Disentangling endogenous and exogenous correlation effects via high frequency information

Thomas Deschatre[1,2]
Joint work with E. Bacry, M. Hoffmann, J.F. Muzy, and R. Ruan

[1]EDF Lab
[2]FiME Lab

October 2024

Context

- Correlation between two time series X^1 and X^2 can be created by:
  - Endogenous events: direct causality of X^1 on X^2 or X^2 on X^1 ;
  - Exogenous events: an event X^3 both affects X^1 and X^2.
- Example: two time series of financial prices:
  - Endo: traders react on market 2 because price on market 1 moved ;
  - Exo: some economical news affects both market.
- Example: intraday electricity prices for two different delivery hours:
  - Endo: traders react on market 2 because price on market 1 moved ;
  - Exo: power plant shutdown for the two hours.

.

- How to quantity the percentage of exogeneity in the correlation ?
  - Without observing exogenous events.

Correlation estimation at a macroscopic scale

- Correlation estimation for semimartingales:

$$dX_t^i = \mu_t^i dt + \sigma^i dW_t^i, \; i = 1, 2,$$

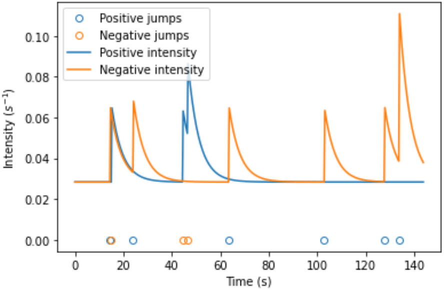
$$d < X^1, X^2 >_t = \rho dt$$

observed on a grid $(i\Delta_n)_{i=0,\dots,\lfloor \frac{T}{\Delta_n} \rfloor}$.

- Estimator $\hat{\rho} = \frac{<X^1,X^2>}{\sqrt{<X^1><X^2>}}$
- $\hat{\rho} \underset{\mathbb{P}}{\rightarrow} \rho$ when $\Delta_n \rightarrow 0$ (speed $\Delta_n^{1/2}$) Aït-Sahalia and Jacod (2014).
- The quantity ρ accounts for both endogenous and exogenous effects.
- How to disentangle them ? Does it even make any sense ?

Hawkes process

- Point process (N_t) with intensity $\lambda_t = \mu + \int_0^t \varphi(t-s)dN_s$.
- μ in $\mathbb{R}^d$ the baseline,
- $(\varphi_{lk})_{1 \leq l,k \leq d}$ the kernel matrix locally integrable.
- To have LLN and CLT, spectral norm of $\|\varphi\|_1 < 1$.

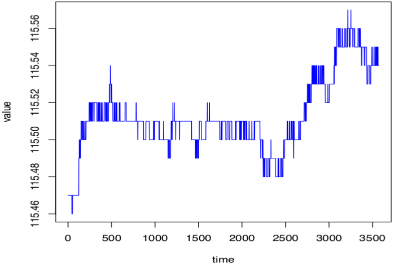


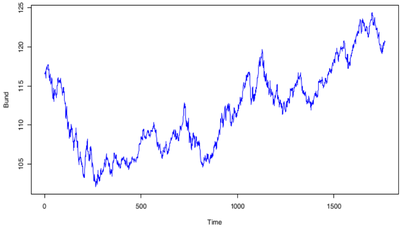
Intensity trajectory in the model $\lambda_t^{\pm} = \mu + \int_0^t \varphi(t-s)dN_s^{\mp}$.

Volatility

- Population point of view (dim 1):
  - A Poisson process gives birth to parents with rate μ.
  - Each parent gives birth to children as an inhomogenous Poisson process with intensity $\varphi(a)$, a being the age of the parent.
  - Each child gives birth to children in the same way.
- Parents = exogenous events, $\mathbb{E}(N_T^{\text{exo}})/T \approx \mu$,
- Children = endogenous events, $\mathbb{E}(N_T^{\text{endo}})/T \approx \frac{\mu\|\varphi\|_1}{1-\|\varphi\|_1}$.
- For T large, $\sqrt{T}\left(N_{tT}/T - \frac{\mu t}{1-\|\varphi\|_1}\right) \to \sigma W_t$, $\sigma^2 = \frac{\mu}{(1-\|\varphi\|_1)^3}$.
- Using microscopic data, one can infer μ and $\|\varphi\|_1$, ...
- then disentangle exo and endo parts of the volatility.

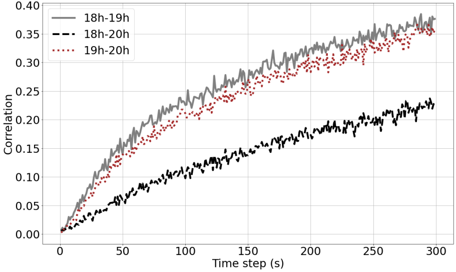
Volatility





German 10Y Bund, 1 data per second (left) and one per day (right), from Hoffmann 2018.

Epps and Hawkes



Correlation with respect to timestep sampling for intraday electricity prices Deschatre and Gruet (2022).

- Correlation depends on scale :

$$\rho(\Delta) = \frac{\sum_{i=1}^{T/\Delta}(X^1_{i\Delta} - X^1_{(i-1)\Delta})(X^2_{i\Delta} - X^2_{(i-1)\Delta})}{\sqrt{\sum_{i=1}^{T/\Delta}(X^1_{i\Delta} - X^1_{(i-1)\Delta})^2 \sum_{i=1}^{T/\Delta}(X^2_{i\Delta} - X^2_{(i-1)\Delta})^2}}.$$

- Null correlation for high frequencies (no events at the same time),
- Then stabilization.
- Hawkes process can represent this feature Bacry et al. (2013b).

Epps and Hawkes

- Population point of view (dim 2):
  - Two PP gives birth to parents of type i with rate μ_i, $i = 1, 2$.
  - Each parent of type j gives birth to children of type i as an non-homogeneous Poisson process with intensity $\varphi_{i,j}(a)$, a being the age of the parent.
  - Each child gives birth to children in the same way.
- N^i is the sum of all the events of type i.
- Parents = exogenous events, $\mathbb{E}(N_T^{exo,1} + N_T^{exo,2})/T \approx \mu_1 + \mu_2$.
- But exo events 1 are not correlated with exo events 2.
- **Correlation is purely endogenous.**
- The Hawkes modeling framework is not sufficient.
- **How to correlate the two exogenous Poisson processes ?**

Outline

Outline

Common shock model

- Powojowski et al. (2002); Lindskog and McNeil (2003)
- Consider three independent PP, M^i with intensities μ_i
- Let $N^i = M^i + M^3$, $i = 1, 2$.
- Then N^1 and N^2 are marginally Poisson processes,
- and are correlated with $\rho = \frac{\mu_3}{\sqrt{(\mu_1+\mu_3)(\mu_2+\mu_3)}}$.
- But no Epps effect : $\rho(\Delta)$ does not depend on Δ (when $T \to \infty$).
- And jumps happen simultaneously for 1 and 2
  - not consistent with null correlation at small time scales.

Definition

- Cox and Lewis (2005)
- Let M^3 be a PP with intensity μ_3 and jumps $(T^3_k)_{k\geq 1}$.
- For $i = 1, 2$, let $M^{3,i}_t = \sum_{k\geq 1} \mathbf{1}_{T^3_k + \epsilon^i_k \leq t}$ with
  - $(\epsilon^i_k)_{k\geq 1}$ two independent iid sequences of positive r.v.,
  - independent from M^3,
  - exponentially distributed with parameter $a > 0$.
- ϵ^i consists in delays that we add to the PP M^3.
- In their own filtration, $M^{3,i}$, $i = 1, 2$, have intensity

$$\mu_3(1 - \exp(-at))$$

and are asymptotically Poisson.

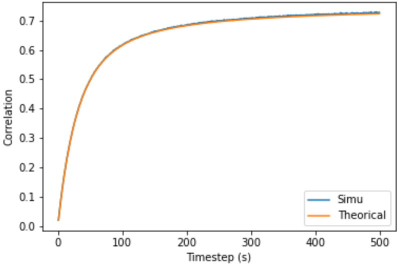
Epps

- Now consider the Common Shock Model

$$N^i = M^i + M^{3,i},\ i = 1, 2.$$

- In their own filtration, N^i, $i = 1, 2$ are asymptotically Poisson processes and
- For T large and $\Delta_T / T \to 0$.

$$\rho(\Delta_T) \sim \frac{\mu_3}{\sqrt{(\mu_1 + \mu_3)(\mu_2 + \mu_3)}} \left(1 - \frac{1 - e^{-a\Delta_T}}{a\Delta_T}\right).$$



Epps effect in the Common Shock delayed model.

The DPP as a Hawkes process

- $(M^3, M^{3,1}, M^{3,2})$ is a point process with
  - no common jumps
  - and intensity (Daley et al., 2003, Example 7.3(a) p.250)

$$\begin{cases} \lambda_t^3 = \mu_3 \\ \lambda_t^{3,1} = a\left(M_t^3 - M_t^{3,1}\right) \\ \lambda_t^{3,2} = a\left(M_t^3 - M_t^{3,2}\right) \end{cases} .$$

- "Hawkes" process with
  - baseline $(\mu_3, 0, 0)$
  - and kernel $\varphi(t) = a \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$.
- **Easy to include in a Hawkes framework.**
- But negative components in the kernel and $\|\varphi\|_1 = \infty$.

Outline

Construction

- We want a model for N^i, $i = 1, 2$, with endo and exo correlation.
- Population point of view (dim 2):
  - Two point processes $N^{\text{exo},i}$ gives birth to parents of type i, $i = 1, 2$.
  - Each parent of type j gives birth to children of type i as an non-homogeneous Poisson process with intensity $\varphi_{i,j}(a)$, a being the age of the parent.
  - Each child gives birth to children in the same way.
- N^i is the sum of all the events of type i.
- $N^{\text{exo},i}$ are now constructed from a Common Shock Delayed
  - Marginally, each is a Poisson process (asymptotically),
  - with intensity $\mu_i + \mu_3$
  - They are correlated,
  - with no common jump times.
- Now exogenous correlation
  - from correlation between the parents (exogenous events) $N^{\text{exo},i}$.

Intensity

- We write $N^{\text{exo},i} = M^i + N^{3,i}$, $i = 1, 2$
- With $N^{3,i}$ a delayed version of a PP N^3.
- Let $N^{H,i} = N^i - N^{3,i}$.
- $(N^{H,1}, N^{H,2}, N^3, N^{3,1}, N^{3,2})$ has intensity

$$\begin{cases} \lambda_t^{H,1} = \mu_1 + \int_0^t \varphi_1(t-s)d(N^{\text{exo},1} + N^{H,1}) + \int_0^t \varphi_{12}(t-s)d(N^{\text{exo},2} + N^{H,2}) \\ \lambda_t^{H,2} = \mu_2 + \int_0^t \varphi_2(t-s)d(N^{\text{exo},2} + N^{H,2})) + \int_0^t \varphi_{21}(t-s)d(N^{\text{exo},1} + N^{H,1}) \\ \lambda_t^3 = \mu_3 \\ \lambda_t^{3,1} = a\left(N_t^3 - N_t^{3,1}\right) \\ \lambda_t^{3,2} = a\left(N_t^3 - N_t^{3,2}\right) \end{cases}.$$

- Still a (degenerated) Hawkes process.

Validity of results on Hawes

- Results of Bacry et al. (2013a) still valid
  - Law of large numbers
  - CLT
  - Convergence of empirical moments.
- Whereas $\|\varphi\|_1 = \infty$ and we have negative component.
- Sketch of the proof:
  - Sufficient condition spectral radius of $\|\varphi\|_1 < 1$ too strong
  - We can replace it by the existence of $\sum_{k\geq 1} \varphi^{(\star k)}$ and its L^1 norm
  - For the sub-matrix of φ, $\tilde{\varphi}(t) = a \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$,

$$\sum_{k\geq 1} \tilde{\varphi}^{(\star k)}(t) = ae^{-at} \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

Outline

Disentangling

- We have a CLT towards a Brownian motion with covariance matrix

$$\mathrm{cov}(N_a, N_b) = \sum_{k=1,2,3} \Lambda_k \left(\sum_{i \in \{a,3\}} R^{ik} \right) \left(\sum_{j \in \{b,3\}} R^{jk} \right), \; a, b = 1, 2.$$

- Depends only on $\|\varphi\|_1$ and μ_i.
- Λ_i corresponds to the mean number of events of type i:
- R^{ij}: mean number of events i triggered by one event j:
- Exogenous part of the covariance:
  - Population interpretation:

$$\underbrace{\Lambda_3}_{\mu_3} \underbrace{\left(\sum_{i \in \{1,3\}} R^{i3} \right)}_{\substack{\text{mean number of events of } N_1 \\ \text{triggered by one exogenous event}}} \underbrace{\left(\sum_{j \in \{2,3\}} R^{j3} \right)}_{\substack{\text{mean number of events of } N_2 \\ \text{triggered by one exogenous event}}} .$$

  - Probabilistic interpretation with law of total covariance:

$$= \mathrm{cov}\left(\mathbb{E}\left(N_1 | \sigma(N_3)\right), \mathbb{E}\left(N_2 | \sigma(N_3)\right)\right).$$

Outline

Method of Moments

- To disentangle macroscopic correlation, we need :
  - μ_1, μ_2, and μ_3,
  - $\|\varphi_{i,j}\|_1$.
- Method of moments from Achab et al. (2017) which is still valid.
- Use of the first three order moments at a macroscopic scale.
- Results on simulation satisfying.
- Results on data : work in progress.

Perspectives

- Estimation on real financial dataset (CAC40),
- Application to intraday electricity prices:
  - Hawkes process for univariate price in Deschatre and Gruet (2022),
  - Common Shock Model in Deschatre and Warin (2023).
- Estimation of φ and not only $\|\varphi\|$ (EM algorithm ?)

Thank you for your attention.

Bibliography I

Achab, M., Bacry, E., Gaıffas, S., Mastromatteo, I., and Muzy, J.-F. (2017). Uncovering causality from multivariate hawkes integrated cumulants. In *International Conference on Machine Learning*, pages 1–10. PMLR.

Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.

Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2013a). Modelling microstructure noise with mutually exciting point processes. *Quantitative finance*, 13(1):65–77.

Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2013b). Some limit theorems for hawkes processes and application to financial statistics. *Stochastic Processes and their Applications*, 123(7):2475–2499.

Cox, D. and Lewis, P. (2005). Multivariate point processes. *Selected Statistical Papers of Sir David Cox*, 1:159.

Daley, D. J., Vere-Jones, D., et al. (2003). *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer.

Deschatre, T. and Gruet, P. (2022). Electricity intraday price modelling with marked hawkes processes. *Applied Mathematical Finance*, 29(4):227–260.

Bibliography II

Deschatre, T. and Warin, X. (2023). A common shock model for multidimensional electricity intraday price modelling with application to battery valuation. *arXiv preprint arXiv:2307.16619*.

Lindskog, F. and McNeil, A. J. (2003). Common poisson shock models: applications to insurance and credit risk modelling. *ASTIN Bulletin: The Journal of the IAA*, 33(2):209–238.

Powojowski, M. R., Reynolds, D., and Tuenter, H. J. (2002). Dependent events and operational risk. *Algo Research Quarterly*, 5(2):65–73.