Cours de Master - Paris 6

Transparents des Parties IV

Tick by tick financial time series

Emmanuel Bacry
Centre de Mathématiques Appliquées
Ecole Polytechnique
emmanuel.bacry@polytechnique
Non uniformly sampled (1d) time-series
Which "tick" to choose?
- Last traded price
- Mid price
- Best bid/ask prices
- ...

It is an arbitrary projection of a complex dynamics.
ACD Autoregressive Conditional Duration model (Engle, Russel 1997)

- Forex rate Mark/Dollar: 51 days May-August 1993
- Best bid/ask prices - 303408 observations
- Average of 15s between each quote
- Strong intraday seasonality
Robert F. Engle, Jeffrey R. Russell

⇒ Simple "Deseasonalizing" by dividing the duration by the average duration

ACD (Engle, Russel 1997) : intraday seasonality
ACD (Engle, Russel 1997) : Autocorrelation function

<table>
<thead>
<tr>
<th>lag</th>
<th>acf</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.083</td>
</tr>
<tr>
<td>2</td>
<td>0.076</td>
</tr>
<tr>
<td>3</td>
<td>0.064</td>
</tr>
<tr>
<td>4</td>
<td>0.053</td>
</tr>
<tr>
<td>5</td>
<td>0.059</td>
</tr>
<tr>
<td>6</td>
<td>0.048</td>
</tr>
<tr>
<td>7</td>
<td>0.050</td>
</tr>
<tr>
<td>8</td>
<td>0.038</td>
</tr>
<tr>
<td>9</td>
<td>0.048</td>
</tr>
<tr>
<td>10</td>
<td>0.040</td>
</tr>
<tr>
<td>11</td>
<td>0.049</td>
</tr>
<tr>
<td>12</td>
<td>0.043</td>
</tr>
<tr>
<td>13</td>
<td>0.039</td>
</tr>
<tr>
<td>14</td>
<td>0.040</td>
</tr>
<tr>
<td>15</td>
<td>0.042</td>
</tr>
</tbody>
</table>

Robert F. Engle, Jeffrey R. Russell
ACD (Engle, Russel 1997) : The model

\( X_n : \text{duration between two ticks} \)

\[ X_n = m_n \epsilon_n \]

where

- \( \epsilon_n \geq 0, \text{iid} \)
- \( E(\epsilon_n) = 1 \)
- \( m_n \text{ independent from } \epsilon_n \)
- \( m_n \mathcal{F}_{n-1} \text{ measurable} \)

\[ m_n = E(X_n|\mathcal{F}_{n-1}). \]
ACD (Engle, Russel 1997) : The model

\[ m_n = E(X_n | F_{n-1}) . \]
\[ m_n = \omega + \alpha X_{n-1} + \beta m_{n-1} \]

Thus

\[ X_n = m_n + X_n - m_n = \omega + \alpha X_{n-1} + \beta m_{n-1} + X_n - m_n \]
\[ = \omega + (\alpha + \beta)X_{n-1} - \beta(X_{n-1} - m_{n-1}) + X_n - m_n \]

Asymptotically stationary if \( \alpha + \beta < 1 \)

\[ \implies E(X_n) = \frac{\omega}{1 - (\alpha + \beta)} = M \]

\[ \implies X_n - M = (\alpha + \beta)(X_{n-1} - M) + X_n - m_n - \beta(X_{n-1} - m_{n-1}) \]
ACD (Engle, Russel 1997) : The model

\[ X_n - M = (\alpha + \beta)(X_{n-1} - M) + X_n - m_n - \beta(X_{n-1} - m_{n-1}) \]

We set

- \( Y_n = X_n - M \)
- \( W_n = X_n - m_n \) : "innovation" (decorrelated with \( Y_{n-1} \))

ARMA(1,1) equation

\[ Y[n] = (\alpha + \beta)Y_{n-1} + W_n - \beta W_{n-1} \]
ACD based models

Many many extensions ...

- ACM-ACD model (Russel, Engle 2004)
- $ACD(m, q)$ model
- EACD model
- log-ACD
- Burr-ACD
- GACD
- GARCH-ACD
- ...

E. Bacry, CMAP Ecole Polytechnique, 2015
Key issue: (historical) variance/covariance estimation

- Diffusion processes: better estimates at fine scales
  ⇒ one should use high frequency data (tick data)

- **However**: microstructure

  - Price processes are point processes
  - Prices "live" on a *tick grid*
  - **Strong mean reversion at very small scales**
  - Some references

  - **In economics**: Roll (1984) [Roll model], Glosten (1987), Glosten et Harris (1988), Harris (1990)
  - **In statistics and econometrics**: Gloter and Jacod (2001), Ait-Sahalia, Myland et Zhang (2002-2006)
**Figure**: Bund 10Y, 6 Feb 2007, 09:00–10:00 (UTC) 1 data per second.
Variance estimators increase when going to high frequency

- $X(t)$: price (last traded price or mid-price or ...)
- Daily "variance" estimator:

$$V_{\Delta t} = \frac{1\text{day}/\Delta t}{\sum_{n=0}^{1\text{day}/\Delta t}} |X((n + 1)\Delta t) - X(n\Delta t)|^2$$

- Bund 10Y 21 days, 9-11 AM - Last Traded Ask - 7000 points
Mutually exciting point process models

Hawkes processes . . .

Back to the black board!
A 2d Hawkes model for microstructure

E.B., S.Delattre, M.Hoffmann, J.F.Muzy
(Quant Finance 2012 + SPA 2013)

General form of the MEP price model

- \( X_t = N_t^+ - N_t^- \) with
  \[ N_t = \begin{pmatrix} N_t^+ \\ N_t^- \end{pmatrix}, \quad \lambda_t^N = \begin{pmatrix} \lambda_t^{N^+} \\ \lambda_t^{N^-} \end{pmatrix}, \quad u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \]

- \( \Phi^N(t) = \begin{pmatrix} \varphi_{N,s}(t) & \varphi_{N,c}(t) \\ \varphi_{N,c}(t) & \varphi_{N,s}(t) \end{pmatrix} \)

- \( \lambda_t^N = \mu.u + \Phi^N \star dN_t. \)

- **Stability** \( \iff \rho(||\Phi^N||) < 1 \), where we defined
  \[ ||\Phi^N(t)|| = \begin{pmatrix} ||\varphi_{N,s}(t)||_1 & ||\varphi_{N,c}(t)||_1 \\ ||\varphi_{N,c}(t)||_1 & ||\varphi_{N,s}(t)||_1 \end{pmatrix} \]
Simulation over 10 hours + Zoom on 1h

![Graph showing microstructure](image)

**Microstructure "Stylized facts"**

- Point processes (Hawkes) diffusing at large scales
- Prices "live" on a *tick grid*
- Strong mean reversion at very small scales

E. Bacry, CMAP Ecole Polytechnique, 2015

Part III- Tick by tick financial time series
What about modeling two correlated assets?

**Figure**: Bund 10Y / Bobl 5Y
A 4d Hawkes process for modeling 2 correlated assets

Diffusive correlation $C_{\Delta t=+\infty} = 10\%$
A 4d Hawkes process for modeling 2 correlated assets

Asymptotic correlation $C_{\Delta t=+\infty} = 60\%$
The mean signature plot in the dimension 1 model

- The signature plot:
  \[ V_{\Delta t} = \sum_{n=0}^{1 \text{day}/\Delta t} |X((n+1)\Delta t) - X(n\Delta t)|^2 \]

- The mean signature plot:
  \[ E(V_{\Delta t}) = \frac{1}{\Delta t} E(|X((n+1)\Delta t) - X(n\Delta t)|^2) = \frac{1}{\Delta t} E(X(\Delta t)^2) \]

with initial condition: \( X(0) = 0 \)

- Closed-form formula for the mean signature plot when \( \Phi(x) = \alpha e^{-\beta x} \) (through the explicit computation of the Bartlett spectrum (1963)).
Closed form for the mean signature plot

\[ \lambda^{\pm}(t) := \frac{\mu}{2} + \int_{[0,t]} \phi(t - s)dN_s^\mp \]
\[ \phi(t) = \alpha e^{-\beta t}1_{[R^+]}(t), \quad \|\phi\|_1 = \frac{\alpha}{\beta} < 1. \]

\[ E(V_{\Delta t}) = \Lambda \left[ \nu^2 + (1 - \nu^2) \frac{1 - e^{-\gamma \Delta t}}{\gamma \Delta t} \right], \]

where

\[ \Lambda = \frac{\mu}{1 - \|\phi\|_1}, \quad \nu = \frac{1}{1 + \|\phi\|_1} \quad \text{and} \quad \gamma = \alpha + \beta \]

\( \Rightarrow \)

\[ E(V_{\Delta t=0}) = \Lambda = 2E(\lambda^{\pm}) = "\text{microstructural}" \text{ variance} \]
\[ E(V_{\Delta t=+\infty}) = \Lambda \nu^2 = "\text{diffusive}" \text{ variance} \]
Signature plot on 11 hours simulated data
Mean signature plot on real data

- Bund 10Y: 21 days, 9-11 AM - Last Traded Ask (7000 points)
Bund 10Y: 21 days, 9-11 AM - Last Traded Ask

Mean square regression fit

⇒ Very good modelization of the 1d microstructure noise.
The mean Epps effect the dimension 2 model

- Daily "correlation" estimator: $C_{\Delta t} = \tilde{C}_{\Delta t}/\tilde{C}_0$

$$\tilde{C}_{\Delta t} = \sum_{n=0}^{1\text{day}/\Delta t} (X((n+1)\Delta t) - X(n\Delta t))(Y((n+1)\Delta t) - Y(n\Delta t))$$

- The mean Epps effect

$$MEpps_{\Delta t} = \frac{E(X(\Delta t)Y(\Delta t))}{\sqrt{E(X(\Delta t)^2)E(Y(\Delta t)^2)}}$$  \hspace{1cm} (1)

with initial condition: $X(0) = 0$

- **Closed-form formula for the mean Epps effect** when \(\Phi_{X,X}, \Phi_{Y,Y}, \Phi_{X,Y}, \Phi_{Y,X}\) are of the form $\alpha e^{-\beta x}$

→ through the explicit computation of the **Bartlett spectrum** (1963).
Closed form formula for the mean Epps effect in dimension 2

- General case $\rightarrow$ too many parameters
- Reducing the parameters
  - $\mu_X, \mu_Y$
  - $\alpha_{\text{same}} = \alpha_{X,X} = \alpha_{X,Y}$
  - $\alpha_{\text{cross}} = \alpha_{X,Y} = \alpha_{Y,X}$
  - $\beta = \beta_{X,Y} = \beta_{Y,X} = \beta_{X,X} = \beta_{Y,Y}$

$\rightarrow$ Sorry : The formula is at least ... 6 slides long!
Mean Epps effect on 50 hours simulated data

\[ ME_{\text{Epps}} \Delta t \]

\[ \Delta t \text{ (seconds)} \]
Mean Epps effect on real data

**Bund 10Y / Bobl 5Y** : 41 days, 9-11 AM - Last Traded

\[
\alpha_{Bobl} = \alpha_{Bund}.
\]
Accounting for market impact of a labeled agent

- Agent at time $t$: $dA^+(t)$ buy orders $dA^-(t)$ sell orders

  \[ dA(t) = \begin{pmatrix} dA^+(t) \\ dA^-(t) \end{pmatrix} \]

- Impacts will be modeled by additive terms on $\lambda^{N^+}, \lambda^{N^-}$

- **Single buy order** at time $t_0$: $dA^+(t) = \delta(t - t_0), dA^- = 0$
  - Impact on upward jumps: $\lambda^+_t \rightarrow \lambda^+_t + \varphi^{l,s}(t - t_0)$
    $\rightarrow$ ”Instantaneous” impact of the trade itself
    $\rightarrow$ delayed upward moves (e.g., cancel orders)
  - Impact on downward jumps: $\lambda^-_t \rightarrow \lambda^-_t + \varphi^{l,c}(t - t_0)$
    $\rightarrow$ delayed downward moves

- **Meta order** starting at time $t_0$

  \[ \lambda^+_t = \mu.u + \Phi^N \star dN_t + \Phi^l \star dA(t), \]

  where $\Phi^l(t) = \begin{pmatrix} \varphi^{l,s} & \varphi^{l,c} \\ \varphi^{l,c} & \varphi^{l,s} \end{pmatrix}$ is the impact kernel
MI(t) = E(X_t - X_{t=0=0}) with X_t = N_t^+ - N_t^-

- Concave impact while trading
  → depend on T, the smaller impact the larger T
- Relaxation after trading
- Is able to reproduce both permanent/non permanent impact
What if all available market orders are anonymous?

Markets generally do not provide labeled data

- Flow of anonymous marker orders $T_t = \left( \begin{array}{c} T_t^+ \\ T_t^- \end{array} \right)$
  
  $T^+$ (resp. $T^-$) : trade arrivals at the best ask (resp. bid)

- No more access to market impact profile!

- Only access to the Response function : $R(t - t_0)$:
  
  Expectation of the price at time $t$ knowing there was a buying market order at time $t_0$, i.e.,

  $$R(t - t_0) = E(N_t^+ - N_t^- | dT_{t_0}^+ = \delta(t - t_0))$$

Towards a model for market impact of anonymous market orders?
Towards a model for market impact of anonymous market orders

Markets generally do not provide labeled data

- Flow of anonymous marker orders $T_t = \begin{pmatrix} T^+_t \\ T^-_t \end{pmatrix}$
  - $T^+$ (resp. $T^-$): trade arrivals at the best ask (resp. bid)
- The Price model with a label agent

$$\lambda^N_t = \mu . u + \Phi^N \star dN_t + \Phi^I \star dA(t)$$

- The Price model with the anonymous market order flow

$$\lambda^N_t = \Phi^N \star dN_t + \Phi^I \star dT(t)$$
Markets generally do not provide labeled data

- Flow of anonymous marker orders \( T_t = \left( \frac{T_t^+}{T_t^-} \right) \)
  \( T^+ \) (resp. \( T^- \)) : trade arrivals at the best ask (resp. bid)

- The Price model with a label agent

\[
\lambda^N_t = \mu.u + \Phi^N \star dN_t + \Phi' \star dA(t)
\]

- The Price model with the anonymous market order flow

\[
\lambda^N_t = \Phi^N \star dN_t + \Phi' \star dT(t)
\]
Towards a model for market impact of anonymous trades

E.B, J.F.Muzy (QF 2014)

- **The anonymous market orders flow** \( T_t = \begin{pmatrix} T_t^+ \\ T_t^- \end{pmatrix} \)
  
  \( T^+ \) (resp. \( T^- \)) : trade arrivals at the best ask (resp. bid)

- **The Price model**

  \[ X_t = N_t^+ - N_t^- \]

  \( N^+ \) (resp. \( N^- \)) : upward (resp. downward) jumps

\[ \lambda^N_t = \Phi^N \star dN_t + \Phi^I \star dT(t) \]

\( \Phi^I \) : "Instantaneous" impact + influence on price moves

\( \Phi^N \) : Influence of past price moves on future price moves
The model for anonymous trades

E.B, J.F.Muzy (QF 2014)

The anonymous trade arrivals model $\rightarrow$ A 2d Hawkes process

$T_t = \begin{pmatrix} T_t^+ \\ T_t^- \end{pmatrix}$

$T^+ \ (\text{resp. } T^-) :$ trade arrivals at the best ask (resp.bid)

$\lambda^T_t = \mu \cdot u + \Phi^T \star dT_t + \Phi^R \star dN_t$

where

$u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\Phi^T = \begin{pmatrix} \varphi^{T,s} & \varphi^{T,c} \\ \varphi^{T,c} & \varphi^{T,s} \end{pmatrix}$ and $\Phi^R = \begin{pmatrix} \varphi^{R,s} & \varphi^{R,c} \\ \varphi^{R,c} & \varphi^{R,s} \end{pmatrix}$

$\rightarrow \mu :$ Anonymous trade intensity

$\rightarrow \Phi^T :$ Auto-correlation of trades

$\rightarrow \Phi^R :$ Retro-influence of price moves on trades
The overall model is a 4 dimensional Hawkes process $P$

E.B, J.F.Muzy (QF 2014)

$$P_t = \begin{pmatrix} T_t \\ N_t \end{pmatrix}$$

whose intensity $\lambda_t = \begin{pmatrix} \lambda^T_t \\ \lambda^N_t \end{pmatrix}$ is given by

$$\lambda_t = M + \Phi \ast dP_t,$$

where

$$M = \begin{pmatrix} \mu \cdot u \\ 0 \end{pmatrix}, \quad \Phi(t) = \begin{pmatrix} \Phi^T(t) & \Phi^R(t) \\ \Phi^I(t) & \Phi^N(t) \end{pmatrix}$$

- $\mu$: Anonymous trade intensity
- $\Phi^T(t)$: Auto-correlation of anonymous trades
- $\Phi^I(t)$: "Instantaneous" impact + influence on price moves
- $\Phi^N(t)$: Influence of past price moves on future price moves
- $\Phi^R(t)$: Retro-influence of price moves on anonymous trades
Non parametric estimation

\[ \lambda_t = M(t) + \Phi \ast dP_t, \]

where

\[ M(t) = \left( \begin{array}{c} \mu \cdot u \\ 0 \end{array} \right), \quad \Phi(t) = \left( \begin{array}{cc} \Phi^T(t) & \Phi^R(t) \\ \Phi^I(t) & \Phi^N(t) \end{array} \right) \]

Non parametric estimation of \( \mu \) and all the kernels \( \Phi^T, \Phi^R, \Phi^N, \Phi^I \), from anonymous market data.
Non parametric estimation of $\Phi^T$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Trade auto-correlation
Non parametric estimation of $\Phi^T$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Zooming ...

Trade auto-correlation :

→ Mainly "positive" correlation : Splitting and Herding
Non parametric estimation of $\Phi'$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Trade "instantaneous" impact + influence on delayed price moves

$\rightarrow$ Mainly instantaneous impact:

$\phi^{I,s}(t) \simeq C\delta(t)$ and $\phi^{I,c} \simeq 0$. 

E. Bacry, CMAP Ecole Polytechnique, 2015
Non parametric estimation of $\Phi^N$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Influence of past price moves on future price moves
Non parametric estimation of $\Phi^N$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Influence of past price moves on future price moves

$\rightarrow$ Mostly mean reverting
Non parametric estimation of $\Phi^R$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Retro-influence of price moves on anonymous trades:

E. Bacry, CMAP Ecole Polytechnique, 2015
Non parametric estimation of $\Phi^R$ for Eurostoxx Futures 10h-12h, 2009-2012 (800 days)

Retro-influence of price moves on anonymous trades:
$\rightarrow \phi^{R,cross}$ large and $\phi^{R,self} \approx 0$

Price goes up $\implies$ more sell market orders
Most kernels are power-law (when non 0): \( \frac{\alpha}{(\delta + x)^\beta} \)

With \( \beta \approx 1 \) : close to unstablility limit!

Except \( \varphi^{1,s} \approx C \delta \) (\( C << 1 \)) and \( \varphi^{R,c} \)
Kernels can be amazingly stable when asset changes

plain line : Eurostoxx Futures,  • : Bund

• No adjustment (no prefactors)!
Intraday seasonalities for Eurostoxx Futures 2009-2012

Log-Log plots of $\varphi^{T,s}$ and $\varphi^{N,c}$ for different intraday slices:
9h-11h, 10h-12h, 11h-13h, 12h-14h, 13h-15h, 14h-16h, 15h-17h

The kernel estimations do not depend on the intraday period
The intraday seasonality is only carried by $\mu$ (U-shape)

Model with intraday seasonality

$$\lambda_t = M(t) + \Phi \ast dP_t,$$

where $M(t) = \begin{pmatrix} \mu_{\text{seasonal}}(t) \\ 0 \end{pmatrix}$, $\Phi(t) = \begin{pmatrix} \Phi^T(t) & \Phi^R(t) \\ \Phi^I(t) & \Phi^N(t) \end{pmatrix}$
Closed analytical formula for many quantities of interest

- Response function
- Diffusive variance of the price
- Auto-correlation function of
  - the trade signs (in practice heavy correlation)
  - the increments of the price (in practice very small correlations)
- Market impact
- ...

E. Bacry, CMAP Ecole Polytechnique, 2015
Analytical formula for the Market Impact of a meta-order

A particular case:

- An "impulsive" Impact kernel: $\varphi^{I,s}(t) = C\delta(t)$, $\varphi^{I,c}(t) = 0$
- A single buy order: $dA^+(t) = \delta(t)$, $dA^-(t) = 0$

$\implies$ the Market impact is

$$MI(t) = E(X_t - X_0) = 1_{[0, +\infty]}(t) - \int_0^t \Delta\xi(u)du,$$

where the Laplace transform of $\Delta\xi(t)$ is given by

$$\hat{\Delta\xi} = 1 - \frac{(1 - \Delta\phi^T)}{(1 - \Delta\phi^T)(1 - \Delta\phi^N) - \Delta\phi^R}$$

where $\Delta\varphi^? = \varphi^{?,s} - \varphi^{?,c}$ measures the "kernel’s imbalance"
Permanent versus non-permanent market impact

Analytical formula for the asymptotic market impact $MI(+\infty)$

In the case of a ”cross-only” Retro-kernel: $\varphi^{R,s}(t) = 0$

$\implies$ The asymptotic market impact is

$$MI(+\infty) = \frac{1}{(1 - \Delta||\varphi^N||_1) + ||\varphi^{R,c}||_1/(1 - \Delta||\varphi^T||_1)},$$

where

$$\Delta||\varphi^T||_1 = ||\varphi^{T,s}||_1 - ||\varphi^{T,c}||_1 (\in ] - 1, 1[ \text{ implied by stability})$$

$$\Delta||\varphi^N||_1 = ||\varphi^{N,s}||_1 - ||\varphi^{N,c}||_1 (\in ] - 1, 1[ \text{ implied by stability})$$

$MI(+\infty)$ decreases when mean reversion increases, i.e. :

- when $\Delta||\varphi^N||_1$ goes to -1
- when $||\varphi^{R,c}||_1$ increases
- when $\Delta||\varphi^T||_1$ goes to 1
Market impact profile estimation from anonymous data on Eurostoxx Futures

- Non parametric estimation of all the kernels: $\Phi^T$, $\Phi^R$, $\Phi^N$, $\Phi^I$

- Setting $\varphi^{T,c} = 0$, $\varphi^{I,c} = 0$ and $\varphi^{R,s} = 0$

- Fitting exponential kernels on $\varphi^{I,s}$ and $\varphi^{R,c}$

- Fitting Power-law kernels on $\varphi^{T,s}$, $\varphi^{N,c}$ and $\varphi^{N,s}$

- Computing the market impact profile from analytical formula
Market impact profile estimation from anonymous data on Eurostoxx Futures
The process \( P_t = \begin{pmatrix} T^-_t \\ T^+_t \\ N^-_t \\ N^+_t \end{pmatrix} \) diffuses at large scales

(from E.B., S.Delattre, M.Hoffmann, J.F.Muzy, preprint 2011)

\[
\frac{1}{\sqrt{h}}(P_{ht} - E(P_{ht})) \rightarrow \text{law} \ (\mathbb{I} - \hat{\Phi}_0)^{-1}\Sigma^{1/2} W_t
\]

where \( W_t \) is a n-dimensional Gaussian process (with stationary increments).

Consequently

- The *Trade process* \( U_t = T^+_t - T^-_t \) diffuses at large scales
- The *Price process* \( X_t = T^+_t - T^-_t \) diffuses at large scales
Trade sign long-range correlations

- $U_t$ diffuses at large scales
- Is it compatible with empirical findings about long range correlations of $U_t$?
  ⇒ Strictly speaking: NO!

However, as long as
- $\Delta \hat{\Phi}^T_0 \simeq 1$ and
- $\varphi^{T,s}_t \sim (c + t)^{-1+\nu}$,
  --- there is a finite range of scales (in practice $\simeq 5$ decades!) on which

$$C^T(\tau) = \text{Cov}(U_t, U_{t+\tau}) \sim \tau^{2\nu-1}$$
Trade sign long-range correlations

\[ C_T(\tau) = \text{Cov}(U_t, U_{t+\tau}) \]

![Graph](image_url)
What about price efficiency?

Price ”long-memory puzzle” (Bouchaud et al. 2004)?

- $U_t$ is long-range correlated on a large range of scales
- How come the price $X_t = N_t^+ - N_t^-$ is not long-range correlated on a large range of scales?

As long as

- $\Delta \hat{\Phi}_{N_0} < 0$ and
- $\Delta \varphi_{N_s} \sim (c' + t)^{-1+\nu'}$, ($\nu' << 1$),

$\implies$ there is a finite range of scales (in practice $\approx 5$ decades!) on which

$$C^N(\tau) = \text{Cov}(X_t, X_{t+\tau}) << 1$$
• (left and right plots) : $C^T(\tau) = \text{Cov}(U_t, U_{t+\tau})$

  ___ (left plot) : $C^N(\tau) = \text{Cov}(X_t, X_{t+\tau})$
Trade sign and price correlations

(after removing the first point of correlation functions)
A microstructure and impact model

- Reproduce microstructure **and** market impact stylized facts (Bund, SP Fut., Euro/$ Fut., Eurostoxx Fut.,...)
- Kernel components can be easily estimated non parametrically
- **Most kernels are heavy-tailed** (as found in K. Al Dayri, E. B, J. F. Muzy, EPJB, 2012)
- Kernel components can be easily interpreted in terms of various dynamics
- **Analytical formula** for many quantities
- **Market impact profile estimation from anonymous data**
- Gives insights about the value of the permanent market impact
- Can be **easily generalized**
  - incorporating trade volumes
  - account for limit/cancel orders
  - Influence of labeled agents on anonymous agents
  - Multiple agents model
  - News model
  - ...

E. Bacry, CMAP Ecole Polytechnique, 2015

Part III- Tick by tick financial time series
Replay of 2 hours of Eurostoxx mid-price from real trades

\[ T_t^+ - T_t^- : \text{True cum. Trades on 3/08/2008 - [10am-12am]} \]
\[ N_t^+ - N_t^- : \text{True mid-price on 3/08/2008 between 10am and 12am} \]

Simulation of the mid-price process \( N \) given the real trades
An 8-dimensionnal model for Level I orderbook events

E.B. T. Jaisson and J.F. Muzy (2014)

- **Database**:  
  - Dax Futures (small tick size)  
  - Bund Futures (large tick size)  
  - 1 year data: 06/2013-06/2014  
  - **time precision = 1 µs**

- $P_t$ is an **8-dimensional counting process**:
  - $PA$ (resp. $PB$): upward (resp. downward) mid-price jumps  
  - $TA$ (resp. $TB$): market orders at the best ask (resp. bid)  
  - $LA$ (resp. $LB$): limit orders at the best ask (resp. bid)  
  - $CA$ (resp. $CB$): cancel orders at the best ask (resp. bid)

<table>
<thead>
<tr>
<th># events/day</th>
<th>PA/PB</th>
<th>TA/TB</th>
<th>LA/LB</th>
<th>CA/CB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dax</strong></td>
<td>72.000</td>
<td>20.000</td>
<td>152.000</td>
<td>184.000</td>
</tr>
<tr>
<td><strong>Bund</strong></td>
<td>14.000</td>
<td>28.000</td>
<td>240.000</td>
<td>212.000</td>
</tr>
</tbody>
</table>
Estimation is based on conditional expectation

$$E(dP_{t}^{LA} \mid dP_{0}^{LA} = 1)$$

MOST of the conditional laws display a peak around $t \simeq 0.25 ms \implies \text{Average Latency}$
Level of exogeneity

Ratio of exogeneous events over all events $R^i = \frac{\mu^i}{\Lambda^i}$

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>PB</th>
<th>TA</th>
<th>TB</th>
<th>LA</th>
<th>LB</th>
<th>CA</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bund</strong></td>
<td>4.4%</td>
<td>4.4%</td>
<td>4.5%</td>
<td>4.5%</td>
<td>1.4%</td>
<td>1.4%</td>
<td>1.6%</td>
<td>1.8%</td>
</tr>
<tr>
<td><strong>Dax</strong></td>
<td>2.7%</td>
<td>2.7%</td>
<td>4.3%</td>
<td>4.5%</td>
<td>1.1%</td>
<td>1.2%</td>
<td>0.7%</td>
<td>0.4%</td>
</tr>
</tbody>
</table>
Kernel $L^1$ Norm

Color coding of the norms $||\Phi\rightarrow?||_1$
Kernel Norm

→ Symmetry upward/downward and ask/bid
Kernel Norm

\[ \Rightarrow \text{Symmetry upward/downward and ask/bid} \]
“Anti-diagonal” shape in the price kernels
⇒ mean reversion of the price
“Anti-diagonal” shape in the price kernels
⇒ mean reversion of the price
• “Anti-diagonal” shape in the price kernels
  ⇒ mean reversion of the price

• “Diagonal” shape in the limit/cancel/trade kernels
  ⇒ splitting/herding
Order flow Kernel Norms

“Anti-diagonal” shape in the price kernels
⇒ mean reversion of the price

“Diagonal” shape in the limit/cancel/trade kernels
⇒ splitting/herding
Shape of some kernels

Power law kernels responsible for

- **price mean reversion**
- **order splitting, herding**

Log-log plots of some kernel estimations on 7 decades
Impact of the order flows on the price

- Trades: main source of impact (diagonal)
- Limits: contrariant
- Cancels: diagonal
Impact of the order flows on the price

- Trades: main source of impact (diagonal)
- Limits: contrariant
- Cancels: diagonal
Price impact of trade flow

Cumulative kernels $\int_0^t \Phi^{T?\rightarrow P?}(s)ds$ as a function of $\log(t)$

- **Impact kernels** $\Phi^{TA\rightarrow PA}$ and $\Phi^{TB\rightarrow PB}$ are very localized
- Localization around “latency value” $\simeq 0.25\text{ms}$
Price impact of limit/cancel flow

Cumulative kernels \( \int_0^t \Phi^{L?\rightarrow P?}(s)ds \) as a function of \( \log(t) \)

- The kernels \( \Phi^{L?\rightarrow P?} \) and \( \Phi^{C?\rightarrow P?} \) are not localized.
- The kernels \( \Phi^{L?\rightarrow P?} \) and \( \Phi^{C?\rightarrow P?} \) are not localized.
Market Price “efficiency”

⇒ Market Price efficiency comes from a “rough” equilibrium between the 4 main power law kernels
Impact on the trades

- Impact of the Price: Bund (contrariant), Dax (diagonal)
- Large tick size: a change in price carries much more information
- Impact of the Limit is very small (actually trades are leading)
- Impact of the Cancel is very small
Impact on the trades

- Impact of the Price: Bund (contrariant), Dax (diagonal)
- Large tick size: a change in price carries much more information
- Impact of the Limit is very small (actually trades are leading)
- Impact of the Cancel is very small
Impact of the price on trade flow of the Bund

Cumulative kernels $\int_{0}^{t} \Phi(s)ds$ as a function of $\log(t)$

- Price goes up $\Rightarrow$ agents buy less and sell more
Impact of the price on trade flow of the Bund

Cumulative kernels $\int_0^t \Phi(s)ds$ as a function of $\log(t)$

- Price goes up $\Rightarrow$ agents buy less and sell more
Impact of the price on order flows

- Impact on Trade: Bund (contrariant), Dax (diagonal)
- Impact on Limit: contrariant
- Impact on Cancel: diagonal
Impact of the price on order flows

- Impact on Trade: Bund (contrariant), Dax (diagonal)
- Impact on Limit: contrariant
- Impact on Cancel: diagonal
Impact of the price on limit flow of the Bund

Cumulative kernels $\int_0^t \Phi(s)ds$ as a function of $\log(t)$

- Price goes up $\Rightarrow$ Market maker reaction
Impact of the price on limit flow of the Bund

Cumulative kernels $\int_0^t \Phi(s)ds$ as a function of $\log(t)$

- **Price goes up $\Rightarrow$ Market maker reaction**
Kernel components can be easily estimated non parametrically
Stable even for slightly negative valued kernels
**Kernel components can be easily interpreted in terms of various dynamics**
- Latency appears clearly in some kernels
- Mean-reversion of price
- Strong localized price impact of trades
- Very weak non-localized price impact of limits and cancels
- Contrariant impact of price changes on trade flow
- Market maker reactions to price change
- ...