Modeling Ultra-High-Frequency Multivariate Financial Data by Monte Carlo Simulation Methods

Ph.D. Student: TingTing Peng
Supervisor: Marco Minozzo

Dipartimento di Scienze Economiche
Università degli Studi di Verona

Ph.D. Student Presentation

Verona, 10 May 2010
1 Ultra-High-Frequency Data (UHF)

2 Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3 Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4 Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Motivation

- Interest in understanding financial markets, which are the source of UHF data.
- Most of the econometric literatures is dealing with regularly spaced data (5-min; 15-min; daily).
- Evidence of strong co-movements in individual trading rate.
- Correlation measures between returns has direct application in portfolio management.

Methodology

Our modeling framework for Multi-UHF data is based on Doubly Stochastic Poisson Processes, developing some ideas from Bauwens(2006). Our model will allow to capture co-movements and common component through the implementation of RJMCMC algorithm.
Motivation

- Interest in understanding financial markets, which are the source of UHF data.
- Most of the econometric literatures is dealing with regularly spaced data (5-min; 15-min; daily).
- Evidence of strong co-movements in individual trading rate.
- Correlation measures between returns has direct application in portfolio management.

Methodology

Our modeling framework for Multi-UHF data is based on Doubly Stochastic Poisson Processes, developing some ideas from Bauwens(2006). Our model will allow to capture co-movements and common component through the implementation of RJMCMC algorithm.
Outline

1. Ultra-High-Frequency Data (UHF)
   2. Marked Doubly Stochastic Poisson Processes (DSPP)
      - Financial Models For Equally spaced data
      - Marked Point Process (MPP) and UHF Data
      - Literature review for MPP
      - Marked DSPP
      - A DSPP model with a common intensity
   3. Modeling framework for logreturns
      - Logreturn Models
      - Simulation studies
      - Brief Conclusion
   4. Filtering by RJMCMC
      - Markov Chain Monte Carlo
      - Reversible jump MCMC
### Definition (UHF)

UHF data also known as tick-by-tick data: each tick is one logical unit of information, like a quote or a transaction price, time of event.

### Data

- **Source**: Borsa di Milano.
- **7 Italian banks**: Banco Popolare (POP), Mediobanca (MED), Banca popolare di Milano (MIL), MPS Banca (MPS), Banca Intesa SanPaolo (ISP), UBI Banca (UBI), Unicredit Banca (UCD).
- **Period**: 15 business days from 27/10/2008 to 14/11/2008; market open from 9:05 to 17:25 (30000s).
The total number and the average transactions for each asset during 27/10/2008 and 14/11/2008 (15 business days).

<table>
<thead>
<tr>
<th>Asset</th>
<th>Total number of transactions</th>
<th>Frequency per business day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banco Popolare</td>
<td>22993</td>
<td>1533</td>
</tr>
<tr>
<td>Medio Banca</td>
<td>14434</td>
<td>962</td>
</tr>
<tr>
<td>Banca popolare di Milano</td>
<td>14411</td>
<td>961</td>
</tr>
<tr>
<td>MPS Banca</td>
<td>14081</td>
<td>938</td>
</tr>
<tr>
<td>Intesa San Paolo Banca</td>
<td>51626</td>
<td>3442</td>
</tr>
<tr>
<td>UBI Banca</td>
<td>15022</td>
<td>1002</td>
</tr>
<tr>
<td>Unicredit Banca</td>
<td>64444</td>
<td>4297</td>
</tr>
</tbody>
</table>
Characteristics of UHF Data

- Irregular time spacing
- Price Discreteness
- Negative first order autocorrelation of logreturns
- Empirical synchronized correlation (Epps effect)
Negative First-Order Autocorrelation

Autocorrelation for Banco Popolare

Autocorrelation for Medio Banca

Autocorrelation for MPS Banca

Autocorrelation for Banca Popolare di Milano

Autocorrelation for Banca di Intesa SanPaolo

Autocorrelation for UBI Banca

Autocorrelation for Unicredit Banca

TingTing Peng
Modeling Multivariate UHF Data
Synchronization Methods

Linear Interpolation and Previous-tick Interpolation

- Linear interpolation: to interpolate time between $t_{j'}$ and $t_{j'+1}$ with the following form:

$$Y(t_0 + i\Delta t) = Y_{j'} + \frac{t_0 + i\Delta t - t_{j'}}{t_{j'+1} - t_{j'}} (Y_{j'+1} - Y_{j'})$$

where the $j'$ refer to the most recent time index of time $t_0 + i\Delta t$.

- Previous-tick interpolation: take the most recent value of the raw data, in formula,

$$Y(t_0 + i\Delta t) = Y_{j'}$$

where the $j'$ refer to the most recent time index of time $t_0 + i\Delta t$. 
A graphic example of interpolation
Previous-tick synchronization-logreturn aggregation

\[ r_{\Delta t} = \log P_3 - \log P_0 = \log P_3 - \log P_2 + \log P_2 - \log P_1 + \log P_1 - \log P_0 = r_3 + r_2 + r_1 \]
Epps effect: Intesa Sanpaolo Banca with others

CORRELATION ISP−POP

CORRELATION ISP−MED

CORRELATION ISP−MPS

CORRELATION ISP−MIL

CORRELATION ISP−UBI

CORRELATION ISP−UCD

TingTing Peng  Modeling Multivariate UHF Data
Linear dependence: Scatter plot of ISP-UCD

TingTing Peng
Modeling Multivariate UHF Data
Epps effect: Medio Banca with others

CORRELATION MED–POP

CORRELATION MED–MPS

CORRELATION MED–MIL

CORRELATION MED–ISP

CORRELATION MED–UBI

CORRELATION MED–UCD
Epps effect: Banco Popolare with others

CORRELATION POP–MED

CORRELATION POP–MPS

CORRELATION POP–MIL

CORRELATION POP–ISP

CORRELATION POP–UBI

CORRELATION POP–UCD
Epps effect: MPS Banca with others

TingTing Peng
Modeling Multivariate UHF Data
Epps effect: Banca Popolare di Milano with others

CORRELATION MIL–POP

CORRELATION MIL–MED

CORRELATION MIL–MPS

CORRELATION MIL–ISP

CORRELATION MIL–UBI

CORRELATION MIL–UCD

Ting Ting Peng  Modeling Multivariate UHF Data
Epps effect: UBI Banca with others

CORRELATION UBI–POP

CORRELATION UBI–MED

CORRELATION UBI–MPS

CORRELATION UBI–MIL

CORRELATION UBI–ISP

CORRELATION UBI–UCD

TingTing Peng
Modeling Multivariate UHF Data
Epps effect: Unicredit Banca with others

TingTing Peng  Modeling Multivariate UHF Data
Epps effect

Epps effect: the correlation of returns decrease dramatically as the time intervals enter the intra-hour level. It is extensively observed in the high frequency data.
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Literature review

Standard Econometric Models

- Engle(1982): ARCH Model
- Bollerslev(1986): GARCH
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Marked Point Processes (MPP) and UHF Data

A marked point process (MPP) is a sequence of pairs
\[ \phi = (T_i, Z_i)_{i \in \mathbb{N}}. \]

Ultra-High-Frequency (UHF) Data
With ultra-high-frequency (UHF) data, all trades (or quotes) of a financial asset are recorded (time of event, (bid-ask) price, volume, etc.). In this case, \( T_i \) and \( Z_i, i = 1, 2, \ldots \), could be the time and the size of the \( i \)th logprice change.
Marked Point Processes (MPP) and UHF Data

Marked Point Process

A marked point process (MPP) is a sequence of pairs

\[ \phi = (T_i, Z_i)_{i \in \mathbb{N}}. \]

Ultra-High-Frequency (UHF) Data

With ultra-high-frequency (UHF) data, all trades (or quotes) of a financial asset are recorded (time of event, (bid-ask) price, volume, etc.). In this case, \( T_i \) and \( Z_i, i = 1, 2, \ldots \), could be the time and the size of the \( i \)th logprice change.
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Literature review for MPP

### Autoregressive Conditional Duration (ACD) Models

- Zhang, Russell and Tsay (2000): 'A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data';
Literature Review for MPP

Marked Doubly Stochastic Poisson Processes (DSPP)

- Frey (2000): 'Risk-minimization with incomplete information in a model for high-frequency data';
- Frey and Runggaldier (2001): 'A nonlinear filtering approach to volatility estimation with a view towards high frequency data';
- Centanni and Minozzo (2006): 'Estimation and filtering by reversible jump MCMC for a doubly stochastic Poisson model for ultra-high-frequency data';
Literature Review for MPP

Multivariate Point Process

- Hall and Hautsch (2006): 'Order Aggressiveness and Order Book Dynamics';
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
     - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Marked Doubly Stochastic Poisson Processes (DSPP)

**Definition (Last and Brandt, 1995)**

Given a probability space \((\Omega, \mathcal{F}, P)\) and a filtration \(\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}\), an MPP \(\phi\) adapted to the filtration is a **DSPP** if there exists a \(\mathcal{F}_0\)-measurable random measure \(\nu\) on \(\mathbb{R}^+ \times \mathbb{R}\) such that

\[
P\left(\mu((s, t] \times A) = k \Big| \mathcal{F}_s\right) = \frac{\left(\nu((s, t] \times A)\right)^k}{k!} e^{-\nu((s, t] \times A)},
\]

almost surely, for every \(A \in \mathcal{B}(\mathbb{R})\), where \(\mu\) denotes the counting measure associated to the MPP \(\phi\), that is,

\[
\mu(\omega, (0, t] \times A) = \sum_{i=1}^{N_t} \mathbf{1}_{\{Z_i \in A\}}.
\]
Marked Doubly Stochastic Poisson Processes (DSPP)

Intensity of a Marked DSPP

DSPPs are point processes in which the number of jumps \( N_t - N_s \) in any time interval \((s, t]\) is Poisson distributed, given another positive stochastic process \( \lambda \), called intensity

\[
P (N_t - N_s = k|\lambda) = \frac{\left( \int_s^t \lambda_u du \right)^k}{k!} \exp \left( - \int_s^t \lambda_u du \right).
\]

Moreover, given \( \lambda \), the number of jumps in disjoint time intervals are independent.

Marked DSPP With Intensity Driven by MPP

We assume the intensity \( \lambda \) depends itself on another MPP, that is, \( \lambda \) is of the form \( \lambda_t = h(t, \psi^t) \), where \( \psi^t \) is the restriction of another MPP \( \psi = (\tau_j, X_j)_{j \in \mathbb{N}} \) to \([0, t] \).
Marked Doubly Stochastic Poisson Processes (DSPP)

Intensity of a Marked DSPP

DSPPs are point processes in which the number of jumps \( N_t - N_s \) in any time interval \((s, t]\) is Poisson distributed, given another positive stochastic process \( \lambda \), called intensity

\[
P(N_t - N_s = k|\lambda) = \frac{\left( \int_s^t \lambda_u du \right)^k}{k!} \exp \left( -\int_s^t \lambda_u du \right).
\]

Moreover, given \( \lambda \), the number of jumps in disjoint time intervals are independent.

Marked DSPP With Intensity Driven by MPP

We assume the intensity \( \lambda \) depends itself on another MPP, that is, \( \lambda \) is of the form \( \lambda_t = h(t, \psi_t^0) \), where \( \psi_t^0 \) is the restriction of another MPP \( \psi = (\tau_j, X_j)_{j \in \mathbb{N}} \) to \([0, t]\).
Marked DSPP With Shot Noise Intensity (Basic Model)

Stochastic Differential Equation

Stochastic noise intensity is assumed as the solution of the following Stochastic Differential Equation:

\[ d\lambda_t = -k\lambda_t dt + dJ_t \]

where \( k > 0 \), \( J = \sum_{j=0}^{N'_t} X_j \), and \( N'_t = \# j : \tau_j \leq t \).

Solution of SDE

\[ \lambda_t = \lambda_0 \exp(-kt) + \sum_{j=0}^{N'_t} X_j \cdot \exp(-k(t - \tau_j)), \quad t \geq 0 \]

where \((\tau_j)_{j=1,2,...}\) is a Poisson process with constant parameter \( \nu \), and \( X_j \) are exponentially distributed with parameter \( \gamma \).
Marked DSPP with Shot Noise Intensity (Basic Model)

Stochastic Differential Equation

Stochastic noise intensity is assumed as the solution of the following Stochastic Differential Equation:

\[ d\lambda_t = -k\lambda_t dt + dJ_t \]

where \( k > 0 \), \( J = \sum_{j=0}^{N_t'} X_j \), and \( N_t' = \# j : \tau_j \leq t \).

Solution of SDE

\[ \lambda_t = \lambda_0 \exp(-kt) + \sum_{j=0}^{N_t'} X_j \cdot \exp(-k(t - \tau_j)), \quad t \geq 0 \]

where \((\tau_j)_{j=1,2,...}\) is a Poisson process with constant parameter \( \nu \), and \( X_j \) are exponentially distributed with parameter \( \gamma \).
Simulated Shot Noise Intensity with parameter $\nu = 0.1$, $k = 0.1$, $\gamma = 0.2$
Outline

1. Ultra-High-Frequency Data (UHF)
   
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Common Factor Intensity Model

Model Description

\[ \tilde{\lambda}_t^0 \infty = \lambda_t^s + \alpha_s \lambda_0^t \quad s = 1, 2, \ldots, m. \]

where \( \lambda_t^s \) denotes s-specific individual intensity and \( \lambda_0^t \) denotes a common component. The coefficient \( s \) is a scaling parameter driving the common intensity impact on the s-individual intensity.

Bivariate case

\[ \tilde{\lambda}_t^{01} = \lambda_t^1 + \alpha_1 \lambda_0^t \]
\[ \tilde{\lambda}_t^{02} = \lambda_t^2 + \alpha_2 \lambda_0^t \]
Simulated common intensity with parameter $\nu_0 = 0.1, k_0 = 0.1, \gamma_0 = 0.2, \nu_1 = 0.5, k_1 = 0.5, \gamma_1 = 1, \nu_2 = 1, k_2 = 0.4, \gamma_2 = 1$. 

TingTing Peng  
Modeling Multivariate UHF Data
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Modelling framework for logreturns

### Independent Wiener Process

\[ R_{T_i} = \mu + \xi_{T_i}, \quad \xi_{T_i} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, (T_i - T_{i-1})\begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}\right) \]

### Concurrently correlated Wiener Process

\[ R_{T_i} = \mu + \xi_{T_{i-1}}, \quad \xi_t \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, (T_i - T_{i-1})\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right) \]

where \( \sigma_{12} = \sigma_{21} \neq 0 \).
Modelling framework for logreturns

### Independent Wiener Process

\[ RT_i = \mu + \xi T_i, \quad \xi T_i \sim N\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, (T_i - T_{i-1}) \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \right) \]

### Concurrently correlated Wiener Process

\[ RT_i = \mu + \xi T_{i-1}, \quad \xi t \sim N\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, (T_i - T_{i-1}) \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right) \]

where \( \sigma_{12} = \sigma_{21} \neq 0 \).
Modelling framework for logreturns

Bi-AR(1) without cross-correlation

\[ R_{Ti} = \mu + \beta R_{Ti-1} + \xi_{Ti}, \quad \xi_{Ti} \sim N \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{array} \right) \right) \]

where \( \beta = \left( \beta_{11} \ 0 \right) \)

Bi-AR(1)

\[ R_{Ti} = \mu + \beta R_{Ti-1} + \xi_{Ti}, \quad \xi_{T-i} \sim N \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{array} \right) \right) \]

where \( \beta_{ij} \neq 0, \ i, j = 1, 2 \).
Modelling framework for logreturns

Bi-AR(1) without cross-correlation

\[ R_{Ti} = \mu + \beta R_{Ti-1} + \xi_{Ti}, \quad \xi_{Ti} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right) \]

where \( \beta = \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix} \)

Bi-AR(1)

\[ R_{Ti} = \mu + \beta R_{Ti-1} + \xi_{Ti}, \quad \xi_{Ti} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right) \]

where \( \beta_{ij} \neq 0, i, j = 1, 2. \)
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Pooled time

TingTing Peng  
Modeling Multivariate UHF Data
$\mu = 0$

i) Independent Wiener Process with $\sigma_{11} = \sigma_{22} = 0.1$, $\sigma_{12} = \sigma_{21} = 0$
ii) Concurrently correlated Wiener Process with
\( \sigma_{11} = 0.1, \sigma_{22} = 0.2, \sigma_{12} = \sigma_{21} = 0.1 \)
iii) Bi-AR(1) without cross correlation: $\sigma_{11} = 0.1, \sigma_{22} = 0.2, \sigma_{12} = \sigma_{21} = 0.1, \beta_{11} = 0.1, \beta_{22} = 0.2, \beta_{12} = \beta_{21} = 0$
iv) Bi-AR(1): $\sigma_{11} = 0.1, \sigma_{22} = 0.2, \sigma_{12} = \sigma_{21} = 0.1, \beta_{11} = 0.2, \beta_{22} = 1.1, \beta_{12} = 0.3, \beta_{21} = -0.6$
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Brief Conclusion

- Common factor model of intensity has no significant impact on 'Epps effect', in general on empirical synchronized correlation function;
- To obtain the empirical correlation, we need not only concurrent correlation but also cross-serial correlation;
- Epps effect caused by asynchronization.

\[
T_{\Delta t} = \log P_3 - \log P_0 = \log P_3 - \log P_2 + \log P_2 - \log P_1 + \log P_1 - \log P_0 = r_3 + r_2 + r_1
\]
Brief Conclusion

- Common factor model of intensity has no significant impact on 'Epps effect', in general on empirical synchronized correlation function;
- To obtain the empirical correlation, we need not only concurrent correlation but also cross-serial correlation;
- Epps effect caused by asynchronization.

\[ r_{\Delta t} = \log P_3 - \log P_0 = \log P_3 - \log P_2 + \log P_2 - \log P_1 + \log P_1 - \log P_0 = r_3 + r_2 + r_1 \]
Cross-Autocorrelation
Outline

1. Ultra-High-Frequency Data (UHF)
2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity
3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion
4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Introduction

MCMC methods are a class of algorithms for sampling from probability distribution based on constructing a Markov chain that has the target distribution ($\pi$) as its stationary distribution. In order for $\pi$ to be the stationary distribution of the chain, the transition probabilities must satisfy the detailed balance condition

$$\pi_i P_{ij} = \pi_j P_{ji}$$
Markov Chain Monte Carlo (MCMC)

Introduction

MCMC methods are a class of algorithms for sampling from a probability distribution based on constructing a Markov chain that has the target distribution ($\pi$) as its stationary distribution. In order for $\pi$ to be the stationary distribution of the chain, the transition probabilities must satisfy the detailed balance condition

$$\pi_i P_{ij} = \pi_j P_{ji}$$

MCMC

Two important algorithms to construct Markov Chain:
- Metropolis-Hastings algorithm;
- Gibbs sampling (which is the special case of M-H algorithm).
Metropolis-Hasting Algorithm

Suppose we have a Markov Chain in state $x$. We want to simulate a draw from the transition kernel $p(x, y)$ with stationary distribution $\pi$, but we do not know the form of $p(x, y)$. We do know how to compute a function proportional to $\pi$, $f(x) = k\pi(x)$. Assume that we can draw $y$ from an arbitrary distribution $q(x, y)$. Consider using this $q$ as a transition kernel. If we construct $\alpha(x, y)$ such that

$$\pi(x)q(x, y)\alpha(x, y) = \pi(y)q(y, x)\alpha(y, x)$$

then we will have a reversible transition kernel with stationary distribution $\pi$. We can take:

$$\alpha(x, y) = \min\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\}$$

Although we do not know $\pi(x)$, we do know $f(x) = k\pi(x)$, so we can calculate $\alpha(x, y)$. 

TingTing Peng
Modeling Multivariate UHF Data
Metropolis-Hasting Algorithm

In summary, the Metropolis-Hasting algorithm proceeds as follows. Given $x^j$, to move to $x^{j+1}$ we:

1. Generate a draw $y$ from $q(x^j, \cdot)$
2. Calculate $\alpha(x^j, y)$
3. Generate $u \sim U[0, 1]$
4. if $\alpha(x^j, y) > u$, then $x^{j+1} = y$. Otherwise $x^{j+1} = x^j$.

After a large number of iterations, the marginal distribution of $x^j$ will converge to $\pi$.

Example

Suppose we have three intensity processes, $\lambda_1 = \nu_1 x_1$, $\lambda_2 = \nu_2 x_2$, $\lambda_0 = \nu_0 x_0$, where $x_i \sim \text{Gamma}(\alpha_i, \beta_i)$, and $\nu_i$, $\alpha_i$ and $\beta_i$ are parameters, $i = 0, 1, 2$. 
Metropolis-Hasting Algorithm

In summary, the Metropolis-Hasting algorithm proceeds as follows. Given \( x^j \), to move to \( x^{j+1} \) we:

1. Generate a draw \( y \) from \( q(x^j, .) \)
2. Calculate \( \alpha(x^j, y) \)
3. Generate \( u \sim U[0, 1] \)
4. if \( \alpha(x^j, y) > u \), then \( x^{j+1} = y \). Otherwise \( x^{j+1} = x^j \).

After a large number of iterations, the marginal distribution of \( x^j \) will converge to \( \pi \).

Example

Suppose we have three intensity processes, \( \lambda_1 = \nu_1 x_1 \), \( \lambda_2 = \nu_2 x_2 \), \( \lambda_0 = \nu_0 x_0 \), where \( x_i \sim Gamma(\alpha_i, \beta_i) \), and \( \nu_i, \alpha_i \) and \( \beta_i \) are parameters, \( i = 0, 1, 2 \).

4.1. MCMC 4.2. RJMCMC

Lambda_1 = v_1 \times x_1

Lambda_2 = v_2 \times x_2

Lambda_0 = v_0 \times x_0

Lambda_{01} = \Lambda_1 + a_1 \times \Lambda_0

Lambda_{02} = \Lambda_2 + a_2 \times \Lambda_0

Return_1

Return_2

TingTing Peng

Modeling Multivariate UHF Data
Run of Metropolis-Hasting algorithm for 100000 iterations with 50000 burn-in period. Parameter $\nu_0 = 1, \nu_1 = 2, \nu_2 = 1, \alpha_0 = 3, \beta_0 = 1, \alpha_1 = 2, \beta_1 = 2, \alpha_2 = 5, \beta_2 = 0.5, a_1 = 1, a_2 = 0.5$.

- $x_0 = 3.8524$
- $x_1 = 6.0323$
- $x_2 = 2.5079$
Outline

1. Ultra-High-Frequency Data (UHF)

2. Marked Doubly Stochastic Poisson Processes (DSPP)
   - Financial Models For Equally spaced data
   - Marked Point Process (MPP) and UHF Data
   - Literature review for MPP
   - Marked DSPP
   - A DSPP model with a common intensity

3. Modeling framework for logreturns
   - Logreturn Models
   - Simulation studies
   - Brief Conclusion

4. Filtering by RJMCMC
   - Markov Chain Monte Carlo
   - Reversible jump MCMC
Reversible jump MCMC

Introduction

RJMCMC algorithm is a generalization of the Metropolis-Hasting algorithm that supplies a sample from a target distribution $\pi$ on the spaces of the form $\mathbf{C} = \bigcup_{n=1}^{\infty} C_n$, where $C_n = \{n\} \bigcup R^{k_n}$, $k_n \in \mathbb{N}$, in which subspaces of different dimension $k_n$ are combined.

RJMCMC

We define $r$ transition moves to choose the next state of the chain, where some of the moves involve vectors of different dimensions. Thus, RJMCMC algorithm update proceeds with the following steps:

i) choose one of the moves;

ii) propose a new state;

iii) accept the new state with a certain probability $A$ such that the Markov chain converges.
# Reversible jump MCMC

## Introduction

RJMCMC algorithm is a generalization of the Metropolis-Hastings algorithm that supplies a sample from a target distribution $\pi$ on the spaces of the form $\mathbf{C} = \bigcup_{n=1}^{\infty} C_n$, where $C_n = \{n\} \cup R^{k_n}$, $k_n \in \mathbb{N}$, in which subspaces of different dimension $k_n$ are combined.

## RJMCMC

We define $r$ transition moves to choose the next state of the chain, where some of the moves involve vectors of different dimensions. Thus, RJMCMC algorithm update proceeds with the following steps:

1. choose one of the moves;
2. propose a new state;
3. accept the new state with a certain probability $A$ such that the Markov chain converges.
Move types of RJMCMC

$$\lambda_t = \lambda_0 \exp(-kt) + \sum_{j=0}^{N_t'} X_j \cdot \exp(-k(t - \tau_j))$$
TingTing Peng  
Modeling Multivariate UHF Data
Bivariate DSPP model filtering by RJMCMC

Filtering of the Intensity by RJMCMC, T=100, run for 100000 iteration with 50000 burn-in a1=1 a2=1

TingTing Peng
Modeling Multivariate UHF Data
Bivariate DSPP model filtering by RJMCMC

Filtering of the Intensity by RJMCMC, $T=100$, run for 100000 iteration with 50000 burn-in, $a_1=1$, $a_2=1$.

- True (simulated) trajectories of the common intensity
- Filtering expectation of the common intensity

True (simulated) trajectories of the first intensity
- Filtering expectation of the first intensity

True (simulated) trajectories of the second intensity
- Filtering expectation of the second intensity

TingTing Peng
Modeling Multivariate UHF Data
A new model for multivariate DSPP based on a common component for the intensities.

The intensities of DSPPs are not observable; RJMCMC methods are applied for filtering.

VAR(1) provides correlations between two assets close to empirical ones (EPPs effect).

Further research: Parameter estimation by Monte Carlo Expection-Maximization (MCEM) methods based on the RJMCMC algorithm.
Bibliography


TingTing Peng Modeling Multivariate UHF Data


