Short-Time Asymptotic Methods in Financial Mathematics

José E. Figueroa-López
Department of Mathematics
Washington University in St. Louis

Probability and Mathematical Finance Seminar
Department of Mathematical Sciences
Carnegie Mellon University
February 20, 2017
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
   - A Calibration Case-Study
   - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
   - A Calibration Case-Study
   - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
     - A Calibration Case-Study
     - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
Short-time ATM Skew Asymptotics

Theorem (F-L & Ólafsson, F&S 2016)

For $Y \in (1, 2)$, the implied volatility $\hat{\sigma}(\kappa, t)$ of a vanilla option with log-moneyness $\kappa$ and time-to-maturity $t$ is such that

$$\left. \frac{\partial \hat{\sigma}(\kappa, t)}{\partial \kappa} \right|_{\kappa=0} \sim \begin{cases} 
\frac{d_1}{t^{1/2}}, & \text{if pure jump model with } C_+ \neq C_-, \\
\frac{d_1}{t^{1/2}} \frac{1}{Y}, & \text{if pure jump model with } C_+ = C_-, \\
\frac{d_1}{t^{1/2}} \frac{1}{2}, & \text{if mixed model with } C_+ \neq C_-, \\
\frac{d_1}{t}, & \text{if mixed model with } C_+ = C_-.
\end{cases}$$

Furthermore, in the asymmetric cases, the sign of $d_1$ is the same as that of $C_+ - C_-$. 
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
   - A Calibration Case-Study
   - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
Calibration of S&P 500 Option Prices

Objectives:

- To assess the plausibility of the divergence of the skew to $-\infty$ as a power of time-to-maturity.
- Calibrate the parameter $Y$, which measures the activity of small jumps in the price process.

Data:

**Implied Volatility Smile and ATM Skew**

Figure: (a) The graphs of the ATM implied volatility skew are consistent with divergence to \(-\infty\) as a power law. (b) The implied volatility smiles on Jan. 15, 2014, corresponding to maturities ranging from 0.032 (8 days) to 0.25 (3 months), show greater steepness as time-to-maturity decreases, which is consistent with the presence of jumps.
Calibration of the Index Of Jump Activity $Y$

Figure: The $Y$-estimates implied by daily and monthly regressions of $\log \left( \frac{\partial \hat{\sigma}(\kappa, t)}{\partial \kappa} \right)_{\kappa=0}$ on $\log t$ assuming a pure-jump model (panel (a)), and a mixed model (panel (b)). The first and third quantiles of the regressions’ $R^2$ are 0.977 and 0.995, respectively.
ATM Implied Volatility

Figure: (a) ATM implied volatility as a function of time-to-maturity in years, for each business day in Jan–Jun, 2014. (b) ATM implied volatility for the shortest outstanding maturity compared to the VIX index. These suggest a mixed model rather than a pure jump model since $\hat{\sigma}(t, 0) \to \sigma_0 > 0$. 
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
   - A Calibration Case-Study
   - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
Ongoing and Future Research

1. Besides the skew, the convexity of the implied volatility, 
\[ \frac{\partial^2 \hat{\sigma}(\kappa,t)}{\partial \kappa^2} \bigg|_{\kappa=0}, \]
is very important in trading and, thus, its asymptotic behavior is of great interest.
Ongoing and Future Research

1. Besides the skew, the convexity of the implied volatility, \( \frac{\partial^2 \hat{\sigma}(\kappa,t)}{\partial \kappa^2} \bigg|_{\kappa=0} \), is very important in trading and, thus, its asymptotic behavior is of great interest.

2. Small-time asymptotics of options written on *Leveraged Exchange-Traded Fund* (LETF). LETF is a managed portfolio, which seeks to multiply the instantaneous returns of a reference Exchange-Traded Fund (ETF).
Ongoing and Future Research

1. Besides the skew, the convexity of the implied volatility, \[ \left. \frac{\partial^2 \hat{\sigma}(\kappa,t)}{\partial \kappa^2} \right|_{\kappa=0} \], is very important in trading and, thus, its asymptotic behavior is of great interest.

2. Small-time asymptotics of options written on *Leveraged Exchange-Traded Fund* (LETF). LETF is a managed portfolio, which seeks to multiply the instantaneous returns of a reference Exchange-Traded Fund (ETF).

3. There is no known results for American options. It is expected that the early exercise feature won’t affect the leading order terms of the asymptotics but they should contribute to the high-order terms.
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
   - A Calibration Case-Study
   - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
Merton Log-Normal Model

Consider the following model for the log-return process $X_t = \log \frac{S_t}{S_0}$ of a financial asset:

$$X_t = at + \sigma W_t + \sum_{i: \tau_i \leq t} \zeta_i, \quad \zeta_i \sim \text{i.i.d. } N(\mu_{jmp}, \sigma_{jmp}^2), \quad \tau_i \sim \text{Poisson}(\lambda)$$

**Goal:** Estimate the volatility $\sigma$ based on a discrete record $X_{t_1}, \ldots, X_{t_n}$
Outline

1. Short-time asymptotics of Option Prices
   Short-time ATM Skew Asymptotics
   A Calibration Case-Study
   Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   Multipower Variations and Truncated Realized Variations
   Optimal Threshold Selection
      via Expected number of jump misclassifications
      via conditional Mean Square Error (cMSE)
   Ongoing and Future Research
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV[X]_n := \frac{1}{T} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \]
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV[X]_n := \frac{1}{T} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \xrightarrow{n \to \infty} \sigma^2 + \sum_{j: \tau_j \leq T} \zeta_j^2 \]
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV[X]_n := \frac{1}{T} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \xrightarrow{n \to \infty} \sigma^2 + \sum_{j: \tau_j \leq T} \zeta_j^2 \]

① Bipower Realized Variations (Barndorff-Nielsen and Shephard):

\[ BPV[X]_n := \sum_{i=0}^{n-1} \left| X_{t_{i+1}} - X_{t_i} \right| \left| X_{t_{i+2}} - X_{t_{i+1}} \right| , \]
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV[X]_n := \frac{1}{T} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \xrightarrow{n \to \infty} \sigma^2 + \sum_{j: \tau_j \leq T} \zeta_j^2 \]

1. Bipower Realized Variations (Barndorff-Nielsen and Shephard):

\[ BPV[X]_n := \sum_{i=0}^{n-1} |X_{t_{i+1}} - X_{t_i}| \left| X_{t_{i+2}} - X_{t_{i+1}} \right|, \]

2. Truncated Realized Variations (Mancini):

\[ TRV_n[X](\varepsilon) := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \mathbf{1}_{\left\{ |X_{t_{i+1}} - X_{t_i}| \leq \varepsilon \right\}}, \quad (\varepsilon \in [0, \infty)). \]
Calibration or Tuning of the Estimator

Problem: How do you choose the threshold parameter $\varepsilon$?
Calibration or Tuning of the Estimator

Problem: How do you choose the threshold parameter $\varepsilon$?

Figure: (left) 5-min Merton observations with $\sigma = 0.4$, $\sigma_{jmp} = 3(h)$, $\mu_{jmp} = 0$, $\lambda = 200$; (right) TRV estimates for all the truncation levels.
Calibration or Tuning of the Estimator

Problem: How do you choose the threshold parameter $\varepsilon$?

Log-Normal Merton Model

Performace of Truncated Realized Variations

Figure: (left) 5 minute Merton observations with $\sigma = 0.2$, $\sigma_{jmp} = 1.5(h)$, $\mu_{jmp} = 0$, $\lambda = 1000$; (right) TRV performance wrt the truncation level
Popular truncations $\varepsilon$

Literature consists of mostly “ad hoc” selection methods for $\varepsilon$, aimed to satisfy sufficient conditions for the consistency and asymptotic normality of the associated estimators. The most popular is the so-called Power Threshold:

$$\varepsilon_{\alpha,\omega}^{Pwr} := \alpha h^\omega, \text{ for } \alpha > 0 \text{ and } \omega \in (0, 1/2).$$

The rule of thumb says to use a value of $\omega$ close to $1/2$ (typically, 0.495).
Outline

1. Short-time asymptotics of Option Prices
   - Short-time ATM Skew Asymptotics
   - A Calibration Case-Study
   - Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   - Multipower Variations and Truncated Realized Variations
   - Optimal Threshold Selection
     - via Expected number of jump misclassifications
     - via conditional Mean Square Error (cMSE)
   - Ongoing and Future Research
General Approach

1. Fix a suitable and sensible metric for estimation error; say,

\[ MSE(\epsilon) = \mathbb{E}\left[\left(\frac{1}{T} TRV_n(\epsilon) - \sigma^2\right)^2\right] \]

2. Show the existence of the optimal threshold \( \epsilon^*_n \) minimizing the error function;

3. Characterize the asymptotic behavior \( \epsilon^*_n \) (when \( n \to \infty \)) to infer
   - qualitative information such as rate of convergence on \( n \) and dependence on the underlying parameters of the model \( (\sigma, \sigma_J, \lambda) \)
   - Devise a plug-in type implementation of \( \epsilon^* \) by estimating those parameters (if possible).
Via Expected number of jump misclassifications
Via Expected number of jump misclassifications

1. Estimation Error: (F-L & Nisen, SPA 2013)

\[ \text{Loss}_n(\varepsilon) := \mathbb{E} \sum_{i=1}^{n} \left( 1[|\Delta^n_j X| > \varepsilon, \Delta^n_j N = 0] + 1[|\Delta^n_j X| \leq \varepsilon, \Delta^n_j N \neq 0] \right). \]

2. Notation:

\[ N_t := \# \text{ of jumps by time } t \]

\[ \Delta^n_j X := X_{t_i} - X_{t_i-1} \]

\[ \Delta^n_j N := N_{t_i} - N_{t_i-1} = \# \text{ of jumps during } (t_{i-1}, t_i) \]
Existence and Infill Asymptotic Characterization

Theorem (F-L & Nisen, SPA 2013)

For $n$ large enough, the loss function $\text{Loss}_n(\epsilon)$ is convex and, moreover, enjoys a unique global minimum $\epsilon^*_n$. As $n \to \infty$, the optimal threshold sequence $(\epsilon^*_n)_n$ is such that $\epsilon^*_n = \sqrt{3} \sigma^2 h_n \log \left( \frac{1}{h_n} \right) + h_n$. o.t., where $h_n := t_i - t_{i-1}$ is the time step.

Remark: The leading order term is proportional to the Lévy's modulus of continuity of the Brownian motion!
Existence and Infill Asymptotic Characterization

Theorem (F-L & Nisen, SPA 2013))

1. For $n$ large enough, the loss function $\text{Loss}_n(\varepsilon)$ is convex and, moreover, enjoys a unique global minimum $\varepsilon^*_n$. 

Remark: The leading order term is proportional to the Lévy’s modulus of continuity of the Brownian motion!
Existence and Infill Asymptotic Characterization

**Theorem (F-L & Nisen, SPA 2013)**

1. **For n large enough**, the loss function $\text{Loss}_n(\varepsilon)$ is **convex** and, moreover, enjoys a **unique global minimum** $\varepsilon^*_n$.

2. **As $n \to \infty$**, the optimal threshold sequence $(\varepsilon^*_n)_n$ is such that

$$
\varepsilon^*_n = \sqrt{3\sigma^2 h_n \log \left( \frac{1}{h_n} \right)} + \text{h.o.t.},
$$

where $h_n := t_i - t_{i-1}$ is the time step.
Existence and Infill Asymptotic Characterization

Theorem (F-L & Nisen, SPA 2013))

1. For $n$ large enough, the loss function $\text{Loss}_n(\varepsilon)$ is convex and, moreover, enjoys a unique global minimum $\varepsilon_n^*$.

2. As $n \to \infty$, the optimal threshold sequence $(\varepsilon_n^*)_n$ is such that

$$
\varepsilon_n^* = \sqrt{3\sigma^2 h_n \log \left( \frac{1}{h_n} \right)} + \text{h.o.t.},
$$

where $h_n := t_i - t_{i-1}$ is the time step.

Remark: The leading order term is proportional to the Lévy’s modulus of continuity of the Brownian motion!
A Feasible Implementation based on $\varepsilon_n^*$

(i) Get a "rough" estimate of $\sigma^2$ via, e.g., realized QV:

$$\hat{\sigma}^2_{n,0} := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_i-1}|^2$$
A Feasible Implementation based on $\varepsilon^*_n$

(i) Get a “rough” estimate of $\sigma^2$ via, e.g., realized QV:

$$\hat{\sigma}^2_{n,0} := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$

(ii) Use $\hat{\sigma}^2_{n,0}$ to estimate the optimal threshold

$$\varepsilon^*_{n,0} := \left(3 \hat{\sigma}^2_{n,0} h_n \log(1/h_n)\right)^{1/2}$$
A Feasible Implementation based on $\varepsilon^*_n$

(i) Get a “rough” estimate of $\sigma^2$ via, e.g., realized QV:

$$\hat{\sigma}^2_{n,0} := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$

(ii) Use $\hat{\sigma}^2_{n,0}$ to estimate the optimal threshold

$$\varepsilon^*_{n,0} := \left(3 \hat{\sigma}^2_{n,0} h_n \log(1/h_n)\right)^{1/2}$$

(iii) Refine $\hat{\sigma}^2_{n,0}$ using thresholding,

$$\hat{\sigma}^2_{n,1} = \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 1[|X_{t_i} - X_{t_{i-1}}| \leq \varepsilon^*_{n,0}]$$
A Feasible Implementation based on $\varepsilon^*_n$

(i) Get a “rough” estimate of $\sigma^2$ via, e.g., realized QV:

$$\hat{\sigma}^2_{n,0} := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$

(ii) Use $\hat{\sigma}^2_{n,0}$ to estimate the optimal threshold

$$\hat{\varepsilon}^*_{n,0} := \left(3 \hat{\sigma}^2_{n,0} h_n \log(1/h_n)\right)^{1/2}$$

(iii) Refine $\hat{\sigma}^2_{n,0}$ using thresholding,

$$\hat{\sigma}^2_{n,1} = \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 1\left[|X_{t_i} - X_{t_{i-1}}| \leq \hat{\varepsilon}^*_{n,0}\right]$$

(iv) Iterate Steps (ii) and (iii):

$$\hat{\sigma}^2_{n,0} \rightarrow \hat{\varepsilon}^*_n \rightarrow \hat{\sigma}^2_{n,1} \rightarrow \hat{\varepsilon}^*_n \rightarrow \hat{\sigma}^2_{n,2} \rightarrow \cdots \rightarrow \hat{\sigma}^2_{n,\infty}$$
Illustration I

**Log-Normal Merton Model**

**Performace of Truncated Realized Variations**

**Figure:** (left) Merton Model with $\sigma = 0.4$, $\sigma_{jmp} = 3(h)$, $\mu_{jmp} = 0$, $\lambda = 200$; (right) TRV performance wrt the truncation level. Red dot is $\hat{\sigma}_{n,1} = 0.409$, while purple dot is the limiting estimator $\hat{\sigma}_{n,\infty} = 0.405$
Illustration II

Log-Normal Merton Model

Performace of Truncated Realized Variations

Figure: (left) Merton Model with $\sigma = 0.2$, $\sigma_{jmp} = 1.5(h)$, $\mu_{jmp} = 0$, $\lambda = 1000$; (right) TRV performance wrt the truncation level. Red dot is $\hat{\sigma}_{n,1} = 0.336$, while purple dot is the limiting estimator $\hat{\sigma}_{n,\infty} = 0.215$
Via conditional Mean Square Error (cMSE)
Via conditional Mean Square Error (cMSE)

1. We now propose a second approach in which we aim to control

\[
MSE_c(\varepsilon) := \mathbb{E} \left[ \left( TRV_n(\varepsilon) - \int_0^T \sigma_s^2 ds \right)^2 \middle| \sigma, J \right].
\]

This is in the more general semimartingale setting:

\[
X_t = X_0 + \int_0^t \sigma_s dW_s + J_t
\]

where \( J \) is a pure-jump process.
We now propose a second approach in which we aim to control

\[ \text{MSE}_c(\varepsilon) := \mathbb{E} \left[ \left( TRV_n(\varepsilon) - \int_0^T \sigma_s^2 ds \right)^2 \right| \sigma, J \] .

This is in the more general semimartingale setting:

\[ X_t = X_0 + \int_0^t \sigma_s dW_s + J_t \]

where \( J \) is a pure-jump process.

Key Assumptions:

\( \sigma_t > 0, \forall \ t \), and \( \sigma \) and \( J \) are independent of \( W \).
Key Relationships

\[ b_i(\varepsilon) := \mathbb{E} \left[ (\Delta_i X)^2 1_{\{|\Delta_i X| \leq \varepsilon\}} \right] \sigma, J \]
Key Relationships

\[ b_i(\varepsilon) := \mathbb{E} \left[ (\Delta_i X)^2 \mathbf{1}_{\{ \Delta_i X \leq \varepsilon \}} | \sigma, J \right] \]

Then, with the notation

\[ m_i := \Delta_i^n J = J_{t_i} - J_{t_{i-1}}, \quad \sigma_i^2 := \int_{t_{i-1}}^{t_i} \sigma_s^2 ds, \]
Key Relationships

\[ b_i(\varepsilon) := \mathbb{E} \left[ (\Delta_i X)^2 \mathbf{1}_{\{\Delta_i X \leq \varepsilon\}} \right| \sigma, J] \]

Then, with the notation

\[ m_i := \Delta_i \eta J = J_{t_i} - J_{t_{i-1}}, \quad \sigma_i^2 := \int_{t_{i-1}}^{t_i} \sigma_s^2 ds, \]

we have

\[
\begin{align*}
  b_i(\varepsilon) &= -\frac{\sigma_i}{\sqrt{2\pi}} \left( e^{-(\varepsilon-m_i)^2/2\sigma_i^2} (\varepsilon + m_i) + e^{-(\varepsilon+m_i)^2/2\sigma_i^2} (\varepsilon - m_i) \right) \\
  &\quad + (m_i^2 + \sigma_i^2) \int_{\frac{\varepsilon-m_i}{\sigma_i}}^{\frac{\varepsilon+m_i}{\sigma_i}} e^{-x^2/2} \sqrt{2\pi} dx
\end{align*}
\]
Key Relationships

\[ b_i(\varepsilon) := \mathbb{E} \left[ (\Delta_i X)^2 \mathbf{1}_{\{\Delta_i X \leq \varepsilon\}} \right] | \sigma, J \]

Then, with the notation

\[ m_i := \Delta_i^\eta J = J_{t_i} - J_{t_{i-1}}, \quad \sigma_i^2 := \int_{t_{i-1}}^{t_i} \sigma_s^2 ds, \]

we have

\[
\begin{align*}
b_i(\varepsilon) &= -\frac{\sigma_i}{\sqrt{2\pi}} \left( e^{-(\varepsilon-m_i)^2/2\sigma_i^2} (\varepsilon + m_i) + e^{-(\varepsilon+m_i)^2/2\sigma_i^2} (\varepsilon - m_i) \right) \\
&\quad + (m_i^2 + \sigma_i^2) \int_{\frac{\varepsilon-m_i}{\sigma_i}}^{\frac{\varepsilon+m_i}{\sigma_i}} e^{-x^2/2} dx
\end{align*}
\]

\[
\frac{db_i(\varepsilon)}{d\varepsilon} =
\]
Key Relationships

\[ b_i(\varepsilon) := \mathbb{E} \left[ (\Delta_i X)^2 \mathbb{1}_{\{\Delta_i X \leq \varepsilon\}} \right] \sigma, J \]

Then, with the notation

\[ m_i := \Delta_i^n J = J_{t_i} - J_{t_{i-1}}, \quad \sigma_i^2 := \int_{t_{i-1}}^{t_i} \sigma_s^2 ds, \]

we have

\[ b_i(\varepsilon) = -\frac{\sigma_i}{\sqrt{2\pi}} \left( e^{-\frac{(\varepsilon - m_i)^2}{2\sigma_i^2}} (\varepsilon + m_i) + e^{-\frac{(\varepsilon + m_i)^2}{2\sigma_i^2}} (\varepsilon - m_i) \right) \]

\[ + \left( m_i^2 + \sigma_i^2 \right) \int_{\frac{\varepsilon - m_i}{\sigma_i}}^{\frac{\varepsilon + m_i}{\sigma_i}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \]

\[ \frac{db_i(\varepsilon)}{d\varepsilon} = \varepsilon^2 a_i(\varepsilon), \quad \text{with} \quad a_i(\varepsilon) := \frac{e^{-\frac{(\varepsilon - m_i)^2}{2\sigma_i^2}} + e^{-\frac{(\varepsilon + m_i)^2}{2\sigma_i^2}}}{\sigma_i \sqrt{2\pi}} \]
Conditional Mean Square Error \((MSE_c)\)

Theorem \((F-L & Mancini, 2017)\)

\[
MSE_c(\varepsilon) := E\left[ (TRV_n(\varepsilon) - IV)^2 \right]_{\sigma, J}
\]

is such that

\[
dMSE_c(\varepsilon) = \varepsilon^2 G(\varepsilon),
\]

where \(G(\varepsilon) := \sum_i a_i(\varepsilon)^2 \varepsilon^2 + 2\sum_j b_j(\varepsilon)^2 - 2\sum_j \sigma_j^2 \).

Furthermore, there exists an optimal threshold \(\varepsilon^\star_n\) that minimizes \(MSE_c(\varepsilon)\) and is such that \(G(\varepsilon^\star_n) = 0\).
Conditional Mean Square Error ($MSE_c$)

Theorem (F-L & Mancini, 2017)

\[
MSE_c(\varepsilon) := \mathbb{E} \left[ (TRV_n(\varepsilon) - IV)^2 \right| \sigma, J] \text{ is such that}
\]

\[
\frac{d}{d\varepsilon} MSE_c(\varepsilon) = \varepsilon^2 G(\varepsilon),
\]

where

\[
G(\varepsilon) := \sum_i a_i(\varepsilon) \left( \varepsilon^2 + 2 \sum_{j \neq i} b_j(\varepsilon) - 2 \sum_j \sigma_j^2 \right).
\]
Theorem (F-L & Mancini, 2017)

1. \( \text{MSE}_c(\varepsilon) := \mathbb{E} \left[ (\text{TRV}_n(\varepsilon) - IV)^2 \right| \sigma, J ] \) is such that

\[
\frac{d}{d\varepsilon} \text{MSE}_c(\varepsilon) = \varepsilon^2 G(\varepsilon),
\]

where \( G(\varepsilon) := \sum_i a_i(\varepsilon) \left( \varepsilon^2 + 2 \sum_{j \neq i} b_j(\varepsilon) - 2 \sum_j \sigma_j^2 \right) \).

2. Furthermore, there exists an optimal threshold \( \varepsilon_{**}^n \) that minimizes \( \text{MSE}_c(\varepsilon) \) and is such that \( G(\varepsilon_{**}^n) = 0 \).
Asymptotics: FA process with constant variance

Theorem (F-L & Mancini, 2017)
Asymptotics: FA process with constant variance

Theorem (F-L & Mancini, 2017)

1. Suppose that $\sigma_t \equiv \sigma$ is constant and $J$ is a finite jump activity process (with or without drift; not necessarily Lévy):

$$X_t = \sigma W_t + \sum_{j=1}^{N_t} \zeta_j$$

$$\epsilon^\star_n \sim \sqrt{\frac{2}{\sigma^2} h_n \log \left(\frac{1}{h_n}\right)}, h_n \rightarrow 0.$$
Asymptotics: FA process with constant variance

Theorem (F-L & Mancini, 2017)

1. Suppose that $\sigma_t \equiv \sigma$ is constant and $J$ is a finite jump activity process (with or without drift; not necessarily Lévy):

$$X_t = \sigma W_t + \sum_{j=1}^{N_t} \zeta_j$$

2. Then, as $n \to \infty$, the optimal threshold $\varepsilon_n^{**}$ that minimizes the cMSE is such that

$$\varepsilon_n^{**} \sim \sqrt{2\sigma^2 h_n \log \left( \frac{1}{h_n} \right)}, \quad h_n \to 0.$$
A Feasible Implementation based on $\varepsilon_n^{**}$

(i) Get a “rough” estimate of $\sigma^2$ via, e.g., the QV:

$$\tilde{\sigma}_{n,0}^2 := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$
A Feasible Implementation based on $\varepsilon^{**}_n$

(i) Get a “rough” estimate of $\sigma^2$ via, e.g., the QV:

$$\tilde{\sigma}^2_{n,0} := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$

(ii) Use $\tilde{\sigma}^2_{n,0}$ to estimate the optimal threshold

$$\tilde{\varepsilon}^{**}_{n,0} := \left(2 \tilde{\sigma}^2_{n,0} h_n \log(1/h_n)\right)^{1/2}$$
A Feasible Implementation based on $\varepsilon^*_{n}$

(i) Get a "rough" estimate of $\sigma^2$ via, e.g., the QV:
\[
\tilde{\sigma}^2_{n,0} := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2
\]

(ii) Use $\tilde{\sigma}^2_{n,0}$ to estimate the optimal threshold
\[
\hat{\varepsilon}^{**}_{n,0} := \left(2 \tilde{\sigma}^2_{n,0} h_n \log(1/h_n)\right)^{1/2}
\]

(iii) Refine $\tilde{\sigma}^2_{n,0}$ using thresholding,
\[
\tilde{\sigma}^2_{n,1} = \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 1\left[|X_{t_i} - X_{t_{i-1}}| \leq \hat{\varepsilon}^{**}_{n,0}\right]
\]
A Feasible Implementation based on $\varepsilon_{n}^{**}$

(i) Get a “rough” estimate of $\sigma^2$ via, e.g., the QV:

$$\tilde{\sigma}_{n,0}^2 := \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$

(ii) Use $\tilde{\sigma}_{n,0}^2$ to estimate the optimal threshold

$$\hat{\varepsilon}_{n,0}^{**} := \left(2 \tilde{\sigma}_{n,0}^2 h_n \log(1/h_n)\right)^{1/2}$$

(iii) Refine $\tilde{\sigma}_{n,0}^2$ using thresholding,

$$\tilde{\sigma}_{n,1}^2 = \frac{1}{T} \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 1[|X_{t_i} - X_{t_{i-1}}| \leq \hat{\varepsilon}_{n,0}^{**}]$$

(iv) Iterate Steps (ii) and (iii):

$$\tilde{\sigma}_{n,0}^2 \rightarrow \hat{\varepsilon}_{n,0}^{**} \rightarrow \tilde{\sigma}_{n,1}^2 \rightarrow \hat{\varepsilon}_{n,1}^{**} \rightarrow \tilde{\sigma}_{n,2}^2 \rightarrow \cdots \rightarrow \tilde{\sigma}_{n,\infty}^2$$
Illustration II. Continued...

**Figure:** (left) Merton Model with $\lambda = 1000$. Red dot is $\hat{\sigma}_{n,1} = 0.336$, while purple dot is the limiting $\hat{\sigma}_{n,k} = 0.215$. Orange square is $\tilde{\sigma}_{n,1} = 0.225$, while brown square is the limiting estimator $\tilde{\sigma}_{n,\infty} = 0.199$.
Monte Carlo Simulations

<table>
<thead>
<tr>
<th>Estimator</th>
<th>$\hat{\sigma}$</th>
<th>std($\hat{\sigma}$)</th>
<th>Loss</th>
<th>$\bar{\varepsilon}$</th>
<th>$\bar{N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQV</td>
<td>0.3921</td>
<td>0.0279</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_{n,1}$</td>
<td>0.29618</td>
<td>0.02148</td>
<td>70.1</td>
<td>0.01523</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{\sigma}_{n,\infty}$</td>
<td>0.23</td>
<td>0.0108</td>
<td>49.8</td>
<td>0.00892</td>
<td>5.86</td>
</tr>
<tr>
<td>$\tilde{\sigma}_{n,1}$</td>
<td>0.265</td>
<td>0.0163</td>
<td>62.6</td>
<td>0.0124</td>
<td>1</td>
</tr>
<tr>
<td>$\tilde{\sigma}_{n,\infty}$</td>
<td>0.211</td>
<td>0.00588</td>
<td>39.1</td>
<td>0.00671</td>
<td>5.10</td>
</tr>
<tr>
<td>BPV</td>
<td>0.2664</td>
<td>0.0129</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table:** 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 $\mu^{Jmp} = 0$. 
Outline

1. Short-time asymptotics of Option Prices
   Short-time ATM Skew Asymptotics
   A Calibration Case-Study
   Ongoing and Future Research

2. High-Frequency Based Estimation Methods
   Multipower Variations and Truncated Realized Variations
   Optimal Threshold Selection
     via Expected number of jump misclassifications
     via conditional Mean Square Error (cMSE)
   Ongoing and Future Research
In principle, we can apply the proposed method for varying volatility \( t \rightarrow \sigma_t \) by localization; i.e., applying it to periods where \( \sigma \) is approximately constant.
In principle, we can apply the proposed method for varying volatility $t \rightarrow \sigma_t$ by localization; i.e., applying it to periods where $\sigma$ is approximately constant.

However, is it possible to analyze the asymptotic behavior of $MSE_c = \varepsilon^2 G(\varepsilon)$ in terms of certain estimable summary measures? Say,

$$G(\varepsilon) \sim G_0(\varepsilon, m_1, \ldots, m_n, \underline{\sigma}, \bar{\sigma}),$$

where $\underline{\sigma} := \inf_{t \leq T} \sigma_t$ and $\bar{\sigma} := \sup_{t \leq T} \sigma_t$
As it turns, for a Lévy process $J$ and constant $\sigma$, the expected $MSE(\varepsilon) := E \left[ (TRV_n(\varepsilon) - IV)^2 \right]$ is such that

$$\frac{d}{d\varepsilon} MSE(\varepsilon) = n\varepsilon^2 E[a_1(\varepsilon)] \left( \varepsilon^2 + 2(n - 1)E[b_1(\varepsilon)] - 2nh_n\sigma^2 \right)$$
As it turns, for a Lévy process $J$ and constant $\sigma$, the expected $MSE(\varepsilon) := \mathbb{E} \left[ (TRV_n(\varepsilon) - IV)^2 \right]$ is such that

$$\frac{d}{d\varepsilon} MSE(\varepsilon) = n\varepsilon^2 \mathbb{E}[a_1(\varepsilon)] \left( \varepsilon^2 + 2(n - 1) \mathbb{E}[b_1(\varepsilon)] - 2nh_n\sigma^2 \right)$$

Therefore, there exists a unique minimum point $\bar{\varepsilon}^\star \star_n$ which is a solution of the equation

$$\varepsilon^2 + 2(n - 1) \mathbb{E}[b_1(\varepsilon)] - 2nh_n\sigma^2 = 0.$$
Ongoing and Future Research

Furthermore,

• As expected, in the finite jump activity case, it turns out
  \[ \bar{\varepsilon}_{n} \sim \sqrt{\frac{2}{\sigma^2_h n \log(1/h_n)}} \]

• But, surprisingly, if \( J \) is a \( Y \)-stable Lévy process (IA),
  \[ \bar{\varepsilon}_{n} \sim \sqrt{\left(2 - Y\right)\sigma^2_h n \log(1/h_n)} \]

Thus, the higher the jump activity is, the lower the optimal threshold
has to be to discard the higher noise represented by the jumps.

• Does this phenomenon holds for the minimizer \( \varepsilon_{n} \) of the cMSE?
• Can we generalize it to Lévy processes with stable like jumps?
Ongoing and Future Research

Furthermore,

- As expected, in the finite jump activity case, it turns out

\[ \bar{\varepsilon}^{**}_n \sim \sqrt{2\sigma^2 h_n \log \left( \frac{1}{h_n} \right)} \]

Thus, the higher the jump activity is, the lower the optimal threshold has to be to discard the higher noise represented by the jumps.

- Does this phenomenon hold for the minimizer \( \varepsilon^{**}_n \) of the cMSE?

- Can we generalize it to Lévy processes with stable-like jumps?
Furthermore,

- As expected, in the finite jump activity case, it turns out
  \[ \bar{\epsilon}^{**}_n \sim \sqrt{2\sigma^2 h_n \log \left( \frac{1}{h_n} \right)} \]

- But, surprisingly, if \( J \) is a \( Y \)-stable Lévy process (IA),
  \[ \bar{\epsilon}^{**}_n \sim \sqrt{(2 - Y)\sigma^2 h_n \log \left( \frac{1}{h_n} \right)} \]

Thus, the higher the jump activity is, the lower the optimal threshold has to be to discard the higher noise represented by the jumps.
Furthermore,

- As expected, in the finite jump activity case, it turns out
  \[ \bar{\varepsilon}_{n}^{**} \sim \sqrt{2\sigma^{2}h_{n}\log \left( \frac{1}{h_{n}} \right)} \]

- But, surprisingly, if \( J \) is a \( Y \)-stable Lévy process (IA),
  \[ \bar{\varepsilon}_{n}^{**} \sim \sqrt{(2 - Y)\sigma^{2}h_{n}\log \left( \frac{1}{h_{n}} \right)} \]

Thus, the higher the jump activity is, the lower the optimal threshold has to be to discard the higher noise represented by the jumps.

- Does this phenomenon holds for the minimizer \( \varepsilon_{n}^{**} \) of the cMSE?
Furthermore,

- As expected, in the finite jump activity case, it turns out
  \[\bar{\varepsilon}^{**}_n \sim \sqrt{2\sigma^2 h_n \log \left(\frac{1}{h_n}\right)}\]

- But, surprisingly, if \(J\) is a \(Y\)-stable Lévy process (IA),
  \[\bar{\varepsilon}^{**}_n \sim \sqrt{(2 - Y)\sigma^2 h_n \log \left(\frac{1}{h_n}\right)}\]

Thus, the higher the jump activity is, the lower the optimal threshold has to be to discard the higher noise represented by the jumps.

- Does this phenomenon holds for the minimizer \(\varepsilon^{**}_n\) of the cMSE?
- Can we generalize it to Lévy processes with stable like jumps?
Further Reading I


Further Reading II

Optimally Thresholded Realized Power Variations for Lévy Jump Diffusion Models.

J.E. Figueroa-López & C. Mancini.
Optimum thresholding using mean and conditional mean square error.

Optimal iterative threshold-kernel estimation of jump diffusion processes.