Applications of short-time asymptotics to the statistical estimation and option pricing of Lévy-driven models

José Enrique Figueroa-López

1Department of Statistics
Purdue University

CIMAT and Universidad de Guanajuato
Guanajuato, Mexico
May 13-17, 2013
Overview of the Lecture Series

1 Lecture I
   - Introduce the elements of continuous-time stochastic modeling.
   - Basic building blocks: Brownian Motion, Poisson Process, etc.
   - **Semiparametric Estimation** of finite-jump activity **Lévy jump-diffusions**
     (Joint work with Jeff Nisen).

2 Lecture II
   - Quick primer of **Lévy processes**
   - Options and derivatives. Elements of arbitrage-free option pricing.
   - Short-time asymptotics for option prices in exponential **Lévy models**
     (Joint work with Martin Forde, Christian Houdré, and Ruoting Gong).

3 Lecture III
   - Exotic options: Asian, lookback, Barrier options, etc.
   - An adaptive **Monte Carlo valuation method** for Barrier options in exponential **Lévy models**.
     (Joint work with Peter Tankov).
Lecture I: Outline

1 Continuous-time stochastic modeling
   Brownian Motion
   Black-Scholes framework
   Poisson Processes
   Compound Poisson Processes
   Merton and Kou Models
   Varying volatility and jump intensity

2 Semiparametric estimation of Lévy Jump-Diffusion Models
   The Statistical Problems and the Main Estimators
   Optimally Thresholded Power Estimators
   Main Results
   Conclusions
Motivation

1. Complex (natural or social) systems (e.g., financial markets, environment, internet or city traffic) are often modeled as random experiments or systems.

2. Key numerical features or variables of complex systems (e.g., stock prices, exchange rates, temperature, ozone levels, etc.) can then be modeled as random measurements evolving continuously in time.

3. (Continuous-Time) Stochastic Process:
   A collection of random variables indexed by a time parameter \( t \geq 0 \), \( \{X_t\}_{t \geq 0} \), whose realized values are the result of a common underlying random experiment.

4. Terminology: For each possible outcome \( \omega \) of the random experiment, the function \( t \rightarrow X_t(\omega) \) is called the sample path of the process. We can see this as a random function.
Outline

1 Continuous-time stochastic modeling
   Brownian Motion
      Black-Scholes framework
   Poisson Processes
   Compound Poisson Processes
      Merton and Kou Models
   Varying volatility and jump intensity

2 Semiparametric estimation of Lévy Jump-Diffusion Models
   The Statistical Problems and the Main Estimators
   Optimally Thresholded Power Estimators
   Main Results
   Conclusions
In physics, Brownian motion refers to the presumably random motion of particles suspended in a fluid resulting from their bombardment by the fast-moving molecules in the fluid.

In probability, Brownian Motion is a type of continuous-time stochastic process $\{W_t\}_{t \geq 0}$ that serves as model for the motion of the previously mentioned type of particles.

Figure: Planar Brownian Motion
Consider a symmetric random walk with unit time step:

\[ R_0 = 0, \quad R_k = \sum_{i=1}^{k} \Delta_i, \quad (k \geq 1), \quad \text{where} \quad \Delta_i \overset{i.i.d.}{\sim} \begin{cases} 1, & \text{w.p. } \frac{1}{2}, \\ -1, & \text{w.p. } \frac{1}{2}. \end{cases} \]

Scale the time and the space as follows:

\[ X_n(t) = \frac{1}{\sqrt{n}} R_{\lfloor nt \rfloor} = R_{k-1}, \quad \text{if} \quad \frac{k-1}{n} \leq t < \frac{k}{n}. \]

What happens to the distributional properties of the process \( \{X_n(t)\}_{t \geq 0} \) when \( n \to \infty \)?
For any \( s \leq t \), the variance \( \text{Var} (X_n(t) - X_n(s)) \) can be written as

\[
\text{Var} \left( \frac{1}{\sqrt{n}} \sum_{i=\lfloor sn \rfloor+1}^{\lfloor nt \rfloor} \Delta_i \right) = \frac{1}{n} \sum_{i=\lfloor sn \rfloor+1}^{\lfloor nt \rfloor} \text{Var} (\Delta_i) = \frac{1}{n} ([nt] - [sn]) \xrightarrow{n \to \infty} t - s.
\]

Asymptotically Stationary Increments:

\[
X_n(t) - X_n(s) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor sn \rfloor+1}^{\lfloor nt \rfloor} \Delta_i \xrightarrow{D} \mathcal{N} (0, t - s).
\]

Memoryless Property:

\[
P \left( X_n(t) - X_n(s) \in (c, d) \mid X_n(u), u \leq s \right) = P \left( X_n(t) - X_n(s) \in (c, d) \right)
\]

\[
x \xrightarrow{n \to \infty} \int_c^d \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{u^2}{2(t-s)}} du.
\]
Definition of (standard) Brownian Motion

1. Asymptotically, as \( n \to \infty \), the scaled random walk \( \{X_n(t)\}_{t \geq 0} \) behaves like a continuous-time process \( \{W_t\}_{t \geq 0} \) with the following properties:
   
   (i) \( W_0 = 0 \)
   
   (ii) Stationary increments:
   
   \[ W_t - W_s \sim \mathcal{N}(0, t - s), \quad \text{for any} \quad s < t; \]

   (iii) Memoryless property:
   
   \[ \mathbb{P}(W_t - W_s \in (c, d)|W_u, u \leq s) = \mathbb{P}(W_t - W_s \in (c, d)), \quad \text{for any} \quad s < t; \]

   (iv) Continuous-paths:
   
   For each possible experiment outcome \( \omega \), the process’ path \( t \to X_t(\omega) \) is continuous;

2. A process satisfying (i)-(iv) is called a (standard) Brownian Motion or a Wiener process.
Some “direct” consequences

1. **Independent Increments**: For any \( t_0 < t_1 < \cdots < t_n \), the increments

\[ W_{t_1} - W_{t_0}, \ W_{t_2} - W_{t_1}, \ldots, \ W_{t_n} - W_{t_{n-1}} \]

are mutually independent.

2. **Markov Property**: For any \( s < t \),

\[ \mathbb{P}(W_t \in (c, d) | W_u, u \leq s) = \mathbb{P}(W_t \in (c, d) | W_s). \]

3. **Self-similarity property**: 

\[ \{\sigma W_t\}_{t \geq 0} \overset{\text{d}}{=} \{W_{\sigma^2 t}\}_{t \geq 0}. \]

4. **Asymptotic quadratic variation**: Recalling that \( W_{\frac{i}{n}} - W_{\frac{i-1}{n}} \sim \text{i.i.d.} \mathcal{N}(0, \frac{1}{n}) \),

\[ \sum_{i=1}^{n} \left( W_{\frac{i}{n}} - W_{\frac{i-1}{n}} \right)^2 = \left( \frac{1}{\sqrt{n}} Z_i \right)^2 = \frac{1}{n} \sum_{i=1}^{n} Z_i^2 \xrightarrow{n \to \infty} \mathbb{E}(Z_1^2) = 1. \]
1 The **Black-Scholes framework** is the most famous mathematical model for pricing options written on a risky tradable asset (say, a stock), evolving continuously in time.

2 A key assumption of the model is that the price process of the underlying risky asset follows a **geometric Brownian motion** (G.B.M.):

\[ S_t = S_0 e^{\sigma W_t + \alpha t} \leftarrow \text{Price of the asset at time } t, \]

for some constants \( \sigma > 0 \) and \( \alpha \in \mathbb{R} \).

3 The constant \( \sigma \) is called the **volatility** of the asset, while \( \alpha \) is the (annualized) expected rate of return of the asset:

\[
R_{[u,v]} = \ln \frac{S_v}{S_u} = \sigma (W_v - W_u) + \alpha (v - u) \leftarrow \text{log return over } [u, v],
\]

\[
\mathbb{E} \left( R_{[u,v]} \right) = \alpha (v - u), \quad \text{Var} \left( R_{[u,v]} \right) = \sigma^2 (v - u).
\]
Outline

1. Continuous-time stochastic modeling
   - Brownian Motion
   - Black-Scholes framework
   - Poisson Processes
      - Compound Poisson Processes
      - Merton and Kou Models
   - Varying volatility and jump intensity

2. Semiparametric estimation of Lévy Jump-Diffusion Models
   - The Statistical Problems and the Main Estimators
   - Optimally Thresholded Power Estimators
   - Main Results
   - Conclusions
Consider the random walk:

\[ M_k^{(p)} = \sum_{i=1}^{k} \Delta_i^{(p)}, \quad \text{where} \quad \Delta_i^{(p)} \sim \begin{cases} 1, & \text{w.p. } p, \\ 0, & \text{w.p. } 1 - p. \end{cases} \]

We think of \( M_k^{(p)} \) as the number of “successes” in \( k \) consecutive independent trials with probability of success \( p \).

Next, scale the time and modify the probability \( p \) so that the trials will take place every \( \frac{1}{n} \) time span and the average number of successes per unit time is held constant to \( \lambda \):

\[ X_n(t) = M_{[tn]}^{(pn)}, \quad \text{with} \quad p_n = \frac{\lambda}{n} \quad (\text{indeed}, \quad \mathbb{E} X_n(1) = \mathbb{E} M_n^{(pn)} = np_n = \lambda); \]

What happen to the distributional properties of the process \( \{X_n(t)\}_{t \geq 0} \) when \( n \to \infty \)?
1. For any \( s \leq t \), the increment \( X_n(t) - X_n(s) = \sum_{i=[sn]+1}^{[nt]} \Delta_i^{(pn)} \) is such that

\[
\mathbb{E}(X_n(t) - X_n(s)) = ([nt] - [ns]) \mathbb{E}\Delta_1^{(pn)} = ([nt] - [ns]) \frac{\lambda}{n} \xrightarrow{n \to \infty} \lambda(t - s)
\]

\[
\text{Var}(X_n(t) - X_n(s)) = ([nt] - [ns]) \text{Var}\left(\Delta_1^{(pn)}\right)
= ([nt] - [ns]) p_n(1 - p_n) \xrightarrow{n \to \infty} \lambda(t - s)
\]

\( X_n(t) - X_n(s) \sim \text{Bin}\left(n, \frac{\lambda}{n}\right) \xrightarrow{D} \text{Poiss}(\lambda(t - s)) \).

2. Memoryless Property:

\[
\mathbb{P}(X_n(t) - X_n(s) = j \mid X_n(u), u \leq s) = P(X_n(t) - X_n(s) = j)
\]

\[
\xrightarrow{n \to \infty} e^{-\lambda(t-s)} \frac{\lambda(t-s)^j}{j!}.
\]
Definition of a Poisson process

1. Asymptotically, as \( n \to \infty \), the counting process \( \{ X_n(t) \}_{t \geq 0} \) behaves like a continuous-time process \( \{ N_t \}_{t \geq 0} \) with the following properties:

   (i) \( N_0 = 0 \)

   (ii) Stationary increments:

   \[ N_t - N_s \sim \text{Poiss}(\lambda(t - s)), \quad \text{for any } s < t; \]

   (iii) Memoryless property:

   \[ \mathbb{P}(N_t - N_s = j \mid N_u, u \leq s) = \mathbb{P}(N_t - N_s = j), \quad \text{for any } s < t; \]

   (iv) Non-decreasing:

   For each possible experiment outcome \( \omega \), the process’ path \( t \to N_t(\omega) \) is piece-wise nondecreasing with size jump’s size 1;

2. A process satisfying (i)-(iv) is called a (homogeneous) Poisson Process with intensity of events, \( \lambda \).
Some important consequences

1. **Independent Increments:** For any \( t_0 < t_1 < \cdots < t_n \), the increments

\[
N_{t_1} - N_{t_0}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}},
\]

are mutually independent.

2. **Markov Property:** For any \( s < t \),

\[
\mathbb{P}(N_t = j \mid N_u, u \leq s) = \mathbb{P}(N_t = j \mid N_s).
\]

3. The time span between consecutive events, namely,

\[
\tau_1 = \inf\{s \geq 0 : N_s = 1\}, \quad \tau_i = \inf\{s \geq 0 : N_{s+\tau_i} = i + 1\} \quad (i \geq 1),
\]

are independent exponentially distributed with mean \( \frac{1}{\lambda} \); indeed, e.g.,

\[
\mathbb{P}(\tau_1 \geq t) = \mathbb{P}(N_t = 0) = e^{-\lambda t} \implies f_{\tau_1}(t) = \lambda e^{-\lambda t} \mathbf{1}_{t \geq 0}.
\]
Outline

1 Continuous-time stochastic modeling
   - Brownian Motion
     - Black-Scholes framework
   - Poisson Processes
   - Compound Poisson Processes
     - Merton and Kou Models
   - Varying volatility and jump intensity

2 Semiparametric estimation of Lévy Jump-Diffusion Models
   - The Statistical Problems and the Main Estimators
   - Optimally Thresholded Power Estimators
   - Main Results
   - Conclusions
Given independent identically distributed random variables \( \{\zeta_j\}_{j \geq 1} \) and an independent Poisson process \( \{N_t\}_{t \geq 0} \), the continuous-time process

\[
X_t := \sum_{j=1}^{N_t} \zeta_j,
\]

is said to be a compound Poisson process (c.p.p.).

Time spans between consecutive jumps are still i.i.d. exponential but now the jump sizes are random themselves.
Finite-jump activity Lévy model

1. The log-returns $\ln S_v/S_u$ of the G.B.M. follow a Gaussian distribution, whose mass probability is highly concentrated around its mean;
2. In turn, this implies that a stock price process following the G.B.M. won’t exhibit jump-like movements (“sharp sudden" price shifts).
3. This is inconsistent with empirical features typically observed in the actual market prices.

Figure: SLM (NYSE): Jan.-Mar. 1993 (left) and Feb. 1993 (right).
Finite-jump activity Lévy model. Cont...

1. The previous remarks motivate the incorporation of jumps via a compound Poisson process:

\[ S_t = S_0 e^{X_t}, \quad X_t = \sigma W_t + \alpha t + \sum_{j=1}^{N_t} \zeta_j. \]

2. Important examples:
   - **The Merton model**: The jumps \( \zeta_j \) are Normally distributed with mean \( \mu \) and standard deviation \( \delta \);
   - **Kou model**: The jumps \( \zeta_j \) follow a Laplace distribution:

\[ \zeta_j \sim \begin{cases} 
\exp(\beta_+), & \text{w.p. } p, \\
-\exp(\beta_-), & \text{w.p. } 1 - p,
\end{cases} \quad \rightarrow \quad f_\zeta(x) = \frac{p}{\beta_+} e^{-\frac{x}{\beta_+}} 1_{x>0} + \frac{1-p}{\beta_-} e^{\frac{x}{\beta_-}} 1_{x<0}. \]
1 Continuous-time stochastic modeling
   Brownian Motion
   Black-Scholes framework
   Poisson Processes
   Compound Poisson Processes
   Merton and Kou Models
   Varying volatility and jump intensity

2 Semiparametric estimation of Lévy Jump-Diffusion Models
   The Statistical Problems and the Main Estimators
   Optimally Thresholded Power Estimators
   Main Results
   Conclusions
Motivation

1. An important issue of Lévy jump-diffusion models is its time-homogeneity:
   - Intensity of jumps is, on average, homogeneous through time.
   - The volatility of the continuous component is constant through time.

2. Empirical studies show actual asset prices exhibit varying volatility and jump intensity:

Figure: Annualized daily open-close returns
Economic Clock Approach

1. An useful approach to introduce variations in volatility and jump intensity is a *time-deformation* or a *random clock*.

2. For positive *deterministic* continuous functions $t \to \sigma_t$ and $t \to \lambda_t$, let

   $$
   \Sigma_t := \int_0^t \sigma_s^2 ds, \quad \Lambda_t := \int_0^t \lambda_s ds.
   $$

3. Next, define the continuous-time process:

   $$
   X_t = \int_0^t \alpha_s ds + W_{\Sigma_t} + \sum_{j=1}^{N\Lambda_t} \zeta_j,
   $$

   where $t \to \alpha_t \in \mathbb{R}$ is some given *deterministic* continuous function.

4. **Remark:** The approach hereafter can readily be modified to allow stochastic $\alpha$, $\sigma$, and $\lambda$ driven by a risky factor $Y$ independent of $W$, $N$, and $\{\zeta_j\}_j$. 

Economic Clock Approach. Interpretation...

The (log) return of the stock price process $S_t = S_0 e^{X_t}$ over $[t, t + dt]$ is

$$R_{[t,t+dt]} := \ln \left( \frac{S_{t+dt}}{S_t} \right) = X_{t+dt} - X_t = \int_t^{t+dt} \alpha_s ds + W_{\Sigma_{t+dt}} - W_{\Sigma_t} + \sum_{j=N_t}^{N_{t+dt}} \zeta_j$$

Hence,

- The number of jumps over $[t, t + dt]$ is $N_{\Lambda_{t+dt}} - N_{\Lambda_t}$, which is Poisson with mean $\Lambda_{t+dt} - \Lambda_t = \int_t^{t+dt} \lambda_s ds \approx \lambda_t dt \implies \lambda_t$ is the instantaneous rate of the expected number of jumps at time $t$.

- In the absence of jumps,

$$\mathbb{E} R_{[t,t+dt]} = \int_t^{t+dt} \alpha_s ds \approx \alpha_t dt, \quad \text{Var} \left( R_{[t,t+dt]} \right) = \int_t^{t+dt} \sigma_s^2 ds \approx \sigma_t^2 dt$$

$$\implies \alpha_t \text{ is the instantaneous expected rate of return at time } t$$

$$\implies \sigma_t \text{ is the instantaneous volatility at time } t$$
Outline

1. Continuous-time stochastic modeling
   - Brownian Motion
     - Black-Scholes framework
   - Poisson Processes
   - Compound Poisson Processes
     - Merton and Kou Models
   - Varying volatility and jump intensity

2. Semiparametric estimation of Lévy Jump-Diffusion Models
   - The Statistical Problems and the Main Estimators
   - Optimally Thresholded Power Estimators
   - Main Results
   - Conclusions
Set-up

1. Continuous-time stochastic process $t \to X_t$ with dynamics

$$X_t = \int_0^t \alpha_s ds + W_{\Sigma_t} + \sum_{j=1}^{N_{\Lambda_t}} \zeta_j,$$

where

- $t \to W_t$ is a standard Brownian motion;
- $t \to N_t$ is a Poisson process with intensity 1;
- $\{\zeta_j\}_{j \geq 1}$ are i.i.d. with density $f_\zeta : \mathbb{R} \to \mathbb{R}^+$;
- The triplet $(\{W_t\}, \{N_t\}, \{\xi_j\})$ are mutually independent;
- $\Sigma_t := \int_0^t \sigma_s^2 ds$ and $\Lambda_t := \int_0^t \lambda_s ds$ with positive continuous deterministic functions $t \to \sigma_t$ and $t \to \lambda_t$;

2. Finite-jump activity Lévy model ($\alpha_t \equiv \alpha, \sigma_t \equiv \sigma, \lambda_t \equiv \lambda$):

$$X_t = \alpha t + W_{\sigma^2 t} + \sum_{j=1}^{N_{\Lambda_t}} \zeta_j.$$
Statistical Problems

Given a discrete record of observations,

\[ X_{t_0}, X_{t_1}, \ldots, X_{t_n} \quad (\pi : 0 = t_0 < t_1 < \cdots < t_n = T), \]

from the process during a fixed finite time-horizon \([0, T]\), the following problems are of interest:

1. Estimate the integrated variance (or quadratic variation):

\[ \Sigma_T = \int_0^T \sigma_t^2 \, dt \]

2. Estimate the jump features of the process:
   - Jump times \( \tau_1 < \tau_2 < \cdots < \tau_{N_{\Lambda T}} \)
   - Jump sizes \( \zeta_1 < \zeta_2 < \cdots < \zeta_{N_{\Lambda T}} \)

3. Jump detection during a given time interval \([s, t] \subset [0, T]\)
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV(X)_\pi := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2, \quad (\pi : 0 = t_0 < \cdots < t_n = T). \]

Under quite general conditions:

\[ QV(X)_\pi \to \Sigma_T + \sum_{j=1}^{N_{\Lambda_T}} \zeta_j^2, \quad \text{when} \quad \text{mesh}(\pi) := \max_i \{t_i - t_{i-1}\} \to 0. \]

1. Realized Multipower Variations (Barndorff-Nielsen and Shephard (2004)):

\[ BPV(X)_\pi := \sum_{i=0}^{n-1} |X_{t_{i+1}} - X_{t_i}||X_{t_{i+2}} - X_{t_{i+1}}|, \]

\[ RMPV(X)^{(r_1, \ldots, r_k)}_\pi := \sum_{i=0}^{n-k} |X_{t_{i+1}} - X_{t_i}|^{r_1} \cdots |X_{t_{i+k}} - X_{t_{i+k-1}}|^{r_k}. \]

2. Threshold Realized Variations (Mancini (2003)):

\[ TRV(X)[B]_\pi := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 1\{|X_{t_{i+1}} - X_{t_i}| \leq B_\pi\}, \quad (B_\pi \in (0, \infty)). \]
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV(X)_\pi := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2, \quad (\pi: 0 = t_0 < \cdots < t_n = T). \]

Under quite general conditions:

\[ QV(X)_\pi \longrightarrow \Sigma_T + \sum_{j=1}^{N^\top} \zeta_j^2, \quad \text{when} \quad \text{mesh}(\pi) := \max_i \{t_i - t_{i-1}\} \to 0. \]

1. Realized Multipower Variations (Barndorff-Nielsen and Shephard (2004)):

\[ BPV(X)_\pi := \sum_{i=0}^{n-1} |X_{t_{i+1}} - X_{t_i}| X_{t_{i+2}} - X_{t_{i+1}} |, \]

\[ RMPV(X)^{(r_1, \ldots, r_k)}_\pi := \sum_{i=0}^{n-k} |X_{t_{i+1}} - X_{t_i}|^{r_1} \cdots |X_{t_{i+k}} - X_{t_{i+k-1}}|^{r_k}. \]

2. Threshold Realized Variations (Mancini (2003)):

\[ TRV(X)[B]_\pi := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 1\{|X_{t_{i+1}} - X_{t_i}| \leq B_\pi\}, \quad (B_\pi \in (0, \infty)). \]
Two main classes of estimators

Precursor. Realized Quadratic Variation:

\[ QV(X)_\pi := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2, \quad (\pi : 0 = t_0 < \cdots < t_n = T). \]

Under quite general conditions:

\[ QV(X)_\pi \longrightarrow \Sigma_T + \sum_{j=1}^{N^\Lambda_T} \zeta_j^2, \quad \text{when} \quad \text{mesh(\pi)} := \max_i \{t_i - t_{i-1}\} \to 0. \]

1 Realized Multipower Variations (Barndorff-Nielsen and Shephard (2004)):

\[ BPV(X)_\pi := \sum_{i=0}^{n-1} |X_{t_{i+1}} - X_{t_i}| |X_{t_{i+2}} - X_{t_{i+1}}|, \]

\[ RMPV(X)^{(r_1, \ldots, r_k)}_\pi := \sum_{i=0}^{n-k} |X_{t_{i+1}} - X_{t_i}|^{r_1} \cdots |X_{t_{i+k}} - X_{t_{i+k-1}}|^{r_k}. \]

2 Threshold Realized Variations (Mancini (2003)):

\[ TRV(X)[B]_\pi := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \mathbf{1}\{ |X_{t_{i+1}} - X_{t_i}| \leq B_\pi \}, \quad (B_\pi \in (0, \infty)). \]
Advantages and Drawbacks

1. **Multipower Realized Variations (MPV)**
   - Easy to implement (minimal parameter tune up);
   - Exhibit “high” bias in the presence of jumps:
     \[
     \mathbb{E} \left[ RMPV(X)^{(r_1, \ldots, r_k)}_{\pi} \right] - C_r \Sigma_T \sim \frac{TC'_r}{n^{1 - \frac{1}{2} \max_i r_i}}, \quad \text{with} \quad C_r = \prod_{i=1}^{k} \mathbb{E}|Z_i|^{r_i} r_1 + \cdots + r_k = 2.
     \]

2. **Threshold Realized Variations (TRV):**
   - Can also be adapted for estimating the jump features of the process:
     \[
     \hat{N}[B]_\pi := \sum_{i=0}^{n-1} 1 \{ |x_{t_{i+1}} - x_{t_i}| > B_\pi \}, \quad \hat{J}[B]_\pi := \sum_{i=0}^{n-1} (x_{t_{i+1}} - x_{t_i})^2 1 \{ |x_{t_{i+1}} - x_{t_i}| > B_\pi \}
     \]
   - Its performance strongly depends on a “good” choice of the threshold level \(B\); e.g., given a sequence \(\pi_n\) of sampling schemes with \(\text{mesh}(\pi_n) \to 0\),
     \[
     \text{TRV}(X)[B]_{\pi_n} \xrightarrow{L_2} \Sigma_T \iff B_{\pi_n} \to 0, \quad \frac{B_{\pi_n}}{\sqrt{\text{mesh}(\pi_n)}} \to \infty.
     \]
   - Thresholds proposed in the literature:
     \[
     B_\pi := \alpha \text{mesh}(\pi) \omega, \quad B_\pi = \alpha \text{mesh}(\pi)^{\frac{1}{2}} \Phi^{-1} (1 - \beta \text{mesh}(\pi)),
     \]
Advantages and Drawbacks

1. Multipower Realized Variations (MPV)
   - Easy to implement (minimal parameter tune up);
   - Exhibit “high” bias in the presence of jumps:

   \[
   \mathbb{E} \left[ RMPV(X)_{\pi}^{(r_1, \ldots, r_k)} \right] - C_r \Sigma_T \sim \frac{TC_r'}{n^{1 - \frac{1}{2} \max_i r_i}}, \quad \text{with} \quad C_r = \prod_{i=1}^{k} \mathbb{E} |Z_i|^{r_i}, \quad r_1 + \cdots + r_k = 2.
   \]

2. Threshold Realized Variations (TRV):
   - Can also be adapted for estimating the jump features of the process:

   \[
   \hat{N}[B]_{\pi} := \sum_{i=0}^{n-1} \mathbb{1}\{ |X_{ti+1} - X_{ti}| > B_{\pi} \}, \quad \hat{J}[B]_{\pi} := \sum_{i=0}^{n-1} (X_{ti+1} - X_{ti})^2 \mathbb{1}\{ |X_{ti+1} - X_{ti}| > B_{\pi} \}
   \]
   - Its performance strongly depends on a “good” choice of the threshold level \( B \); e.g., given a sequence \( \pi_n \) of sampling schemes with \( \text{mesh}(\pi_n) \to 0 \),

   \[
   \text{TRV}(X)[B]_{\pi_n} \xrightarrow{\mathbb{L}_2} \Sigma_T \iff B_{\pi_n} \to 0, \quad \frac{B_{\pi_n}}{\sqrt{\text{mesh}(\pi_n)}} \to \infty.
   \]
   - Thresholds proposed in the literature:

   \[
   B_{\pi} := \alpha \text{mesh}(\pi)^\omega, \quad B_{\pi} = \alpha \text{mesh}(\pi)^{\frac{1}{2}} \Phi^{-1} (1 - \beta \text{mesh}(\pi))
   \]
Advantages and Drawbacks

1. Multipower Realized Variations (MPV)
   - Easy to implement (minimal parameter tune up);
   - Exhibit “high” bias in the presence of jumps:
     \[
     \mathbb{E} \left[ RMPV(X)_{\pi}^{(r_1, \ldots, r_k)} \right] - C_r \Sigma_T \sim \frac{TC'_r}{n^{1-\frac{1}{2} \max_i r_i}}, \quad \text{with} \quad C_r = \prod_{i=1}^k \mathbb{E} |Z_i|^{r_i} r_1 + \cdots + r_k = 2.
     \]

2. Threshold Realized Variations (TRV):
   - Can also be adapted for estimating the jump features of the process:
     \[
     \hat{N}[B]_{\pi} := \sum_{i=0}^{n-1} \mathbf{1}\{ |X_{t_{i+1}} - X_{t_i}| > B_{\pi} \}, \quad \hat{J}[B]_{\pi} := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \mathbf{1}\{ |X_{t_{i+1}} - X_{t_i}| > B_{\pi} \}
     \]
   - Its performance strongly depends on a “good” choice of the threshold level \( B \); e.g., given a sequence \( \pi_n \) of sampling schemes with \( \text{mesh}(\pi_n) \to 0 \),
     \[
     TRV(X)[B]_{\pi_n} \xrightarrow{L_2} \Sigma_T \iff B_{\pi_n} \to 0, \quad \frac{B_{\pi_n}}{\sqrt{\text{mesh}(\pi_n)}} \to \infty.
     \]
   - Thresholds proposed in the literature:
     \[
     B_{\pi} := \alpha \text{mesh}(\pi)^{\omega}, \quad B_{\pi} = \alpha \text{mesh}(\pi)^{\frac{1}{2}} \Phi^{-1} (1 - \beta \text{mesh}(\pi))
     \]
Advantages and Drawbacks

1. Multipower Realized Variations (MPV)
   - Easy to implement (minimal parameter tune up);
   - Exhibit “high” bias in the presence of jumps:

\[
\mathbb{E} \left[ RMPV(X)^{(r_1, \ldots, r_k)}_{\pi} \right] - C_r \sum_T \sim \frac{TC'_r}{n^{1-\frac{1}{2} \max_i r_i}}, \quad \text{with} \quad C_r = \prod_{i=1}^k \mathbb{E}|Z_i|^{r_i} \quad r_1 + \cdots + r_k = 2.
\]

2. Threshold Realized Variations (TRV):
   - Can also be adapted for estimating the jump features of the process:

\[
\hat{N}[B]_{\pi} := \sum_{i=0}^{n-1} 1\left\{ |X_{t_{i+1}} - X_{t_i}| > B_{\pi} \right\}, \quad \hat{J}[B]_{\pi} := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 1\left\{ |X_{t_{i+1}} - X_{t_i}| > B_{\pi} \right\}
\]

   - Its performance strongly depends on a “good” choice of the threshold level \( B \); e.g., given a sequence \( \pi_n \) of sampling schemes with \( \text{mesh}(\pi_n) \to 0 \),

\[
TRV(X)[B]_{\pi_n} \xrightarrow{L^2} \Sigma_T \iff B_{\pi_n} \to 0, \quad \frac{B_{\pi_n}}{\sqrt{\text{mesh}(\pi_n)}} \to \infty.
\]

   - Thresholds proposed in the literature:

\[
B_{\pi} := \alpha \text{mesh}(\pi)^\omega, \quad B_{\pi} = \alpha \text{mesh}(\pi)^{\frac{1}{2}} \Phi^{-1} (1 - \beta \text{mesh}(\pi))
\]
Advantages and Drawbacks

1. Multipower Realized Variations (MPV)
   - Easy to implement (minimal parameter tune up);
   - Exhibit “high” bias in the presence of jumps:
     \[
     \mathbb{E} \left[ RMPV(X)_{\pi}^{(r_1, \ldots, r_k)} \right] - C_r \Sigma_T \sim \frac{TC_r'}{n^{1-\frac{1}{2} \max_i r_i}}, \quad \text{with} \quad C_r = \prod_{i=1}^{k} \mathbb{E} |Z_i|^{r_i} r_1 + \cdots + r_k = 2.
     \]

2. Threshold Realized Variations (TRV):
   - Can also be adapted for estimating the jump features of the process:
     \[
     \hat{N}[B]_{\pi} := \sum_{i=0}^{n-1} \mathbf{1}\{|x_{t_{i+1}} - x_{t_i}| > B_{\pi}\}, \quad \hat{J}[B]_{\pi} := \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 \mathbf{1}\{|x_{t_{i+1}} - x_{t_i}| > B_{\pi}\}
     \]
   - Its performance strongly depends on a “good” choice of the threshold level \(B\); e.g., given a sequence \(\pi_n\) of sampling schemes with \(\text{mesh}(\pi_n) \to 0\),
     \[
     TRV(X)[B]_{\pi_n} \xrightarrow{L^2} \Sigma_T \iff B_{\pi_n} \to 0, \quad \frac{B_{\pi_n}}{\sqrt{\text{mesh}(\pi_n)}} \to \infty.
     \]
   - Thresholds proposed in the literature:
     \[
     B_{\pi} := \alpha \text{mesh}(\pi)^{1/2}, \quad B_{\pi} = \alpha \text{mesh}(\pi)^{1/2} \Phi^{-1} (1 - \beta \text{mesh}(\pi))
     \]
Figure: Box Plots of several estimators based on 2500 MC simulations with $T = 1$ year, $\text{mesh} = 5$ min sample observations. Parameters: $\sigma = 0.3$, $\lambda = 20$, $\zeta_j \sim \mathcal{N}(-0.1, 0.1^2)$. 

Numerical illustration
Outline

1 Continuous-time stochastic modeling
   Brownian Motion
      Black-Scholes framework
   Poisson Processes
   Compound Poisson Processes
      Merton and Kou Models
   Varying volatility and jump intensity

2 Semiparametric estimation of Lévy Jump-Diffusion Models
   The Statistical Problems and the Main Estimators
   Optimally Thresholded Power Estimators
   Main Results
   Conclusions
Optimal Threshold Realized Estimators

Aims

• Develop a well-posed “optimal” selection criterion for the threshold $B$, that minimizes a suitable loss function of estimation.
• Characterize the optimal threshold $B^*$ asymptotically when $\text{mesh}(\pi) \to 0$.
• Develop a feasible implementation method for the optimal threshold sequence.

Assumptions

• Finite activity Lévy model: $X_t = \alpha t + W_{\sigma^2 t} + \sum_{j=1}^{N_{\xi t}} \zeta_j, \zeta_j \overset{\text{i.i.d.}}{\sim} f_{\zeta}$.
• Regular sampling scheme with mesh $h_n := \frac{1}{n}$; i.e., $\pi : t_i = \frac{i}{n}$.
• The jump density function $f_{\zeta}$ takes the mixture form:
  
  $f_{\zeta}(x) = p f_+(x) \mathbf{1}_{\{x \geq 0\}} + q f_-(-x) \mathbf{1}_{\{x < 0\}}$

  with $p + q = 1$ and $f_\pm : [0, \infty) \to \mathbb{R}_+ \in C_1_b(0, \infty)$.

Notation:

• $C(f_{\zeta}) := p f_+(0) + q f_-(0)$.
• $\Phi$ and $\phi$ are the cdf and pdf of a standard Normal variable, respectively.
• $\Delta^n_i X := \Delta_i X := X_{t_i} - X_{t_{i-1}}$, $\Delta^n_i N := \Delta_i N := N_{t_i} - N_{t_{i-1}}$.
Optimal Threshold Realized Estimators

1 Aims

- Develop a well-posed "optimal" selection criterion for the threshold $B$, that minimizes a suitable loss function of estimation.
- Characterize the optimal threshold $B^*$ asymptotically when $\text{mesh}(\pi) \to 0$.
- Develop a feasible implementation method for the optimal threshold sequence.

2 Assumptions

- Finite activity Lévy model: $X_t = \alpha t + W_{\sigma^2 t} + \sum_{j=1}^{N_t} \zeta_j$, $\zeta_j \overset{i.i.d.}{\sim} f_\zeta$.
- Regular sampling scheme with mesh $h_n := \frac{1}{n}$; i.e., $\pi : t_i = \frac{i}{n}$.
- The jump density function $f_\zeta$ takes the mixture form:
  
  $f_\zeta(x) = pf_+(x)1_{\{x \geq 0\}} + qf_-(-x)1_{\{x < 0\}}$

  with $p + q = 1$ and $f_\pm : [0, \infty) \to \mathbb{R}^+ \in C^1_b(0, \infty)$.

3 Notation:

- $C(f_\zeta) := pf_+(0) + qf_-(-0)$.
- $\Phi$ and $\phi$ are the cdf and pdf of a standard Normal variable, respectively.
- $\Delta^n_i X := \Delta_i X := X_{t_i} - X_{t_{i-1}}$, $\Delta^n_i N := \Delta_i N := N_{t_i} - N_{t_{i-1}}$.
Optimal Threshold Realized Estimators

1 Aims

- Develop a well‐posed “optimal” selection criterion for the threshold $B$, that minimizes a suitable loss function of estimation.
- Characterize the optimal threshold $B^*$ asymptotically when mesh $(\pi) \to 0$.
- Develop a feasible implementation method for the optimal threshold sequence.

2 Assumptions

- Finite activity Lévy model: $X_t = \alpha t + W_{\sigma^2 t} + \sum_{j=1}^{N_{t\lambda t}} \zeta_j$, $\zeta_j \sim \text{i.i.d.} f_\zeta$.
- Regular sampling scheme with mesh $h_n := \frac{1}{n}$; i.e., $\pi : t_i = \frac{i}{n}$.
- The jump density function $f_\zeta$ takes the mixture form:
  $$f_\zeta(x) = pf_+(x)1_{\{x \geq 0\}} + qf_-(−x)1_{\{x < 0\}}$$
  with $p + q = 1$ and $f_\pm : [0, \infty) \to \mathbb{R}_+ \in C^1_b(0, \infty)$.

3 Notation:

- $C(f_\zeta) := pf_+(0) + qf_-(0)$.
- $\Phi$ and $\phi$ are the cdf and pdf of a standard Normal variable, respectively.
- $\Delta_n^X := \Delta_i X := X_{t_i} - X_{t_{i−1}}$, $\Delta_n^N := \Delta_i N := N_{t_i} - N_{t_{i−1}}$.
Loss Functions

1. Natural Loss Function

\[ \text{Loss}_{n}^{(1)}(B) := \mathbb{E} \left[ | TRV(X)[B]_n - T\sigma^2 |^2 \right] + \mathbb{E} \left[ | \hat{N}[B]_n - N_{\lambda T} |^2 \right] . \]

2. Alternative Loss Function

\[ \text{Loss}_{n}^{(2)}(B) := \mathbb{E} \left( \sum_{i=1}^{\lfloor nT \rfloor} \left( 1_{|\Delta^\eta_i X| > B, \Delta^\eta_i N = 0} + 1_{|\Delta^\eta_i X| \leq B, \Delta^\eta_i N \neq 0} \right) \right) . \]

3. Interpretation

- \( \text{Loss}_{n}^{(1)}(B) \) favors sequences that minimizes the estimation errors of both the continuous and the jump component.
- \( \text{Loss}_{n}^{(2)}(B) \) favors sequences that minimizes the total number of miss-classifications; i.e., flagging the presence of a jump when there is no jump, and failing to detect a jump when there is a jump.
- \( \text{Loss}_{n}^{(2)}(B) \) is much more tractable than \( \text{Loss}_{n}^{(1)}(B) \).
Outline

1. Continuous-time stochastic modeling
   - Brownian Motion
     - Black-Scholes framework
   - Poisson Processes
   - Compound Poisson Processes
     - Merton and Kou Models
   - Varying volatility and jump intensity

2. Semiparametric estimation of Lévy Jump-Diffusion Models
   - The Statistical Problems and the Main Estimators
   - Optimally Thresholded Power Estimators
   - Main Results
   - Conclusions
Asymptotic Comparison of Loss Functions

Theorem (FL & Nisen (2013))

Given a threshold sequence \((B_n)\) satisfying \(B_n \to 0, \ B_n\sqrt{n} \to \infty\), there exists a positive sequence \((C_n)_n\), with \(\lim_{n \to \infty} C_n = 0\), such that

\[
\text{Loss}_n^{(2)}(B) + R_n(B) \leq \text{Loss}_n^{(1)}(B) \leq (1 + C_n(B))\text{Loss}_n^{(2)}(B) + R_n(B) + \overline{R}_n(B),
\]

where, as \(n \to \infty\),

\[
R_n(B) \sim \frac{T \lambda^2}{2n} + T^2 \left( \frac{2\sigma \sqrt{n}}{B_n} \phi \left( \frac{\sqrt{n}B_n}{\sigma} \right) - \lambda B_n C(f) \right)^2,
\]

\[
\overline{R}_n(B) \sim \left[ \frac{6 T \sigma^4}{n} + 3B_n^6 T^2 \lambda^2 C(f)^2 \right].
\]

Furthermore,

\[
\lim_{n \to \infty} \frac{\inf_{B>0} \text{Loss}_n^{(1)}(B)}{\inf_{B>0} \text{Loss}_n^{(2)}(B)} = 1. \tag{1}
\]
Well-posedness and asymptotic characterization

Theorem (FL & Nisen (2013))

There exists an $N \in \mathbb{N}$ such that for all $n \geq N$, the loss function $Loss_n^{(2)}(B)$ is quasi-convex and possesses a unique global minimum $B_n^*$:

$$B_n^* := \arg \inf_{B > 0} Loss_n^{(2)}(B).$$

Furthermore, the optimal threshold sequence $(B_n^*)_n$ is such that

$$B_n^* = \sqrt{\frac{3\sigma^2 \ln(n)}{n}} + o \left( \sqrt{\ln(n)/n} \right), \quad (n \to \infty).$$
Remarks

1. The leading term of the optimal threshold is proportional to the Lévy modulus of Brownian motion:

\[ \limsup_{h \to 0} \frac{1}{\sqrt{2h \ln(1/h)}} \sup_{|t-s|<h, s,t \in [0,1]} |W_t - W_s| = 1, \quad \text{a.s.} \]

2. The leading term of the optimal threshold,

\[ B_n^{*,1} := \sqrt{\frac{3\sigma^2 \ln(n)}{n}}, \]

give us a “blueprint” to “fabricate” a suitable threshold level.

3. Merton Model: \( \zeta \sim \mathcal{N}(0, \delta^2) \). With \( \delta_n^2 := \frac{\sigma^2}{n} + \delta^2 \),

\[ (B_n^*)^2 = \frac{3\sigma^2 \ln(n)}{n} - \frac{2\sigma^2 \ln \left( \frac{\sigma \lambda}{\delta_n} \right)}{n} + \frac{3\sigma^4 \ln(n)}{n^2 \delta^2} - \frac{2\sigma^4 \ln \left( \frac{\sigma \lambda}{\delta_n} \right)}{n^2 \delta^2}. \]

4. The performance of the leading term (compared to the optimal threshold) will depend on the quantity: \( \frac{\sigma \lambda}{\delta} \).
A Feasible Iterative Algorithm to Find $B^*_n$

1. **Key Issue:** The optimal threshold $B^*$ would allow us to find an optimal estimate $\hat{\sigma}$ for $\sigma^2$ of the form

$$\hat{\sigma}^2 := \frac{1}{T} TRV(X)[B^*(\sigma^2)]_n,$$

but $B^*$ precisely depends on $\sigma^2$.

2. The previous issue suggests a fixed-point type of implementation:

Set $\hat{\sigma}_{n,0}^2 := \frac{1}{T} \sum_{i=1}^{\lfloor nT \rfloor} |X_{t_i} - X_{t_{i-1}}|^2$ and $\hat{B}_{n,0}^* := \left( \frac{3 \hat{\sigma}_{n,0}^2 \ln(n)}{n} \right)^{1/2}$

while $\hat{\sigma}_{n,k-1}^2 > \hat{\sigma}_{n,k}^2$ do

$$\hat{\sigma}_{n,k+1}^2 \leftarrow \frac{1}{T} TRV(X)[\hat{B}_{n,k}^*]_n \text{ and } \hat{B}_{n,k+1}^* \leftarrow \left( \frac{3 \hat{\sigma}_{n,k+1}^2 \ln(n)}{n} \right)^{1/2}$$

end while

Let $k^*_n := \inf \left\{ k \geq 1 : \hat{\sigma}_{n,k+1}^2 = \hat{\sigma}_{n,k}^2 \right\}$ and take $\hat{\sigma}_{n,k^*_n}^2$ as the final estimate for $\sigma$ and the corresponding $\hat{B}_{n,k^*_n}^*$ as an estimate for $B_n^*$.

3. The previous algorithm generates a non-increasing sequence of estimators $\{\hat{\sigma}_{n,k}^2\}_k$ and finish in finite time.
A numerical illustration

Merton Model: 4-year / 1-day
\[ \sigma = 0.3 \quad \lambda = 5 \]
\[ \mu = 0, \delta = 0.6 \]

<table>
<thead>
<tr>
<th>Method</th>
<th>( \overline{TRV} )</th>
<th>( S_{TRV} )</th>
<th>( \overline{Loss} )</th>
<th>( S_{Loss} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \overline{B}_{n,k}^* )</td>
<td>0.2985</td>
<td>0.0070</td>
<td>2.0588</td>
<td>1.4267</td>
</tr>
<tr>
<td>( Pow(0.495) )</td>
<td>0.2967</td>
<td>0.0066</td>
<td>2.2992</td>
<td>1.4972</td>
</tr>
<tr>
<td>( BF )</td>
<td>0.2983</td>
<td>0.0071</td>
<td>2.1756</td>
<td>1.4749</td>
</tr>
</tbody>
</table>

Table: Finite-sample performance of the threshold realized variation (TRV) estimators based on \( K = 5,000 \) sample paths for the Merton model \( \zeta_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \delta^2) \).

\( Loss \) represents the total number of Jump Misclassification Errors, while \( \overline{TRV}, \overline{Loss}, S_{TRV}, \) and \( S_{Loss} \) denote the corresponding sample means and standard deviations, respectively.
A numerical illustration

<table>
<thead>
<tr>
<th>Method</th>
<th>$\bar{TRV}$</th>
<th>$S_{TRV}$</th>
<th>Loss</th>
<th>$S_{Loss}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{B}_{n,k_n^*}$</td>
<td>0.5004</td>
<td>0.0186</td>
<td>0.2232</td>
<td>0.4706</td>
</tr>
<tr>
<td>$Pow(0.495)$</td>
<td>0.4407</td>
<td>0.0142</td>
<td>13.5302</td>
<td>3.6392</td>
</tr>
<tr>
<td>$BF$</td>
<td>0.4917</td>
<td>0.0193</td>
<td>1.180</td>
<td>1.0775</td>
</tr>
</tbody>
</table>

Table: Finite-sample performance of the threshold realized variation (TRV) estimators based on $K = 5,000$ sample paths for the Kou model:

$$f_{Kou}(x) = \frac{p}{\beta_+} e^{-x/\beta_+} 1_{[x \geq 0]} + \frac{(1-p)}{\beta_-} e^{-|x|/\beta_-} 1_{[x < 0]}.$$  

$Loss$ represents the total number of Jump Misclassification Errors, while $\bar{TRV}$, $Loss$, $S_{TRV}$, and $S_{Loss}$ denote the corresponding sample means and standard deviations, respectively.
A numerical illustration

(S3) Kou Model: 1-year / 5-minute

\[ \sigma = 0.4 \quad \lambda = 1000 \quad p = 0.5, \beta_+ = \beta_- = 0.1 \]

<table>
<thead>
<tr>
<th>Method</th>
<th>TRV</th>
<th>STRV</th>
<th>Loss</th>
<th>S_{Loss}</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{B}_{n,k_n^*})</td>
<td>0.4039</td>
<td>0.0028</td>
<td>139.6776</td>
<td>12.2193</td>
</tr>
<tr>
<td>Pow</td>
<td>0.3767</td>
<td>0.0019</td>
<td>230.0170</td>
<td>15.0308</td>
</tr>
<tr>
<td>BF</td>
<td>0.6495</td>
<td>0.0315</td>
<td>375.5850</td>
<td>24.3999</td>
</tr>
</tbody>
</table>

**Table:** Finite-sample performance of the threshold realized variation (TRV) estimators based on \(K = 5,000\) sample paths for the Kou model:

\[ f_\zeta(x) = \frac{p}{\beta_+} e^{-x/\beta_+} \mathbf{1}_{[x \geq 0]} + \frac{q}{\beta_-} e^{-|x|/\beta_-} \mathbf{1}_{[x < 0]} \cdot \]

**Loss** represents the total number of Jump Misclassification Errors, while \(\overline{TRV}\), \(\overline{Loss}\), \(\overline{STRV}\), and \(\overline{S_{Loss}}\) denote the corresponding sample means and standard deviations, respectively.
Additive Processes

1. The model

\[ X_s := \int_0^s \alpha(u) du + W_{\Sigma_s} + \sum_{j=1}^{N_{\Lambda_s}} \zeta_j =: X^c_s + J_s, \]

where \((N_s)_{s \geq 0} \sim \text{Poiss}(1)\), independent of \(W\), and \(\Sigma_s := \int_0^s \sigma_u^2 du\) and \(\Lambda_s := \int_0^s \lambda u du\) with some deterministic smooth functions \(\sigma, \lambda : [0, \infty) \to (0, \infty)\) and \(\alpha : [0, \infty) \to \mathbb{R}\).

2. Optimal Threshold Problem

Given a sampling scheme \(\pi : t_0 < \cdots < t_n = T\), determine the vector \(\vec{B}_{\pi,*} = (\vec{B}_{t_1,*}, \ldots, \vec{B}_{t_n,*})\) that minimizes the problem

\[
\inf_{\vec{B} = (B_{t_1}, \ldots, B_{t_n}) \in \mathbb{R}_+^m} \mathbb{E} \sum_{i=1}^n \left( 1[|X_{t_i} - X_{t_i-1}| > B_{t_i}, N_{t_i} - N_{t_i-1} = 0] + 1[|X_{t_i} - X_{t_i-1}| \leq B_{t_i}, N_{t_i} - N_{t_i-1} \neq 0] \right)
\]

\[= \sum_{i=1}^n \inf_{B_{t_i}} \{\mathbb{P}(|\Delta_i X| > B_{t_i}, \Delta_i N = 0) + \mathbb{P}(|\Delta_i X| \leq B_{t_i}, \Delta_i N \neq 0)\},\]

\[(\Delta_i X := X_{t_i} - X_{t_{i-1}}, \quad \Delta_i N := N_{t_i} - N_{t_{i-1}})\]
Optimal Threshold Spot Volatility Estimation

Notation: $h_i = t_i - t_{i-1}$ (Mesh), $K_\theta(t) = \frac{1}{\theta} K(t/\theta)$ (Kernel), $\theta = $ Bandwidth

Algorithm:

For each $i \in \{1, 2, \ldots, n\}$, set

$$\hat{\sigma}^2_0(t_i) := \sum_{j=-\ell}^{\ell} \frac{1}{h_{i+j}} |\triangle_{i+j}X|^2 K_\theta(t_i - t_{i+j})$$

and

$$\hat{B}_{t_i}^{*,0} := \left[3\hat{\sigma}^2_0(t_i)h_i \ln(1/h_i)\right]^{1/2}$$

while there exists $i \in \{1, 2, \ldots, m\}$ such that $\hat{\sigma}^2_k(t_i) > \hat{\sigma}^2_{k-1}(t_i)$ do

$$\hat{\sigma}^2_{k+1}(t_i) \leftarrow \sum_{j=-\ell}^{\ell} \frac{1}{h_{i+j}} |\triangle_{i+j}X|^2 \mathbf{1}_{[|\triangle_{i+j}X| \leq \hat{B}_{t_i}^{*,k}]} K_\theta(t_i - t_{i+j})$$

and

$$\hat{B}_{t_i}^{*,k+1} \leftarrow \left[3\hat{\sigma}^2_{k+1}(t_i)h_i \ln(1/h_i)\right]^{1/2}$$

end while

Let $k^*(\pi) := \inf \{k \geq 1 : \hat{\sigma}^2_{k+1}(t_i) = \hat{\sigma}^2_k(t_i); \text{ for all } i = 1, 2, \ldots, n\}$ and take $\hat{\sigma}^2_{k^*_m}(t_i)$ as the final estimate for $\sigma(t_i)$ and the corresponding $\hat{B}_{t_i}^{*,k^*_m}$ as an estimate for $B_{t_i}^*$. 

The previous algorithm generates a non-increasing sequence of estimators $\{\hat{\sigma}^2_k(t_i)\}_{k,i}$ and finish in finite time.
Figure: Spot Volatility Estimation using Adaptive Kernel Weighted Realized Volatility.
Figure: Spot Volatility Estimation using Adaptive Kernel Weighted Realized Volatility.
Figure: Spot Volatility Estimation using Adaptive Kernel Weighted Realized Volatility.
Numerical Illustration

Figure: Spot Volatility Estimation using Adaptive Kernel Weighted Realized Volatility. 50 simulations
Outline

1 Continuous-time stochastic modeling
   Brownian Motion
     Black-Scholes framework
   Poisson Processes
   Compound Poisson Processes
     Merton and Kou Models
   Varying volatility and jump intensity

2 Semiparametric estimation of Lévy Jump-Diffusion Models
   The Statistical Problems and the Main Estimators
   Optimally Thresholded Power Estimators
   Main Results
   Conclusions
Conclusions

1. Introduced an objective threshold selection procedure based on statistical optimality reasoning via a well-posed optimization problem.
2. Characterized precisely the infill asymptotic behavior of the optimal threshold sequence.
3. Proposed an iterative algorithm to find the optimal threshold sequence.
4. Extended the approach to more general stochastic models, which allows time-varying volatility and jump intensity.
For Further Reading I

- O.E. Barndorff-Nielsen and N. Shephard.  
  Power and bipower variation with stochastic volatility and jumps. 

- Figueroa-López & Nisen.  
  Optimally Thresholded Realized Power Variations for Lévy Jump Diffusion Models  
  Available at www.stat.purdue.edu/~figueroa.

- C. Mancini  
  Estimation of the characteristics of the jumps of a general Poisson-diffusion model 
For Further Reading II

O.E. Barndorff-Nielsen, N. Shephard, and M. Winkel
Limit theorems for multi-power variation in the presence of jumps

C. Mancini
Non-parametric threshold estimation for models w/ stochastic diffusion coefficient

R. Cont and P. Tankov.
Financial modelling with jump processes.