Statistical Analysis of the Filtering Model for Financial Ultra-High Frequency (UHF) Data

Yong Zeng
zeng@mendota.umkc.edu

Department of Mathematics and Statistics
University of Missouri at Kansas City
Outline

- UHF data, and Marked Poisson Process (MPP)
- Lit. Review: Two Different Views of UHF data
  - An Irregularly-Spaced Time Series
  - A Realized Sample Path of Marked Point Process
- A New Model with two Equivalent Representations
  - Filtering with MPP (counting process) Observations
  - Random-arrival-time State Space Model
- Bayesian Inference via Filtering
  - Continuous-time Likelihoods, Posterior, and Bayes Factors
  - Filtering Equations and Evolution Equations for Bayes Factors
  - Two Computational Approaches and their Consistency
    - MC Approx. Method, nearly likelihood, and an example.
    - Particle Filtering (or Sequential Monte Carlo)
- Examples - Conclusions and Future Works
UHF Data and Marked Poisson Process

Two Characteristics of UHF Data
- Observations occur at varying random time intervals
- Market microstructure noises (frictions) are in price data

Marked Point Process and Marked Poisson Process
- Marked Point Process: \( \{(T_n, X_n)\} \), where \( \{T_n\} \) increasing, i.e. \( T_n \leq T_{n+1} \).
- Marked Poisson Process: \( \{(T_n, X_n)\} \) with \( \{T_n\} \) from a conditional Poisson Process (or Cox process, or doubly stochastic Poisson).
Durations or Inter-trading Times

Duration of Trade for MSFT (Jan. & Feb. 1994)

Duration (in seconds)

Density/Probability

0.00 0.01 0.02 0.03 0.04 0.05 0.06

0 100 200 300 400

Duration of US Treasury 10-year note

Total #
An Irregularly-Spaced Time Series

- Engle (2000) – Data: \( \{(\Delta t_i, y_i), i = 1, \ldots, N\} \) where \( \Delta t_i = t_i - t_{i-1} \).

- Conditional density:

\[
(\Delta t_i, y_i) | \mathcal{F}_{i-1} \sim f(\Delta t_i, y_i | \check{\Delta}t_{i-1}, \check{y}_{i-1}; \theta)
\]

where \( \check{z}_i = \{z_i, z_{i-1}, \ldots, z_1\} \).

\[
f(\Delta t_i, y_i | \check{\Delta}t_{i-1}, \check{y}_{i-1}; \theta) = g(\Delta t_i | \check{\Delta}t_{i-1}, \check{y}_{i-1}; \theta) q(y_i | \Delta t_i, \check{\Delta}t_{i-1}, \check{y}_{i-1}; \theta) \quad (1)
\]

- Log likelihood:

\[
l(\Delta, Y; \theta) = \sum_{i=1}^{N} \log g(\Delta t_i | \check{\Delta}t_{i-1}, \check{y}_{i-1}; \theta) + \sum_{i=1}^{N} \log q(y_i | \Delta t_i, \check{\Delta}t_{i-1}, \check{y}_{i-1}; \theta)
\]
Related Developments

- Bivariate Point process model by Engle and Lunde (2003), and Autoregressive Conditional Intensity model by Russell (1999)

The Second View:

A Realized Sample Path of MPP: Zeng (2003)
Construction of Price from Value

Three-Step Construction:

- **Intrinsic value process:** $X(t)$

**Assumption 1.1:** Markov process, $(\theta, X)$, is the solution of a martingale problem for a generator $A$ such that

$$M_f(t) = f(\theta(t), X(t)) - \int_0^t A f(\theta(s), X(s)) \, ds$$

is a $\mathcal{F}_t^{\theta,X}$-martingale.

- **Trading times:** $t_1, t_2, \ldots, t_i, \ldots$ follows a conditional Poisson process with $a(\theta(t), X(t), t)$.

- **Price at $t_i$:** $Z(t_i) = F(X(t_i)) = b_i \left( R[X(t_i) + V_i, \frac{1}{M}] \right)$

where $F$, a random transformation modeling market microstructure noise, is specified by the transition probability $p(Z(t_i) | X(t_i); \theta)$.

**Remark:** Closely related to (1) the structure models in Hasbrouck (1996) and (2) the TS Models for RV in ZMA (2005), Bandi and Russell (2006), Fan and Wang (2007), Li and Mykland (2007) and others.
Price Clustering in Treasury Notes

Price fraction of US Treasury 10-year note

Conference on the Econometrics of UHF Data in Finance, University of Waterloo, Mar. 7, 2008 – p. 8/50
A Collection of Counting Processes

About One-Half Day’s Transaction Data of Microsoft
A Collection of Counting Processes

About One-Half Day’s Transaction Data of Microsoft

\[
\bar{Y}(t) = \begin{pmatrix}
N_1(\int_0^t \lambda_1(\theta(s), X(s), s)ds) \\
N_2(\int_0^t \lambda_2(\theta(s), X(s), s)ds) \\
\vdots \\
N_n(\int_0^t \lambda_n(\theta(s), X(s), s)ds)
\end{pmatrix},
\]

where \( Y_j(t) = N_j(\int_0^t \lambda_j(\theta(s), X(s), s)ds) \) records the cumulative # of trades that have occurred at the \( j \)th price level up to time \( t \).
Other Assumptions of Model I

Filtering with counting process observations

- **Assumption 1.2:** \((N_1, \ldots, N_n)\) are unit Poisson processes under measure \(P\).
- **Assumption 1.3:** \((\theta, X), N_1, \ldots, N_n\) are independent under \(P\).
- **Assumption 1.4:** \(0 \leq a(\theta(t), X(t), t) \leq C\) for some \(C > 0\) and all \(\theta(t), X(t), t > 0\).
- **Assumption 1.5:** Intensities: \(\lambda_j(\theta, x, t) = a(\theta, x, t)p(y_j|x)\), where \(a(x, \theta, t)\) is the total trading intensity, and \(p_j = p(y_j|x)\) is the transition probability from \(x\) to \(y_j\).

**Signal:** \((\theta, X)\)     **Observation:** \(\vec{Y}\) or \(Z\).
Price and Order Flow

Price and cumulative log(order flow) of US Treasury 10-year note

Conférence on the Econometrics of UHF Data in Finance, University of Waterloo, Mar. 7, 2008 – p. 11/50
Price and Economic Announcements

10years 2nd Instrument

- No News
- CPI
- PPI
- UnEmployment
- FOMC

Conference on the Econometrics of UHF Data in Finance, University of Waterloo, Mar. 7, 2008 – p. 12/50
A Motivating Example

$V_t$: observable variable such as *economic news, signed order flow, signed trade and others* in Treasury market.

- **Intrinsic value process:**

  $$\frac{dX_t}{X_t} = (\mu + \kappa \mu V_t)dt + (\sigma + \kappa \sigma V_t)dB_t$$

- **Trading intensity:** $a(V^t, Z^t, t)$ is *non-anticipating*, where $V^t(\cdot) = V(\cdot \land t)$. E.g. $a(V(t_{i-1}), Z(t_{i-1}), \Delta_{i-1}, t)$ for $t_{i-1} \leq t < t_i$. This allows *Exponential ACD Model*.

- **Price** at time $t_i$:

  $$Z(t_i) = F(X(t_i))$$

  where $F$ is a random transformation specified by a *non-anticipating* transition probability $p(Z(t_i)|X(t_i); \theta^{t_i-}, X^{t_i-}, Z^{t_i-}, V^{t_i-}, \cdot)$. E.g. $p(Z(t_i)|X(t_i); \theta, Z(t_{i-1}) - X(t_{i-1}))$.

  This further allows *serial dependent microstructure noise*. 
Random-arrival-time State-Space Model

Three-Step Construction:

- **State process:** $X(t)$.

Assumption 2.1: is the same as Assumption 1.1, but both $\theta$ and $X$ can be vector processes, and $M_f(t)$ is a $\mathcal{F}_{\theta,X}^t$-martingale.

- **Event times:** $t_1, t_2, \ldots, t_i, \ldots$ follows a conditional Poisson process with a non-anticipating intensity $a(\theta^t, X^t, V^t, Z^t, t)$.

- **Observation at $t_i$:**

$$Z(t_i) = F(X(t_i))$$

where $F$ is a random transformation with the non-anticipating transition probability $p(Z(t_i)|X(t_i); \theta^{t_i-}, X^{t_i-}, V^{t_i-}, Z^{t_i-}, t))$. 
Filtering with MPP Observations

**Setup:** Mark space: \( U \); measure space: \((U, \mathcal{U}, \mu)\), \( \mu \): finite measure; \( \xi \) is a Poisson Random Measure (PRM) on \( U \times \mathcal{B}[0, \infty) \times \mathcal{B}[0, \infty) \) with mean measure \( \mu \times m \times m \). For \( A \in \mathcal{U} \),

\[
Y(A, t) = \int_{A \times [0, t] \times [0, +\infty)} I_{[0, \lambda(\theta^s, X^s, V^s, Z^s; u, s)]}(v) \xi(du \times ds \times dv),
\]
where \( Y(A, t) \) is a counting process recording the cumulative number of events that have occurred in the set \( A \) up to time \( t \).

\[
\tilde{Y}(A, t) = Y(A, t) - \int_{A \times [0, t]} \lambda(\theta^s, X^s, V^s, Z^s; u, s) \mu(du)ds
\]
is a martingale.

**Signal:** \((\theta, X)\)  
**Observation:** \((Y, V)\) or \((Z, V)\).
The Reference Measure

- Under the reference measure $Q$, For $A \in \mathcal{U}$,

$$Y(A, t) = \int_{A \times [0,t] \times [0,\infty)} I_{[0,1]}(v) \xi(du \times ds \times dv),$$

where $Y(A, t)$ is a counting process recording the cumulative number of events that have occurred in the set $A$ up to time $t$.

$$\tilde{Y}(A, t) = Y(A, t) - \int_{A \times [0,t]} \mu(du)ds = Y(A, t) - \mu(A)t$$

is a martingale.

- For Model I: $U = \{1, 2, \ldots, n\}$, $A = j$ and $V$ and $Z$ are not in the intensity,

Under $P$, $Y(A, t) = Y(j, t) = N_j(\int_0^t \lambda_j(\theta(s), X(s), s)ds)$.

Under $Q$, $Y(A, t) = Y(j, t) = Y_j(t)$ is a unit Poisson process.
Other Assumptions of Model II

- **Assumption 2.2:** $\xi$ is a PRM with the mean measure $\mu \times m \times m$ under $P$, where $\mu$ is finite measure.
- **Assumption 2.3:** $(\theta, X)$ and $\xi$ are independent under measure $P$.
- **Assumption 2.4:** $0 \leq a(\theta^t, X^t, V^t, Z^t, t) \leq C$ for some $C > 0$ and all possible $\theta^t, X^t, V^t, Z^t, t$.
- **Assumption 2.5:** Stochastic intensity kernel (non-anticipating):

$$
\lambda(\theta^{t-}, x^{t-}, v^{t-}, z^{t-}; u, t-) = a(\theta^{t-}, x^{t-}, v^{t-}, z^{t-}, t-)p(u|x; \theta^{t-}, x^{t-}, v^{t-}, z^{t-})
$$

(4)

where $p(u|x; \theta^{t-}, x^{t-}, v^{t-}, z^{t-})$ is the transition probability from $X(t)$ to $u$.

**Remark:** Note the similarity between Eq.(4) and Eq.(1).
Examples I

Group I: without continuous-time latent $X_t$

- $U = \{1\}$: Cox process and Exponential ACD model (Engle and Russell 1998).
- $U = \{\frac{a}{M}, \frac{a+1}{M}, \cdots, \frac{b}{M}\}$, or $R$, or $R^+$: UHF-GARCH model and many previously reviewed models under the framework of Engle (2000).

Group II: with latent $X_t$ not depending on $V$

- Zeng (2003) and its extension to multi-stocks (Scott and Zeng 2008).
- Classical examples: Segall, Davis and Kailath (1975), Bremaud (1981), Liptser and Shiryayev (1978), Kliemann, Koch and Marchetti (1990), and Last and Brandt (1995).
Examples II

Group III: with latent $X_t$ depending on $V$

- $V_t$ can change at random times (such as trading time) or deterministic times (such as every 5 minutes or every day).
- $V_t$ can takes values such as $\pm 1$ for the indicator of buyer or seller initiating trade; or 0 or 1 for the indicator of a special period; or other values such as order flow or a function of order flow.
- When $X_t$ is GBM, $V_t$ can be added in instantaneous expected return, or instantaneous volatility, or in the trading noise. Or find the “best" $f(V_t)$ for the model.
- $X_t$ can be O-U process; CIR model; CES model; SV models; plus jumps; plus regime-switching; with spikes; $\alpha$-stable process; and others. Then, $V_t$ can be added in the related parameters with different economic interpretations.

- $Z$ can be trading prices or ask and bid quotes.
An Integral Form of Price

Let $Z(t)$ be the price of the most recent transaction at or before time $t$.

$$Z(t) = Z(0) + \int_{[0,t] \times U} (u - Z(s-))Y(du \times ds).$$

Remarks:

- This is the telescoping sum: $Z(t) = Z(0) + \sum_{t_i \leq t} (Z(t_i) - Z(t_{i-1})$.
- This form is similar to that of Skorohod (1989) for continuous time Markov chain.

- This form is essential for the risk minimization hedging (Lee and Zeng 2006, for Model I), and the mean-variance portfolio selection problem of the model (Xiong and Zeng 2007, for Model I).
Joint Likelihood Function

Continuous-time joint likelihood function of \((\theta, X, Y)\):

For Model I,

\[
L(t) = \frac{dP}{dQ}(t) = \exp \left\{ \sum_{k=1}^{n} \int_{0}^{t} \log \lambda_k(\theta(s^-), X(s^-), s^-) dY_k(s) \right. \\
- \left. \sum_{k=1}^{n} \int_{0}^{t} \left[ \lambda_k(\theta(s^-), X(s^-), s) - 1 \right] ds \right\}.
\]

For Model II,

\[
L(t) = \exp \left\{ \int_{0}^{t} \int \log \lambda(\theta^s, X^s, V^s, Z^s; u, s^-) Y(du \times ds) \\
- \int_{0}^{t} \int \left[ \lambda(\theta^s, X^s, V^s, Z^s; u, s) - 1 \right] \mu(du)ds \right\}
\]
Likelihoods and Posterior

Define: \( \phi(f, t) = E^Q[f(\theta(t), X(t))L(t)|\mathcal{F}^Y,V_t] \). Then, \( \phi(1, t) = E^Q[L(t)|\mathcal{F}^Y,V_t] \) is the \textit{likelihood} of \( Y \) or the \textit{integrated (marginal) likelihood} of \( Y \) after assigning a prior to \((\theta(0), X(0))\).

Define: \( \pi_t \) is the conditional distribution of \((\theta(t), X(t))\) given \( \mathcal{F}^Y,V_t \). \( \pi_t \) becomes the \textit{posterior} after a prior is assigned.

Define: \( \pi(f, t) = E^P[f(\theta(t), X(t))|\mathcal{F}^Y,V_t] = \int f(\theta, x) \pi_t(d\theta, dx) \).

\( \bullet \) Bayes theorem gives:

\[
\pi(f, t) = \frac{\phi(f, t)}{\phi(1, t)}.
\]

SPDE for \( \phi_t \) – unnormalized filtering equation
SPDE for \( \pi_t \) – normalized filtering equation
Suppose there are two models: Model 1 and Model 2.

Define: two conditional ratio processes:

\[ q_1(f_1, t) = \frac{\phi_1(f_1, t)}{\phi_2(1, t)} \quad \text{and} \quad q_2(f_2, t) = \frac{\phi_2(f_2, t)}{\phi_1(1, t)} \]

The Bayes Factors: (BF: the ratio of two integrated likelihoods)

\[ BF_{12} = \frac{\phi_1(1, t)}{\phi_2(1, t)} = q_1(1, t) \quad \text{and} \quad BF_{21} = \frac{\phi_2(1, t)}{\phi_1(1, t)} = q_2(1, t) \]

- **Strongly Reject Model 1** if \( BF_{21} \) is larger than 20.
- **Decisively Reject Model 1** if \( BF_{21} \) is larger than 150.
Unnormalized Filtering Equation

Theorem 1: (Zeng 2003, for Model I) Under Assumptions 2.1–2.5, \( \phi_t \) is the unique measure-valued solution of the following SPDE, called the unnormalized filtering equation:

\[
\phi(f, t) = \phi(f, 0) + \int_0^t \phi(Af, s)ds - \int_0^t \int_U \phi(f(\lambda(u) - 1), s)\mu(du)ds
\]

\[
+ \int_0^t \int_U \phi(f(\lambda(u) - 1), s-\lambda)\mu(du \times ds),
\]

for every \( t > 0 \) and \( f \in D(A) \).
Theorem 1: (continued) \( \pi_t \) is the unique measure-valued solution of the following SPDE, called the normalized filtering equation:

\[
\begin{align*}
\pi(f, t) &= \pi(f, 0) + \int_0^t \pi(Af, s) ds + \int_0^t \pi(f, s) \int_U \pi(\lambda(u), s) \mu(du) ds \\
&\quad - \int_0^t \int_U \pi(f \lambda(u), s) \mu(du) ds + \int_0^t \int_U \left[ \frac{\pi(f \lambda(u), s-)}{\pi(\lambda(u), s-)} - \pi(f, s-) \right] dY(du \times ds)
\end{align*}
\]

Remark: When \( a(\theta^t, X^t, V^t, Z^t, t) = a(V^t, Z^t, t) \) (including the case of EACD model), it can be simplified as:

\[
\begin{align*}
\pi(f, t) &= \pi(f, 0) + \int_0^t \pi(Af, s) ds + \int_0^t \int_U \left[ \frac{\pi(fp(u), s-)}{\pi(p(u), s-)} - \pi(f, s-) \right] dY(du \times ds)
\end{align*}
\]

where \( p(u) = p(u|X(t); \theta^t, X^t, V^t, Z^t, t) \) allowing dependent noise.
Theorem 2: (Kouritzin and Zeng 2005, for Model I) Assume Model 1 has \((A_1, \lambda_1, \mu_1)\) and Model 2 has \((A_2, \lambda_2, \mu_2)\). Both models satisfy Assumptions 2.1–2.5. Then, \((q_1(t), q_2(t))\) is the unique pair measure-valued solution of the following system of SPDEs,

\[
q_1(f_1, t) = q_1(f_1, 0) + \int_0^t q_1(A_1 f_1, s) ds \\
+ \int_0^t \frac{q_1(f_1, s)}{q_2(1, s)} \int_U q_2(\lambda_2(u), s) \mu_2(du) ds - \int_0^t \int_U q_1(f_1 \lambda_1(u), s) \mu_1(du) ds \\
+ \int_0^t \int_U \left[ \frac{q_1(f_1 \lambda_1(u), s-)}{q_2(\lambda_2(u), s-)} q_2(1, s-) - q_1(f_1, s-) \right] dY(du \times ds)
\]

and

\[
q_2(f_2, t) = ...
\]
A Consistency Theorem

Theorem 3: (Zeng 2003, and Kouritzin and Zeng 2005, for Model I)
Suppose that Assumptions 2.1 to 2.5 hold for $(\theta, X, Y)$ and $(\theta_\epsilon, X_\epsilon, Y_\epsilon)$. If $(\theta_\epsilon, X_\epsilon) \Rightarrow (\theta, X)$ as $\epsilon \to 0$, then for bounded continuous functions, $f$, (i) $Y_\epsilon \Rightarrow Y$, (ii) $\phi_\epsilon(f, t) \Rightarrow \phi(f, t)$, (iii) $\pi_\epsilon(f, t) \Rightarrow \pi(f, t)$.

In the two-model case for model selection, then (iv) $q_{k,\epsilon}(f_k, t) \Rightarrow q_k(f_k, t)$ for $k = 1, 2$ simultaneously.

Sketch of Proof:
First, use Kurtz and Protter (1991)’s theorem on convergence of stochastic integral and the Continuous Mapping theorem to prove $L_\epsilon \Rightarrow L$. Then, $((\theta_\epsilon, X_\epsilon), Y_\epsilon, L_\epsilon) \Rightarrow ((\theta, X), Y, L)$.
Second, use Goggin (1994)’s or Kouritzin and Zeng (2005)’s theorems on convergence of conditional expectations and the Continuous Mapping theorem to prove (ii), (iii) and (iv).
Markov Chain Approximation Method

Construction of Parallel Recursive Online Algorithms – for computing nearly posterior, integrated likelihood and Bayes factors

For Example, to compute the nearly posterior: Three Steps:

- Construct a Markov chain \((\theta_\epsilon, X_\epsilon)\) to approximate \((\theta, X)\).
- Derive the filtering (or evolution) equations for \((\theta_\epsilon, X_\epsilon, Y_\epsilon)\).
  - **Propagation Equation:**
    \[
    \pi_\epsilon(f, t_{i+1}) = \pi_\epsilon(f, t_i) + \int_{t_i}^{t_{i+1}} \pi_\epsilon(A_\epsilon f, s) \, ds.
    \]
  - **Updating Equation:**
    \[
    \pi_\epsilon(f, t_{i+1}) = \frac{\pi_\epsilon(fp(u), t_{i+1})}{\pi_\epsilon(p(u), t_{i+1})}
    \]
- Convert the equation for \((\theta_\epsilon, X_\epsilon, Y_\epsilon)\) to recursive algorithms by
  - (a) representing \(\pi_\epsilon(\cdot, t)\), for example, as a finite array with components being \(\pi_\epsilon(f, t)\) for lattice-point indicator \(f\);
  - (b) approximating the time integral with an Euler scheme.
Two FM Models for Stock Price

Value Processes of the two Filtering Micromovement (FM) Models

1. GBM: (Zeng 2003)

\[
\frac{dX_t}{X_t} = \mu dt + \sigma dW_t,
\]

\[
Af(x) = \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} f(x) + \mu x \frac{\partial}{\partial x} f(x).
\]

2. JSV-GBM: (Zeng 2004)

\[
\frac{dX_t}{X_t} = \mu dt + \sigma(t) dW_t,
\]

\[
d\sigma(t) = (U_{N(t)} - \sigma(t-))dN(t)
\]

where \(N(t)\) is a Poisson process with intensity \(\lambda_\sigma\) and the jump size, \(\{U_i\}\), are i.i.d random variables with uniform distribution on \([\alpha_\sigma, \beta_\sigma]\).
Noise for the Models

\[ Y(t_i) = F(X(t_i)) = b_i(R[X(t_i), \frac{1}{M}] + D_i) \]

- **Discrete noise**: \( R[x, \frac{1}{M}] \), rounding function, where \( M = 8, 64, 128 \).
- **Non-clustering noise**: \( \{D_i\} \), has a doubly-geometric distribution:

\[
P\{D = d\} = \begin{cases} 
1 - (\rho + \kappa \rho v) & \text{if } d = 0 \\
\frac{1}{2}[1 - (\rho + \kappa \rho v)](\rho + \kappa \rho v)^{64|d|} & \text{if } d = \pm \frac{i}{M} \text{ for } i = 1, 2, 3 \ldots .
\end{cases}
\]

- **Clustering noise**: \( b_i(\cdot) \), a random biasing function

**biasing rule**: Set \( y' = R[X(t_i), \frac{1}{M}] + D_i \) and \( y = Y(t_i) = b(y') \).

- If fractional part of \( y' \) is an even \( M \)th, \( y \) stays on \( y' \) w. p. 1.
- If the fractional part of \( y' \) is an odd \( M \)th, then \( y' \) moves to the closest \( M/2 \)th w. p. \( \alpha \), or \( y' \) moves to the closest odd \( M/4 \)th or integer w. p. \( \beta \), or \( y \) stays on \( y' \) w. p. \( 1 - \alpha - \beta \).

- Parameters in the Model: \((\mu, \sigma, \rho, \kappa_\mu, \kappa_\sigma, \kappa_\rho, \alpha, \beta)\).
Theorem 4: (Zeng 2003) For the simple filtering model (I) of GBM, assume that the clustering parameters \((\alpha, \beta)\) are known, and \((\mu, \sigma, \rho)\) has a prior. Then, the Bayes estimates are consistent almost surely. Namely, \(E[f(\mu, \sigma, \rho)|\mathcal{F}_t^Z] \to f(\mu, \sigma, \rho)\) as \(t \to \infty\) for bounded continuous function \(f\) when \(\mu - \frac{1}{2}\sigma^2 > 0\).

Remark: The standard condition \(\mu - \frac{1}{2}\sigma^2 > 0\) rules out bankruptcy.
Bayesian estimates of MU and their two-SDs Bounds

Bayesian estimates of SIGMA and their two-SDs Bounds

Day 32,500 simulated data
Bayes Estimates II: Simulated Data

Bayes estimates of volatility and their true values in simulated data

Bayes estimates of volatility and two–SE bounds: last 25,000 simulated data
Bayes Factor: Simulated Data

Bayes Factors of JSV-GBM vs GBM: first 2550 simulated data

Bayes Factors of JSV-GBM vs GBM: Among the Second Sigma

Conference on the Econometrics of UHF Data in Finance, University of Waterloo, Mar. 7, 2008 – p. 34/50
One FM Model for Treasury UHF Data

(Joint with D. Kuipers (UMKC, Finance) and X. Hu (Princeton, Econ))

- **Intrinsic value process:**

\[
\frac{dX_t}{X_t} = \mu dt + (\sigma + \kappa_V V_t) dB_t
\]

**Generator:** \( \theta = (\mu, \sigma, \kappa_\sigma, \rho, \kappa_\rho) \)

\[
A_v f(x, \theta) = \mu x \frac{\partial}{\partial x} f(\theta, x) + \frac{1}{2} (\sigma + \kappa_\sigma v)^2 x^2 \frac{\partial^2}{\partial x^2} f(\theta, x).
\]

where
1. \( V_t \) is the indicator of buyer (+1) or seller (-1) initiating;
2. \( V_t \) is signed sqrt(Vol) with sign +(-) when buyer(seller) initiating;
3. \( V_t \) is 1 when \( t \) is in a a specific time interval, otherwise 0.

- **Trading intensity:** \( a(V^t, Z^t, t) \) independent of \( X_t \) can be modeled by an Exponential ACD Model.
Theorem 5: For the above simple filtering model (III) of GBM with regression factors, assume that $V_t = \pm 1$ has infinite non-alternating $\pm 1$ and the clustering parameters $(\alpha, \beta)$ are known, and $\theta = (\mu, \sigma, \rho, \kappa_\mu, \kappa_\sigma, \kappa_\rho)$ has a prior. Then, the Bayes estimates are consistent almost surely. Namely, $E[f(\theta)|F^Z_t] \to f(\theta)$ as $t \to \infty$ for bounded continuous function $f$ when $\mu - \frac{1}{2}(\sigma + \kappa_\sigma)^2 > 0$.

Remark: The standard condition "$\mu - \frac{1}{2}(\sigma + \kappa_\sigma)^2 > 0$" rules out bankruptcy.
Estimation Results: Cases (1) and (2)

Table 1: Annualized Bayes estimates for 10yr. Treasury UHF Data, 8/15/-12/31, 2000

<table>
<thead>
<tr>
<th>Models</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\kappa_\sigma$</th>
<th>$\kappa_\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case I:</td>
<td>34.48%</td>
<td>4.92%</td>
<td>0.0503</td>
<td>-0.29%</td>
<td>-0.0109</td>
</tr>
<tr>
<td>$V_t = \pm 1$</td>
<td>(4.27%)</td>
<td>(0.05%)</td>
<td>(0.0045)</td>
<td>(0.04%)</td>
<td>(.0045)</td>
</tr>
<tr>
<td>Case II:</td>
<td>28.64%</td>
<td>4.94%</td>
<td>0.0486</td>
<td>-0.10%</td>
<td>n.a.</td>
</tr>
<tr>
<td>$V_t = \pm \sqrt{Vol}$</td>
<td>(8.17%)</td>
<td>(0.05%)</td>
<td>(0.0044)</td>
<td>(0.03%)</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
Simulated Data: Case (3)
Table 2: Bayes estimates for 2050 simulated data when 50 data have extra volatility,

<table>
<thead>
<tr>
<th>Models</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\kappa_\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>5.00E-8</td>
<td>2.00E-5</td>
<td>0.05</td>
<td>3.00E-5</td>
</tr>
<tr>
<td>Bayes Est.</td>
<td>4.27E-8</td>
<td>2.00E-5</td>
<td>0.0477</td>
<td>2.86E-5</td>
</tr>
<tr>
<td></td>
<td>(1.08E-8)</td>
<td>(2.89E-14)</td>
<td>(0.0056)</td>
<td>(6.58E-6)</td>
</tr>
</tbody>
</table>
Real Data: FOMC Period I

Time Series for 5Year Notes Aug 17 – Aug 30 2000

Before Aug 22nd
Aug 22nd, 12 – 3 PM
After Aug 22nd

Price

Time (seconds)

Conference on the Econometrics of UHF Data in Finance, University of Waterloo, Mar. 7, 2008 – p. 40/50
Real Data: FOMC Period II

Time Series for Aug 22nd

Before 12PM 12 - 3 PM After 3PM
Table 3: Bayes estimates for August 17 - 30, 2000, 5yr notes

<table>
<thead>
<tr>
<th>Models</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\kappa_\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayes Est.</td>
<td>8.51%</td>
<td>1.68%</td>
<td>0.022</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>(7.21%)</td>
<td>(0.08%)</td>
<td>(0.0054)</td>
<td>(0.58%)</td>
</tr>
</tbody>
</table>

(1247 data with 40 marking V=1, for 12-3pm of August 22 when FOMC announcement time is around 2pm)

These results are consistent with Fleming and Remolona (1999), Green (2004) and Brandt and Kavacejz (2004): Primary impact: Economic News; Secondary impact: order flow and liquidity.
Price and Economic Announcements

![Graph showing price movements and economic announcements](image)
Bayes Est. (I) for Treasury data: Case (3)

<table>
<thead>
<tr>
<th>Data</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\kappa_\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2yr: 7/2000</td>
<td>0.26%</td>
<td>0.93%</td>
<td>0.0054</td>
<td>3.32%</td>
</tr>
<tr>
<td></td>
<td>(0.10%)</td>
<td>(0.03%)</td>
<td>(0.0073)</td>
<td>(0.16%)</td>
</tr>
<tr>
<td>2yr: 8/2000</td>
<td>0.06%</td>
<td>0.74%</td>
<td>0.0008</td>
<td>1.41%</td>
</tr>
<tr>
<td></td>
<td>(0.05%)</td>
<td>(0.03%)</td>
<td>(0.0003)</td>
<td>(0.15%)</td>
</tr>
<tr>
<td>2yr: 9/2000</td>
<td>0.06%</td>
<td>1.04%</td>
<td>0.0014</td>
<td>2.16%</td>
</tr>
<tr>
<td></td>
<td>(0.01%)</td>
<td>(0.02%)</td>
<td>(0.0002)</td>
<td>(0.29%)</td>
</tr>
<tr>
<td>2yr: 10/2000</td>
<td>0.09%</td>
<td>1.49%</td>
<td>0.0011</td>
<td>1.53%</td>
</tr>
<tr>
<td></td>
<td>(0.02%)</td>
<td>(0.03%)</td>
<td>(0.0003)</td>
<td>(0.16%)</td>
</tr>
<tr>
<td>2yr: 11/2000</td>
<td>0.09%</td>
<td>0.99%</td>
<td>0.0009</td>
<td>1.95%</td>
</tr>
<tr>
<td></td>
<td>(0.02%)</td>
<td>(0.03%)</td>
<td>(0.0001)</td>
<td>(0.30%)</td>
</tr>
</tbody>
</table>
### Bayes Est. (II) for Treasury data: Case (3)

<table>
<thead>
<tr>
<th>Data</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\kappa_\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2yr: 12/2000</td>
<td>0.12%</td>
<td>1.50%</td>
<td>0.0001</td>
<td>1.31%</td>
</tr>
<tr>
<td></td>
<td>(0.04%)</td>
<td>(0.03%)</td>
<td>(0.0001)</td>
<td>(0.27%)</td>
</tr>
<tr>
<td>5yr: 7/1 - 11/16/2000</td>
<td>2.78%</td>
<td>2.71%</td>
<td>0.0204</td>
<td>4.19%</td>
</tr>
<tr>
<td></td>
<td>(1.86%)</td>
<td>(0.03%)</td>
<td>(0.0029)</td>
<td>(0.03%)</td>
</tr>
<tr>
<td>5yr: 11/17 - 12/31/2000</td>
<td>27.66%</td>
<td>3.37%</td>
<td>0.0008</td>
<td>4.96%</td>
</tr>
<tr>
<td></td>
<td>(4.62%)</td>
<td>(0.03%)</td>
<td>(0.0006)</td>
<td>(0.69%)</td>
</tr>
<tr>
<td>10yr: 7/1 - 8/12/2000</td>
<td>11.82%</td>
<td>4.81%</td>
<td>0.0276</td>
<td>8.60%</td>
</tr>
<tr>
<td></td>
<td>(9.01%)</td>
<td>(0.14%)</td>
<td>(0.0051)</td>
<td>(0.16%)</td>
</tr>
<tr>
<td>10yr: 8/15 - 12/31/2000</td>
<td>13.77%</td>
<td>4.86%</td>
<td>0.0479</td>
<td>4.03%</td>
</tr>
<tr>
<td></td>
<td>(7.33%)</td>
<td>(0.05%)</td>
<td>(0.0044)</td>
<td>(0.30%)</td>
</tr>
</tbody>
</table>
Conclusions and Future Works

- Financial applications on market microstructure theory
- Statistics and information theory:
  - Consistency, CLT for the estimators of parameters.
  - Mutual information and its rate and its applications
- Mathematical finance:
  - Option pricing and hedging, portfolio optimization, and utility maximization.
- Particle filtering or sequential Monte Carlo
  - Convergence, convergence rate and large deviation
- To allow long-range dependence by using fractional signal.

Related papers, real data examples, Fortran codes are available at
http://mendota.umkc.edu/paper-tick.html
Particle Filtering (or SMC)

Suppose parameters are known. To estimate the value process, \( X \),

- **Simulate** 100 independent sample paths of \( X \) following GBM: \( V_j(t) \),
  \( j = 1, 2, \cdots, 100 \), at each trading times.

- At some time, \( t_i \), calculate the **importance weight** for each \( V_j \):

  \[
  w_i^j(V_j(t_i)) = L_i^j(t_i) = \prod_{k=1}^{n} \exp \left\{ \int_{t_{i-1}}^{t_i} \log \lambda_k(V_j(s-), s-) dY_k(s) \right\} \\
  \times \prod_{k=1}^{n} \exp \left\{ - \int_{t_{i-1}}^{t_i} [\lambda_k - 1] ds \right\}.
  \]

- **Resample** the sample paths according to the distribution proportional to the importance weights at times.

  Or

- **Branch** each particle to a random number of particles proportional to the importance weights at times. (Xiong and Zeng 2006)
Simulation: No Resampling

Resampling Particle Simulation
Simulate Price/Estimated Price

# of Obs: 10000 .  # of Particles: 100 .   Resample every 10000 intervals

Particle (V)
Stock price (X)
Stock price w/noise (Y)
Simulation: Resampling Every 100 Trades

Resampling Particle Simulation
Simulate Price/Estimated Price

Particle (V)
Stock price (X)
Stock price w/noise (Y)

Time: seconds
# of Obs: 10000 . # of Particles: 100 . Resample every 100 intervals
Simulation: Resampling Every 10 Trades

Resampling Particle Simulation
Simulate Price/Estimated Price

# of Obs: 10000 . # of Particles: 100 . Resample every 10 intervals