Semi-Markov model for market microstructure and HF trading

Huyên PHAM

LPMA, University Paris Diderot
and JVN Institute, VNU, Ho-Chi-Minh City

NUS-UTokyo Workshop on Quantitative Finance
Singapore, 26-27 september 2013

Joint work with Pietro FODRA
EXQIM and LPMA, University Paris Diderot
Financial data modelling

- Continuous time price process \( (P_t)_t \) over \([0, T]\) observed at \( P_0, P_{\tau}, \ldots, P_{n\tau} \)

- **Different modelling** of \( P \) according to scales \( \tau \) and \( T \):
  - Macroscopic scale (hourly, daily observation data): **Itô semimartingale**
  - Microscopic scale (tick data) → **High frequency**
Euribor contract, 2010, for different observation scales
Stylized facts on HF data

Microstructure effects

- Discreteness of prices: jump times and prices variations (tick data)
- **Mean-reversion**: negative autocorrelation of consecutive variation prices
- Irregular spacing of jump times: clustering of trading activity
Limit order book

- Most of modern equities exchanges organized through a mechanism of *Limit Order Book* (LOB):

![Graph of Limit Order Book](image)

**Figure**: Instantaneous picture of a LOB
Two main streams in literature:

- **Models of intra-day asset price**
  - Latent process approach: Gloter and Jacod (01), Ait Sahalia, Mykland and Zhang (05), Robert and Rosenbaum (11), etc
  - Point process approach: Bauwens and Hautsch (06), Cont and de Larrard (10), Bacry et al. (11), Abergel, Jedidi (11)
  → Sophisticated models intended to reproduce microstructure effects, often for purpose of volatility estimation

- **High frequency trading problems**
  - Liquidation and market making in a LOB: Almgren, Cris (03), Alfonsi and Schied (10, 11), Avellaneda and Stoikov (08), etc
  → Stochastic control techniques for optimal trading strategies based on classical models of asset price (arithmetic or geometric Brownian motion, diffusion models)
Objective

• Make a “bridge” between these two streams of literature:
  ▶ Construct a “simple” model for asset price in Limit Order Book (LOB)
    • realistic: captures main stylized facts of microstructure
      • Diffuses on a macroscopic scale
      • Easy to estimate and simulate
    • tractable (simple to analyze and implement) for dynamic optimization problem in high frequency trading

→ Markov renewal and semi-Markov model approach
References

- P. Fodra and H. Pham (2013b): “High frequency trading in a Markov renewal model”.
Model-free description of asset mid-price (constant bid-ask spread)

**Marked point process**

Evolution of the univariate mid price process \((P_t)\) determined by:

- **The timestamps** \((T_k)_k\) of its jump times \(\iff\) \(N_t\) counting process: \(N_t = \inf\{n : \sum_{k=1}^n T_k \leq t\}\): modeling of **volatility clustering**, i.e. presence of spikes in intensity of market activity.

- **The marks** \((J_k)_k\) valued in \(\mathbb{Z} \setminus \{0\}\), representing (modulo the tick size) the price increment at \(T_k\): modeling of the **microstructure noise** via mean-reversion of price increments.

\[
P_t = P_0 + 2\delta \sum_{k=1}^{N_t} J_k
\]
Semi-Markov model approach

Markov Renewal Process (MRP) to describe \((T_k, J_k)_k\).

- Largely used in reliability
- Independent paper by d’Amico and Petroni (13) using also semi Markov model for asset prices
Jump side modeling

For simplicity, we assume $|J_k| = 1$ (on data, this is true 99.9% of the times):

- $J_k$ valued in $\{+1, -1\}$: side of the jump (upwards or downwards)

$$J_k = J_{k-1}B_k$$

$(B_k)_k$ i.i.d. with law: $\mathbb{P}[B_k = \pm 1] = \frac{1 \pm \alpha}{2}$ with $\alpha \in [-1, 1)$.

$\leftrightarrow (J_k)_k$ irreducible Markov chain with symmetric transition matrix:

$$Q_\alpha = \begin{pmatrix}
\frac{1+\alpha}{2} & \frac{1-\alpha}{2} \\
\frac{1-\alpha}{2} & \frac{1+\alpha}{2}
\end{pmatrix}$$

**Remark**: arbitrary random jump size can be easily considered by introducing an i.i.d. multiplication factor in (1).
Mean reversion

• Under the stationary probability of $(J_k)_k$, we have:

$$\alpha = \text{correlation}(J_k, J_{k-1})$$

• Estimation of $\alpha$:

$$\hat{\alpha}_n = \frac{1}{n} \sum_{k=1}^{n} J_k J_{k-1}$$

$\rightarrow \alpha \simeq -87.5\%, \ (\text{Euribor3m}, \ 2010, \ 10h-14h)$

$\rightarrow \text{Strong mean reversion of price returns}$
Timestamp modeling

Conditionally on \( \{J_k J_{k-1} = \pm 1\} \), the sequence of inter-arrival jump times \( \{S_k = T_k - T_{k-1}\} \) is i.i.d. with distribution function \( F_{\pm} \) and density \( f_{\pm} \):

\[
F_{\pm}(t) = \mathbb{P}[S_k \leq t | J_k J_{k-1} = \pm 1].
\]

Remarks

• The sequence \( (S_k)_k \) is (unconditionally) i.i.d with distribution:

\[
F = \frac{1 + \alpha}{2} F_+ + \frac{1 - \alpha}{2} F_-.
\]

• \( h_+ = \frac{1+\alpha}{2} \frac{f_+}{1-F} \) is the intensity function of price jump in the same direction, \( h_- = \frac{1-\alpha}{2} \frac{f_-}{1-F} \) is the intensity function of price jump in the opposite direction.
Non parametric estimation of jump intensity

Figure: Estimation of $h_{\pm}$ as function of the renewal quantile
Simulated price

**Figure : 30 minutes simulation**

**Figure : 1 day simulation**
Diffusive behavior at macroscopic scale

Scaling:

\[ P_t^{(T)} = \frac{P_{tT}}{\sqrt{T}}, \quad t \in [0, 1]. \]

**Theorem**

\[
\lim_{T \to \infty} P^{(T)} \overset{(d)}{=} \sigma_{\infty} W,
\]

where \( W \) is a Brownian motion, and \( \sigma^2 \) is the macroscopic variance:

\[
\sigma^2_{\infty} = \lambda \left( \frac{1 + \alpha}{1 - \alpha} \right).
\]

with \( \lambda^{-1} = \int_0^{\infty} tdF(t) \).
We consider the case of \textit{delayed renewal} process:

- $S_n \sim F$, $n \geq 1$, with finite mean $1/\lambda$, and $S_1 \sim$ density $\lambda(1-F)$

$\rightarrow$ Price process $P$ has \textit{stationary} increments

\begin{prop}
\begin{align*}
\tilde{V}(\tau) & := \frac{1}{\tau} \mathbb{E}[(P_{\tau} - P_0)^2] = \sigma^2 + \left(\frac{-2\alpha}{1-\alpha}\right) \frac{1 - G_\alpha(\tau)}{(1-\alpha)\tau}, \\
\end{align*}
\end{prop}

where $G_\alpha(t) = \mathbb{E}[\alpha^N_t]$ is explicitly given via its Laplace-Stieltjes transform $\hat{G}_\alpha$ in terms of $\hat{F}(s) := \int_0^\infty e^{-st} dF(t)$.

$$\tilde{V}(\infty) = \sigma_{\infty}^2,$$

and $\tilde{V}(0^+) = \lambda$.

\textbf{Remark}: Similar expression as in Robert and Rosenbaum (09) or Bacry et al. (11).
Figure : Mean signature plot for $\alpha < 0$
Markov embedding of price process

- Define the **last price jump direction**:

\[ I_t = J_{N_t}, \quad t \geq 0, \quad \text{valued in } \{+1, -1\} \]

and the **elapsed time** since the last jump:

\[ S_t = t - \sup_{T_k \leq t} T_k, \quad t \geq 0. \]

- Then the price process \((P_t)\) valued in \(2\delta \mathbb{Z}\) is embedded in a Markov process with three **observable** state variables \((P_t, I_t, S_t)\) with generator:

\[
\mathcal{L} \varphi(p, i, s) = \frac{\partial \varphi}{\partial s} + h_+(s)[\varphi(p + 2\delta i, i, 0) - \varphi(p, i, s)] \\
+ h_-(s)[\varphi(p - 2\delta i, -i, 0) - \varphi(p, i, s)],
\]
Trading issue

Problem of an agent (market maker) who submits limit orders on both sides of the LOB: limit buy order at the best bid price and limit sell order at the best ask price, with the aim to gain the spread.

▶ We need to model the market order flow, i.e. the counterpart trade of the limit order
Market trades

- A market order flow is modelled by a marked point process $(\theta_k, Z_k)_k$:
  - $\theta_k$: arrival time of the market order $\leftrightarrow M_t$ counting process
  - $Z_k$ valued in $\{-1, +1\}$: side of the trade.
    - $Z_k = -1$: trade at the best BID price (market sell order)
    - $Z_k = +1$: trade at the best ASK price (market buy order)

<table>
<thead>
<tr>
<th>index $n$</th>
<th>$\theta_k$</th>
<th>best ask</th>
<th>best bid</th>
<th>traded price</th>
<th>$Z_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9:00:01.123</td>
<td>98.47</td>
<td>98.46</td>
<td>98.47</td>
<td>+1</td>
</tr>
<tr>
<td>2</td>
<td>9:00:02.517</td>
<td>98.47</td>
<td>98.46</td>
<td>98.46</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>9:00:02.985</td>
<td>98.48</td>
<td>98.47</td>
<td>98.47</td>
<td>-1</td>
</tr>
</tbody>
</table>

- Dependence modeling between market order flow and price in LOB: Cox marked point process
Trade timestamp modeling

- The counting process \((M_t)\) of the market order timestamps \((\theta_k)_k\) is a Cox process with conditional intensity \(\lambda_M(S_t)\).

Examples of parametric forms reproducing intensity decay when \(s\) is large:

\[
\lambda^{\text{exp}}_M(s) = \lambda_0 + \lambda_1 s^r e^{-ks}
\]

\[
\lambda^{\text{power}}_M(s) = \lambda_0 + \frac{\lambda_1 s^r}{1 + s^k}.
\]

with positive parameters \(\lambda_0, \lambda_1, r, k\), estimated by MLE.
Strong and weak side of LOB

- We call **strong** side (+) of the LOB, the side in the **same direction** than the last jump, e.g. best ask when price jumped upwards.
- We call **weak** side (−) of the LOB, the side in the **opposite direction** than the last jump, e.g. best bid when price jumped upwards.

▶ We observe that trades (market order) arrive mostly on the weak side of the LOB.
Trade side modeling

- The trade sides are given by:

\[ Z_k = \Gamma_k I_{\theta_k^-}, \]

\((\Gamma_k)_k\) i.i.d. valued in \([-1, 1]\) with law:

\[ \mathbb{P}[\Gamma_k = \pm 1] = \frac{1 \pm \rho}{2} \]

for \(\rho \in [-1, 1]\).
Interpretation of $\rho$

$$\rho = \text{corr}(Z_k, l_{\theta_k^-})$$

- $\rho = 0$: market order flow arrive **independently** at best bid and best ask (usual assumption in the existing literature)
- $\rho > 0$: market orders arrive more often in the **strong** side of the LOB
- $\rho < 0$: market orders arrive more often in the **weak** side of the LOB

- Estimation of $\rho$: $\hat{\rho}_n = \frac{1}{n} \sum_{k=1}^{n} Z_k l_{\theta_k^-}$ leads to $\rho \simeq -50\%$: about 3 over 4 trades arrive on the weak side.

$\rho$ related to **adverse selection**
Market making strategy

- **Strategy control**: predictable process \((\ell_t^+, \ell_t^-)_t\) valued in \(\{0, 1\}\)
  - \(\ell_t^+ = 1\): limit order of fixed size \(L\) on the strong side: \(+I_t^-\)
  - \(\ell_t^- = 1\): limit order of fixed size \(L\) on the weak side: \(-I_t^-\)

- **Fees**: any transaction is subject to a fixed cost \(\varepsilon \geq 0\)

- **Portfolio process**:
  - **Cash** \((X_t)_t\) valued in \(\mathbb{R}\),
  - **inventory** \((Y_t)_t\) valued in a set \(\mathcal{Y}\) of \(\mathbb{Z}\)
Agent execution

- Execution of limit order occurs when:
  - A **market trade** arrives at $\theta_k$ on the **strong** (resp. **weak**) side if $Z_k l_{\theta_k} = +1$ (resp. $-1$), and with an executed quantity given by a distribution (price time priority/prorata) $\vartheta_+^L$ (resp. $\vartheta_-^L$) on $\{0, \ldots, L\}$
  - The **price jumps** at $T_k$ and crosses the limit order price

**Remark**

$\vartheta_\pm^L$ cannot be estimated on historical data. It has to be evaluated by a backtest with a zero intelligence strategy.

**Risks:**

- Inventory $\leftrightarrow$ price jump
- Adverse selection in market order trade
Market making optimization

- **Value function** of the market making control problem:

\[
v(t, s, p, i, x, y) = \sup_{(\ell^+, \ell^-)} \mathbb{E}[PNL_T - \text{CLOSE}(Y_T) - \eta \cdot RISK_{t,T}]
\]

where \(\eta \geq 0\) is the agent risk aversion and:

\[
P NL_t = X_t + Y_t \cdot P_t, \quad \text{ (ptf valued at the mid price)}
\]

\[
\text{CLOSE}(y) = -(\delta + \varepsilon) \cdot |y|, \quad \text{ (closure market order)}
\]

\[
RISK_{t,T} = \int_t^T Y_u^2 \cdot d[P]_u, \quad \text{ (no inventory imbalance)}
\]
Variable reduction to strong inventory and elapsed time

Theorem

The value function is given by:

\[ v(t, s, p, i, x, y) = x + yp + \omega_{yi}(t, s) \]

where \( \omega_q(t, s) = \omega(t, s, q) \) is the unique viscosity solution to the integro ODE:

\[
\left[ \partial_t + \partial_s \right] \omega + 2\delta(h^+ - h^-)q - 4\delta^2\eta(h^+ + h^-)q^2 \\
+ \max_{\ell \in \{0, 1\}, q - \ell \ell L \in \mathbb{Y}} \mathcal{L}_+^\ell \omega + \max_{\ell \in \{0, 1\}, q + \ell \ell L \in \mathbb{Y}} \mathcal{L}_-^\ell \omega = 0
\]

\[ \omega_q(T, s) = -|q|(\delta + \epsilon) \]

in \([0, T] \times \mathbb{R}_+ \times \mathbb{Y}\).
\[ \mathcal{L}_\pm^\ell = \mathcal{L}_{\pm,M}^\ell + \mathcal{L}_{\pm,jump}^\ell \]

- **favorable** execution of random size \( \leq L \) by market order
  \[ \mathcal{L}_{\pm,M}^\ell \omega := \lambda_{\pm,M}(s) \int \left[ \omega(t, s, q \mp k\ell) - \omega(t, s, q) \right. \\
  \left. + (\delta - \epsilon)k\ell \right] \nu_{L}^\pm(dk) \]

- **unfavorable** execution of maximal size \( L \) due to price jump
  \[ \mathcal{L}_{\pm,jump}^\ell \omega := h_{\pm}(s) \left[ \omega(t, 0, \pm q - L\ell) - \omega(t, s, q) + (-\delta - \epsilon)L\ell \right] \]

with
\[ \lambda_{\pm,M}(s) := \frac{1 \pm \rho}{2} \cdot \lambda_M(s) \quad (\text{trade intensities}) \]
Optimal policy shape: $\rho = 0$, execution probability = 10%
Optimal policy shape: $\rho = 0$, execution probability $= 5\%$
Optimal policy shape: $\rho = -0.33$, execution probability $= 5\%$
Concluding remarks

- **Markov renewal** approach for market microstructure
  - Easy to understand and simulate
  - **Non parametric estimation** based on i.i.d. sample data
  - Dependency between price return $J_k$ and jump time $T_k$
  - Reproduces well microstructure effects, diffuses on macroscopic scale
  - Markov embedding with observable state variables ($\neq$ Hawkes process approach)
  - Develop stochastic control algorithm for HF trading
    - MRP forgets correlation between inter-arrival jump times $\{S_k = T_k - T_{k-1}\}_k$

- Extension to multivariate price model
- Model with **market impact** for liquidation problem