tends to zero and, under each criterion, an asymptotic expansion with respect to h is possible for some terms within the equation, which in turn implies an asymptotic expansion of the optimal TH. Under the MSE criterion, when X is Lévy and J has either finite activity jumps or the activity is infinite but J is symmetric strictly stable, the leading term of the expansion is explicit in h, and in both cases is proportional to the modulus of continuity of the Brownian motion paths and to the spot volatility of X, the proportionality constant being $\sqrt{2-Y}$, where Y is the jump activity index of X Thus the higher the jump activity is, the lower the optimal threshold has to be to discard the higher noise represented by the jumps, in order to catch information about IV.

The leading term of the optimal TH does not satisfy the classical assumptions under which the truncation method has been shown in [6] to consistently estimate IV, however at least in the finite activity jumps case it turns out that the threshold estimator of IV constructed with the optimal TH is still consistent.

The assumptions needed for the cMSE criterion are a little bit less restrictive, and we find that, for constant σ and FA jumps, the leading term of the optimal TH still has to be proportional to the modulus of continuity of the Brownian motion paths and to σ. One of the main motivations for considering the cMSE arises from a novel application of this to tuneup the threshold parameter. The idea consists in iteratively updating the optimal TH and estimates of the increments of the continuous and jump components $X_t^c = \int_0^t \sigma_s dW_s$ and $\{J_t\}_{t\geq 0}$, respectively. We illustrate this method on simulated data. Minimization of the conditional mean square estimation error in the presence of infinite activity jumps in X is object of further research.

An outline of the paper is as follows. Section 2 deals with the MSE: the existence of an optimal threshold $\varepsilon^\star(h)$ is established for a quite general SM X; for a Lévy process X, uniqueness is also established (Subsection 2.1) and the asymptotic expansion for the optimal TH is found in Section 2.3, in both the cases of a finite jump activity Lévy X and of an infinite activity symmetric strictly stable X. In Section 3, for any finite jump activity SM X, consistency of $\hat{IV}_n$ is verified even when the threshold function consists of the leading term of the optimal threshold, which does not satisfy the classical hypothesis. Section 4 deals with the cMSE in the case where X is a SM with constant volatility and FA jumps: existence of an optimal TH $\bar{\varepsilon}(h)$ is established, its asymptotic expansion is found, then uniqueness is obtained. In Section 5 the results of Section 4 are used to construct a new method for iteratively determine the optimal threshold value in finite samples, and a reliability check is executed on simulated data. Section 6 concludes and Section 7 contains the proofs not given in the main text.

Acknowledgements. José Figueroa-López's research was supported in part by the National Science Foundation grants: DMS-114969 and DMS-1613016. Cecilia Mancini's work has benefited from support by GNAMPA (Italian Group for research in Analysis, Probability and their Applications. It is a subunit of the INdAM group, the Itaian Group for research in High Mathematics, with site in Rome) and EIF (Institut Europlace de Finance, subunit of the Institut Louis Bachelier in Paris).

2 MEAN SQUARE ERROR: general results

We compute and optimize the mean square error (MSE) of $\hat{IV}_n$ passing through the *conditional* expectation with respect to the paths of σ and J:

$$MSE := E[(\hat{IV}_n - IV)^2] = E\Big[E[(\hat{IV}_n - IV)^2|\sigma, J]\Big].$$

Conditioning on σ, as well as assuming no drift in X, is standard in papers where MSE-optimality is looked for, in the absence of jumps (see e.g. [1]). We assume evenly spaced observation over a fixed time horizon $[0, T]$, so that $t_i = t_{i,n} = ih_n$, for any $i = 1 \dots n$, with $h = h_n = T/n$. Denote by ε the square root of a given threshold function: $\varepsilon := \sqrt{r(\sigma, h)}$. $\hat{IV}_n$ and MSE are in fact functions of ε (other than of h), and we indicate them by $\hat{IV}_n(\varepsilon)$, $MSE(\varepsilon)$. Note that for $\varepsilon = 0$ we have $\hat{IV}_n = 0$, so $MSE(\varepsilon) = E[IV^2]$; as ε increases some squared increments $(\Delta_i X)^2$ are included within $\hat{IV}_n$, so $\hat{IV}_n$ becomes closer to IV and $MSE(\varepsilon)$ decreases. However, if $J \not\equiv 0$, for $\varepsilon \to +\infty$ the quantity $MSE(\varepsilon)$ increases again, since $\hat{IV}_n$ includes all the squared increments $(\Delta_i X)^2$ and thus $\hat{IV}_n$ estimates the global quadratic variation $IV + \sum_{s \le T} \Delta X_s^2$ of X at time T, and $MSE(\varepsilon)$ becomes close to $E[(\sum_{s\le T} \Delta X_s^2)^2]$. We look for a threshold $\varepsilon^\star$ giving

$$MSE(\varepsilon^\star) = \min_{\varepsilon \in [0,\infty[} MSE(\varepsilon).$$

In this section we analyze the first derivative $MSE'(\varepsilon)$ and we find that an optimal threshold exists, in the general framework where X is a semimartingale satisfying **A1** below, and we furnish an equation to which $\varepsilon^\star$ is a solution, while in Section 2.1, we find that $\varepsilon^\star$ is even unique. The equation has no explicit solution, but $\varepsilon^\star$ is a function of h and we can explicitly characterize the first order term of its asymptotic expansion in h. Clearly we can always find an approximation of the optimal threshold with arbitrary precision making use of numerical methods.

Let us denote

$$\Delta_i X_\star := \Delta_i X I_{\{(\Delta_i X)^2 \le \varepsilon^2\}}, \quad \sigma_i^2 := \int_{t_{i-1}}^{t_i} \sigma_s^2 ds, \quad m_i := \Delta_i J.$$

To guarantee that W remains a Brownian motion conditionally to σ and J, we need to assume the following

A1. A.s. $\sigma_s^2 > 0$ for all s, and σ, J are independent on W.

Theorem 1. *Under A1 and the finiteness of the expectation of the terms below, for fixed h and $\varepsilon > 0$, we have $MSE'(\varepsilon) = \varepsilon^2 G(\varepsilon)$, where*

$$G(\varepsilon) := \sum_i E\Big[a_i(\varepsilon)\Big(\varepsilon^2 + 2\sum_{j\neq i} b_j(\varepsilon) - 2IV\Big)\Big], \quad a_i(\varepsilon) := \frac{e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}} + e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}}{\sigma_i\sqrt{2\pi}},$$

$$b_i(\varepsilon) := E[(\Delta_i X_\star)^2|\sigma, J] = -\Big(e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}(\varepsilon + m_i) + e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}(\varepsilon - m_i)\Big)\frac{\sigma_i}{\sqrt{2\pi}} + \frac{m_i^2 + \sigma_i^2}{\sqrt{2\pi}} \int_{\frac{m_i-\varepsilon}{\sigma_i}}^{\frac{m_i+\varepsilon}{\sigma_i}} e^{-\frac{x^2}{2}} dx.$$

It clearly follows that $MSE'(\varepsilon) > 0$ if and only if $G(\varepsilon) > 0$ and, thus, to our aim of finding an optimal threshold, it suffices to study the sign of $G(\varepsilon)$ as ε varies.

Notation. For brevity we sometimes omit to precise the dependence on ε of $a_i(\varepsilon)$ and $b_i(\varepsilon)$.
For a function $f(\varepsilon)$ we sometimes use $f(+\infty)$ for $\lim_{\varepsilon\to+\infty} f(\varepsilon)$.
For two functions $f(x), g(x)$ of a non-negative variable x which tends to 0 (respectively to $+\infty$), by $f \ll g$, or $g \gg f$ we mean that $f = o(g)$ as $x \to 0$ (respectively $x \to +\infty$), by $f \asymp g$ we mean that both $f = O(g)$ and $g = O(f)$ as $x \to 0$ (respectively $x \to +\infty$), while by $f \sim g$ we mean that f and g are asymptotically equivalent (i.e. $f/g \to 1$).
We denote $\phi(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $\quad \bar{\Phi}(x) = \int_x^{+\infty} \phi(s)ds$.
h.o.t means *higher order terms*

Proof of Theorem 1. Under **A1** we have that conditionally to (σ, J) the increment $\Delta_i X = \int_{t_{i-1}}^{t_i} \sigma_s dW_s + \Delta_i J$ is a Gaussian r.v. with law $\mathcal{N}(m_i, \sigma_i^2)$, which allows to compute the conditional expectation $E[\hat{IV}_n|\sigma, J]$. We have

$$E[\hat{IV}_n|\sigma, J] = \sum_{i=1}^n b_i(\varepsilon) = \sum_{i=1}^n -\Big(e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}(\varepsilon+m_i) + e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}(\varepsilon-m_i)\Big)\frac{\sigma_i}{\sqrt{2\pi}}$$
$$+\frac{m_i^2+\sigma_i^2}{\sqrt{\pi}}\Big(\int_0^{\frac{\varepsilon-m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt + \int_0^{\frac{\varepsilon+m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt\Big),$$

and

$$E[(\hat{IV}_n(\varepsilon))^2|\sigma, J] = \sum_i E[(\Delta_i X_\star)^4|\sigma, J] + 2\sum_i\sum_{j>i} E[(\Delta_i X_\star)^2(\Delta_j X_\star)^2|\sigma, J]$$
$$= \sum_i \Big[-e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}\sigma_i\Big(\varepsilon^3 + m_i\varepsilon^2 + m_i^2\varepsilon + m_i^3 + 5m_i\sigma_i^2 + 3\sigma_i^2\varepsilon\Big)$$
$$-e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}\sigma_i\Big(\varepsilon^3 - m_i\varepsilon^2 + m_i^2\varepsilon - m_i^3 - 5m_i\sigma_i^2 + 3\sigma_i^2\varepsilon\Big)$$
$$+\Big(\int_0^{\frac{\varepsilon-m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt + \int_0^{\frac{\varepsilon+m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt\Big)\sqrt{2}\Big(m_i^4 + 6m_i^2\sigma_i^2 + 3\sigma_i^4\Big)\Big]\frac{1}{\sqrt{2\pi}} + 2\sum_i\sum_{j>i} b_i b_j, \tag{3}$$

having used that conditionally to σ and J, $\Delta_i X_\star$ and $\Delta_j X_\star$ are independent. It follows that

$$MSE(\varepsilon) = E\Bigg[\sum_i \Big[-e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}\sigma_i\Big(\varepsilon^3 + m_i\varepsilon^2 + m_i^2\varepsilon + m_i^3 + 5m_i\sigma_i^2 + 3\sigma_i^2\varepsilon\Big)$$
$$-e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}\sigma_i\Big(\varepsilon^3 - m_i\varepsilon^2 + m_i^2\varepsilon - m_i^3 - 5m_i\sigma_i^2 + 3\sigma_i^2\varepsilon\Big)$$
$$+\Big(\int_0^{\frac{\varepsilon-m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt + \int_0^{\frac{\varepsilon+m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt\Big)\sqrt{2}\Big(m_i^4 + 6m_i^2\sigma_i^2 + 3\sigma_i^4\Big)\Big]\frac{1}{\sqrt{2\pi}}$$
$$+2\sum_i\sum_{j>i} b_i(\varepsilon)b_j(\varepsilon) - 2IV\sum_{i=1}^n \Big[-\Big(e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}(\varepsilon+m_i) + e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}(\varepsilon-m_i)\Big)\frac{\sigma_i}{\sqrt{2\pi}}$$
$$+\frac{m_i^2+\sigma_i^2}{\sqrt{\pi}}\Big(\int_0^{\frac{\varepsilon-m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt + \int_0^{\frac{\varepsilon+m_i}{\sqrt{2}\sigma_i}} e^{-t^2}dt\Big)\Big] + IV^2\Bigg].$$

$MSE(\varepsilon)$ is a differentiable functions of ε, therefore to find the minimum on $[0, +\infty[$ of $MSE(\varepsilon)$ we can study the sign of its first derivative $MSE'(\varepsilon)$. Since $MSE'(\varepsilon) = \frac{d}{d\varepsilon}E[(\hat{IV}_n(\varepsilon))^2] - 2IV\frac{d}{d\varepsilon}E[\hat{IV}_n(\varepsilon)]$, we begin to compute $\frac{d}{d\varepsilon}E[\hat{IV}_n(\varepsilon)|\sigma, J]$. Note that

$$\frac{d}{d\varepsilon}b_i(\varepsilon) = \Big(e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}} + e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}\Big)\frac{(\varepsilon+m_i)(\varepsilon-m_i)}{\sigma_i\sqrt{2\pi}}$$

$$-\Big(e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}+e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}\Big)\frac{\sigma_i}{\sqrt{2\pi}}+\frac{m_i^2+\sigma_i^2}{\sigma_i\sqrt{2\pi}}\Big(e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}+e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}\Big)=\varepsilon^2\frac{e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}+e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}}{\sigma_i\sqrt{2\pi}}=\varepsilon^2 a_i(\varepsilon),$$

so that

$$\frac{d}{d\varepsilon}E[I\hat{V}_n(\varepsilon)|\sigma,J]=\varepsilon^2\sum_{i=1}^n a_i(\varepsilon) \tag{4}$$

is strictly greater than zero for all values of $\varepsilon>0$. As for $\frac{d}{d\varepsilon}E[(I\hat{V}_n(\varepsilon))^2|\sigma,J]$, note that the term $2\sum_i\sum_{j>i}b_ib_j$ in (3) can be written as $\sum_i\sum_{j\neq i}b_ib_j$, so its derivative coincides with $\sum_i\sum_{j\neq i}(\varepsilon^2a_ib_j+b_i\varepsilon^2a_j)$, however

$$\sum_i b_i\sum_{j\neq i}a_j=\Big(\sum_i b_i\sum_j a_j-\sum_i b_ia_i\Big)$$

$$=\Big(\sum_i a_i\sum_j b_j-\sum_i a_ib_i\Big)=\sum_i a_i\sum_{j\neq i}b_j$$

so that $\sum_i\sum_{j\neq i}(\varepsilon^2a_ib_j+b_i\varepsilon^2a_j)=2\sum_{i=1}^n\sum_{j\neq i}\varepsilon^2a_ib_j$,

$$\frac{d}{d\varepsilon}E[(I\hat{V}_n(\varepsilon))^2|\sigma,J]=\varepsilon^4\sum_i a_i(\varepsilon)+2\Big(\sum_{i=1}^n\sum_{j>i}b_i(\varepsilon)b_j(\varepsilon)\Big)' \tag{5}$$

$$=\sum_i\Big[\varepsilon^4a_i+2\varepsilon^2a_i\sum_{j\neq i}b_j\Big]$$

and

$$\frac{d}{d\varepsilon}MSE(\varepsilon)=\varepsilon^2\sum_i E\Big[\varepsilon^2a_i+2a_i\sum_{j\neq i}b_j-2IVa_i\Big].$$

$$=\varepsilon^2\textstyle\sum_i E\Big[a_i\Big(\varepsilon^2+2\sum_{j\neq i}b_j-2IV\Big)\Big] \tag{6}$$

$$=\varepsilon^2G(\varepsilon).$$

□

Remark 1. If also $J\not\equiv 0$, we have

$$MSE(0)=E[IV^2]>0\quad\text{ and, for small } h,\ \lim_{\varepsilon\to+\infty}MSE(\varepsilon)>0.$$

Corollary 1. *Under the same assumptions of Theorem 1, even in the absence of jumps, an optimal threshold exists and is solution of* $G(\varepsilon)=0$.

Proof. Note that $a_i(\varepsilon)$ and $b_i(\varepsilon)$ are continuously differentiable functions of ε, and, with fixed $h=\frac{T}{n}$,

$$a_i(0)=\frac{2e^{-\frac{m_i^2}{2\sigma_i^2}}}{\sigma_i\sqrt{2\pi}},\quad b_i(0)=0,$$

$$a_i(+\infty)=0,\quad b_i(+\infty)=E[(\Delta_iX_\star)^2|\sigma,J]=m_i^2+\sigma_i^2,$$

$$a_i'(\varepsilon)=-\frac{1}{\sigma_i^3\sqrt{2\pi}}\Big[e^{-\frac{(\varepsilon-m_i)^2}{2\sigma_i^2}}(\varepsilon-m_i)+e^{-\frac{(\varepsilon+m_i)^2}{2\sigma_i^2}}(\varepsilon+m_i)\Big],\quad b_i'(\varepsilon)=\varepsilon^2a_i(\varepsilon),$$

so we find that $G(0)=-\frac{4}{\sigma_i\sqrt{2\pi}}\sum_i E\Big[e^{-\frac{m_i^2}{2\sigma_i^2}}\cdot IV\Big]<0$, and $\lim_{\varepsilon\to+\infty}G(\varepsilon)=0^+$, so there exists $\varepsilon_+>0$: $MSE'(\varepsilon)>0$ on $[\varepsilon_+,+\infty)$. On the compact set $[0,\varepsilon_+]$ the continuous function MSE has necessarily absolute minimum value $\underline{MSE}$, and since on $[\varepsilon_+,+\infty)$ MSE is increasing we have that on $[0,+\infty)$ the absolute minimum is $\underline{MSE}$.

$MSE'(\varepsilon)$ is continuous and assumes both negative and positive values, thus equation $G(\varepsilon)=0$ has a solution. Any minimum point of MSE on $[0,+\infty)$ has to be a stationary point, so it has to solve the equation. □

Remark 2. In principle $MSE(\varepsilon)$ could even have many points ε where the absolute minimum value $\underline{MSE}$ of MSE on $[0, +\infty)$ is reached; MSE could even have an infinite number of local not absolute minima.

To determine the number of solutions to $G(\varepsilon) = 0$, we need to study the sign of $G'(\varepsilon)$ (corresponding to the convexity properties of $MSE(\varepsilon)$), but this is not easy. Define

$$g_i(\varepsilon) := \varepsilon^2 + 2\sum_{j\neq i} b_j - 2IV,$$

so that

$$G(\varepsilon) = \sum_i E[a_i(\varepsilon)g_i(\varepsilon)].$$

We can easily study the functions g_i, since we know that $g_i(0) = -2IV < 0$, $\lim_{\varepsilon\to+\infty} g_i(\varepsilon) = +\infty$ and $g_i'(\varepsilon) = 2\varepsilon(1 + \varepsilon\sum_{j\neq i} a_i) > 0$ for all $\varepsilon > 0$. However within the joint function $G(\varepsilon)$ the presence of the terms $a_i(\varepsilon)$ makes it difficult even to know whether $(a_i g_i)'$ is positive.

2.1 When X is Lévy

Let us assume

A2. X is Lévy.

We now have that $\sigma > 0$ is constant and $\Delta_i X_\star$ are i.i.d., so the equation characterizing $MSE'(\varepsilon) = 0$ is much simpler: from (6), since within $a_i \sum_{j\neq i} b_j$ the term m_i of a_i is independent on the terms m_j of b_j, we have

$$MSE'(\varepsilon) = \varepsilon^2 G(\varepsilon) = \varepsilon^2 n E[a_1(\varepsilon)]\Big(\varepsilon^2 + 2(n-1)E[b_1(\varepsilon)] - 2IV\Big).$$

Theorem 2. *If X is Lévy, equation*

$$\varepsilon^2 + 2(n-1)E[b_1(\varepsilon)] - 2IV = 0 \tag{7}$$

has a unique solution $\varepsilon^\star$ and, thus, there exists a unique optimal threshold, which is $\varepsilon^\star$.

Proof. For $\varepsilon > 0$ we have $MSE'(\varepsilon) > 0$ if and only if $G(\varepsilon) > 0$, which in turn is true if and only if

$$g(\varepsilon) := \varepsilon^2 + 2(n-1)E[b_1] - 2IV > 0$$

where, setting $m := m_1 = \Delta_1 J$, we recall that we have

$$\begin{aligned} E[b_1] = E\Big[&- \Big(e^{-\frac{(\varepsilon-m)^2}{2\sigma^2 h}}(\varepsilon+m) + e^{-\frac{(\varepsilon+m)^2}{2\sigma^2 h}}(\varepsilon-m)\Big)\frac{\sigma\sqrt{h}}{\sqrt{2\pi}} \\ &+ \frac{m^2+\sigma^2 h}{\sqrt{\pi}}\Big(\int_0^{\frac{\varepsilon-m}{\sqrt{2}\sigma\sqrt{h}}} e^{-t^2}dt + \int_0^{\frac{\varepsilon+m}{\sqrt{2}\sigma\sqrt{h}}} e^{-t^2}dt\Big)\Big]. \end{aligned}$$

The sign of $g(\varepsilon)$ is studied as follows:

$$g(0) = -2\sigma^2 T < 0,$$

$$\lim_{\varepsilon\to+\infty} g(\varepsilon) = +\infty,$$

$$g'(\varepsilon) = 2\varepsilon(1 + (n-1)\varepsilon E[a_1])$$

so that $g'(\varepsilon) > 0$ for all $\varepsilon > 0$, $n > 1$. That implies that $g(\varepsilon)$ starts at $\varepsilon = 0$ from a negative value and strictly increases towards $+\infty$, as ε increases, so that there exists a unique $\varepsilon^\star$ such that $g(\varepsilon) < 0$ for $\varepsilon \in [0, \varepsilon^\star[$, $g(\varepsilon^\star) = 0$ and $g(\varepsilon) > 0$ for $\varepsilon \in]\varepsilon^\star, +\infty[$. That implies in turn that $MSE(\varepsilon)$ has a unique minimum point in $\varepsilon^\star$, which is then the optimal threshold we were looking for: $\varepsilon^\star$ is the unique solution of equation (7), corresponding to $g(\varepsilon) = G(\varepsilon) = 0$. □

The equation in (7) has no explicit solution, however we can give some important indications to approximate $\varepsilon^\star$.

2.2 Asymptotic behavior of $\mathbb{E}\left(b_i(\varepsilon)\right)$

For the rest of Section 2 we assume that $\varepsilon := \varepsilon(h) = \varepsilon_h$, even when for brevity we omit to indicate the dependence on h. We still are under **A2**, so recall that

$$\mathbb{E}\left[b_i(\varepsilon)\right] = \mathbb{E}\left[\left|\sigma\Delta_i^n W + \Delta_i^n J\right|^2 \mathbf{1}_{\{|\sigma\Delta_i^n W+\Delta_i^n J|\le\varepsilon\}}\right],$$

is constant in i. Note that $\mathbb{E}[b_i(\varepsilon)]$ is finite for any Lévy process J, regardless of whether J has bounded first moment or not. We consider two cases: the case where J is a finite jump activity process and the one where this is a symmetric strictly stable process. The asymptotic characterization of $\mathbb{E}\left[b_i(\varepsilon)\right]$ will be used in Subsection 2.3 to deduce the asymptotic behavior of the optimal threshold $\varepsilon^\star$.

We anticipate that in Subsection 2.3 we will also see that an optimal threshold $\varepsilon^\star$ has to tend to 0 as $h\to 0$ and in such a way that $\frac{\varepsilon^\star}{\sqrt{h}} \to +\infty$.

2.2.1 Finite Jump Activity Lévy process

Theorem 3. *Let X be a finite jump activity Lévy process with jump size density f and with jump intensity λ. Suppose also that the restrictions of f on $(0,\infty)$ and $(-\infty,0)$ admit C_1 extensions on $[0,\infty)$ and $(-\infty,0]$, respectively. Then, for any $\varepsilon=\varepsilon(h)$ such that $\varepsilon\to 0$ and $\varepsilon \gg \sqrt{h}$, as $h\to 0$, we have*

$$\mathbb{E}\left[b_1(\varepsilon)\right] = \sigma^2 h - \frac{2}{\sqrt{2\pi}}\sigma\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2 h}} + \lambda h\frac{\varepsilon^3}{3}C(f) + O\left(h^2\right) + o\left(\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\right) + o\left(h\varepsilon^3\right),$$

where above $C(f) := f(0^+) + f(0^-)$.

Proof. By definition,

$$\mathbb{E}\left[b_1(\varepsilon)\right] = \mathbb{E}\left[(\Delta_1^n X)^2\mathbf{1}_{\{|\Delta_1^n X|<\varepsilon,\Delta_1^n N=0\}}\right] + \mathbb{E}\left[(\Delta_1^n X)^2\mathbf{1}_{\{|\Delta_1^n X|<\varepsilon,\Delta_1^n N\neq 0\}}\right] =: \mathcal{G}+\mathcal{L}. \quad (8)$$

By Lemma S.2 and Lemma S.5 with $k=2$ in [3], provided that $\varepsilon\to 0$, we have

$$\mathcal{L} := \mathbb{E}\left[(\Delta_1^n X)^2\mathbf{1}_{\{|\Delta_1^n X|<\varepsilon,\Delta_1^n N\neq 0\}}\right] \sim \lambda h\frac{\varepsilon^3}{3}C(f), \quad (h\to 0), \quad (9)$$

$$\mathcal{G} := \sigma^2 h - \frac{2}{\sqrt{2\pi}}\sigma\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2 h}} + O\left(h^2\right) + o\left(\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\right),$$

which shows the result. □

2.2.2 Strictly stable symmetric Lévy process

Let us start by noting that

$$\begin{aligned}\mathbb{E}[b_1(\varepsilon)] &= \mathbb{E}\left[\left(\sigma W_h + J_h\right)^2 \mathbf{1}_{\{|\sigma W_h + J_h|\le\varepsilon\}}\right]\\ &= \sigma^2\mathbb{E}\left[W_h^2\mathbf{1}_{\{|\sigma W_h+J_h|\le\varepsilon\}}\right] + 2\sigma\mathbb{E}\left[W_h J_h\mathbf{1}_{\{|\sigma W_h+J_h|\le\varepsilon\}}\right] + \mathbb{E}\left[J_h^2\mathbf{1}_{\{|\sigma W_h+J_h|\le\varepsilon\}}\right]\\ &=: C_h(\varepsilon) + D_h(\varepsilon) + E_h(\varepsilon).\end{aligned}$$

The first term above can be written as

$$C_h(\varepsilon) = \sigma^2 h - \sigma^2\mathbb{E}\left[W_h^2\mathbf{1}_{\{|\sigma W_h+J_h|>\varepsilon\}}\right] = \sigma^2 h - \sigma^2 h\left(C_h^+(\varepsilon) + C_h^-(\varepsilon)\right),$$

where

$$C_h^+(\varepsilon) = \mathbb{E}\left[W_1^2\mathbf{1}_{\{W_1+\sigma^{-1}h^{-1/2}J_h>\sigma^{-1}h^{-1/2}\varepsilon\}}\right], \quad C_h^-(\varepsilon) = \mathbb{E}\left[W_1^2\mathbf{1}_{\{W_1+\sigma^{-1}h^{-1/2}J_h<-\sigma^{-1}h^{-1/2}\varepsilon\}}\right].$$

By conditioning on J and using the fact that $\mathbb{E}[W_1^2 \mathbf{1}_{\{W_1>x\}}] = x\phi(x) + \bar{\Phi}(x)$, for all $x \in \mathbb{R}$, we have

$$C_h^{\pm}(\varepsilon) = \mathbb{E}\left[\left(\frac{\varepsilon}{\sigma\sqrt{h}} \mp \frac{J_h}{\sigma\sqrt{h}}\right)\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} \mp \frac{J_h}{\sigma\sqrt{h}}\right) + \bar{\Phi}\left(\frac{\varepsilon}{\sigma\sqrt{h}} \mp \frac{J_h}{\sigma\sqrt{h}}\right)\right].$$

In what follows, we determine the behavior of the above quantities under the assumption that $\varepsilon \gg \sqrt{h}$. The proofs of the following Lemma 1 and Lemma 2 are in an Appendix.

Lemma 1. Suppose that $\{J_t\}_{t\geq 0}$ is a symmetric Y-stable process with $Y \in (0, 2)$. Then, there exist constants K_1 and K_2 such that:

$$\mathbb{E}\left[\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\right] = \frac{1}{\sqrt{2\pi}} e^{-\frac{\varepsilon^2}{2\sigma^2 h}} - K_1 \varepsilon^{-1-Y} h^{\frac{3}{2}} + \text{h.o.t.} \tag{10}$$

$$\mathbb{E}\left[J_h \phi\left(\frac{\varepsilon}{\sqrt{h}} - \frac{J_h}{\sqrt{h}}\right)\right] = K_2 h \varepsilon^{1-Y} + \text{h.o.t.}. \tag{11}$$

Lemma 2. Suppose that $\{J_t\}_{t\geq 0}$ is a symmetric strictly stable process with Lévy measure $C|x|^{-Y-1}dx$. Then, the following asymptotics hold:

$$\mathbb{E}\left[\bar{\Phi}\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\right] = \frac{C}{Y} h \varepsilon^{-Y} + O\left(\varepsilon^{-2Y} h^2\right) + O\left(\mathbb{E}\left[\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\right]\right), \tag{12}$$

$$\mathbb{E}\left[J_h^2 \mathbf{1}_{\{|\sigma W_h + J_h| \leq \varepsilon\}}\right] = \frac{2C}{2-Y} h \varepsilon^{2-Y} + O\left(h^2 \varepsilon^{2-2Y}\right) + O\left(h^{\frac{4-Y}{2}}\right) + O\left(h^{\frac{2}{Y}}\right). \tag{13}$$

We are ready to show our main result in this part:

Theorem 4. *Let $X_t = \sigma W_t + J_t$, where W is a Wiener process and J is a symmetric strictly stable Lévy process with Lévy measure $C|x|^{-Y-1}$. Then, for any $\varepsilon = \varepsilon(h)$ such that $\varepsilon \to 0$ and $\varepsilon \gg \sqrt{h}$, as $h \to 0$, we have*

$$\mathbb{E}\left[b_1(\varepsilon)\right] = \sigma^2 h - \frac{2\sigma}{\sqrt{2\pi}} \sqrt{h} \varepsilon e^{-\frac{\varepsilon^2}{2\sigma^2 h}} + \frac{2C}{2-Y} h \varepsilon^{2-Y} + \text{h.o.t.}.$$

Proof. From Lemmas 1 and 2,

$$\begin{aligned}
C_h^{+}(\varepsilon) &= \mathbb{E}\left[\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right) + \bar{\Phi}\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\right] \\
&= \frac{\varepsilon}{\sigma\sqrt{h}}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{\varepsilon^2}{2\sigma^2 h}} - K_1 \varepsilon^{-1-Y} h^{\frac{3}{2}}\right) - \frac{1}{\sigma\sqrt{h}}\left(K_2 h \varepsilon^{1-Y}\right) + \frac{C}{Y} h \varepsilon^{-Y} + \text{h.o.t.} \\
&= \frac{\varepsilon}{\sigma\sqrt{h}\sqrt{2\pi}} e^{-\frac{\varepsilon^2}{2\sigma^2 h}} - \frac{K_2}{\sigma} h^{1/2} \varepsilon^{1-Y} + \text{h.o.t.},
\end{aligned}$$

where above we used that $\varepsilon^{-Y} h \ll h^{1/2}\varepsilon^{1-Y}$. Therefore, using that $D_h = 0$ and Lemma 2, with $K_3 = \frac{2C}{2-Y}$,

$$\begin{aligned}
\mathbb{E}[b_1(\varepsilon)] &= \mathbb{E}\left[(\sigma W_h + J_h)^2 \mathbf{1}_{\{|\sigma W_h + J_h| \leq \varepsilon\}}\right] = C_h(\varepsilon) + D_h(\varepsilon) + E_h(\varepsilon) \\
&= \sigma^2 h - 2\sigma^2 h\left(\frac{\varepsilon}{\sigma\sqrt{h}\sqrt{2\pi}} e^{-\frac{\varepsilon^2}{2\sigma^2 h}} - \frac{K_2}{\sigma} h^{1/2} \varepsilon^{1-Y}\right) + K_3 h \varepsilon^{2-Y} + \text{h.o.t.} \\
&= \sigma^2 h - \frac{2\sigma}{\sqrt{2\pi}} \sqrt{h} \varepsilon e^{-\frac{\varepsilon^2}{2\sigma^2 h}} + K_3 h \varepsilon^{2-Y} + \text{h.o.t.},
\end{aligned}$$

where above we used that $h\varepsilon^{2-Y} \gg h^{3/2}\varepsilon^{1-Y}$. □

2.3 Asymptotic behavior of $\varepsilon^{\star}$

We now assume

A3. $J \not\equiv 0$ and the support of any ΔJ_t is $\mathbb{R}$.

We firstly see that an optimal threshold $\varepsilon^{\star} = \varepsilon^{\star}(h)$ has to tend to 0 as $h \to 0$ and in such a way that $\frac{\varepsilon^{\star}}{\sqrt{h}} \to +\infty$. Then we will show the asymptotic behavior of $\varepsilon^{\star}$ in more detail.

Remark 3. Note that under **A3**, if $\varepsilon^\star(h)$ minimizes MSE, then necessarily $\varepsilon^\star(h) \to 0$ as $h \to 0$. Indeed, if $\liminf \varepsilon^\star(h) = c > 0$, then on a sequence $\varepsilon^\star(h)$ converging to c we would have $\hat{IV}_n - IV \to \sum_{s\le T} \Delta J_s^2 I_{|\Delta J_s|\le c}$ in probability, rather than $\hat{IV}_n - IV \to 0$; since $P\{\sum_{s\le T} \Delta J_s^2 I_{|\Delta J_s|\le c} > 0\} > 0$, the MSE could not be minimized.

Lemma 3. Suppose $X_t = \sigma W_t + J_t$, where W is a Brownian motion and J is a pure-jump Lévy process of bounded variation or, more generally, such that, for some $Y \in (0,2)$, $h_n^{-1/Y} J_{h_n} \xrightarrow{P} J$, for a real-valued random variable J. Then, $\varepsilon_n^\star/\sqrt{h_n} \to \infty$, as $n \to \infty$.

Remark. If J has FA jumps and drift η, $J_t = \eta t + \sum_{k=1}^{N_t} \gamma_k$, then we have $h^{-1} J_h \xrightarrow{P} \eta$ and, thus, the above assumption is satisfied with $Y = 1$.

Proof. We show the result by contradiction. Suppose that $\liminf_{n\to\infty} \frac{\varepsilon_n^\star}{\sqrt{h_n}} < \infty$. For simplicity and without loss of generality, we further assume that $\lim_{n\to\infty} \frac{\varepsilon_n^\star}{\sqrt{h_n}} =: L < \infty$ as all the statements below are valid on a subsequence $\{n_k\}_{k\ge 0}$. Let $M \in (0,\infty)$ be such that $\sup_n \frac{\varepsilon_n^\star}{\sqrt{h_n}} \le M$. Also, for simplicity, let us write ε_n for $\varepsilon_n^\star$ and assume that $T = 1$ so that $h_n = 1/n$. Consider the decomposition

$$\begin{aligned}
\mathbb{E}[b_1(\varepsilon)] &= \mathbb{E}\left[(\sigma W_h + J_h)^2 \mathbf{1}_{\{|\sigma W_h + J_h|\le\varepsilon\}}\right] \\
&= \sigma^2 \mathbb{E}\left[W_h^2 \mathbf{1}_{\{|\sigma W_h + J_h|\le\varepsilon\}}\right] + 2\sigma \mathbb{E}\left[W_h J_h \mathbf{1}_{\{|\sigma W_h + J_h|\le\varepsilon\}}\right] + \mathbb{E}\left[J_h^2 \mathbf{1}_{\{|\sigma W_h + J_h|\le\varepsilon\}}\right] \\
&=: c_h(\varepsilon) + d_h(\varepsilon) + e_h(\varepsilon).
\end{aligned}$$

Note that dominated convergence implies that

$$\frac{1}{h_n} c_{h_n}(\varepsilon_n) = \sigma^2 \mathbb{E}\left[W_1^2 \mathbf{1}_{\{|\sigma W_1 + h_n^{-1/2} J_{h_n}|\le h_n^{-1/2}\varepsilon_n\}}\right] \stackrel{n\to\infty}{\longrightarrow} \sigma^2 \mathbb{E}\left[W_1^2 \mathbf{1}_{\{|W_1|\le L/\sigma\}}\right] < \sigma^2,$$

since $h_n^{-1/2} J_{h_n} = h_n^{\frac{1}{Y}-\frac{1}{2}} (h_n^{-1/Y} J_{h_n}) \to 0$, in probability. For d_h note that

$$\sigma |W_1 h_n^{-1/2} J_{h_n}| \mathbf{1}_{\{|\sigma W_1 + h_n^{-1/2} J_{h_n}|\le h_n^{-1/2}\varepsilon_n\}} \le \sigma^2 |W_1|^2 + \sigma |W_1| h_n^{-1/2} \varepsilon_n \le \sigma^2 |W_1|^2 + \sigma |W_1| M,$$

therefore, again by dominated convergence

$$h_n^{-1} d_{h_n}(\varepsilon_n) = 2\sigma \mathbb{E}\left[W_1 h_n^{-1/2} J_{h_n} \mathbf{1}_{\{|\sigma W_1 + h_n^{-1/2} J_{h_n}|\le h_n^{-1/2}\varepsilon_n\}}\right] \stackrel{n\to\infty}{\longrightarrow} 0.$$

Similarly, since $(h_n^{-1/2} J_{h_n})^2 \mathbf{1}_{\{|\sigma W_1 + h_n^{-1/2} J_{h_n}|\le h_n^{-1/2}\varepsilon_n\}} \le 2\sigma^2 W_1^2 + 2h_n^{-1}\varepsilon_n^2 \le 2W_1^2 + 2M^2$,

$$h_n^{-1} e_{h_n}(\varepsilon_n) = \mathbb{E}\left[\left(h_n^{-1/2} J_{h_n}\right)^2 \mathbf{1}_{\{|\sigma W_1 + h_n^{-1/2} J_{h_n}|\le h_n^{-1/2}\varepsilon_n\}}\right] \stackrel{n\to\infty}{\longrightarrow} 0.$$

Finally, let us write the equation $\varepsilon_n^2 + 2(n-1)\mathbb{E}[b_1(\varepsilon_n)] - 2nh_n\sigma^2 = 0$ as

$$\varepsilon_n^2 + 2\frac{n-1}{n}\left(\frac{d_{h_n}(\varepsilon_n)}{h_n} + \frac{e_{h_n}(\varepsilon_n)}{h_n}\right) = 2\sigma^2 - 2\frac{n-1}{n}\frac{c_{h_n}(\varepsilon_n)}{h_n}. \tag{14}$$

The right-hand side of the equation converges to $2\sigma^2\left(1 - \mathbb{E}\left[W_1^2 \mathbf{1}_{\{|W_1|\le L/\sigma\}}\right]\right) > 0$, while the left hand side converges to 0 and this leads to a contradiction and therefore $\lim_{n\to\infty} \frac{\varepsilon_n^\star}{\sqrt{h_n}} = \infty$. □

We are now ready to show more precisely the asymptotic behavior of $\varepsilon^\star$. The following result covers the FA case.

Proposition 1. Let J have FA jumps and let $\varepsilon^\star = \varepsilon^\star(h)$ be the optimal threshold. Then,

$$\varepsilon^\star \sim \sqrt{2\sigma^2 h \ln \frac{1}{h}}, \quad \text{as} \quad h \to 0.$$

Proof. For simplicity, in what follows we take $T=1$ so that $h=1/n$. Again, recall that $\varepsilon^{\star}$ is the solution of

$$(\varepsilon^{\star})^2+2(n-1)\mathbb{E}[b_1(\varepsilon^{\star})]-2nh\sigma^2=0.$$

Throughout, we shall use that $\varepsilon^{\star}\gg\sqrt{h}$, as proved in the above lemma. For simplicity, we write ε instead of $\varepsilon^{\star}$. By the asymptotic behavior of $\mathbb{E}\left[b_1(\varepsilon)\right]$ described above,

$$\varepsilon^2+2(n-1)\left(\sigma^2h-\frac{2}{\sqrt{2\pi}}\sigma\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2h}}+\lambda h\frac{\varepsilon^3}{3}C(f)+O\left(h^2\right)+o\left(\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2h}}\right)+o\left(h\varepsilon^3\right)\right)-2nh\sigma^2=0,$$

and, thus, using that $h=1/n$,

$$\varepsilon^2-2\sigma^2h-\frac{4}{\sqrt{2\pi}}\sigma\frac{\varepsilon}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}+2\lambda\frac{\varepsilon^3}{3}C(f)+O\left(h\right)+o\left(\frac{\varepsilon}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}\right)+o\left(\varepsilon^3\right)=0. \tag{15}$$

Now, since $h=o(\varepsilon^2)$ (as assumed at the beginning), we can write the previous equation as

$$\varepsilon^2-\frac{4}{\sqrt{2\pi}}\sigma\frac{\varepsilon}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}+o\left(\frac{\varepsilon}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}\right)+o\left(\varepsilon^2\right)=0.$$

Dividing by ε and rearranging the terms,

$$\varepsilon\left(1+o(1)\right)=\frac{4}{\sqrt{2\pi}}\sigma\frac{1}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}\left(1+o(1)\right). \tag{16}$$

Then, taking logarithms of both sides and since $\ln(1+o(1))=o(1)$,

$$\ln\varepsilon+o(1)=-\frac{\varepsilon^2}{2\sigma^2h}-\frac{1}{2}\ln h+\ln\left(\frac{4\sigma}{\sqrt{2\pi}}\right)+o(1). \tag{17}$$

which can be written as

$$\ln\left(\frac{\varepsilon^2}{\sigma^2h}\right)+o(1)=-\frac{\varepsilon^2}{\sigma^2h}-2\ln h+\ln\left(\frac{8}{\pi}\right)+o(1)$$

Defining $\varpi=\varepsilon^2/(\sigma^2h)$, we can write

$$-\frac{\ln\varpi}{\varpi}+\frac{2\ln\frac{1}{h}}{\varpi}-\frac{\ln\frac{\pi}{8}}{\varpi}-\frac{o(1)}{\varpi}=1+\frac{o(1)}{\varpi}.$$

Therefore, making $h\to0$ and using that $\varpi\to\infty$ (since $\varepsilon\gg\sqrt{h}$),

$$\frac{2\ln\frac{1}{h}}{\varpi}\overset{h\to0}{\longrightarrow}1.$$

Recalling that $\varpi=\varepsilon^2/(\sigma^2h)$, we conclude the result. □

The following result specifies the asymptotic behavior of $\varepsilon^{\star}$ for symmetric strictly stable processes.

Proposition 2. Under the conditions of Theorem 4, the optimal threshold $\varepsilon^{\star}=\varepsilon^{\star}(h)$ is such that

$$\varepsilon^{\star}\sim\sqrt{(2-Y)\sigma^2h\ln\frac{1}{h}},\quad\text{as}\quad h\to0.$$

Proof. For simplicity, we again take $T=1$ so that $h=1/n$ and write ε instead of $\varepsilon^{\star}$. By the asymptotic behavior of $\mathbb{E}\left[b_1(\varepsilon)\right]$ described in Theorem 4, we can write $(\varepsilon^{\star})^2+2(n-1)\mathbb{E}[b_1(\varepsilon^{\star})]-2nh\sigma^2=0$ as

$$\varepsilon^2+2(n-1)\left(\sigma^2h-\frac{2\sigma}{\sqrt{2\pi}}\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2h}}+\frac{2C}{2-Y}h\varepsilon^{2-Y}+\text{h.o.t.}\right)-2nh\sigma^2=0,$$

and, thus, using that $h=o(\varepsilon^2)$ and $\varepsilon^2=o\left(\varepsilon^{2-Y}\right)$, we have

$$\frac{4C}{2-Y}\varepsilon^{2-Y}-\frac{4}{\sqrt{2\pi}}\sigma\frac{\varepsilon}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}+o\left(\frac{\varepsilon}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2h}}\right)+o\left(\varepsilon^{2-Y}\right)=0. \tag{18}$$

Dividing by ε and rearranging the terms,

$$\varepsilon^{1-Y}\left(1+o(1)\right)=\frac{2-Y}{C\sqrt{2\pi}}\sigma\frac{1}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\left(1+o(1)\right).$$

Then, taking logarithms of both sides and since $\ln(1+o(1))=o(1)$,

$$(1-Y)\ln\varepsilon+o(1)=-\frac{\varepsilon^2}{2\sigma^2 h}-\frac{1}{2}\ln h+\ln\left(\frac{(2-Y)\sigma}{C\sqrt{2\pi}}\right)+o(1),$$

which can be written as

$$\frac{1-Y}{2}\ln\left(\frac{\varepsilon^2}{\sigma^2 h}\right)+\frac{1-Y}{2}\ln\left(2\sigma^2\right)+\frac{1-Y}{2}\ln\left(h\right)+o(1)=-\frac{\varepsilon^2}{2\sigma^2 h}-\frac{1}{2}\ln h+\ln\left(\frac{(2-Y)\sigma}{C\sqrt{2\pi}}\right)+o(1).$$

Equivalently, writing $\varpi=\varepsilon^2/(\sigma^2 h)$ and dividing by $-\varpi$,

$$-(1-Y)\frac{\ln\varpi}{\varpi}+\frac{(2-Y)\ln\frac{1}{h}}{\varpi}-\frac{K}{-\varpi}=1+\frac{o(1)}{\varpi}.$$

and using that $\varpi\to\infty$ (since $\varepsilon\gg\sqrt{h}$), we get

$$\frac{(2-Y)\ln\frac{1}{h}}{\varpi}\overset{h\to 0}{\longrightarrow}1.$$

Recalling that $\varpi=\varepsilon^2/(\sigma^2 h)$, we conclude the result. □

3 Threshold criterion when $\varepsilon_h=\sqrt{2Mh\log\frac{1}{h}}$

Under the framework described in [6], in the case of equally spaced observations, the threshold criterion allows convergence of

$$\hat{IV}_n:=\sum_{i=1}^{n}(\Delta_i X)^2 I_{\{(\Delta_i X)^2\le r(\sigma_{t_{i-1}},h)\}}$$

to $IV_T=\int_0^T\sigma_s^2 ds$ when, for all $i=1,\ldots,n$, we have $r(\sigma_{t_{i-1}},h)=r(h)$ and $r(h)$ is a deterministic function of h s.t. $r(h)\to 0$, $\frac{r(h)}{h\log\frac{1}{h}}\to\infty$, as $h\to 0$. Here we show that, under finite activity jumps, the same estimator is also consistent in the case $r(\sigma,h)=2M_i h\log\frac{1}{h}$, where M_i are proper random numbers. Concretely, assume the following

A4. Let

$$dX_t=a_t dt+\sigma_t dW_t+dJ_t, \tag{19}$$

where $J_t=\sum_{i=1}^{N_t}\gamma_i$ for a non-explosive counting process N and real-valued random variables γ_j, a,σ are càdlàg and a.s. $\underline{\sigma}^2:=\inf_{s\in[0,T]}\sigma_s^2>0$.

Recall that a.s., the paths of a and of σ are bounded on $[0,T]$. Define $\bar{\sigma}^2:=\sup_{s\in[0,T]}\sigma_s^2$, then, the following Proposition and Corollary hold true.

Proposition 3. Under **A4**, if we choose $r_i(h)=2M_i h\log\frac{1}{h}$, with any $M_i(\omega)$ such that $M_i(\omega)\in[\inf_{s\in[t_{i-1},t_i]}\sigma_s^2(\omega),\bar{\sigma}]$, we have:

a.s. $\forall\eta>0$, for sufficiently small h: $\forall\, i=1,\ldots,n,\quad I_{\{(\Delta_i X)^2\le(1+\eta)r_i(h)\}}=I_{\{\Delta_i N=0\}}$.

Corollary 2. *For all $\eta>0$, we have $\sum_{i=1}^{n}(\Delta_i X)^2 I_{\{(\Delta_i X)^2\le(1+\eta)r_i(h)\}}\overset{P}{\to}IV$, as $h\to 0$.*

Proof of Proposition 3. In order to prove the proposition, we follow and modify the proof of Theorem 1 in [6], in that we show that a.s., for all $\eta>0$, for sufficiently small h, we have
1) $\forall i=1,\ldots,n, I_{\{\Delta_i N=0\}}\le I_{\{(\Delta_i X)^2\le(1+\eta)r_i(h)\}}$

2) $\forall i=1,\dots,n, I_{\{\Delta_i N=0\}} \geq I_{\{(\Delta_i X)^2\leq(1+\eta)r_i(h)\}}$.
Then the thesis follows.

Call $\Delta_i X_0 = \int_{t_{i-1}}^{t_i} a_s ds + \int_{t_{i-1}}^{t_i} \sigma_s dW_s$, $\bar{a} = \sup_{s\in[0,T]} |a_s|$, $\bar{\sigma} = \sup_{s\in[0,T]} \sigma_s$ and $\underline{\gamma}(\omega) = \min_{\ell:\Delta N_\ell\neq 0} |\gamma_\ell(\omega)|$, and note that under our assumptions $P(\underline{\gamma}\neq 0)=1$. To show 1) a 2) we use the following key fact:

$$\sup_{i\in\{1,\dots,n\}} \frac{|\Delta_i X_0|}{\sqrt{2M_i h\log\frac{1}{h}}} \leq \sup_i \frac{\bar{a}\sqrt{h}}{\sqrt{2M_i\log\frac{1}{h}}} +$$

$$\sup_i \frac{|B_{IV_{t_i}} - B_{IV_{t_{i-1}}}|}{\sqrt{2\Delta_i IV\log\frac{1}{\Delta_i IV}}} \sup_i \frac{\sqrt{2\Delta_i IV\log\frac{1}{\Delta_i IV}}}{\sqrt{2M_i h\log\frac{1}{M_i h}}} \sup_{i\in\{1,\dots,n\}} \frac{\sqrt{2h\log\frac{1}{M_i h}}}{\sqrt{2h\log\frac{1}{h}}},$$

where B is a standard Brownian motion and we used the fact that $\sigma\cdot W$ is a time changed Brownian motion ([7], theorems 1.9 and 1.10), meaning that we can represent $\Delta_i(\sigma\cdot W) = B_{IV_{t_i}} - B_{IV_{t_{i-1}}}$. By the Paul Lévy law on the modulus of continuity of the BM paths ([5], theorem 9.25) and the monotonicity of the function $x\ln(1/x)$ on $(0,1/e)$, it follows that for sufficiently small h the first two factors of the last line of last display are bounded above by 1, so that

$$\sup_i \frac{|\Delta_i X_0|}{\sqrt{2M_i h\log\frac{1}{h}}} \leq \sup_i \frac{\bar{a}\sqrt{h}}{\sqrt{2M_i\log\frac{1}{h}}} + \sup_i \sqrt{\frac{\log\frac{1}{M_i}}{\log\frac{1}{h}} + 1}$$

$$\leq M_h := \frac{\bar{a}\sqrt{h}}{\sqrt{2\underline{\sigma}^2\log\frac{1}{h}}} + \sqrt{\frac{\log\frac{1}{\underline{\sigma}^2}}{\log\frac{1}{h}} + 1}$$

which tends to 1, as $h\to 0$.

Now, in order to show 1), we define $\{J\} = \{i\in\{1,2,...,n\}:\ \Delta_i N\neq 0\}$, and it is sufficient to prove that for h small enough $\sup_{i\notin\{J\}} \frac{|\Delta_i X|}{\sqrt{r_i(h)}} \leq 1+\eta$. Indeed, $\sup_{i\notin\{J\}} \frac{|\Delta_i X|}{\sqrt{r_i(h)}} = \sup_{i\notin\{J\}} \frac{|\Delta_i X_0|}{\sqrt{r_i(h)}} \leq \sup_{i\in\{1,..,n\}} \frac{|\Delta_i X_0|}{\sqrt{r_i(h)}} \leq M_h \to 1$, thus for all $\eta>0$ for sufficiently small h, it is ensured that $\sup_{i\notin\{J\}} \frac{|\Delta_i X|}{\sqrt{r_i(h)}} < 1+\eta$, that is: for all i, if $\Delta_i N = 0$ then necessarily we have $|\Delta_i X| < (1+\eta)\sqrt{r_i(h)}$, and 1) follows.

In order to show 2) we prove that, for sufficiently small h, $\inf_{i\in\{J\}} \frac{|\Delta_i X|}{\sqrt{r_i(h)}} > 1+\eta$. In fact firstly note that for sufficiently small h all the increments of N are either 0 or 1. It follows that if $\Delta_i N\neq 0$, then $\Delta_i N = 1$, and $\Delta_i J$ coincides with the size, say γ_{ℓ_i}, of a single jump $\Delta_i J = \gamma_{\ell_i}$. Then $\frac{|\Delta_i X|}{\sqrt{r_i(h)}} \geq \frac{|\gamma_{\ell_i}|}{\sqrt{r_i(h)}} - \frac{|\Delta_i X_0|}{\sqrt{r_i(h)}}$ and

$$\inf_{i\in\{J\}} \frac{|\Delta_i X|}{\sqrt{r_i(h)}} \geq \frac{\underline{\gamma}}{\bar{\sigma}\sqrt{2h\log\frac{1}{h}}} - \sup_{i\in\{J\}} \frac{|\Delta_i X_0|}{\sqrt{2M_i h\log\frac{1}{h}}} \geq \frac{\underline{\gamma}}{\bar{\sigma}\sqrt{2h\log\frac{1}{h}}} - (1+\eta)$$

and this tends to $+\infty$ when $h\to 0$, thus $\inf_{i\in\{J\}} \frac{|\Delta_i X|}{\sqrt{r_i(h)}} > 1+\eta$, meaning that if $\Delta_i N\neq 0$ then necessarily $|\Delta_i X| > \sqrt{r_i(h)}(1+\eta)$, as we needed. □

Proof of Corollary 2. The proof of the Corollary is straightforward, in that a.s. we fix any $\eta>0$, and for sufficiently small h we have

$$\sum_{i=1}^n (\Delta_i X)^2 I_{\{(\Delta_i X)^2\leq(1+\eta)r_i(h)\}} = \sum_{i=1}^n (\Delta_i X)^2 I_{\{\Delta_i N=0\}} = \sum_{i=1}^n (\Delta_i X_0)^2 - \sum_{i=1}^n (\Delta_i X_0)^2 I_{\{\Delta_i N\neq 0\}} \stackrel{P}{\to} IV_T,$$

since the last term tends to 0 in probability, as $E[\sum_{i=1}^n (\Delta_i X_0)^2 I_{\{\Delta_i N\neq 0\}}] \leq N_T O(h)\to 0$. □

4 CONDITIONAL MEAN SQUARE ERROR: FA jumps case

We now put ourselves under **A1**. The quantity of our interest here, $cMSE(\varepsilon) \doteq E[(\hat{IV} - IV)^2|\sigma, J]$, is such that $\forall\omega,\ cMSE(0) = IV^2$ and as soon as $J \not\equiv 0$ then $cMSE(+\infty) > 0$, because $\hat{IV} \overset{\varepsilon\to+\infty}{\to} QV$. Further, from the proof of Theorem 1, we have

$$cMSE'(\varepsilon) = \varepsilon^2 F(\varepsilon), \text{ with } F(\varepsilon) \doteq \sum_{i=1}^n a_i g_i, \quad g_i = \varepsilon^2 + 2\sum_{j\neq i} b_j - 2IV.$$

We analyze the sign of $F(\varepsilon)$: for n, h fixed, σ_i^2 and m_i also are fixed, and we have $F(0) = -2IV\sum_{i=1}^n a_i < 0$, since $b_j(0) = 0$. Further we have $F(+\infty) = 0^+$: to see it, first note that, from the expression of $b_i(\varepsilon)$, $b_i(+\infty) = m_i^2 + \sigma_i^2$, then $g_i(\varepsilon) \sim \varepsilon^2 + 2\sum_{j\neq i} m_j^2 - 2\sigma_i^2 \sim \varepsilon^2$, as $\varepsilon \to +\infty$. Moreover, each $a_i \sim 2(2\pi)^{-1/2}\sigma_i^{-1}\exp\left(-\frac{\varepsilon^2}{2\sigma_i^2}\right)$, thus, for sufficiently large ε, $F = \sum_{i=1}^n a_i g_i$ is a finite sum of n positive terms $a_i g_i \le K(2\pi)^{-1/2}\sigma_i^{-1}\varepsilon^2\exp\left(-\frac{\varepsilon^2}{2\sigma_i^2}\right)$ for some constant K and fixed σ_i, so $F(\varepsilon) \to 0^+$, as $\varepsilon \to +\infty$. Since F is continuous, it follows that, even in the absence of jumps, an optimal threshold exists and solves $F(\varepsilon) = 0$.

We now assume also **A3**.

Remark 4. As in Remark 3, if $\bar{\varepsilon} = \bar{\varepsilon}(h)$ minimizes cMSE, then it has to be true that $\bar{\varepsilon} \to 0$, as $h \to 0$. In what follows we again also find that necessarily $\frac{\bar{\varepsilon}(h)}{\sqrt{h}} \to +\infty$.

A4'. We assume **A4** with $a \equiv 0$, constant $\sigma > 0$ and $nh = 1$.

When considering $h \to 0$, we assume to have a sufficiently small h so that a.s. the number of jumps occurring during $]t_{i-1}, t_i]$ is at most 1; note that for any t we have $m_i I_{t\in]t_{i-1},t_i]} \to \Delta J_t$, so when considering a jump time t we assume that h is sufficiently small so that the sign of any $m_i I_{t\in]t_{i-1},t_i]}$ is the same as the one of ΔJ_t, in particular if $\Delta J_t \neq 0$ then the increments m_i approaching it are non-zero.

4.1 Asymptotic behavior of $b_i(\varepsilon)$ and F

Proposition 4. Under **A1, A3, A4'**, if $\bar{\varepsilon} = \bar{\varepsilon}(h)$ solves $F(\varepsilon) = 0$ and $\bar{\varepsilon} = \bar{\varepsilon}(h) \to 0$, then $\frac{\bar{\varepsilon}(h)}{\sqrt{h}} \to +\infty$.

Proof. $\bar{\varepsilon}$ is such that $\sum_{i=1}^n a_i g_i = 0$, i.e. $\sum_{i=1}^n a_i(\bar{\varepsilon}^2 + 2\sum_{j\neq i} b_j - 2IV) = 0$. For simplicity let us rename $\bar{\varepsilon}$ by ε. If $\liminf_{h\to 0}\frac{\varepsilon(h)}{\sqrt{h}} = L \in [0, +\infty)$ we can find a subsequence (that we recall $\varepsilon(h)$) such that $\lim \frac{\varepsilon(h)}{\sqrt{h}} = L$. Note that

$$0 = \sum_{i=1}^n a_i\left(\frac{\varepsilon^2}{h} + \frac{2\sum_{j\neq i} b_j}{h} - 2\sigma^2 n\right) = \frac{\varepsilon^2}{h}\sum_{i=1}^n a_i + \frac{2}{h}\sum_{i=1}^n a_i\sum_{j\neq i} b_j - 2\sigma^2 n\sum_{i=1}^n a_i,$$

i.e.

$$\frac{\varepsilon^2}{h} = 2\sigma^2 n - \frac{2}{h}\frac{\sum_{i=1}^n a_i\sum_{j\neq i} b_j}{\sum_{i=1}^n a_i} = 2n\Big[\sigma^2 - \frac{\sum_{i=1}^n a_i\sum_{j\neq i} b_j}{\sum_{i=1}^n a_i}\Big]. \quad (20)$$

Now we show that $\sigma^2 - \frac{\sum_{i=1}^n a_i\sum_{j\neq i} b_j}{\sum_{i=1}^n a_i}$ tends to a strictly positive constant, which in turn means that equality (20) is impossible, since on any sequence $\varepsilon(h)$ such that $\frac{\varepsilon(h)}{\sqrt{h}} \to L$ the left term tends to L^2, while the right one tends to $+\infty$.

Let us then check that $\sigma^2 - \frac{\sum_{i=1}^n a_i\sum_{j\neq i} b_j}{\sum_{i=1}^n a_i}$ tends to a strictly positive constant. Since J has FA, a.s. we only have finitely many $\Delta J_t \neq 0$, and, for small h, N_T coincides with $\sum_{i=1}^n I_{m_i\neq 0}$. Recalling the explicit expression of b_j (also reported below), we have

$$\sum_{j\neq i} b_j = \sum_{j\neq i, m_j=0} b_j + \sum_{j\neq i, m_j\neq 0} b_j \le -(n - N_T)\frac{\sigma\sqrt{h}}{\sqrt{2\pi}}2\varepsilon e^{-\frac{\varepsilon^2}{2\sigma^2 h}} + (n - N_T)\frac{\sigma^2 h}{\sqrt{2\pi}}\int_{-\frac{\varepsilon}{\sigma\sqrt{h}}}^{\frac{\varepsilon}{\sigma\sqrt{h}}} e^{-\frac{x^2}{2}}dx$$

$$-\sum_{j\neq i, m_j\neq 0}\frac{\sigma\sqrt{h}}{\sqrt{2\pi}}\left(\varepsilon\Big(e^{-\frac{(\varepsilon-|m_j|)^2}{2\sigma^2 h}}+e^{-\frac{(\varepsilon+|m_j|)^2}{2\sigma^2 h}}\Big)+|m_j|\Big(e^{-\frac{(\varepsilon-|m_j|)^2}{2\sigma^2 h}}-e^{-\frac{(\varepsilon+|m_j|)^2}{2\sigma^2 h}}\Big)\right)+\sum_{j\neq i, m_j\neq 0}\frac{m_j^2+\sigma^2 h}{\sqrt{2\pi}}\int_{\frac{m_j-\varepsilon}{\sigma\sqrt{h}}}^{\frac{m_j+\varepsilon}{\sigma\sqrt{h}}} e^{-\frac{x^2}{2}}dx.$$

Now, the factors $\varepsilon e^{-\frac{\varepsilon^2}{2\sigma^2 h}}$ and $\varepsilon\Big(e^{-\frac{(\varepsilon-|m_j|)^2}{2\sigma^2 h}}+e^{-\frac{(\varepsilon+|m_j|)^2}{2\sigma^2 h}}\Big)+|m_j|\Big(e^{-\frac{(\varepsilon-|m_j|)^2}{2\sigma^2 h}}-e^{-\frac{(\varepsilon+|m_j|)^2}{2\sigma^2 h}}\Big)$ of $\frac{\sigma\sqrt{h}}{\sqrt{2\pi}}$ are strictly positive, so

$$\sum_{j\neq i} b_j \le (n-N_T)\frac{\sigma^2 h}{\sqrt{2\pi}}\int_{-\frac{\varepsilon}{\sigma\sqrt{h}}}^{\frac{\varepsilon}{\sigma\sqrt{h}}} e^{-\frac{x^2}{2}}dx+\sum_{j\neq i, m_j\neq 0}\frac{m_j^2+\sigma^2 h}{\sqrt{2\pi}}\int_{\frac{m_j-\varepsilon}{\sigma\sqrt{h}}}^{\frac{m_j+\varepsilon}{\sigma\sqrt{h}}} e^{-\frac{x^2}{2}}dx,$$

where if $\frac{\varepsilon(h)}{\sqrt{h}}\to L$ as $h\to 0$ then the first term of the rhs above tends to $d:=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\frac{L}{\sigma}}^{\frac{L}{\sigma}} e^{-\frac{x^2}{2}}dx<\sigma^2$, while each term of the latter finite sum tends to 0, since $\frac{|m_j|}{\sqrt{h}}\to\infty$, so the finite sum tends to 0. It follows that, for all i, $\sum_{j\neq i} b_j\le d+o(1)$, where $d<\sigma^2$, so $\frac{\sum_{i=1}^n a_i\sum_{j\neq i} b_j}{\sum_{i=1}^n a_i}\le d+o(1)$, and $\sigma^2-\frac{\sum_{i=1}^n a_i\sum_{j\neq i} b_j}{\sum_{i=1}^n a_i}\ge\sigma^2-d+o(1)\to\sigma^2-d>0$, as we wanted. □

We now check the asymptotic behavior of b_i and a_i when $\varepsilon=\varepsilon(h)$ tends to 0 as $h\to 0$ in such a way that $\frac{\varepsilon}{\sqrt{h}}\to+\infty$. To this end, for fixed σ, we define

$$b(\varepsilon,m,h):=-\frac{\sigma\sqrt{h}}{\sqrt{2\pi}}\left(e^{-\frac{(\varepsilon-m)^2}{2\sigma^2 h}}(\varepsilon+m)+e^{-\frac{(\varepsilon+m)^2}{2\sigma^2 h}}(\varepsilon-m)\right)+\frac{m^2+\sigma^2 h}{\sqrt{2\pi}}\int_{\frac{m-\varepsilon}{\sigma\sqrt{h}}}^{\frac{m+\varepsilon}{\sigma\sqrt{h}}} e^{-x^2/2}dx \tag{21}$$

$$a(\varepsilon,m,h):=\frac{e^{-\frac{(\varepsilon-m)^2}{2\sigma^2 h}}+e^{-\frac{(\varepsilon+m)^2}{2\sigma^2 h}}}{\sigma\sqrt{h}\sqrt{2\pi}}, \tag{22}$$

so that $b_j(\varepsilon)=b(\varepsilon,m_j,h)$ and $a_j(\varepsilon)=a(\varepsilon,m_j,h)$, and note that, as $h\to 0$, we have (see the Appendix for the simple proof),

$$b(\varepsilon,m,h)=\begin{cases}\sigma^2 h-\frac{2\sigma}{\sqrt{2\pi}}\varepsilon\sqrt{h}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}+\text{h.o.t.}, & \text{if}\quad m=0,\\[2ex] \frac{\sigma}{|m|\sqrt{2\pi}}\varepsilon^2\sqrt{h}e^{-\frac{(|m|-\varepsilon)^2}{2\sigma^2 h}}+\text{h.o.t.}, & \text{if}\quad m\neq 0.\end{cases} \tag{23}$$

$$a(\varepsilon,m,h)=\begin{cases}\frac{2}{\sigma\sqrt{2\pi}}\frac{1}{\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}, & \text{if}\quad m=0,\\[2ex] \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\sqrt{h}}e^{-\frac{(|m|-\varepsilon)^2}{2\sigma^2 h}}+\text{h.o.t.}, & \text{if}\quad m\neq 0,\end{cases} \tag{24}$$

It follows that

$$g_i(\varepsilon)=\varepsilon^2+2\sum_{j\neq i} b_j(\varepsilon)-2IV=\varepsilon^2+2\sum_{j\neq i:m_j\neq 0} b_j(\varepsilon)+2\sum_{j\neq i:m_j=0}\big(b_j(\varepsilon)-\sigma^2 h\big)-2\sum_{j\neq i:m_j\neq 0}\sigma^2 h-2\sigma^2 h$$

$$=\varepsilon^2+\frac{2}{\sqrt{2\pi}}\sqrt{h}\varepsilon\left(\sum_{j\neq i:m_j\neq 0}\frac{\sigma}{|m_j|}\varepsilon e^{-\frac{(|m_j|-\varepsilon)^2}{2\sigma^2 h}}-2\sum_{j\neq i:m_j=0}\sigma e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\right)-2\sum_{j\neq i:m_j\neq 0}\sigma^2 h-2\sigma^2 h+\text{h.o.t.}. \tag{25}$$

Given any sequence $\varepsilon=\varepsilon(h)=\varepsilon_h$, which tends to 0 as $h\to 0$ in such a way that $\frac{\varepsilon(h)}{\sqrt{h}}\to+\infty$, we now show that $F(\varepsilon_h)=F_0(\varepsilon_h)+R(\varepsilon_h)$, where $F_0(\varepsilon_h)$ is constituted by the leading terms of F, while $R(\varepsilon_h)$ gives the remainder higher order terms. A solution $\bar{\bar{\varepsilon}}$ of $F=0$ non necessarily is such that $F_0(\bar{\bar{\varepsilon}})=0$, however if with the ε_h above we have $F_0(\varepsilon_h)\to 0$ then the whole $F(\varepsilon_h)\to 0$, so it has to be true that ε_h is close (in a way that will become explicit later) to one of the solutions $\bar{\bar{\varepsilon}}$ of $F=0$.

Proposition 5. Under **A4'**, if $\varepsilon_h\to 0$ as $h\to 0$ in such a way that $\frac{\varepsilon(h)}{\sqrt{h}}\to+\infty$ then $F(\varepsilon_h)=F_0(\varepsilon_h)+\text{h.o.t.}$, where

$$F_0(\varepsilon_h):=\frac{\varepsilon_h}{h\sqrt{h}}e^{-\frac{\varepsilon_h^2}{2\sigma^2 h}}\Big(\varepsilon_h-\frac{e^{-\frac{\varepsilon_h^2}{2\sigma^2 h}}}{\sqrt{h}}\frac{4\sigma}{\sqrt{2\pi}}\Big)\frac{1}{\sigma\sqrt{2\pi}}.$$

Proof. For simplicity, in what follows, we omit the dependence on h in the functions $a(\varepsilon, m, h)$ and $b(\varepsilon, m, h)$ defined in (21-22). Let us recall that, under Assumption **A4'**, N_t is the number of jumps by time t, $\{\gamma_\ell\}_{\ell\geq 1}$ are the consecutive jumps of J and $\{J\} = \{J\}_{(n)} := \{i : \Delta_i^n N \neq 0\}$. It follows that, for h is small enough,

$$\begin{aligned}
F(\varepsilon_h) &= \sum_{i=1}^{n} a(\varepsilon, m_i)\left(\varepsilon^2 + 2\sum_{j\neq i} b(\varepsilon, m_j) - 2IV\right)\\
&= \sum_{i\notin\{J\}} a(\varepsilon, m_i)\left(\varepsilon^2 + 2\sum_{j\neq i: j\in\{J\}} b(\varepsilon, m_j) + 2\sum_{j\neq i: j\notin\{J\}} b(\varepsilon, m_j) - 2IV\right)\\
&\quad + \sum_{i\in\{J\}} a(\varepsilon, m_i)\left(\varepsilon^2 + 2\sum_{j\neq i: j\in\{J\}} b(\varepsilon, m_j) + 2\sum_{j\neq i: j\notin\{J\}} b(\varepsilon, m_j) - 2IV\right)\\
&= (n - N_T)a(\varepsilon, 0)\left[\varepsilon^2 - 2h\sigma^2(N_T+1) + 2\left(\sum_{k=1}^{N_T} b(\varepsilon, \gamma_k) + (n - N_T - 1)(b(\varepsilon, 0) - \sigma^2 h)\right)\right] +\\
&\quad + \sum_{\ell=1}^{N_T} a(\varepsilon, \gamma_\ell)\left[\varepsilon^2 - 2h\sigma^2 N_T + 2\left(\sum_{k\neq\ell} b(\varepsilon, \gamma_k) + (n - N_T)(b(\varepsilon, 0) - \sigma^2 h)\right)\right]\\
&= (n - N_T)\frac{2}{\sigma\sqrt{h}\sqrt{2\pi}}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\left[\varepsilon^2 - 2h\sigma^2(N_T+1) - 4(n - N_T - 1)\frac{\sigma\varepsilon\sqrt{h}}{\sqrt{2\pi}}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\right.\\
&\qquad\qquad \left. + 2\sum_{k=1}^{N_T}\frac{\sigma}{|\gamma_k|}\frac{\varepsilon^2\sqrt{h}}{\sqrt{2\pi}}e^{-\frac{(|\gamma_k|-\varepsilon)^2}{2\sigma^2 h}}\right] +\\
&\quad + \sum_{\ell=1}^{N_T}\frac{1}{\sigma\sqrt{h}\sqrt{2\pi}}e^{-\frac{(|\gamma_k|-\varepsilon)^2}{2\sigma^2 h}}\left[\varepsilon^2 - 2h\sigma^2 N_T - 4(n - N_T)\frac{\sigma\varepsilon\sqrt{h}}{\sqrt{2\pi}}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\right.\\
&\qquad\qquad \left. + 2\sum_{k\neq\ell}\frac{\sigma}{|\gamma_k|}\frac{\varepsilon^2\sqrt{h}}{\sqrt{2\pi}}e^{-\frac{(|\gamma_k|-\varepsilon)^2}{2\sigma^2 h}}\right] + \text{h.o.t.}.
\end{aligned}$$

In what follows we use the following notation:

$$v_h = \frac{\varepsilon_h}{\sqrt{h}}, \quad u_{\ell h} = \frac{1}{\sqrt{2\pi}}e^{-\frac{\left(v_h - \frac{|\gamma_\ell|}{\sqrt{h}}\right)^2}{2\sigma^2}}, \quad s_h = \frac{1}{\sqrt{2\pi}}e^{-\frac{v_h^2}{2\sigma^2}}, \quad p_{\ell h} = e^{-\frac{|\gamma_\ell|}{\sigma^2 h}\left(\frac{|\gamma_\ell|}{2} - \sqrt{h}v_h\right)}.$$

Now, since $u_{\ell h} = s_h p_{\ell h}$ and $p_{\ell h} \to 0$, as $h \to 0$,

$$\begin{aligned}
F(\varepsilon_h) &= (n - N_T)\frac{2}{\sigma}\sqrt{h}s_h\left[v_h^2 - 2\sigma^2(N_T+1) + 2\sigma v_h s_h\left(\varepsilon\sum_{k=1}^{N_T}\frac{1}{|\gamma_k|}p_{kh} - 2(n - N_T - 1)\right)\right] +\\
&\quad + \frac{1}{\sigma}\sqrt{h}s_h\sum_{\ell=1}^{N_T}p_{\ell h}\left[v_h^2 - 2\sigma^2 N_T + 2\sigma v_h s_h\left(\varepsilon\sum_{k\neq\ell}\frac{1}{|\gamma_k|}p_{kh} - 2(n - N_T)\right)\right] + \text{h.o.t.}\\
&= (n - N_T)\frac{2}{\sigma}\sqrt{h}s_h\left[v_h^2 - 4\sigma v_h s_h n\right] + \frac{1}{\sigma}\sqrt{h}s_h\sum_{\ell=1}^{N_T}p_{\ell h}\left[v_h^2 - 4\sigma v_h s_h n\right] + \text{h.o.t.}\\
&= \left(n - N_T + \frac{1}{2}\sum_{\ell=1}^{N_T}p_{\ell h}\right)\frac{2}{\sigma}\sqrt{h}s_h\left[v_h^2 - 4\sigma v_h s_h n\right] + \text{h.o.t.}\\
&= \frac{2n}{\sigma}\sqrt{h}s_h v_h\left[v_h - 4\sigma s_h n\right] + \text{h.o.t.}\\
&= \frac{\varepsilon_h}{h\sqrt{h}}e^{-\frac{\varepsilon_h^2}{2\sigma^2 h}}\left(\varepsilon_h - \frac{e^{-\frac{\varepsilon_h^2}{2\sigma^2 h}}}{\sqrt{h}}\frac{4\sigma}{\sqrt{2\pi}}\right)\frac{1}{\sigma\sqrt{2\pi}}. \quad \square
\end{aligned} \tag{26}$$

Note that $v_h \ll n$, but $s_h \to 0$, so which is the leading term between v_h and ns_h depends on the choice of v_h.

Remark 5. The asymptotic behavior (26) also holds for any drift process $\{a_t\}_{t\geq 0}$ that has almost surely locally bounded paths (recall that any cadlag a satisfies such a requirement) and that is independent on W. Indeed, for nonzero drift, by conditioning also on a, we have that

$$F(\varepsilon_h) = \sum_{i\notin\{J\}} a(\varepsilon, h\bar{a}_i)\left[\varepsilon^2 - 2h\sigma^2(N_T+1) + 2\left(\sum_{k=1}^{N_T} b(\varepsilon,\gamma_k + h\bar{a}_{i_k}) + \sum_{j\neq i: j\notin\{J\}} (b(\varepsilon, h\bar{a}_j) - \sigma^2 h)\right)\right] +$$
$$+ \sum_{\ell=1}^{N_T} a(\varepsilon, \gamma_\ell + h\bar{a}_{i_\ell})\left[\varepsilon^2 - 2h\sigma^2 N_T + 2\left(\sum_{k\neq\ell} b(\varepsilon,\gamma_k + h\bar{a}_{i_k}) + \sum_{j\neq i_\ell: j\notin\{J\}} (b(\varepsilon, h\bar{a}_j) - \sigma^2 h)\right)\right],$$

where $\bar{a}_i = \int_{t_{i-1}}^{t_i} a_s ds/h$ and the indices $i_1 < i_2 < \cdots < i_{N_T}$ are defined such that $\Delta_{i_k} J \neq 0$, while $\Delta_i J = 0$ for any other $i \notin \{i_1, i_2, \ldots, i_{N_T}\}$. Next, we can follow the same arguments as above using the facts that, if a has locally bounded paths, for any i and k

$$a(\varepsilon, h\bar{a}_i) = \frac{2}{\sigma}h^{-1/2}\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}}\right) + \text{h.o.t.}, \quad a(\varepsilon,\gamma_k + h\bar{a}_{i_k}) = \frac{1}{\sigma}h^{-1/2}\phi\left(\frac{|\gamma_k| - \varepsilon}{\sigma\sqrt{h}}\right)e^{-\frac{\gamma_k \bar{a}_{i_k}}{\sigma^2}} + \text{h.o.t.}.$$
$$b(\varepsilon, h\bar{a}_i) = \sigma^2 h - 2\sigma\varepsilon\sqrt{h}\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}}\right) + \text{h.o.t.}, \quad b(\varepsilon,\gamma_k + h\bar{a}_{i_k}) = \frac{\sigma}{|\gamma_k|}\varepsilon^2\sqrt{h}\phi\left(\frac{|\gamma_k| - \varepsilon}{\sigma\sqrt{h}}\right)e^{-\frac{\gamma_k \bar{a}_{i_k}}{\sigma^2}} + \text{h.o.t.}.$$

4.2 Asymptotic behavior of $\bar{\varepsilon}$

Corollary 3. *Under* **A1, A3, A4'** *we have that*

$$\bar{\varepsilon} \sim \sqrt{2\sigma^2 h \ln\frac{1}{h}}, \quad \text{as} \quad h \to 0.$$

Proof. In fact, from Proposition 4 and (26), we have that

$$F(\bar{\varepsilon}_h) = \frac{2}{\sigma} n\bar{s}_h \bar{v}_h \sqrt{h}\Big(\bar{v}_h - n\bar{s}_h \cdot 4\sigma\Big) + \text{h.o.t.} = 0,$$

where $\bar{v}_h := \bar{\varepsilon}_h/\sqrt{h}$ and $\bar{s}_h = \frac{e^{-\frac{\bar{\varepsilon}_h^2}{2h\sigma^2}}}{\sqrt{2\pi}}$. Thus,

$$\bar{v}_h - n\bar{s}_h \cdot 4\sigma + \text{h.o.t.} = 0, \tag{27}$$

or, equivalently,

$$\bar{\varepsilon}_h - \frac{e^{-\frac{\bar{\varepsilon}_h^2}{2h\sigma^2}}}{\sqrt{h}}\frac{4\sigma}{\sqrt{2\pi}} + \text{h.o.t.} = 0,$$

which is exactly the condition in (16), entailing that

$$\bar{\varepsilon}_h \sim \sqrt{2\sigma^2 h \ln\frac{1}{h}}, \quad \text{as} \quad h \to 0. \qquad \square$$

Now we aim at approximating any optimal $\bar{\varepsilon} := \bar{\varepsilon}_h$, which is such that $F(\bar{\varepsilon}) = 0$, using a sequence $\varepsilon_h = \sqrt{h}v_h$. To this end, we aim at making $F(\varepsilon_h) \to 0$ as quickly as possible, the only possible way being rendering v_h and ns_h (26) of the same order. So we want to choose v_h such that

$$v_h = ns_h \cdot \frac{4\sigma}{\sqrt{2\pi}} + \text{h.o.t.}, \tag{28}$$

which is exactly the condition in (27).

Remark 6. There exists a deterministic function w_h of h such that $w : (0,1] \to (0,+\infty)$ and

$$\begin{aligned} &1)\ w_h \to +\infty \\ &2)\ w_h\sqrt{h} \to 0 \\ &3)\ \frac{e^{-w_h^2}}{w_h h} \to \frac{\sqrt{\pi}}{2} \end{aligned} \tag{29}$$

as $h \to 0$. In fact, for example a function of type $w_h = \sqrt{\ln\frac{1}{h} - \frac{1}{2}\ln\ln\frac{1}{h} - \ln y_h}$, with any continuous function y_h tending to $\frac{\sqrt{\pi}}{2}$ as $h \to 0$, satisfies the 3 conditions[1]. Then $v_h = \sqrt{2}\sigma w_h$ satisfies (28). However the quickest convergence speed of F to 0 would be reached by choosing a function w_h which satisfies the following three more restrictive conditions, as $h \to 0$,

$$\begin{aligned} &1)\ w_h \to +\infty \\ &2)\ w_h\sqrt{h} \to 0 \\ &3')\ \frac{e^{-w_h^2}}{w_h h} \equiv \frac{\sqrt{\pi}}{2}, \end{aligned} \tag{30}$$

where condition 3') means that $F_0(\varepsilon_h) \equiv 0$. In fact such a w_h exists, since the following holds true[2].

Theorem 5. *There exists a unique deterministic function w_h of h such that $w_h : (0,1] \to (0,+\infty)$ and the three conditions 1), 2) and 3') are satisfied. Such a w_h turns out to be differentiable and to satisfy also the ODE $w_h' = \frac{w_h h}{1+2w_h^2}$, which entails that $w_h \le w_1 + \frac{1}{2\sqrt{2}}\log\frac{1}{h}$.*

We finally reach the uniqueness of the optimal threshold $\bar\varepsilon$ as a consequence of the following Proposition, whose proof is in Appendix.

Proposition 6. The first derivative $\frac{d}{d\varepsilon}F(\varepsilon)$ of F is such that, when evaluated at a function ε_h of h such that $\varepsilon_h \to 0$, $\frac{\varepsilon_h}{\sqrt{h}} \to +\infty$, and $\varepsilon_h = 4\sigma\frac{s_h}{\sqrt{h}} + h.o.t.$, as $h \to 0$, then

$$F'(\varepsilon_h) = F_1(\varepsilon_h) + \text{h.o.t., as } h \to 0, \quad \text{where} \quad F_1(\varepsilon_h) = \frac{4}{\sigma^2\pi}e^{-\frac{\varepsilon_h^2}{\sigma^2 h}}\frac{\varepsilon_h^2}{h^3}.$$

Remark 7. Uniqueness of $\bar\varepsilon$. Since $F_1(\varepsilon_h) > 0$ for any ε_h, we reach that for sufficiently small h we have $\frac{d}{d\varepsilon}F(\varepsilon_h) > 0$ on any sequence ε_h as in the above Proposition. That entails that for any sufficiently small h the cMSE optimal $\bar\varepsilon$ is unique. In fact if there existed two optimal $\bar\varepsilon_h^{(1)} < \bar\varepsilon_h^{(2)}$ we would necessarily have that $\bar\varepsilon_h^{(i)} \to 0$, $\frac{\bar\varepsilon_h^{(i)}}{\sqrt{h}} \to +\infty$ and $\bar\varepsilon_h^{(i)} = \frac{4\sigma}{\sqrt{2\pi}}\frac{\bar s_h}{\sqrt{h}} + h.o.t.$, but then, for small h, on such sequences $F' > 0$, and then on such sequences F is strictly increasing, and thus $F(\bar\varepsilon_h^{(1)}) < F(\bar\varepsilon_h^{(2)})$, which is a contradiction, because in order to be optimal both sequences have to satisfy $F(\bar\varepsilon_h^{(i)}) = 0$.

Remark 8. The asymptotic behavior of the optimal threshold $\bar\varepsilon = \bar\varepsilon(h)$ for the cMSE criterion is the same as the one of the optimal threshold $\varepsilon^\star$ for the MSE criterion under FA jumps.
This is due to the fact that $\bar\varepsilon$ solves $F = 0$, $\varepsilon^\star$ solves $G = 0$, $F = F_0 + h.o.t.$, $G = G_0 + h.o.t.$, and the leading terms in F are the ones with $m_i = 0$, which do not depend on ω, thus they are the same as for G. It follows that, in the case of Lévy FA jumps, we have $F = F_0 + h.o.t. = E[F_0] + h.o.t. = G + h.o.t.$. Also, an alternative heuristic justification is that we expect that $F(\varepsilon) = \frac{\sum_{i=1}^n a_i g_i}{n} \cdot n \sim nE[a_i g_i]$, thus the asymptotic behavior of the $\varepsilon^\star$ satisfying $G = nE[a_i g_i] = 0$ is the same as any $\bar\varepsilon$ satisfying $F(\varepsilon) = 0$.

Remark 9. Comparison with the results in [2]. In [2] a FA jumps process X is considered, either of Lévy type, with jumps sizes having distribution density satisfying given conditions, or of Itô SM type, with deterministic absolutely

[1] We thank Andrey Sarychev for having provided such nice examples.
[2] We thank Salvatore Federico for having provided a such nice result. The proof is available upon request.

continuous local characteristics (additive process). The estimators

$$\hat{J}_n = \sum_{i=1}^{n} \Delta_i X I_{\{|\Delta_i X| > \varepsilon_h\}}, \quad \hat{N}_n = \sum_{i=1}^{n} I_{\{|\Delta_i X| > \varepsilon_h\}}$$

are considered, and, as $h \to 0$, firstly it is shown that the condition $\frac{\varepsilon_h}{\sqrt{h}} \to +\infty$ is necessary and sufficient for the convergence to 0 of both $MSE(\hat{IV}_n - IV)$ (stronger condition implying consistency of $\hat{IV}_n$) and $MSE(\hat{J}_n - J_T)$. Secondly, the authors show that

$$MSE(\hat{N}_n - N_T) \to 0 \Leftrightarrow \frac{e^{-\frac{\varepsilon_h^2}{2\sigma^2 h}}}{\sqrt{h}\varepsilon_h} \to 0,$$

meaning that in order to have $L^2(\Omega, P)$ convergence to 0 of the estimation error $\hat{N}_n - N_T$ a stronger condition on ε_h is needed, implying $\frac{\varepsilon_h}{\sqrt{h}} \to \infty$. Thirdly, existence and uniqueness of an optimal threshold $\breve{\varepsilon}(h)$ minimizing

$$E[|\hat{IV}_n - IV|^2 + |\hat{N}_n - N_T|^2]$$

for fixed h is obtained, and the asymptotic expansion in h of $\breve{\varepsilon}(h)$ has leading term $\sqrt{3\sigma^2 h \log \frac{1}{h}}$. The factor 3 is higher than the factor 2 of the leading terms of $\bar{\varepsilon}$ and $\varepsilon^{\star}$: that is due to the fact that the minimization criterion for $\breve{\varepsilon}(h)$ includes also the error on N_T, which requires that $\frac{\breve{\varepsilon}(h)}{\sqrt{h}}$ is higher than $\frac{\bar{\varepsilon}}{\sqrt{h}}$, and thus $\breve{\varepsilon}(h) > \bar{\varepsilon}(h)$ is necessary.

5 A NEW METHOD FOR FINITE JUMP ACTIVITY PROCESSES

In this section, we propose a new method to tuneup the threshold parameter $\varepsilon := \sqrt{r(\sigma, h)}$ of the Threshold Realized Variance (TRV) introduced in (2). This is based on the conditional mean square error $cMSE(\varepsilon) = E[(\hat{IV} - IV)^2 | \sigma, J]$ studied in Section 4. We illustrate the method for a driftless FA process with constant volatility σ. As proved therein, the optimal threshold $\bar{\varepsilon}$ is such that

$$F(\bar{\varepsilon}) = \sum_{i=1}^{n} a_i(\bar{\varepsilon}) g_i(\bar{\varepsilon}) = 0, \quad g_i(\bar{\varepsilon}) = \bar{\varepsilon}^2 + 2\sum_{j \neq i} b_j(\bar{\varepsilon}) - 2nh\sigma^2,$$

where $a_i(\varepsilon)$ and $b_i(\varepsilon)$ are rewritten here for easy reference:

$$a_i(\varepsilon) := a(\varepsilon, m_i, \sigma) := \frac{e^{-\frac{(\varepsilon - m_i)^2}{2\sigma^2 h}} + e^{-\frac{(\varepsilon + m_i)^2}{2\sigma^2 h}}}{\sigma\sqrt{2\pi h}},$$

$$b_i(\varepsilon) := b(\varepsilon, m_i, \sigma) := -\frac{\sigma\sqrt{h}}{\sqrt{2\pi}} \left(e^{-\frac{(\varepsilon - m_i)^2}{2\sigma^2 h}} (\varepsilon + m_i) + e^{-\frac{(\varepsilon + m_i)^2}{2\sigma^2 h}} (\varepsilon - m_i) \right) + \frac{m_i^2 + \sigma^2 h}{\sqrt{2\pi}} \int_{\frac{m_i - \varepsilon}{\sigma\sqrt{h}}}^{\frac{m_i + \varepsilon}{\sigma\sqrt{h}}} e^{-x^2/2} dx.$$

For future reference we set $\mathbf{m} = (m_1, \ldots, m_n)$ and

$$F(\varepsilon; \sigma, \mathbf{m}) := \sum_{i=1}^{n} a(\varepsilon, m_i, \sigma) \left(\varepsilon^2 + 2\sum_{j \neq i} b(\varepsilon, m_j, \sigma) - 2nh\sigma^2 \right)$$

The main issue with the optimal threshold $\bar{\varepsilon}$ lies on the fact that this depends on σ and the increments $\mathbf{m} = (m_1, \ldots, m_n)$ of the jump process, which we don't know. Note also that, for h small enough, each m_i will be either 0 or one of the jumps of the process and a good proxy of m_i is actually $(\Delta_i^n X)\mathbf{1}_{\{|\Delta_i^n X| > \bar{\varepsilon}\}}$. The idea is then to iteratively estimating $\bar{\varepsilon}$, σ, and $\mathbf{m}$ as follows: [3]

1. Start with some initial 'guesses' of σ and $\mathbf{m}$, which we call $\hat{\sigma}_0$ and $\hat{\mathbf{m}}_0$. In the sequel, we obtain $\hat{\sigma}_0$ by assuming that there is no jump; that is, we set $\hat{\mathbf{m}}_0 = (0, \ldots, 0)$ and $\hat{\sigma}_0^2 = T^{-1} \sum_{i=1}^{n} (\Delta_i^n X)^2$.

[3] To be consistent with section 4, I corrected $\varepsilon^{\star}$ here with $\bar{\varepsilon}$ and put $\breve{\varepsilon}$ for the optimal threshold of [2]. Check whether you approve.

2. Using $\hat{\sigma}_0$ and $\hat{\mathbf{m}}_0$, we then find an initial estimate for the optimum $\bar{\varepsilon}$ that we denote $\bar{\varepsilon}_0$. Thus, under the no-jump initial guess of the previous item, $\bar{\varepsilon}_0$ is such that $F(\bar{\varepsilon}_0; \hat{\sigma}_0, \hat{\mathbf{m}}_0) = 0$ or, more specifically, $\bar{\varepsilon}_0$ solves the equation:
$$\varepsilon^2 + 2(n-1)\left(-\frac{2\hat{\sigma}_0\sqrt{h}}{\sqrt{2\pi}}e^{-\frac{\varepsilon^2}{2\hat{\sigma}_0^2 h}}\varepsilon + \frac{\hat{\sigma}_0^2 h}{\sqrt{2\pi}}\int_{\frac{-\varepsilon}{\hat{\sigma}_0\sqrt{h}}}^{\frac{\varepsilon}{\hat{\sigma}_0\sqrt{h}}} e^{-x^2/2}dx\right) - 2nh\hat{\sigma}_0^2 = 0. \tag{31}$$
It is easy to see that $\bar{\varepsilon}_0$ is of the form $v_n\hat{\sigma}_0\sqrt{h}$, where v_n is the unique solution of the equation:
$$v_n^2 + 4(n-1)\left(-v_n\frac{1}{\sqrt{2\pi}}e^{-\frac{v_n^2}{2}} + \frac{1}{\sqrt{2\pi}}\int_0^{v_n} e^{-x^2/2}dx\right) - 2n = 0. \tag{32}$$
Figure 1 shows that v_n ranges from about 3 to 4 when n ranges from 100 to 10000.

3. Once we have an initial estimate of $\bar{\varepsilon}_0$, we can update our estimates of σ and $\mathbf{m}$ using the estimators:
$$\hat{\sigma}_1^2 := \hat{IV}_n(\bar{\varepsilon}_0) := \sum_{i=1}^n (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X| \le \bar{\varepsilon}_0\}}, \quad \hat{\mathbf{m}}_1 := ((\Delta_1^n X)\mathbf{1}_{\{|\Delta_1^n X| > \bar{\varepsilon}_0\}}, \dots, (\Delta_n^n X)\mathbf{1}_{\{|\Delta_n^n X| > \bar{\varepsilon}_0\}}) \tag{33}$$

4. We continue this procedure iteratively by setting $\bar{\varepsilon}_k$ such that $F(\bar{\varepsilon}_k; \hat{\sigma}_k, \hat{\mathbf{m}}_k) = 0$, which is then used to get
$$\hat{\sigma}_{k+1}^2 := \sum_{i=1}^n (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X| \le \bar{\varepsilon}_k\}}, \quad \hat{\mathbf{m}}_{k+1} := ((\Delta_1^n X)\mathbf{1}_{\{|\Delta_1^n X| > \bar{\varepsilon}_k\}}, \dots, (\Delta_n^n X)\mathbf{1}_{\{|\Delta_n^n X| > \bar{\varepsilon}_k\}}). \tag{34}$$
We stop when the sequence of estimates $\hat{\sigma}_{k+1}$ stabilizes (e.g., when $|\hat{\sigma}_{k+1} - \hat{\sigma}_k| \le \text{tol}$, for some desired small tolerance tol).

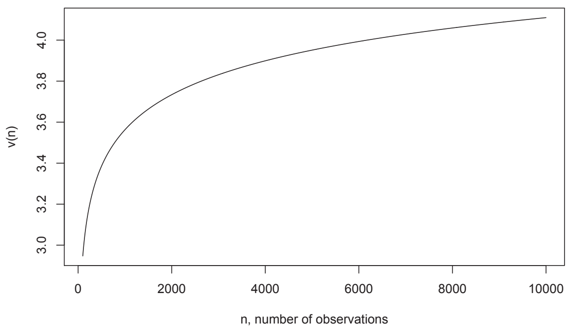

Figure 1: The solution v_n of equation (32) as a function of n.

The previous procedure resembles the one introduced in [2], which is based on choosing the threshold ε so to minimize $E[|\hat{IV}_n - IV|^2 + |\hat{N}_n - N_T|^2]$, or equivalently the expected number of jumps miss-classifications:
$$\text{Loss}(\varepsilon) := E\left[\sum_{i=1}^n \left(\mathbf{1}_{\{|\Delta_i^n X| > \varepsilon, \Delta_i^n N = 0\}} + \mathbf{1}_{\{|\Delta_i^n X| \le \varepsilon, \Delta_i^n N > 0\}}\right)\right]. \tag{35}$$
It was proved therein that, for a FA Lévy processes, the optimal threshold, denoted $\check{\varepsilon}_h$, is asymptotically equivalent to $\sqrt{3\sigma^2 h \ln(1/h)}$, as $h \to 0$. Using this information, an iterative method was proposed, in which, given an initial estimate $\check{\sigma}_0$ of σ, it was set
$$\check{\varepsilon}_k := \sqrt{3\check{\sigma}_k^2 h \ln\frac{1}{h}}, \quad \check{\sigma}_{k+1}^2 := \sum_{i=1}^n (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X| \le \check{\varepsilon}_k\}}, \quad k \ge 0. \tag{36}$$

In the light of the procedure used in [2], we adopt here also the following simpler one, other than the procedure (31)-(34) described above. Since, as proved in Section 4, the optimal threshold $\bar{\varepsilon}_h$ has the asymptotic behavior $\sqrt{2\sigma^2 h \ln(1/h)}$, as $h \to 0$, it is natural to consider the following iterative method to estimate $\bar{\varepsilon}$:

$$\bar{\varepsilon}_k^\star := \sqrt{2\bar{\sigma}_k^2 h \ln\frac{1}{h}}, \quad \bar{\sigma}_{k+1}^2 := \sum_{i=1}^n (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X| \le \bar{\varepsilon}_k^\star\}}, \quad k \ge 0, \tag{37}$$

starting again from an initial guess $\bar{\sigma}_0$ of σ. It can be proved that if we take both $\check{\sigma}_0^2$ and $\bar{\sigma}_0^2$ equal to the realized quadratic variation $T^{-1}\sum_{i=1}^n (\Delta_i^n X)^2$ in both (36) and (37), then the sequences of estimates $\{\bar{\sigma}_k\}_{k\ge 0}$, $\{\check{\sigma}_k\}_{k\ge 0}$ is nonincreasing and, thus, eventually $\check{\sigma}_k = \check{\sigma}_{k+1}$ and $\bar{\sigma}_k = \bar{\sigma}_{k+1}$, for some k. So, we can (and will) set the tolerance tol to 0.

5.1 Simulation results

We now proceed to assess the methods introduced above. We take a Lévy Merton's log-normal model of the form:

$$X_t = at + \sigma W_t + \sum_{j=1}^{N_t} \gamma_j,$$

where N is a Poisson process with intensity λ and $\{\gamma_i\}_{i\ge 1}$ is an independent sequence of independent normally distributed variables with mean and standard deviation μ^{Jmp} and σ^{Jmp}, respectively.

We consider the following estimators:

1. $\sigma_0 := \sqrt{T^{-1}\sum_{i=1}^n (\Delta_i^n X)^2}$;
2. The estimator $\hat{\sigma}_1$ as defined in (33) with initial guesses $\hat{\sigma}_0^2 = T^{-1}\sum_{i=1}^n (\Delta_i^n X)^2$ and $\hat{\mathbf{m}}_0 = (0, \ldots, 0)$;
3. $\hat{\sigma}_k$ found with the new method described by the iterative formulas (34). We stop when $|\hat{\sigma}_k - \hat{\sigma}_{k-1}| \le \text{tol} = 10^{-5}$;
4. The estimator $\check{\sigma}_1^2$ as in (36) with $k = 0$, using the threshold $\sqrt{3\check{\sigma}_0^2 h \log(1/h)}$ with $\check{\sigma}_0 = \sqrt{T^{-1}\sum_{i=1}^n (\Delta_i^n X)^2}$;
5. The estimator $\check{\sigma}_k$ defined by (36) with k such that $\check{\sigma}_k = \check{\sigma}_{k-1}$, $k \ge 1$;
6. The estimator $\bar{\sigma}_1^2$ as in (37) with $k = 1$, using the threshold $\bar{\varepsilon}_0^\star = \sqrt{2\bar{\sigma}_0^2 h \log(1/h)}$ with $\bar{\sigma}_0 = \sqrt{T^{-1}\sum_{i=1}^n (\Delta_i^n X)^2}$;
7. The estimator $\bar{\sigma}_k$ defined by the iterative formulas (37) and with k such that $\bar{\sigma}_k = \bar{\sigma}_{k-1}$, $k \ge 1$;
8. Threshold Realized Variance using the threshold $\varepsilon = h^\omega$ with $\omega = 0.495$
9. Threshold Realized Variance using the threshold $\varepsilon = 2h^\omega$ with $\omega = 0.495$
10. Realized Bipower Variation (BPV)
11. Threshold Realized Variance using a threshold of the form $4h^\omega\sqrt{\text{BPV}/T}$ with $\omega = 0.49$ (this is used in the recent work [4]);
12. The estimator $\hat{\sigma}_1^2 := \sum_{i=1}^n (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X| \le \bar{\varepsilon}_0\}}$ given in (33) where $\bar{\varepsilon}_0$ is such that $F(\bar{\varepsilon}_0; \hat{\sigma}_0, \hat{\mathbf{m}}_0) = 0$, but this time taking $\hat{\sigma}_0^2 = T^{-1}\sum_{i=1}^n (\Delta_i^n X)^2 \mathbf{1}_{\{|\Delta_1^n X| \le \bar{\varepsilon}_0^\star\}}$ and $\hat{\mathbf{m}}_0 := ((\Delta_1^n X)\mathbf{1}_{\{|\Delta_1^n X| > \bar{\varepsilon}_0^\star\}}, \ldots, (\Delta_n^n X)\mathbf{1}_{\{|\Delta_n^n X| > \bar{\varepsilon}_0^\star\}})$ with $\bar{\varepsilon}_0^\star$ as defined in the item 6 above.
13. The estimator $\hat{\sigma}_k^2$ defined in (34), where $\bar{\varepsilon}_k$ is such that $F(\bar{\varepsilon}_k; \hat{\sigma}_k, \hat{\mathbf{m}}_k) = 0$, where $\bar{\varepsilon}_0$ is given as in the item 12 above, and k is such that $|\hat{\sigma}_k - \hat{\sigma}_{k-1}| \le \text{tol} = 10^{-5}$.

The adopted unit of measure is 1 year (250 days) and we consider 5 minute observations over a 1 month time horizon with a 6.5 hours per day open market. For our first simulation, we use the following parameters:

$$\sigma = 0.4, \quad \sigma^{Jmp} = 3\sqrt{h}, \quad \mu^{Jmp} = 0, \quad \lambda = 100, \quad h = \frac{1}{250(6.5)(12)}. \tag{38}$$

The dependence of σ^{Jmp} on σ^{Jmp} on $\sqrt{h}$ was done for easier comparison with standard deviation of the increments of the continuous component, which is $0.4\sqrt{h}$. So, the standard deviation of the jumps is about 7.5 times the standard deviation of the continuous component increment. The parameter values in (38) yield an expected annualized volatility of 0.45, which is reasonable. Table 1 below shows the sample means and standard deviations based on 10000 simulations (below Loss equals the number of jump misclassifications as defined by (35), while N is the number of iterations needed to find the estimator's value). As shown therein, the new proposed estimator (items 3, 13) performs the best, followed by the iterative method 7 based on (37). It takes on average 2 iterations to finish if we take as an initial guess for the threshold the solution of Eq. (31). However, if we take advantage of the asymptotic behavior of $\bar{\varepsilon}$ as in the method 12 above, one iteration suffices.

Method	$\overline{\hat{\sigma}}$	std($\hat{\sigma}$)	$\overline{\text{Loss}}$	std(Loss)	$\bar{\varepsilon}$	std(ε)	$\overline{N}$	std(N)
1	0.45311689	0.03104886						
2	0.40132	0.00732	3.66530	1.92646	0.01228	0.00084	1	0
3	0.40029	0.00727	3.48050	1.86772	0.01099	0.00048	2.39300	0.55909
4	0.4058	0.0085	4.9251	2.2672	0.0176	0.0012	1	0
5	0.40398	0.00789	4.49490	2.14551	0.01569	0.00031	2.31500	0.53796
6	0.40288	0.00765	4.16110	2.07624	0.01437	0.00098	1	0
7	0.40166	0.00741	3.75950	1.95909	0.01274	0.00024	2.31960	0.50723
8	0.3842	0.0062	16.2724	4.0291	0.0075	0	1	0
9	0.4033	0.0075	4.2897	2.0800	0.0150	0	1	0
10	0.413	0.011						
11	0.40181	0.00743	3.81300	1.96485	0.01301	0.00034	1	0
12	0.400429	0.007206	3.468400	1.873967	0.011118	0.000517	1	0
13	0.400282	0.007218	3.464100	1.876983	0.010968	0.000466	1.711700	0.587041

Table 1: Estimation of the volatility $\sigma = 0.4$ for a log-normal Merton model based on 10000 simulations of 5-minute observations over a 1 month time horizon. The jump parameters are $\lambda = 100$, $\sigma^{Jmp} = 3\sqrt{h}$ and $\mu^{Jmp} = 0$.

We now double the intensity of jumps and consider the following parameter setting:

$$\sigma = 0.4, \quad \sigma^{Jmp} = 3\sqrt{h}, \quad \mu^{Jmp} = 0, \quad \lambda = 200, \quad h = \frac{1}{250(6.5)(12)},$$

which yields an expected annualized volatility of 0.5. The results are shown in Table 2. We again notice that the methods 3 and 13 outperforms all the others, followed by method 7 based on the asymptotic behavior $\bar{\varepsilon}_h \sim \sqrt{2\sigma^2 h \ln(1/h)}$.

Finally, we consider a jump intensity of 1000 jumps per year but we reduce σ and σ^{Jmp} in order to obtain an expected annualized volatility of 0.39. Concretely, we set:

$$\sigma = 0.2, \quad \sigma^{Jmp} = 1.5\sqrt{h}, \quad \mu^{Jmp} = 0, \quad \lambda = 1000, \quad h = \frac{1}{250(6.5)(12)},$$

The results are shown in Table 3. In spite of being a tough setting, the new method does a good job and outperforms all others, except method 7, which is based on the asymptotics $\bar{\varepsilon}_h \sim \sqrt{2\sigma^2 h \ln(1/h)}$. Note that in this case it takes on average 5 iterations for the iterative methods to converge.

Est	$\overline{\hat{\sigma}}$	std($\hat{\sigma}$)	$\overline{\text{Loss}}$	std(Loss)	$\overline{\varepsilon}$	std(ε)	$\overline{N}$	std(N)
1	0.5002	0.0385						
2	0.40482	0.00792	7.80670	2.83657	0.01356	0.00104	1	0
3	0.402181	0.007588	6.917800	2.655174	0.011623	0.000908	2.718900	0.723834
4	0.4159	0.0111	10.5434	3.3448	0.0194	0.0015	1	0
5	0.408570	0.008844	8.980800	3.038413	0.015871	0.000344	2.852900	0.613921
6	0.40858	0.00884	8.94540	3.05446	0.01586	0.00122	1	0
7	0.403786	0.007761	7.449100	2.754445	0.012807	0.000246	2.813000	0.557012
8	0.38401	0.00628	18.46560	4.25603	0.00749	0	1	0
9	0.40682	0.00792	8.48660	2.89431	0.01499	0	1	0
10	0.4265	0.0128						
11	0.404555	0.007859	7.730600	2.794102	0.013432	0.000404	1	0
12	0.402950	0.007601	7.205800	2.721909	0.012204	0.000934	1	0
13	0.402160	0.007587	6.965600	2.661903	0.011617	0.000905	2.105500	0.709380

Table 2: Estimation of the volatility $\sigma = 0.4$ for a log-normal Merton model based on 10000 simulations of 5-minute observations over a 1 month time horizon. The jump parameters are $\lambda = 200$, $\sigma^{Jmp} = 3\sqrt{h}$ and $\mu^{Jmp} = 0$.

Est	$\overline{\hat{\sigma}}$	std($\hat{\sigma}$)	$\overline{\text{Loss}}$	std(Loss)	$\overline{\varepsilon}$	std(ε)	$\overline{N}$	std(N)
1	0.3921	0.0279						
2	0.246	0.0127	56.3	8.29	0.0106	0.000756	1	0
3	0.21563	0.00860	41.72400	7.80543	0.00728	0.00083	5.60410	1.57455
4	0.29618	0.02148	70.17440	9.20290	0.01523	0.00108	1	0
5	0.23	0.0108	49.8	8.39	0.00892	0.00042	5.86	1.33
6	0.265	0.0163	62.6	8.74	0.0124	0.00088	1	0
7	0.211	0.00588	39.1	6.79	0.00671	0.00018	5.10	0.910
8	0.21663	0.00518	42.74350	6.58275	0.00749	0	1	0
9	0.293	0.014	69.497	8.293	0.015	0	1	0
10	0.2664	0.0129						
11	0.224	0.00779	47	7.39	0.00839	0.000405	1	0
12	0.241	0.0121	54.7	8.16	0.0102	0.000801	1	0
13	0.216	0.00863	41.8	7.74	0.00728	0.000835	5.36	1.64

Table 3: Estimation of the volatility $\sigma = 0.2$ for a log-normal Merton model based on 10000 simulations of 5-minute observations over a 1 month time horizon. The jump parameters are $\lambda = 1000$, $\sigma^{Jmp} = 1.5\sqrt{h}$ and $\mu^{Jmp} = 0$.

6 Conclusions

We consider the problem of estimating the integrated variance IV of a semimartingale model X with jumps for the log price of a financial asset. In view of adopting the truncated realized variance of X, we look for a theoretical and practical way to select an optimal threshold in finite samples. We consider the following two optimality criteria: minimization of MSE, the expected quadratic error in the estimation of IV; and minimization of cMSE, the expected quadratic error conditional to the realized paths of the jump process J and of the volatility process $(\sigma_s)_{s\geq 0}$. Under given assumptions, we find that for each criterion an optimal TH exists, is unique and is a solution of an explicitly given equation, the equation being different under the two criteria. Also, under each criterion, an asymptotic expansion with respect to the step h between the observations is possible for the optimal TH . The leading terms of the two expansions turn out to be proportional to the modulus of continuity of the Brownian motion paths and to the spot volatility of X, with proportionality constant $\sqrt{2-Y}$, Y being the jump activity index of X. It turns out that the threshold estimator of IV constructed with the optimal TH is consistent, at least in the finite activity jumps case. The results obtained for the cMSE criterion allow for a novel numerical way to tuneup the threshold parameter in finite samples. We illustrate the superiority of the new method on simulated data. Minimization of the conditional mean square estimation error in the presence of infinite activity jumps in X is object of further research.

7 Appendix: additional proofs

Proof of Lemma 1. Throughout, p_t denotes the density of J_t and recall that the characteristic function of J_t is of the form $\mathbb{E}\left[e^{iuJ_t}\right] = e^{-ct|u|^Y}$. Let us also recall that the Fourier transform and its inverse are defined by $\mathcal{F}g(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(z)e^{-izx}dz$ and $\mathcal{F}^{-1}G(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} G(z)e^{izx}dz$. In what follows, we set

$$h(u) := \left(\mathcal{F}^{-1}\phi\left(\frac{\cdot}{\sigma\sqrt{h}} - \frac{\varepsilon}{\sigma\sqrt{h}}\right)\right)(u) = \frac{1}{\sqrt{2\pi}}\int \phi\left(\frac{x}{\sigma\sqrt{h}} - \frac{\varepsilon}{\sigma\sqrt{h}}\right)e^{iux}dx.$$

Let us start by noting that

$$\mathbb{E}\left[\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\right] = \int \phi\left(\frac{x}{\sigma\sqrt{h}} - \frac{\varepsilon}{\sigma\sqrt{h}}\right)p_h(x)dx = \int (\mathcal{F}h)(x)\,p_h(x)dx = \int h(u)\,(\mathcal{F}p_h)(u)du,$$

where, since J is a symmetric stable process, $(\mathcal{F}p_h)(u) = (2\pi)^{-1/2}e^{-ch|u|^Y}$. Therefore, we obtain the representation

$$\mathbb{E}\left[\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} \pm \frac{J_h}{\sigma\sqrt{h}}\right)\right] = \frac{\sigma h^{1/2}}{2\pi}\int e^{-ch|u|^Y - \frac{\sigma^2 hu^2}{2} + i\varepsilon u}du. \tag{39}$$

In order to prove (10), let us make the change of variables $w = \sigma h^{1/2}u$ and, then, expand in a Taylor's expansion $\exp(-c\sigma^{-Y}h^{1-Y/2}|w|^Y)$ as follows:

$$\frac{1}{2\pi}\int e^{-c\sigma^{-Y}h^{1-Y/2}|w|^Y - \frac{w^2}{2} + i\frac{\varepsilon}{\sigma h^{1/2}}w}dw = \frac{1}{2\pi}\int e^{-\frac{w^2}{2} + i\frac{\varepsilon}{\sigma h^{1/2}}w}dw + \sum_{k=1}^{\infty} I_{k,n},$$

where

$$\begin{aligned} I_{k,n} &:= \frac{1}{k!}(-c)^k\sigma^{-kY}h^{k(1-Y/2)}\frac{1}{\sqrt{2\pi}}\int |w|^{kY}e^{-\frac{w^2}{2} + i\frac{\varepsilon}{\sigma h^{1/2}}w}dw \\ &= \frac{1}{k!}(-c)^k\sigma^{-kY}h^{k(1-Y/2)}\frac{2}{\sqrt{2\pi}}\int_0^{\infty} w^{kY}e^{-\frac{w^2}{2}}\cos\left(\frac{\varepsilon}{\sigma h^{1/2}}w\right)dw. \end{aligned}$$

The first term of (10) is then clear. For the subsequent terms, let us apply the formula for the cosine integral transformation of $w^{kY}e^{-w^2/2}$ as well as the asymptotics for the generalized hypergeometric series or Kummer's function $M(a,b,z)$:

$$\begin{aligned} I_{k,n} &= \frac{1}{k!}(-c)^k\sigma^{-kY}h^{k(1-Y/2)}\frac{2}{\sqrt{2\pi}}\left\{\frac{1}{2}2^{\frac{1}{2}(1+kY)}\Gamma\left(\frac{1}{2} + \frac{kY}{2}\right)M\left(\frac{1}{2} + \frac{kY}{2};\frac{1}{2};-\frac{\varepsilon^2}{2\sigma^2 h}\right)\right\} \\ &= \frac{1}{k!}(-c)^k\sigma^{-kY}h^{k(1-Y/2)}\frac{2}{\sqrt{2\pi}}\left(\frac{1}{2}2^{\frac{1}{2}(1+kY)}\Gamma\left(\frac{1}{2} + \frac{kY}{2}\right)\right) \\ &\quad\times\left(\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(-\frac{kY}{2}\right)}\left(\frac{\varepsilon^2}{2\sigma^2 h}\right)^{-\frac{1}{2}-\frac{kY}{2}} + \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}+\frac{kY}{2}\right)}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\left(\frac{\varepsilon^2}{2\sigma^2 h}\right)^{\frac{kY}{2}}\right) + \text{h.o.t.}. \end{aligned}$$

In the asymptotic formula for the Kummer's function above, the first term (respectively, second term) vanishes if $\Gamma(-kY/2)$ (respectively, $\Gamma(1/2+kY/2)$) are infinity. This happens when $-kY/2$ or $1/2$ +kY/2 are nonpositive integers. It is now evident that there exists nonzero constants a_k and b_k such that

$$I_{k,n}=\frac{a_k}{\Gamma\left(-\frac{kY}{2}\right)}\varepsilon^{-1-kY}h^{k+\frac{1}{2}}+\frac{b_k}{\Gamma\left(\frac{1}{2}+\frac{kY}{2}\right)}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\varepsilon^{kY}h^{k(1-Y)}+\text{h.o.t.}.$$

Note that

$$\varepsilon^{-1-kY}h^{k+\frac{1}{2}}\gg\varepsilon^{-1-(k+1)Y}h^{k+1+\frac{1}{2}}\iff\varepsilon\gg h^{1/Y}\Longleftarrow\varepsilon\gg h^{1/2},$$
$$\varepsilon^{-1-Y}h^{1+\frac{1}{2}}\gg\varepsilon^{-1-kY}h^{k+\frac{1}{2}}\gg e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\varepsilon^{kY}h^{k(1-Y)}.$$

Therefore, $\varepsilon^{-1-Y}h^{1+\frac{1}{2}}\gg I_{k,n}$, for all $k>1$.

We now show (11). Note that

$$\mathbb{E}\left[J_h\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}}-\frac{J_h}{\sigma\sqrt{h}}\right)\right]=\int\phi\left(\frac{x}{\sigma\sqrt{h}}-\frac{\varepsilon}{\sigma\sqrt{h}}\right)xp_h(x)dx=\int h\left(u\right)\mathcal{F}(xp_h(x))(u)du,$$

where

$$\mathcal{F}(xp_h(x))(u)=i\frac{d}{du}(\mathcal{F}p_h)(u)=\frac{i}{\sqrt{2\pi}}\frac{d}{du}e^{-ch|u|^Y}=\frac{-i}{\sqrt{2\pi}}e^{-ch|u|^Y}Y\text{sign}(u)ch|u|^{Y-1}.$$

Therefore, we have the following representation:

$$\mathbb{E}\left[J_h\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}}-\frac{J_h}{\sigma\sqrt{h}}\right)\right]=\sigma\frac{-iYc}{\sqrt{2\pi}}h^{3/2}\int\text{sign}(u)|u|^{Y-1}e^{-ch|u|^Y-\frac{\sigma^2hu^2}{2}+i\varepsilon u}du.$$

Furthermore,

$$\begin{aligned}\mathbb{E}\left[J_h\phi\left(\frac{\varepsilon}{\sqrt{h}}-\frac{J_h}{\sqrt{h}}\right)\right]&=2\sigma\frac{Yc}{\sqrt{2\pi}}h^{3/2}\int_0^\infty u^{Y-1}e^{-chu^Y-\frac{\sigma^2hu^2}{2}}\sin\left(\varepsilon u\right)du\\&=2\sigma^{-(Y-1)}\frac{Yc}{\sqrt{2\pi}}h^{\frac{3-Y}{2}}\int_0^\infty w^{Y-1}e^{-c\sigma^{-Y}h^{1-Y/2}w^Y-\frac{w^2}{2}}\sin\left(\sigma^{-1}\varepsilon h^{-1/2}w\right)dw.\end{aligned}$$

Next, we expand in a Taylor's expansion $\exp(-c\sigma^{-Y}h^{1-Y/2}w^Y)$ as follows:

$$\frac{1}{\sqrt{2\pi}}\int_0^\infty w^{Y-1}e^{-c\sigma^{-Y}h^{1-Y/2}w^Y-\frac{w^2}{2}}\sin\left(\sigma^{-1}\varepsilon h^{-1/2}w\right)dw=\sum_{k=0}^\infty I_{k,n},$$

where

$$I_{k,n}:=\frac{1}{k!}(-c)^k\sigma^{-Yk}h^{k(1-Y/2)}\frac{1}{\sqrt{2\pi}}\int_0^\infty w^{(k+1)Y-1}e^{-\frac{w^2}{2}}\sin\left(\varepsilon h^{-1/2}w\right)dw.$$

Then, we again apply the following formula for the sine integral transformation of $w^{(k+1)Y-1}e^{-w^2/2}$:

$$I_{k,n}=\frac{1}{k!}(-c)^k\sigma^{-Yk}h^{k(1-Y/2)}\frac{1}{\sqrt{2\pi}}\left\{\frac{1}{2}2^{\frac{1}{2}(1+(k+1)Y)}\Gamma\left(\frac{1}{2}+\frac{(k+1)Y}{2}\right)\left(\frac{\varepsilon^2}{h}\right)M\left(\frac{1}{2}+\frac{(k+1)Y}{2};\frac{3}{2};-\frac{\varepsilon^2}{2h}\right)\right\}.$$

Finally, we use the relationship

$$\begin{aligned}M\left(\frac{1}{2}+\frac{(k+1)Y}{2};\frac{3}{2};-\frac{\varepsilon^2}{2h}\right)=&\frac{\Gamma(\frac{3}{2})}{\Gamma\left(1-\frac{(k+1)Y}{2}\right)}\left(\frac{\varepsilon^2}{2h}\right)^{-\frac{1}{2}-\frac{(k+1)Y}{2}}\\&+\frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}+\frac{(k+1)Y}{2}\right)}e^{-\frac{\varepsilon^2}{2\sigma^2h}}\left(-\frac{\varepsilon^2}{2\sigma^2h}\right)^{-1+\frac{(k+1)Y}{2}}+\text{h.o.t.},\end{aligned}$$

which, in turn shows that,

$$I_{k,n}\ll I_{1,n}\ll h\varepsilon^{1-Y}.$$

We then conclude the result of the Lemma. □

Proof of Lemma 2. Let

$$I_n^{\pm} := \mathbb{E}\left[\bar{\Phi}\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\mathbf{1}_{\left\{\pm\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\geq 0\right\}}\right]$$

For I_n^+, let us note that for a constant K, $\bar{\Phi}(z) \leq K\phi(z)$ for all $z \geq 0$ and, thus,

$$I_n^+ \leq K\mathbb{E}\left[\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\mathbf{1}_{\left\{\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\geq 0\right\}}\right] = O\left(\mathbb{E}\left[\phi\left(\frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right)\right]\right).$$

For the other term, we decompose it as follows:

$$\begin{aligned}
I_n^- &= \int_{\mathbb{R}} \phi(u)\mathbb{P}\left[0 \geq \frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}, u \geq \frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right] du \\
&= \int_0^{\infty} \phi(u)\mathbb{P}\left[0 \geq \frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right] du + \int_{-\infty}^0 \phi(u)\mathbb{P}\left[u \geq \frac{\varepsilon}{\sigma\sqrt{h}} - \frac{J_h}{\sigma\sqrt{h}}\right] du \\
&= \frac{1}{2}\mathbb{P}\left[J_1 \geq h^{-\frac{1}{Y}}\varepsilon\right] + \int_{-\infty}^0 \phi(u)\mathbb{P}\left[J_1 \geq h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right] du.
\end{aligned}$$

The first term above is well-known to be $\mathbb{P}\left[J_1 \geq h^{-1/Y}\varepsilon\right] = Y^{-1}C\left(h^{-1/Y}\varepsilon\right)^{-Y} + O\left(\varepsilon^{-2Y}h^2\right)$. For the second term, let us first recall that there exists a constant K such that for all $x > 0$,

$$|\mathcal{E}(x)| := \left|\mathbb{P}\left[J_1 \geq x\right] - \frac{C}{Y}x^{-Y}\right| \leq Kx^{-2Y}. \tag{40}$$

Therefore,

$$\begin{aligned}
\int_{-\infty}^0 \phi(u)\mathbb{P}\left[J_1 \geq h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right] du &= \frac{C}{Y}\int_{-\infty}^0 \phi(u)\left(h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right)^{-Y} du \\
&\quad + \int_{-\infty}^0 \phi(u)\mathcal{E}\left(h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right) du.
\end{aligned}$$

For the first term above, note that

$$\frac{1}{h\varepsilon^{-Y}}\int_{-\infty}^0 \phi(u)\left(h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right)^{-Y} du = \int_{-\infty}^0 \phi(u)\left(1 - \sigma u\varepsilon^{-1}h^{1/2}\right)^{-Y} du,$$

which, by the dominated convergence theorem, converges to $1/2$, because $\varepsilon^{-1}h^{1/2} \to 0$, as $n \to \infty$. Similarly, using (40), we have

$$\left|\int_{-\infty}^0 \phi(u)\mathcal{E}\left(h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right) du\right| \leq K\int_{-\infty}^0 \phi(u)\left(h^{-\frac{1}{Y}}\varepsilon - \sigma u h^{\frac{1}{2}-\frac{1}{Y}}\right)^{-2Y} du = O\left(\varepsilon^{-2Y}h^2\right).$$

Therefore, we finally conclude that $I_n^- = Y^{-1}Ch\varepsilon^{-Y} + O\left(\varepsilon^{-2Y}h^2\right)$, which implies (12).

We now show (13). To this end, let us first consider

$$\begin{aligned}
E_{1,h}(\varepsilon) &:= \mathbb{E}\left[J_h^2\mathbf{1}_{\{0\leq\sigma W_h+J_h\leq\varepsilon, J_h\geq 0, W_h\geq 0\}}\right] \\
&= h^{2/Y}\int_0^{\varepsilon\sigma^{-1}h^{-\frac{1}{2}}} \phi(x)\int_0^{h^{-\frac{1}{Y}}\varepsilon - \sigma h^{\frac{1}{2}-\frac{1}{Y}}x} u^2 p_1(u)\,du\,dx \\
&= h^{\frac{2}{Y}}\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}\right)\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\int_0^{h^{-\frac{1}{Y}}\varepsilon(1-w)} u^2 p_1(u)\,du\,dx.
\end{aligned}$$

Let $\mathcal{E}(u) := p_1(u) - Cu^{-Y-1}$ and let us recall that, for a constant K, $|\mathcal{E}(u)| \leq K\left(u^{-Y-1} \wedge u^{-2Y-1}\right) \leq Ku^{-2Y-1}$, for all $u > 0$. Next,

$$\begin{aligned}
E_{1,h}(\varepsilon) &= Ch^{\frac{2}{Y}}\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}\right)\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\int_0^{h^{-\frac{1}{Y}}\varepsilon(1-w)} u^{1-Y}\,du\,dx \\
&\quad + h^{\frac{2}{Y}}\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}\right)\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\int_0^{h^{-\frac{1}{Y}}\varepsilon(1-w)} u^2\mathcal{E}(u)\,du\,dw
\end{aligned}$$

For the first term above, note that

$$\frac{1}{2-Y}\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\left(h^{-\frac{1}{Y}}\varepsilon(1-w)\right)^{2-Y}dw = \frac{h^{-\frac{2-Y}{Y}}\varepsilon^{2-Y}}{2-Y}\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)(1-w)^{2-Y}dw$$
$$\sim 2^{-1}\frac{h^{-\frac{2-Y}{Y}}\varepsilon^{2-Y}}{2-Y}\left(\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}\right).$$

We divide the second term in two cases. If $Y \leq 1$, then

$$\left|\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\int_0^{h^{-\frac{1}{Y}}\varepsilon(1-w)} u^2\mathcal{E}(u)dudx\right| \leq K\frac{1}{2-2Y}\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\left(h^{-\frac{1}{Y}}\varepsilon(1-w)\right)^{2-2Y}dw$$
$$\leq K\frac{h^{-\frac{2-2Y}{Y}}\varepsilon^{2-2Y}}{2-2Y}\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)(1-w)^{2-2Y}dw$$
$$\sim 2K\frac{h^{-\frac{2-2Y}{Y}}\varepsilon^{2-2Y}}{2-2Y}\left(\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}\right).$$

Note that the last limit is valid provided that $\int_0^1 (1-w)^{2-2Y}dw < \infty$, which holds true when $Y \leq 1$. For $Y > 1$, let us first observe that

$$\int_0^z u^2\left(u^{-Y-1}\wedge u^{-2Y-1}\right)du \leq \frac{1}{2-Y} + \mathbf{1}_{\{z>1\}}\frac{1-z^{2(1-Y)}}{2(Y-1)} \leq \frac{1}{2-Y} + \frac{1}{2(Y-1)}. \tag{41}$$

Therefore, for a constant K,

$$\left|\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)\int_0^{h^{-\frac{1}{Y}}\varepsilon(1-w)} u^2\mathcal{E}(u)dudx\right| \leq K\int_0^1 \phi\left(\frac{\varepsilon}{\sigma h^{\frac{1}{2}}}w\right)dw \sim K\left(\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}\right).$$

We conclude that

$$E_{1,h}(\varepsilon) = \frac{2^{-1}C}{2-Y}h\varepsilon^{2-Y} + O\left(h^2\varepsilon^{2-2Y}\right) + O\left(h^{\frac{2}{Y}}\right).$$

Next, we consider

$$\begin{aligned}
E_{2,h}(\varepsilon) &:= \mathbb{E}\left[J_h^2\mathbf{1}_{\{0\leq\sigma W_h+J_h\leq\varepsilon, J_h\geq 0, W_h\leq 0\}}\right] \\
&= h^{2/Y}\int_{-\infty}^0 \phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x} u^2 p_1(u)dudx \\
&= Ch^{2/Y}\int_{-\infty}^0 \phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x} u^{1-Y}dudx \\
&\quad + h^{2/Y}\int_{-\infty}^0 \phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x} u^2\mathcal{E}(u)dudx.
\end{aligned}$$

The first term on the right-hand side above can be written as

$$\frac{C}{2-Y}h^{2/Y}\left(h^{-\frac{1}{Y}}\varepsilon\right)^{2-Y}\int_{-\infty}^0 \phi(x)\left\{\left(1-\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}x\right)^{2-Y} - \left(-\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}x\right)^{2-Y}\right\}dx \sim 2^{-1}\frac{C}{2-Y}h\varepsilon^{2-Y},$$

where the last asymptotic relationship follows from dominated convergence theorem and the facts that $h^{1/2}/\varepsilon \to 0$ and $\int_{-\infty}^0 (1-x)^{2-Y}\phi(x)dx < \infty$. For the second term of $E_{2,h}(\varepsilon)$, we have two cases. For $Y \leq 1$, we have

$$h^{\frac{2}{Y}}\left|\int_{-\infty}^0 \phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x} u^2\mathcal{E}(u)dudx\right| \leq Kh^{\frac{2}{Y}}\int_{-\infty}^0 \phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x} u^{1-2Y}dudx$$
$$= \frac{K}{2(1-Y)}h^{\frac{2}{Y}}\left(h^{-\frac{1}{Y}}\varepsilon\right)^{2-2Y}\int_{-\infty}^0 \phi(x)\left\{\left(1-\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}x\right)^{2(1-Y)} - \left(-\frac{\sigma h^{\frac{1}{2}}}{\varepsilon}x\right)^{2(1-Y)}\right\}dx \sim Kh^2\varepsilon^{2-2Y},$$

where again we used dominated convergence and use the fact that $\int_{-\infty}^{0}\phi(x)(1-x)^{2(1-Y)}dx<\infty$. For $Y>1$, we just use (41) to deduce that

$$h^{\frac{2}{Y}}\int_{-\infty}^{0}\phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}u^2|\mathcal{E}(u)|dudx\le K'h^{\frac{2}{Y}}\int_{-\infty}^{0}\phi(x)dx,$$

for a constant K'. Finally, we conclude that

$$E_{2,h}=2^{-1}\frac{C}{2-Y}h\varepsilon^{2-Y}+O\left(h^2\varepsilon^{2-2Y}\right)+O\left(h^{\frac{2}{Y}}\right).$$

Finally, let us consider

$$\begin{aligned}E_{3,h}(\varepsilon)&:=\mathbb{E}\left[J_h^2\mathbf{1}_{\{0\le\sigma W_h+J_h\le\varepsilon,J_h\le0,W_h\ge0\}}\right]\\&=h^{2/Y}\int_0^{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}\phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{0}u^2p_1(u)dudx\\&\quad+h^{2/Y}\int_{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}^{\infty}\phi(x)\int_{-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}^{h^{-\frac{1}{Y}}\varepsilon-\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}u^2p_1(u)dudx.\end{aligned}$$

Using the fact that $p_1(u)\le Ku^{-Y-1}$ for a constant K and all $u>0$, the first term above is such that

$$\begin{aligned}h^{2/Y}\int_0^{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}\phi(x)\int_0^{\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}u^2p_1(u)dudx&\le Kh^{2/Y}\int_0^{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}\phi(x)\int_0^{\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}u^{1-Y}dudx\\&=\frac{K}{2-Y}\left(\sigma h^{\frac{1}{2}-\frac{1}{Y}}\right)^{2-Y}\int_0^{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}\phi(x)x^{2-Y}dx\\&\sim\frac{K}{2-Y}h^{\frac{4-Y}{2}}\int_0^{\infty}\phi(x)x^{2-Y}dx=o\left(h\varepsilon^{2-Y}\right).\end{aligned}$$

Similarly, the second term can be written as

$$\begin{aligned}h^{2/Y}\int_{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}^{\infty}\phi(x)\int_{\sigma h^{\frac{1}{2}-\frac{1}{Y}}x-h^{-\frac{1}{Y}}\varepsilon}^{\sigma h^{\frac{1}{2}-\frac{1}{Y}}x}u^2p_1(u)dudx&\le\frac{K}{2-Y}\left(\sigma h^{\frac{1}{2}-\frac{1}{Y}}\right)^{2-Y}\int_{\sigma^{-1}h^{-\frac{1}{2}}\varepsilon}^{\infty}\phi(x)x^{2-Y}dx\\&=o\left(h^{\frac{4-Y}{2}}\right)=o\left(h\varepsilon^{2-Y}\right).\end{aligned}$$

Putting together the previous results, we obtain that

$$\begin{aligned}E_h(\varepsilon)&=2\mathbb{E}\left[J_h^2\mathbf{1}_{\{0\le\sigma W_h+J_h\le\varepsilon\}}\right]=2E_{1,h}(\varepsilon)+2E_{2,h}(\varepsilon)+2E_{3,h}(\varepsilon)\\&=\frac{2C}{2-Y}h\varepsilon^{2-Y}+O\left(h^2\varepsilon^{2-2Y}\right)+O\left(h^{\frac{4-Y}{2}}\right)+O\left(h^{\frac{2}{Y}}\right).\end{aligned}$$

□

Proof of (23). Let

$$\bar{N}(x)=\int_x^{\infty}\phi(z)dz,\quad R(x)=\int_x^{\infty}\phi(z)dz-\frac{\phi(x)}{x}$$

and recall that, for $x>0$,

$$\bar{N}(x)\le\frac{1}{x}\phi(x),\quad |R(x)|\le\frac{\phi(x)}{x^3}.$$

Then, for fixed $m > 0$ and h small enough such that $\varepsilon_h < m$, we have

$$
\begin{aligned}
b(\varepsilon, m, h) &= \sigma\sqrt{h}\phi\left(\frac{m-\varepsilon}{\sigma\sqrt{h}}\right)\left(\frac{m^2}{m-\varepsilon} - m\right) - \sigma\sqrt{h}\phi\left(\frac{m+\varepsilon}{\sigma\sqrt{h}}\right)\left(\frac{m^2}{m+\varepsilon} - m\right) \\
&\quad - \sigma\sqrt{h}\phi\left(\frac{m-\varepsilon}{\sigma\sqrt{h}}\right)\varepsilon - \sigma\sqrt{h}\phi\left(\frac{m+\varepsilon}{\sigma\sqrt{h}}\right)\varepsilon \\
&\quad + \sigma^3 h^{3/2}\phi\left(\frac{m-\varepsilon}{\sigma\sqrt{h}}\right)\left(\frac{1}{m-\varepsilon}\right) - \sigma^3 h^{3/2}\phi\left(\frac{m+\varepsilon}{\sigma\sqrt{h}}\right)\left(\frac{1}{m+\varepsilon}\right) \pm (m^2+\sigma^2 h)R\left(\frac{m \mp \varepsilon_h}{\sigma\sqrt{h}}\right) \\
&= \frac{\sigma}{m}\sqrt{h}\phi\left(\frac{m-\varepsilon}{\sigma\sqrt{h}}\right)\varepsilon^2 - \frac{\sigma}{m}\sqrt{h}\phi\left(\frac{m+\varepsilon}{\sigma\sqrt{h}}\right)\varepsilon^2 \\
&\quad + \frac{\sigma}{m(m-\varepsilon)}\sqrt{h}\phi\left(\frac{m-\varepsilon}{\sigma\sqrt{h}}\right)\varepsilon^3 - \frac{\sigma}{m(m+\varepsilon)}\sqrt{h}\phi\left(\frac{m+\varepsilon}{\sigma\sqrt{h}}\right)\varepsilon^3 \\
&\quad + \frac{\sigma^3}{m-\varepsilon}h^{3/2}\phi\left(\frac{m-\varepsilon}{\sigma\sqrt{h}}\right) - \frac{\sigma^3}{m+\varepsilon}h^{3/2}\phi\left(\frac{m+\varepsilon}{\sigma\sqrt{h}}\right) \pm (m^2+\sigma^2 h)R\left(\frac{m \mp \varepsilon_h}{\sigma\sqrt{h}}\right)
\end{aligned}
$$

It is now clear that (23) holds true. We can similarly deal with the case $m < 0$. The asymptotic behavior for $a(\varepsilon, m, h)$ is direct. □

Proof of Proposition 6. Let us fix h, and $nh = 1$, then $\frac{d}{d\varepsilon}F(\varepsilon) = \sum_{i=1}^n [a_i' g_i + a_i g_i']$

$$
= -\sum_{i=1}^n \frac{1}{\sigma^3 h^{\frac{3}{2}}\sqrt{2\pi}}\left[e^{-\frac{(\varepsilon-|m_i|)^2}{2\sigma^2 h}}(\varepsilon - |m_i|) + e^{-\frac{(\varepsilon+|m_i|)^2}{2\sigma_i^2}}(\varepsilon + |m_i|)\right]g_i + \sum_{i=1}^n \frac{e^{-\frac{(\varepsilon-|m_i|)^2}{2\sigma^2 h}} + e^{-\frac{(\varepsilon+|m_i|)^2}{2\sigma^2 h}}}{\sigma\sqrt{h}\sqrt{2\pi}}\left[2\varepsilon + 2\sum_{j\neq i}\varepsilon^2 a_j\right].
$$

We now evaluate $F'(\varepsilon)$ at ε_h such that $\varepsilon_h \to 0$ with $\varepsilon_h \gg \sqrt{h}$, as $h \to 0$. Since again when $m_i \neq 0$ we have $e^{-\frac{(\varepsilon-|m_i|)^2}{2\sigma^2 h}} \gg e^{-\frac{(\varepsilon+|m_i|)^2}{2\sigma^2 h}}$ and $\varepsilon \ll m_i$, then

$$
F'(\varepsilon)\sqrt{2\pi} = \sum_{i\in\{J\}} \frac{1}{\sigma\sqrt{h}}e^{-\frac{(\varepsilon-|m_i|)^2}{2\sigma^2 h}}\left[\frac{|m_i|}{\sigma^2 h}g_i + 2\varepsilon(1 + \varepsilon\sum_{j\neq i}a_j)\right] + \sum_{i\notin\{J\}} \frac{2\varepsilon}{\sigma\sqrt{h}}e^{-\frac{\varepsilon^2}{2\sigma^2 h}}\left[-\frac{g_i}{\sigma^2 h} + 2(1 + \varepsilon\sum_{j\neq i}a_j)\right] + h.o.t.
$$

Note that within g_i in (25) we have that the finite sum $\frac{1}{\sqrt{2\pi}}\sum_{j\neq i: j\in\{J\}} \frac{\sigma}{|m_j|}\varepsilon e^{-\frac{(|m_j|-\varepsilon)^2}{2\sigma^2 h}} = \sum_{j\neq i: j\in\{J\}} \frac{\sigma}{|m_j|}\varepsilon u_{jh}$ $= s_h\varepsilon \sum_{j\neq i: j\in\{J\}} \frac{\sigma}{|m_j|}p_{jh}$ is negligible wrt $s_h \ll \frac{1}{\sqrt{2\pi}}\sum_{j\neq i: j\notin\{J\}} e^{-\frac{\varepsilon^2}{2\sigma^2 h}} = [(n-N_T)I_{\{i\in\{J\}\}} + (n-N_T-1)I_{\{i\notin\{J\}\}}]s_h$, since $\varepsilon \sum_{j\neq i: j\in\{J\}} \frac{p_{jh}}{m_j} \overset{a.s.}{\to} 0$. Therefore

$$
g_i = \varepsilon^2 - \frac{4\sigma}{\sqrt{2\pi}}\sqrt{h}\varepsilon s_h[(n-N_T)I_{\{i\in\{J\}\}} + (n-N_T-1)I_{\{i\notin\{J\}\}}] - 2\sigma^2 h[N_T I_{\{i\in\{J\}\}} + (N_T+1)I_{\{i\notin\{J\}\}}] + h.o.t..
$$

Further, $N_T \ll n$ and $h \ll \varepsilon^2$, then for all i

$$
g_i = \varepsilon^2 - \frac{4\sigma}{\sqrt{2\pi}}\frac{\varepsilon s_h}{\sqrt{h}} + h.o.t..
$$

Moreover from (27) we reach that $\sum_{j\neq i} a_j = \sum_{j\neq i, j\notin\{J\}} 2\frac{s_h}{\sigma\sqrt{h}} + \sum_{j\neq i, j\in\{J\}} \frac{u_{jh}}{\sigma\sqrt{h}} + h.o.t.$, and again the second sum is negligible wrt the first one, thus, for all i,

$$
\varepsilon\sum_{j\neq i} a_j = 2\frac{s_h\varepsilon}{\sigma\sqrt{h}}[(n-N_T)I_{\{m_i\neq 0\}} + (n-N_T-1)I_{\{m_i=0\}}] + h.o.t. = \frac{2}{\sigma}\frac{s_h\varepsilon}{h\sqrt{h}} + h.o.t..
$$

Now, using (28), from

$\sum_{i\in\{J\}} a_i g_i = \frac{1}{\sigma}\sqrt{h}s_h \sum_{\ell=1}^{N_T} p_{\ell h}\left[v_h^2 - 2\sigma^2 N_T + 2\sigma v_h s_h\left(\varepsilon\sum_{k\neq\ell}\frac{1}{|\gamma_k|}p_{kh} - 2(n-N_T)\right)\right] + h.o.t. =$

$\frac{1}{\sigma}\sqrt{h}s_h \sum_{\ell=1}^{N_T} p_{\ell h}\left[v_h^2 - 4\sigma v_h s_h n\right] + h.o.t.$

we reach that
$\sum_{i\in\{J\}} a_i \frac{g_i|m_i|}{\sigma^2 h} = \frac{1}{\sigma}\sqrt{h}s_h \sum_{\ell=1}^{N_T} p_{\ell h}\left[v_h^2 - 4\sigma v_h s_h n\right]\frac{|\gamma_\ell|}{\sigma^2 h} + h.o.t.$
and from
$\sum_{i\notin\{J\}} a_i g_i = (n-N_T)\frac{2}{\sigma}\sqrt{h}s_h\left[v_h^2 - 2\sigma^2(N_T+1) + 2\sigma v_h s_h\left(\varepsilon\sum_{k=1}^{N_T}\frac{1}{|\gamma_k|}p_{kh} - 2(n-N_T-1)\right)\right] =$
$(n-N_T)\frac{2}{\sigma}\sqrt{h}s_h\left[v_h^2 - 4\sigma v_h s_h n\right] + h.o.t.$
we reach that
$\sum_{i\notin\{J\}} a_i \frac{g_i\varepsilon}{\sigma^2 h} = (n-N_T)\frac{2}{\sigma}\sqrt{h}s_h\left[v_h^2 - 4\sigma v_h s_h n\right]\frac{\varepsilon}{\sigma^2 h} + h.o.t..$

Thus

$$F'\sqrt{2\pi} = v_h\Big[v_h - 4\sigma s_h n\Big]\left[\frac{1}{\sigma}\sqrt{h}\sum_{\ell=1}^{N_T} u_{\ell h}\frac{|\gamma_\ell|}{\sigma^2 h} - (n-N_T)\frac{2}{\sigma}\sqrt{h}s_h\frac{\varepsilon}{\sigma^2 h}\right]$$
$$+2\varepsilon\Big(1+\frac{2}{\sigma}\frac{s_h\varepsilon}{h\sqrt{h}}\Big)\Big(\sum_{i\in J}\frac{u_{ih}}{\sigma\sqrt{h}} + \sum_{i\notin J}\frac{2s_h}{\sigma\sqrt{h}}\Big) + h.o.t..$$

If now our sequence ε_h is such that $v_h = 4\sigma n s_h + h.o.t.$, and noting that also $\sum_{j\in J} p_{jh}|\gamma_j| \overset{a.s.}{\to} 0$ and that $n\varepsilon = n\sqrt{h}v_h = \frac{v_h}{\sqrt{h}} \to +\infty$ then

$$F'(\varepsilon_h)\sqrt{2\pi} = v_h\cdot o(ns_h)\frac{s_h}{\sigma^3\sqrt{h}}\left[\sum_{\ell=1}^{N_T} p_{\ell h}|\gamma_\ell| - 2n\varepsilon\right] + \frac{2\varepsilon s_h}{\sigma\sqrt{h}}\Big(1+\frac{2}{\sigma}\frac{s_h\varepsilon}{h\sqrt{h}}\Big)\cdot 2(n-N_T) + h.o.t.$$
$$= -2n\varepsilon v_h\cdot o(ns_h)\frac{s_h}{\sigma^3\sqrt{h}} + \frac{4n\varepsilon s_h}{\sigma\sqrt{h}}\Big(1+\frac{2}{\sigma}\frac{s_h\varepsilon}{h\sqrt{h}}\Big) + h.o.t.$$

now $v_h = 4\sigma n s_h + o(v_h)$ means also $s_h = \varepsilon_h\sqrt{h} + o(\varepsilon_h\sqrt{h})$, and thus $\frac{s_h\varepsilon_h}{h\sqrt{h}} = \frac{\varepsilon_h^2}{h} + o(\frac{\varepsilon_h^2}{h}) \to +\infty$, therefore

$$F'(\varepsilon_h)\sqrt{2\pi} = -2\varepsilon\frac{\varepsilon}{\sqrt{h}}\cdot o(ns_h)\frac{s_h}{\sigma^3 h\sqrt{h}} + \frac{4n\varepsilon s_h}{\sigma\sqrt{h}}\frac{2}{\sigma}\frac{s_h\varepsilon}{h\sqrt{h}} + h.o.t. =$$
$$\frac{\varepsilon}{\sqrt{h}}\frac{s_h\varepsilon}{h\sqrt{h}}ns_h\frac{8}{\sigma^2}(1+o(1)) + h.o.t. = \frac{8}{\sigma^2}\Big(\frac{s_h\varepsilon_h}{h\sqrt{h}}\Big)^2 + h.o.t..$$

□

References

[1] Bandi, F. & Russel, J.R. (2007). Realized covariation, realized beta, and microstructure noise. Working paper of the Graduate School of Business, The University of Chicago, January 2005

[2] Figueroa-López J. & Nisen, J. (2013): Optimally thresholded realized power variations for Lévy jump diffusion models. *Stochastic Processes and their Applications* 123(7), 2648-2677.

[3] Figueroa-López J. & Nisen, J. (2016): Second-order properties of thresholded realized power variations of FJA additive processes. *Preprint*, available at https://pages.wustl.edu/figueroa/publications.

[4] Jacod, J. & Todorov, V. (2014). Efficient Estimation of Integrated Volatility in Presence of Infinite Variation Jumps. *The Annals of Statistics* 42(3):1029-1069.

[5] Karatzas, I. & Shreve, S.E. (1999). *Brownian motion and stochastic calculus*, Springer.

[6] Mancini, C.: Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics, 36, 270-296 (2009)

[7] Revuz, D. & Yor, M. (2001). *Continuous martingales and Brownian motion*, Springer.