



CentraleSupélec

FINANCIAL DATA ANALYSIS VIA NON UNIFORM FOURIER TRANSFORM

Imane EL BOUZID
Richard LIN
Anas HAIMOUD

4 Juin 2021

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. Estimation de la covariance instantanée
 6. Application sur des données réelles
 7. Conclusion

INTRODUCTION

Estimation de variances et covariances :

- Trading Haute Fréquence : $113 \mu\text{s}$ soit 27612 données en moyenne par jour
 - Bruits de microstructure
 - Dérivation numérique
- ⇒ **Biais d'estimation**

Estimateur de Malliavin-Mancino :

- Estimateur pour données synchrones et asynchrones
- Estimateur non-paramétrique
- Intégration numérique

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. Estimation de la covariance instantanée
 6. Application sur des données réelles
 7. Conclusion

ESTIMATEUR DE MALLIAVIN-MANCINO

Théorème

Sous certaines hypothèses :

$$\frac{1}{2\pi} \mathcal{F}\left(\Sigma^{l,j}\right) = \mathcal{F}\left(dp^l\right) *_B \mathcal{F}\left(dp^j\right)$$

Estimateurs avec inversion de Dirichlet

$$\Sigma_{M,N}^{l,j}(t) = \sum_{k=-M}^{k=M} \frac{2\pi}{2N+1} \sum_{s=-N}^{s=N} \mathcal{F}(dp_l)(s) \mathcal{F}(dp_j)(k-s) \exp(ikt)$$

$$\Sigma_{M,N}^{l,j} = \frac{(2\pi)^2}{2N+1} \sum_{s=-M}^{s=M} \mathcal{F}(dp_l)(s) \mathcal{F}(dp_j)(-s)$$

ESTIMATEUR DE MALLIAVIN-MANCINO

Estimateurs avec inversion de Fejér

$$\Sigma_{M,N}^{l,j}(t) = \sum_{k=-M}^{k=M} \left(1 - \frac{|k|}{M}\right) \frac{2\pi}{N+1} \sum_{s=-N}^{s=N} \mathcal{F}(dp_l)(s) \mathcal{F}(dp_j)(k-s) \exp(ikt)$$

$$\Sigma_{M,N}^{l,j} = \frac{(2\pi)^2}{N+1} \sum_{s=-M}^{s=M} \left(1 - \frac{|s|}{M}\right) \mathcal{F}(dp_l)(s) \mathcal{F}(dp_j)(-s)$$

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. Estimation de la covariance instantanée
 6. Application sur des données réelles
 7. Conclusion

CALCUL DES COEFFICIENTS DE FOURIER

Calcul initial proposé par Malliavin et Mancino

$$\begin{aligned}\mathcal{F}(dp_j)(k) &= \frac{1}{2\pi} \sum_{l=0}^{n_j-1} \exp(-ikt_l) \int_{[t_l, t_{l+1}]} dp_j(t) \\ &= \frac{1}{2\pi} \sum_{l=0}^{n_j-1} e^{-ikt_l^j} \delta_l(p_j)\end{aligned}$$

⇒ Complexité en $O(n_j^2)$!

CALCUL DES COEFFICIENTS DE FOURIER

Calcul par transformée de Fourier non uniforme rapide

$$\begin{aligned}\mathcal{F}(f)(k) &= \frac{1}{2\pi} \sum_{n=0}^{N-1} f_n e^{-ikx_n} = \frac{1}{2\pi} \int \left(\sum_{n=0}^{N-1} f_n \delta(t - x_n) \right) e^{-ikt} dt \\ &:= \mathcal{F}(g)(k) \quad \text{où } \forall x, g(x) := \sum_{j=0}^{N-1} f_j \delta(x - x_j)\end{aligned}$$

⇒ On applique la FFT à $\tilde{g} := g * \tilde{\varphi}$ où $\tilde{\varphi}(x) := \sum_{r \in \mathbb{Z}} \varphi(x - rp)$ et φ est un noyau périodique et régulier de période p

⇒ La complexité est de $O(N \log(N) + \frac{\sigma}{\epsilon} N)$!

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. Estimation de la covariance instantanée
 6. Application sur des données réelles
 7. Conclusion

DISTRIBUTION ASYMPTOTIQUE DE L'ERREUR

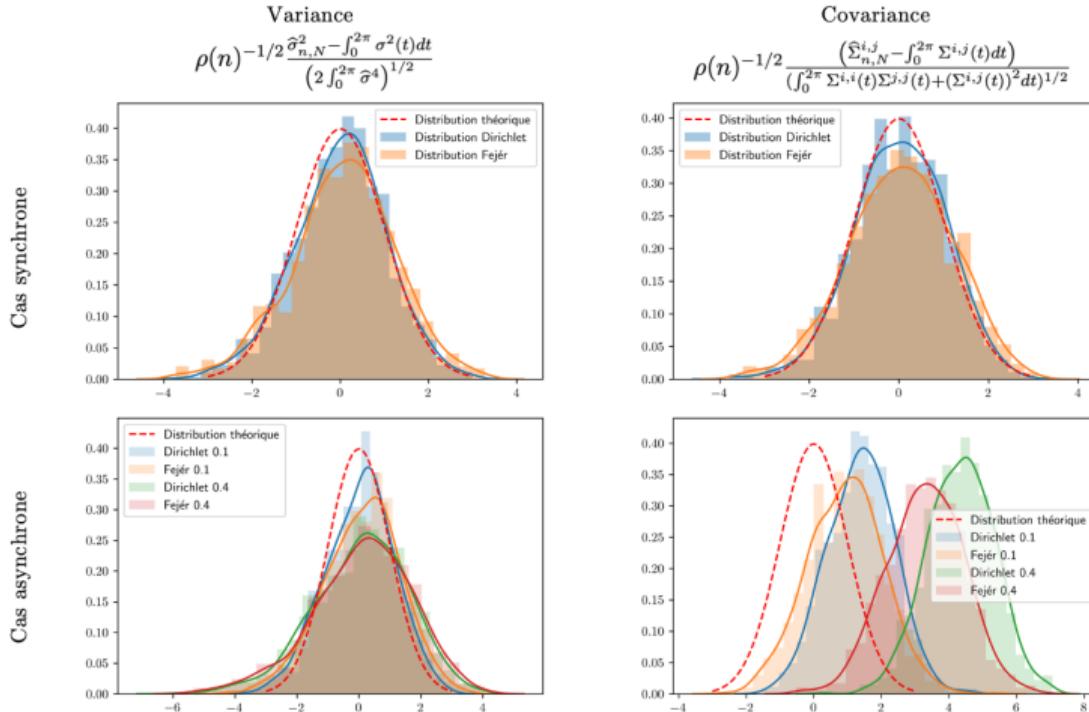
Théorème

Si $N = O(n_i^\gamma)$ et $M = O(n_j^\gamma)$ et $N\Delta(n_i, n_j) \rightarrow 0$ en notant $\Delta(n_i, n_j)$ le pas maximal entre les observations alors :

$$\rho(n_i)^{-1/2\gamma} \left(\widehat{\sigma}_{n_i, M, N}^2 - \int_0^{2\pi} \sigma^2(t) dt \right) \xrightarrow{n_i, N, M \rightarrow \infty} \mathcal{N} \left(0, 2 \int_0^{2\pi} \sigma^4(t) dt \right),$$

$$\rho(n)^{-\frac{1}{2\gamma}} \left(\widehat{\Sigma}_{n, N}^{i,j} - \int_0^{2\pi} \Sigma^{i,j}(t) dt \right) \xrightarrow{n, N \rightarrow \infty} \mathcal{N} \left(0, \int_0^{2\pi} \Sigma^{i,i}(t) \Sigma^{j,j}(t) + (\Sigma^{i,j}(t))^2 dt \right).$$

DISTRIBUTIONS ASYMPTOTIQUE DE L'ERREUR



OPTIMISATION DU CALCUL VIA NUFFT

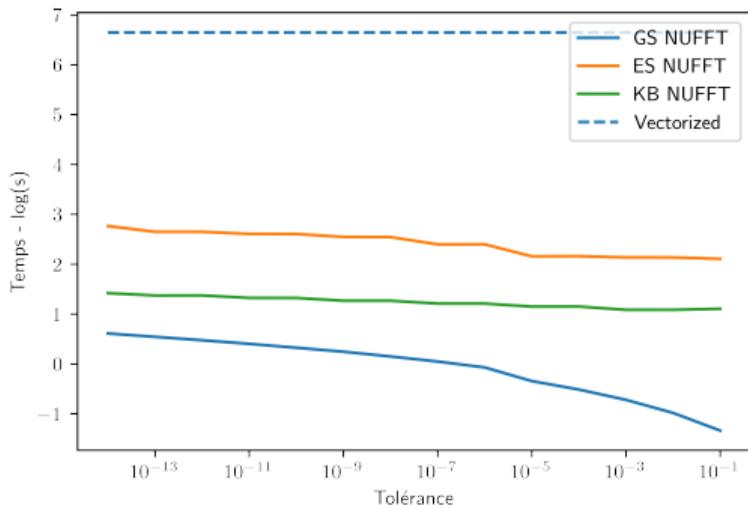
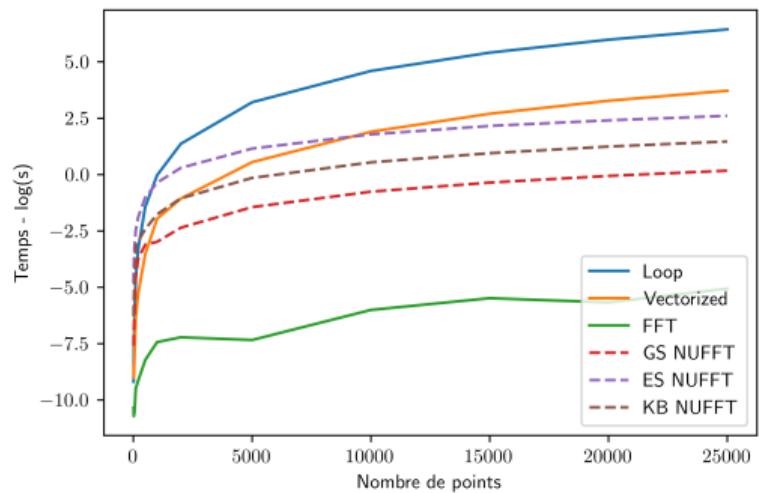


Figure: Temps de calcul en fonction de la tolérance et du nombre de points

OPTIMISATION DU CALCUL VIA NUFFT

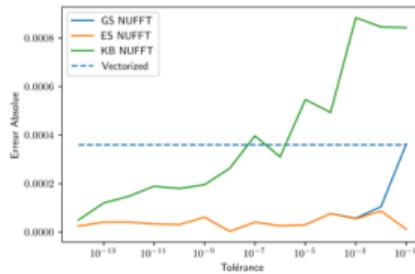
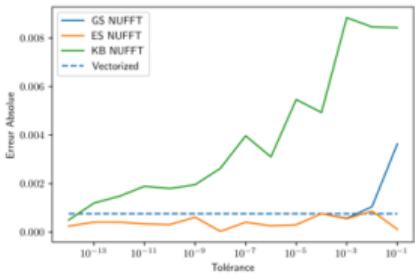


Figure: Erreur selon la tolérance sur la variance (à droite) et la covariance - Cas synchrone

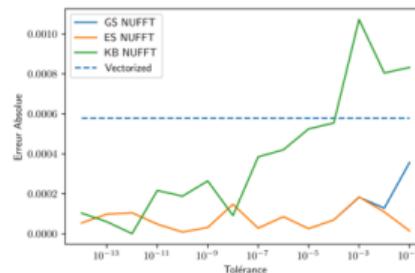
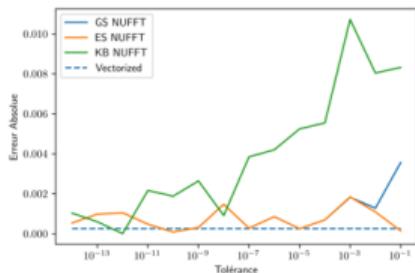


Figure: Erreur selon la tolérance sur la variance (à droite) et la covariance - Downsample 40%

COMPARAISON DES PERFORMANCES

Sous-échantillonnage uniforme

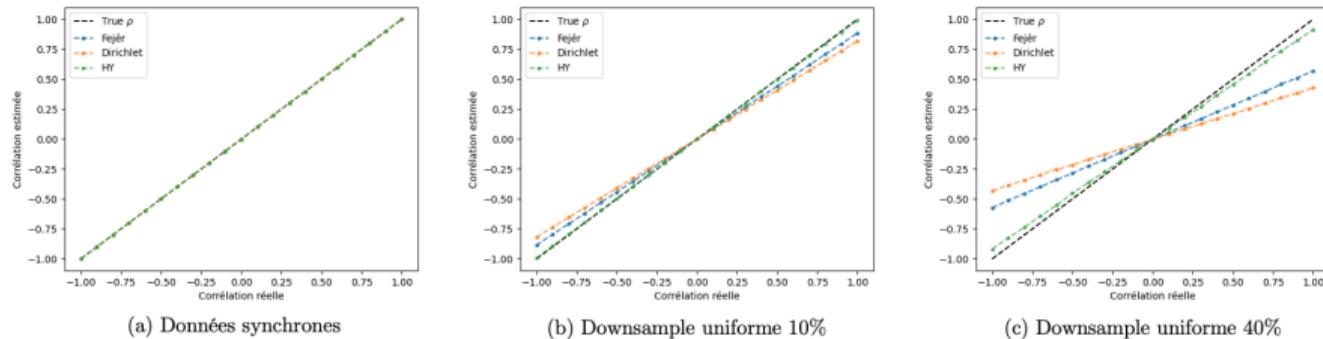
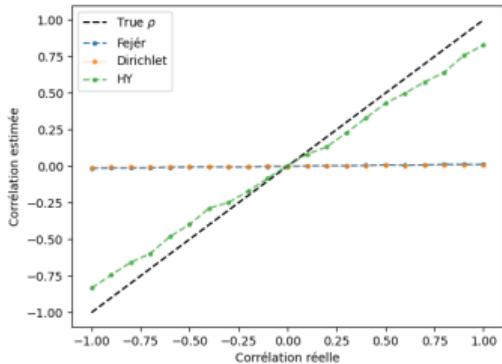


Figure: Corrélations estimées sous différentes conditions de sous-échantillonnage uniforme

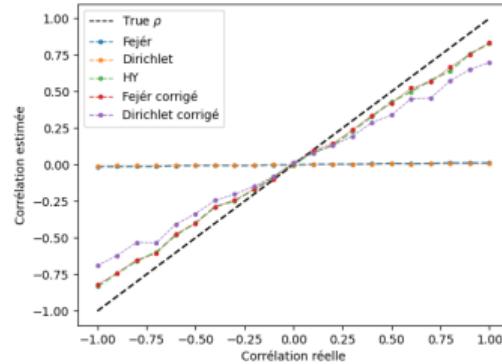
⇒ Baisse de la précision des estimations à mesure que le nombre de données manquantes grandit

COMPARAISON DES PERFORMANCES

Sous-échantillonnage exponentiel



(a) $N = \text{Nyquist}$



(b) $N = 0.85 \min(n_1, n_2)^{3/4}$

Figure: Corrélation estimée après sous-échantillonnage exponentiel, $\lambda_1 = 30, \lambda_2 = 40$

⇒ Nécessité de choisir une fréquence de coupure N adaptée

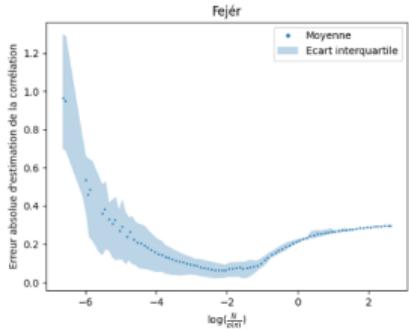
EFFET DES PARAMÈTRES

Erreur de l'estimation

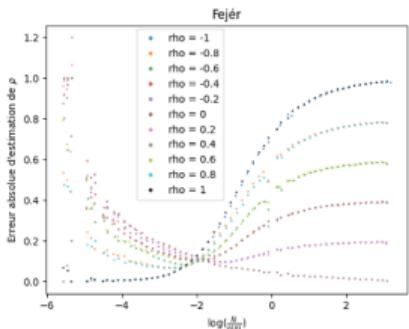
$$\mathbb{E} \left[\widehat{\Sigma}_{n_1, n_2, N}^{1,2} - \int_0^{2\pi} \Sigma^{1,2}(t) dt \right] = \sum_{l=1}^{n_1-1} \sum_{r=1}^{n_2-1} (D_N(t_l^1 - t_r^2) - 1) \mathbb{E} \left[\int_{I_l^1 \cap I_r^2} \Sigma^{1,2}(t) dt \right]$$

- ⇒ L'estimation est toujours consistante dans le cas synchrone
- ⇒ L'estimation est toujours consistante pour la variance d'un même prix
- ⇒ Pour la covariance, parmi les paramètres influant l'erreur : N , n et ρ

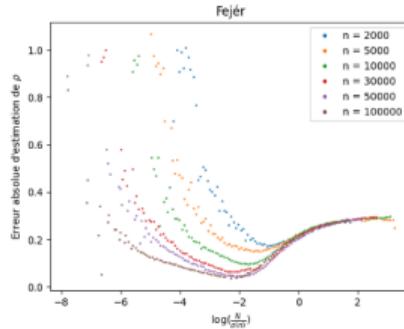
EFFET DES PARAMÈTRES



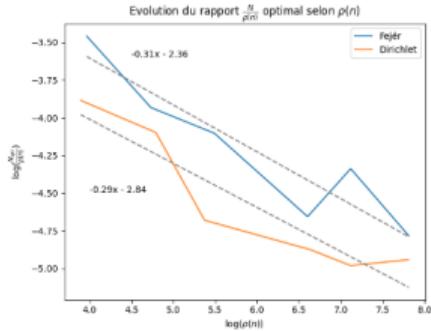
(a) Effet de N



(c) Effet de ρ

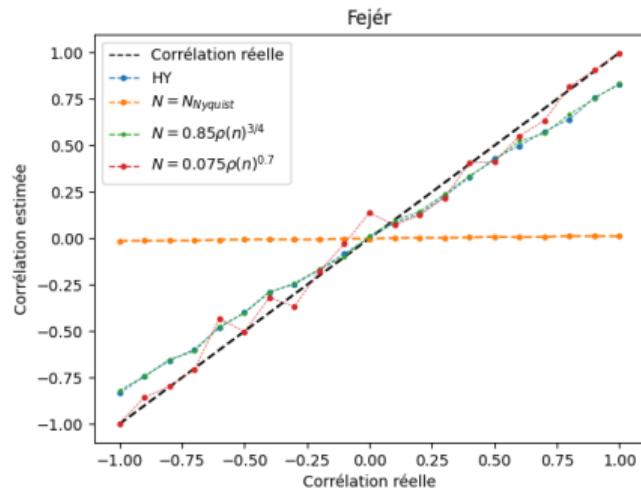


(b) Effet de n

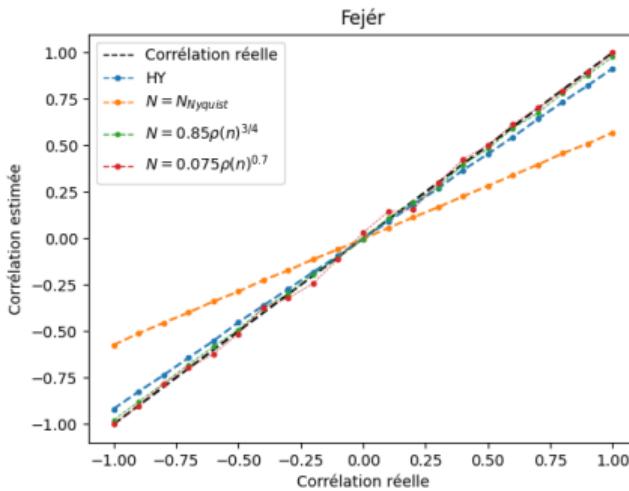


(d) Evolution de $\frac{N}{\rho(n)}$ selon $\log(\rho(n))$

EFFET DES PARAMÈTRES



(a) Echantillonnage exponentiel $\lambda_1 = 30$ et $\lambda_2 = 40$



(b) Downsample uniforme à 40% de données manquantes

Figure: Comparaison d'estimations de corrélations pour plusieurs fréquences de coupures

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. **Estimation de la covariance instantanée**
 6. Application sur des données réelles
 7. Conclusion

EXEMPLE DANS LE CAS SYNCHRONE

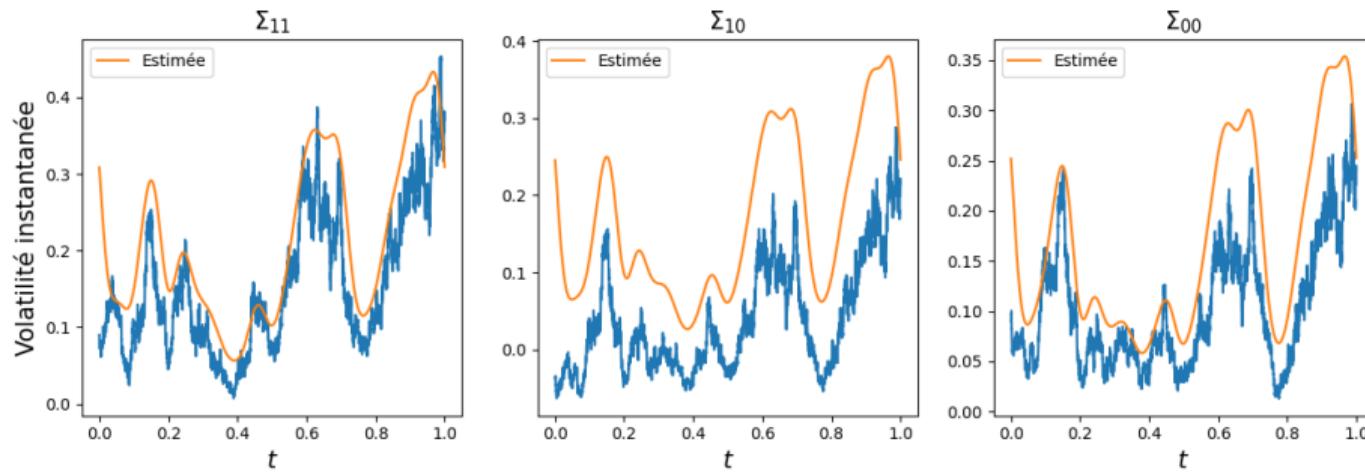


Figure: Exemple d'estimation dans le cas synchrone

⇒ Restitution des fluctuations majeures de la variance

DISTRIBUTION LIMITÉE DE L'ERREUR

Théorème

Sous les conditions $\frac{N}{n} \rightarrow 0$ et $\frac{M}{n} \rightarrow 0$ quand $n, N, M \rightarrow \infty$, la convergence en loi suivante est vérifiée :

$$\sqrt{\frac{n}{M}} (\hat{\sigma}_{n,N,M}^2(t) - \sigma^2(t)) \xrightarrow[n,N,M \rightarrow \infty]{} \mathcal{N}\left(0, \frac{4}{3}\sigma^4(t)\right).$$

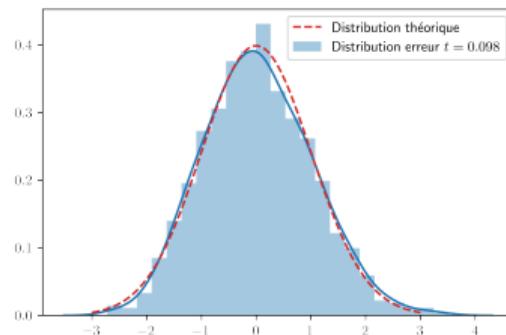


Figure: Distribution de $\sqrt{\frac{3n}{4M}} \frac{(\hat{\sigma}_{n,N,M}^2(t) - \sigma^2(t))}{\sigma^2(t)}$ à $t = 0.098$

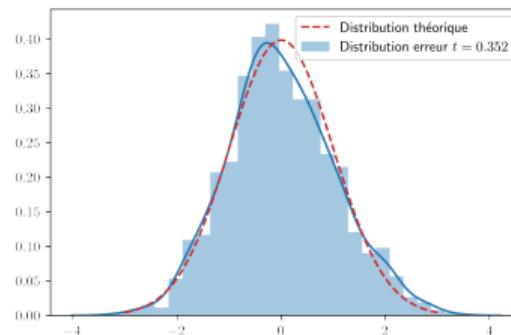


Figure: Distribution de $\sqrt{\frac{3n}{4M}} \frac{(\hat{\sigma}_{n,N,M}^2(t) - \sigma^2(t))}{\sigma^2(t)}$ à $t = 0.352$

EXEMPLE DANS LE CAS ASYNCHRONE

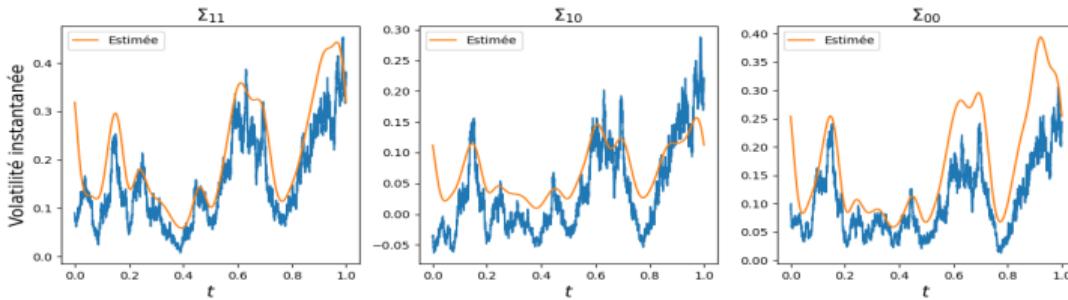


Figure: Cas sous-échantillonnage à 40%

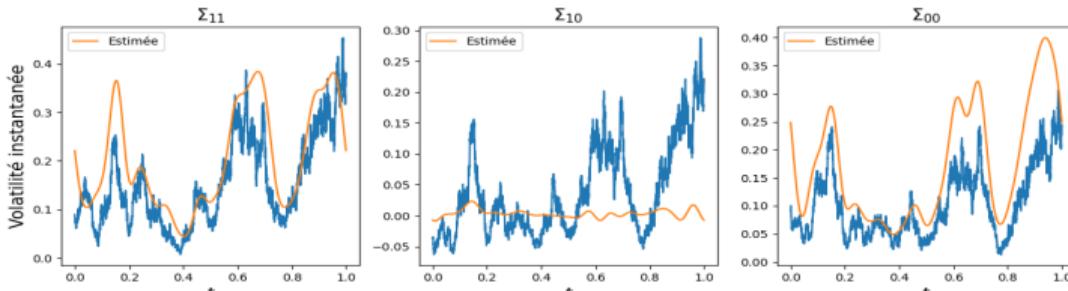


Figure: Cas échantillonnage exponentiel $\lambda_1 = 4$, $\lambda_2 = 6$

EFFET DE LA FRÉQUENCE DE COUPURE M

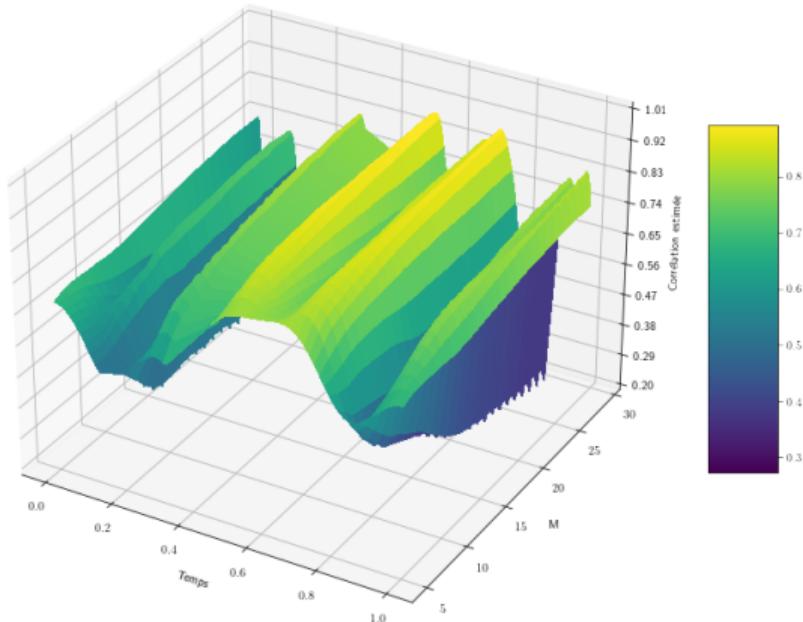


Figure: Estimation de la corrélation instantanée à différents M

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. Estimation de la covariance instantanée
 6. Application sur des données réelles
 7. Conclusion

ETUDE PRÉLIMINAIRE DES DONNÉES

- 5 actifs (BNP, EDF, LVMH, PRTP, SOGN) sur Juin 2015
- Etude sur 10h-16h pour éviter les effets de bord
- Dernier prix dans le cas de transactions simultanées
- Estimation des paramètres des exponentielles modélisant les temps d'attentes :

	Nombre de transactions	$\frac{1}{\lambda}$ (s)	Intervalle de confiance à 95%
BNPP	96765	5.72	[5.69, 5.76]
EDF	27048	20.44	[20.20, 20.69]
LVMH	51957	10.65	[10.56, 10.74]
PRTP	25144	21.98	[21.71, 22.25]
SOGN	108493	5.10	[5.08, 5.14]

MISE EN ÉVIDENCE DE L'EFFET EPPS

Corrélation de Pearson et effet Epps pour le 2015-06-01

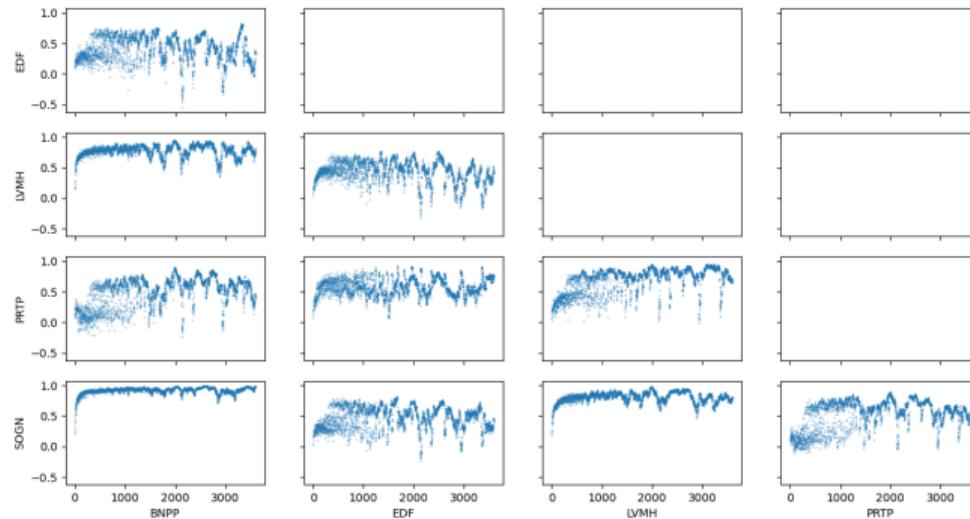


Figure: Indices de Pearson le 01/01/21

MISE EN ÉVIDENCE DE L'EFFET EPPS

- Relation entre N et Δt : $N = \lfloor \frac{1}{2} \left(\frac{T}{\Delta t} - 1 \right) \rfloor$

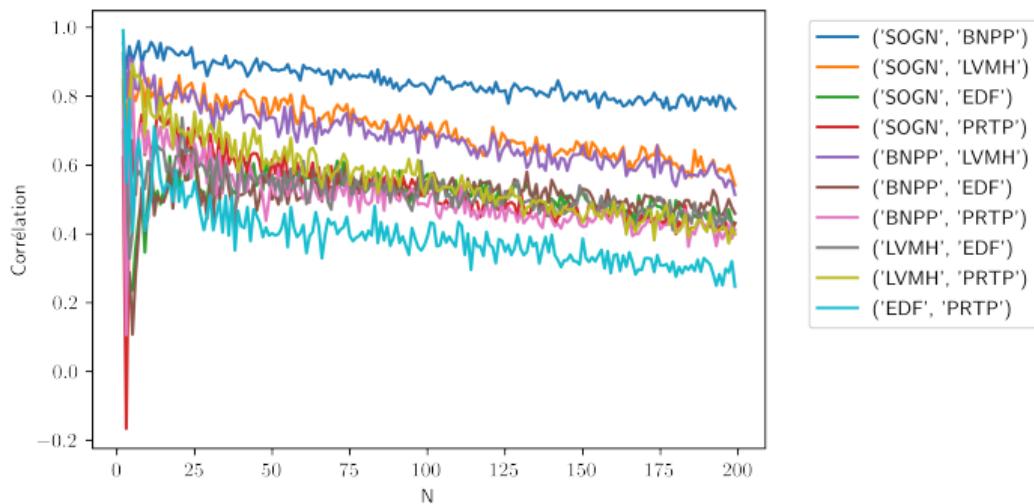
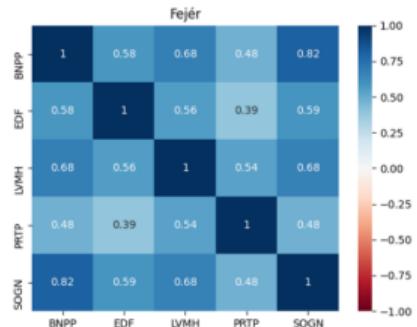
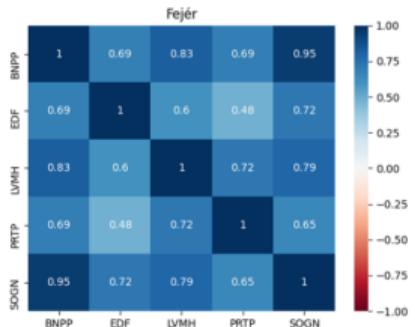


Figure: Effet Epps à l'aide de l'estimateur de Malliavin-Mancino

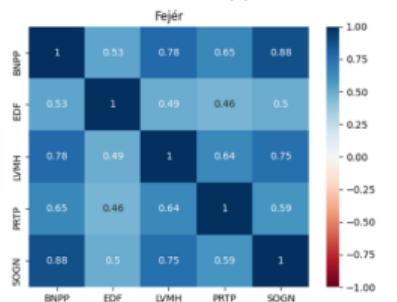
MATRICES DE CORRÉLATION



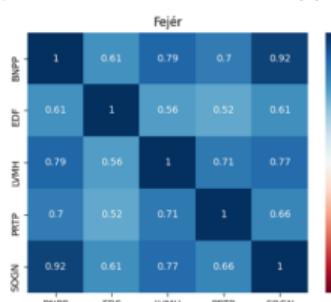
(a) Corrélations moyennées



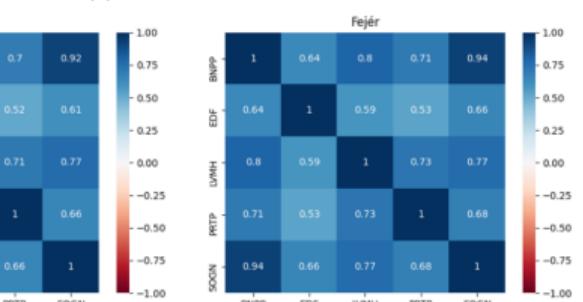
(b) Données concaténées



(c) $\Delta t = 60s$



(d) $\Delta t = 100s$

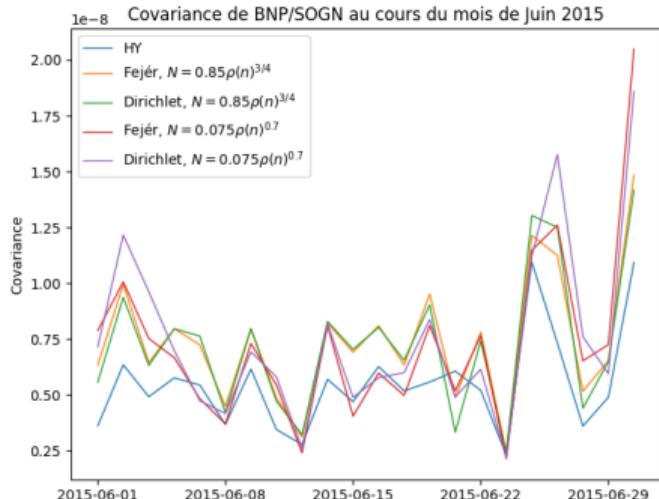


(e) $\Delta t = 150s$

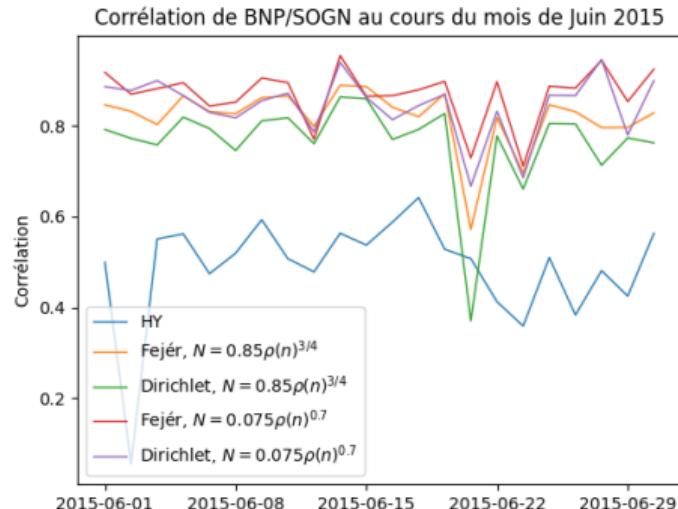
Figure: Matrices de corrélations par inversion de Fejér

ETUDE DE LA PAIRE BNP / SOGN

Etude mensuelle



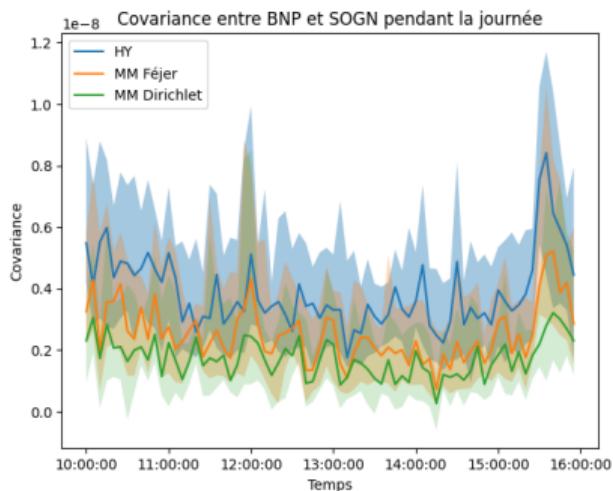
(a) Covariance



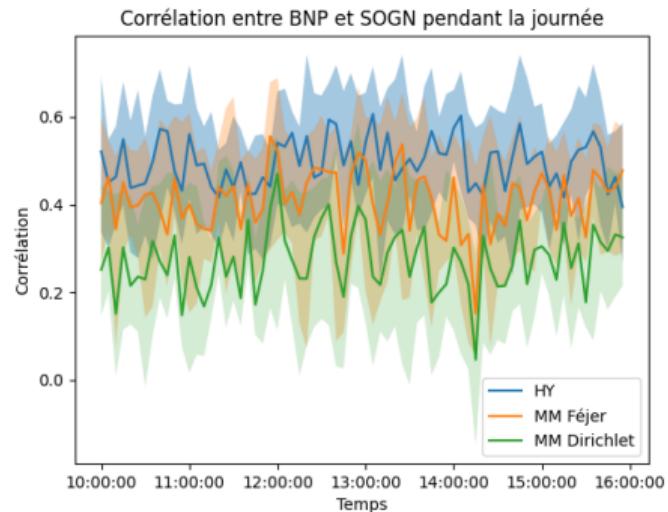
(b) Corrélation

ETUDE DE LA PAIRE BNP / SOGN

Etude intra-day



(a) Covariance

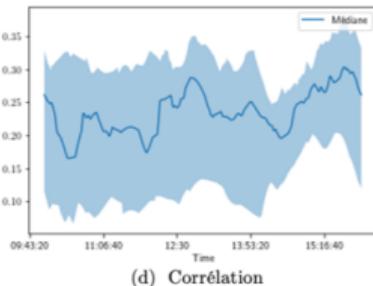
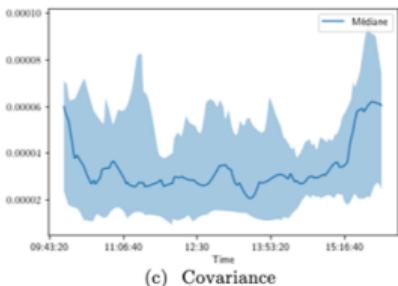
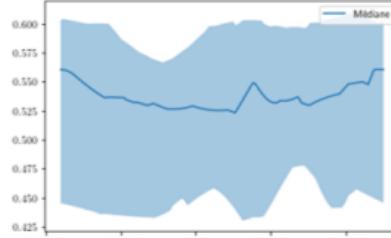
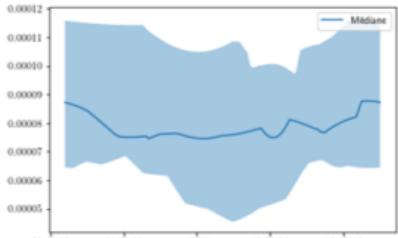


(b) Corrélation

ETUDE DE LA PAIRE BNP / SOGN

Etude intra-day

$$\begin{cases} N^* = \lfloor \frac{n}{2} \rfloor \\ M^* = \lfloor \frac{1}{16\pi} \sqrt{n} \log n \rfloor \end{cases}$$

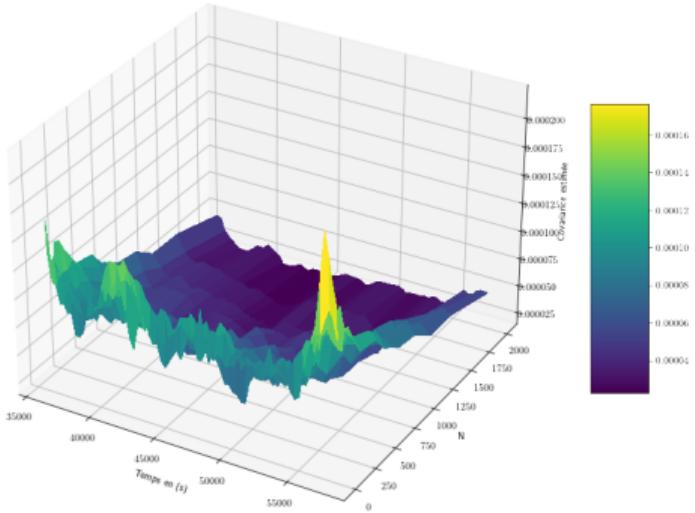


$$\begin{cases} N^{**} = \lfloor 0.85n^{3/4} \rfloor \\ M^{**} = \lfloor \frac{1}{16\pi} \sqrt{n^{3/4}} \log n^{3/4} \rfloor \end{cases}$$

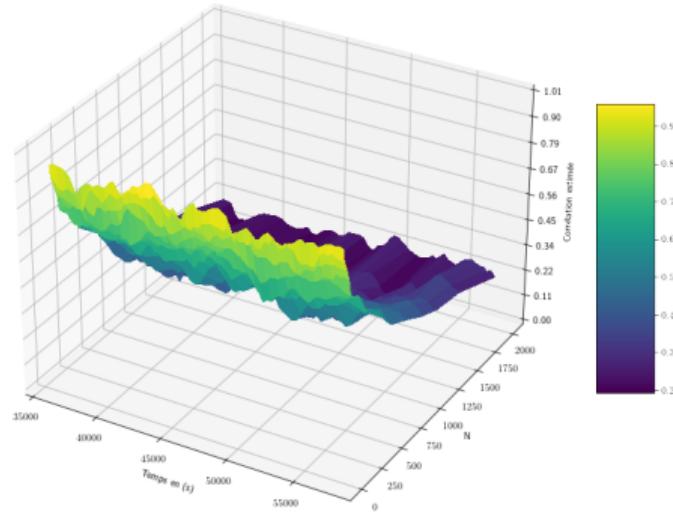
Figure: Corrélation et covariance instantanées sous différentes fréquences de coupure

ETUDE DE LA PAIRE BNP / SOGN

Etude intra-day



(a) Covariance



(b) Corrélation

-
1. Introduction
 2. Estimateur de Malliavin-Mancino
 3. Calcul des coefficients de Fourier
 4. Estimation de la covariance intégrée
 5. Estimation de la covariance instantanée
 6. Application sur des données réelles
 7. Conclusion

CONCLUSION

En résumé :

- Caractérisation de l'estimateur de Malliavin-Mancino
- Comparaisons avec l'estimateur de Hayashi-Yoshida
- Méthode d'accélération du calcul
- Mise en évidence de résultats sur des données empiriques

Pistes d'ouverture :

- Etude des performances sous des conditions plus variées
- Comparaison à d'autres estimateurs (Cuchiero-Teichmann, Maximum de vraisemblance...)
- Etude théorique et expérimentale des paramètres optimaux pour les estimations intégrées et instantanées



Fin

Merci pour votre attention

8. Annexes

9. Bibliographie

ESTIMATEUR DE HAYASHI-YOSHIDA

Estimateur de la Covariance Réalisée

$$V := \sum_{i=1}^{n-1} (P_{t_{i+1}}^1 - P_{t_i}^1)(P_{t_{i+1}}^2 - P_{t_i}^2).$$

Estimateur de Hayashi-Yoshida

$$V := \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (P_{t_{i+1}^1}^1 - P_{t_i^1}^1)(P_{t_{j+1}^2}^2 - P_{t_j^2}^2) \mathbb{1}_{]t_i^1, t_{i+1}^1] \cap]t_j^2, t_{j+1}^2] \neq \emptyset}$$

MODÈLE DE HESTON GÉNÉRALISÉ

$$\begin{cases} X(t) &= X(0) + \int_0^t -\frac{1}{2}\Sigma^{\text{diag}}(s)ds + \int_0^t \sqrt{\Sigma(s)}dZ \\ \Sigma(t) &= \Sigma(0) + \int_0^t (b + M\Sigma(t) + \Sigma(t)M^\top) dt + \sqrt{\Sigma(t)}dBH + HdB^\top\sqrt{\Sigma(t)} \end{cases}$$

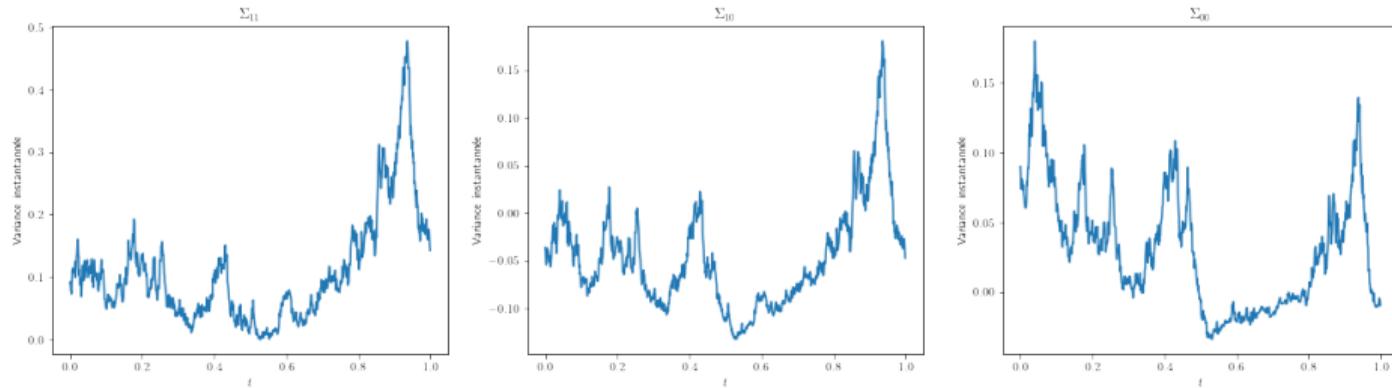
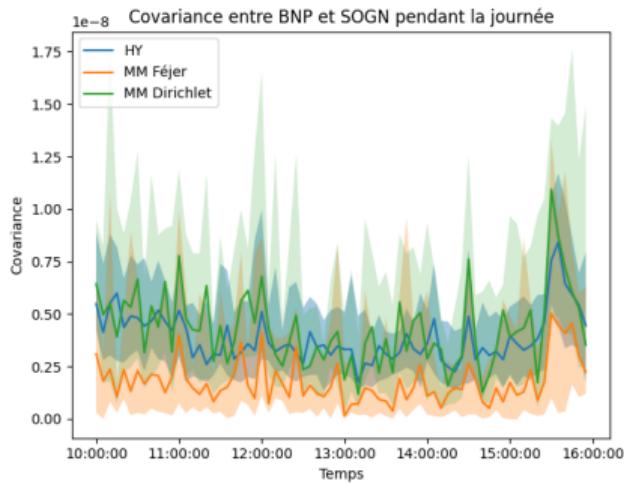


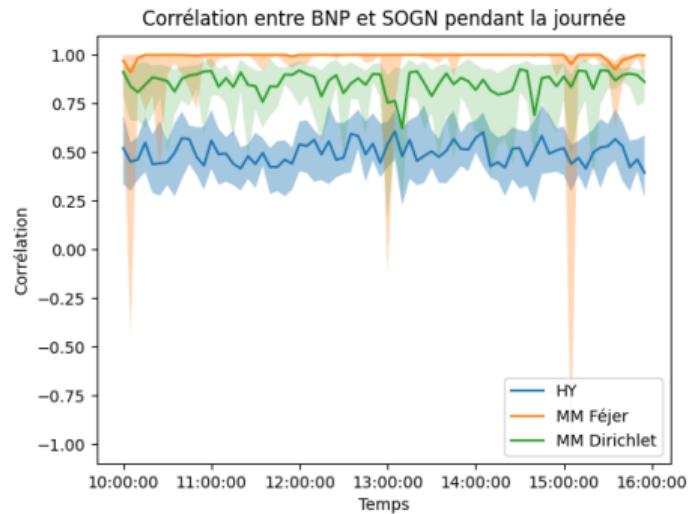
Figure: Exemple de volatilités simulées par le modèle de Heston

ETUDE DE LA PAIRE BNP / SOGN

Etude intra-day



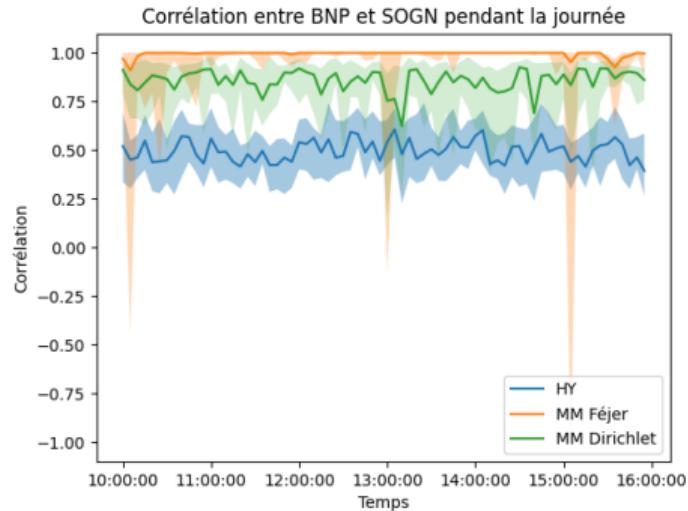
(a) Covariance



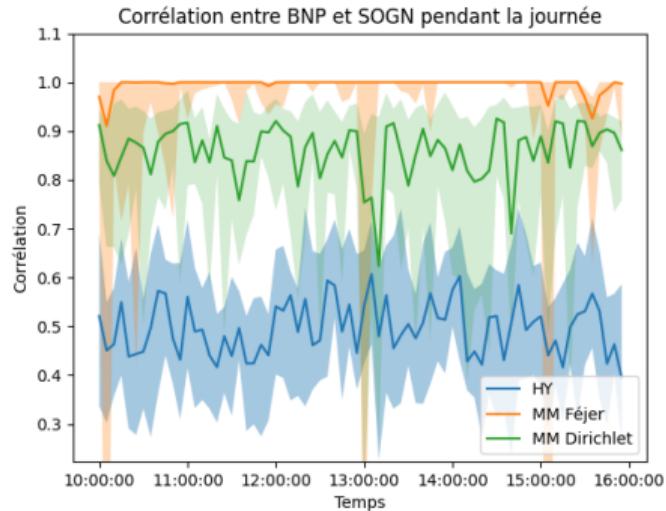
(b) Corrélation

ETUDE DE LA PAIRE BNP / SOGN

Etude intra-day



(a) Corrélation



(b) Corrélation (zoom)

8. Annexes

9. Bibliographie

-  Yacine Aït-Sahalia, Per A. Mykland, and Lan Zhang.
How often to sample a continuous-time process in the presence of market microstructure noise.
The Review of Financial Studies, 18:351–416, 2005.
-  H. Alexander Barnett, Jeremy Magland, and Ludvig Af Klinteberg.
A parallel nonuniform fast fourier transform library based on an exponential of semicircle kernel.
2019.
-  L. Bergomi.
Correlations in asynchronous markets.
Risk, pages 76–82, 2010.
-  A. Chakraborti, I. Muni Toke, M. Patriarca, and Abergel F.
Econophysics review: I. empirical facts.
Quantitative Finance, 11(7):991–1012, 2011.

 Patrick Chang.
Fourier instantaneous estimators and the epps effect.
2020.

 Patrick Chang and Roger Bukuru.
An exercise in R : High frequency covariance estimation using malliavin-mancino
and hayashi-yoshida estimators.
2019.

 Patrick Chang, Etienne Pienaar, and Tim Gebbie.
Malliavin-mancino estimators implemented with non-uniform fast fourier
transforms.
2020.

 Emmanuelle Clement and Arnaud Gloter.
Weak limit theorem in the Fourier tranform method for the estimation of
multivariate volatility.
Stochastic Processes and their Applications, 121:1097–1124, 2011.

-
-  A. Dutt and V. Rokhlin.
Fast fourier transforms for nonequispaced data.
Applied and Computational Harmonic Analysis, 2:85–100, 1995.
 -  T. W. Epps.
Comovements in stock prices in the very short run.
Journal of American Statistics Association, 74:291–298, 1979.
 -  Torben G.Andersen and Tim Bollerslev.
Intraday periodicity and volatility persistence in financial markets.
Empirical Finance, 4:115–158, 1997.
 -  Paul Glasserman.
Monte Carlo Methods in Financial Engineering.
Springer, 2003.

-
-  Leslie Greengard and June-Yub Lee.
Accelerating the nonuniform fast fourier transform.
SIAM review, 46:443–454, 2004.
 -  T. Hayashi and N. Yoshida.
On covariance estimation of nin-synchronously observed diffusion processes.
Bernoulli, 11(2):359–379, 2005.
 -  Paul Malliavin and Maria Elvira Mancino.
A fourier transform method for non parametric estimation of multivariate volatility.

The Annals of Statistics, 37(4):4–10, 2009.
 -  Maria Elvira Mancino, Maria Cristina Recchioni, and Simona Sanfelici.
Fourier-Malliavin Volatility Estimation: Theory and Practice.
SpringerBriefs in Quantitative Finance. Springer International Publishing, 2017.

-
- Vanessa Mattiussi, Michele Tumminello, Giulia Iori, and Rosario N. Mantegna.
Comparing correlation matrix estimators via kullback-leibler divergence.
2011.
 - Roberto Renò.
A closer look at the epps effect.
International Journal of Theoretical and Applied Finance, 6(1):87–102, 2003.
 - Gayatri Tilak, Tamás Szell, Rémy Chicheportiche, and Anirban Chakraborti.
Study of statistical correlations in intraday and daily financial return time series.
Econophys-Kolkata, IV:77, 2011.