

# **RAPPORT DE L'ENSEIGNEMENT D'APPROFONDISSEMENT RECHERCHE**

## **APM\_511\_EP-4.3: STATISTIQUE HAUTE FRÉQUENCE ET MICROSTRUCTURE DES MARCHÉS**

Ouattara ABOUBAKAR / Borel DOMGUE / Coralie TONLE

---

Sous la supervision de M. Mathieu ROSENBAUM



# SOMMAIRE

<b>Introduction</b>	<b>3</b>
<b>1 Almgren et Chriss, 2000</b>	<b>4</b>
1.1 Présentation de l'article	4
1.1.1 Résumé	4
1.1.2 Description du modèle de trading	4
1.1.3 Dynamique des prix	4
1.1.4 Impacts de marché temporaire et permanent	4
1.1.5 Trajectoires optimales - Formules d'Almgren-Chriss	5
1.2 Simulations numériques	6
1.2.1 trajectoires des prix	7
1.2.2 Effet de l'aversion sur les trajectoires optimales	7
1.2.3 Effet de la volatilité estimée sur l'utilité	8
1.3 Discussions	9
1.4 Conclusion 1	9
<b>2 Zhang et al., 2005</b>	<b>10</b>
2.1 Présentation de l'article	10
2.1.1 Description du modèle	10
2.1.2 Dynamique des prix	10
2.1.3 Estimateurs de la volatilité réalisée	10
2.2 Simulations numériques	11
2.2.1 Vérifications empiriques des résultats théoriques	12
2.2.2 Trajectoires de prix	12
2.3 Discussions	12
2.4 Conclusion 2	14
<b>3 Robert et Rosenbaum, 2010</b>	<b>15</b>
3.1 Présentation de l'article	15
3.1.1 Description du modèle	15
3.1.2 Algorithme de construction des prix	15
3.1.3 Estimation de la volatilité réalisée	16
3.2 Simulations numériques	16
3.2.1 Trajectoires de prix	16
3.2.2 Estimation de la volatilité	17
3.3 Discussions	17
3.4 Conclusion 3	18
<b>4 Hayashi et Yoshida, 2005</b>	<b>19</b>
4.1 Présentation de l'article	19
4.1.1 Description du modèle	19
4.1.2 Dynamique des prix et effet de Epps	19
4.1.3 Estimateurs de HY de la covariance	19
4.2 Simulations numériques	20
4.2.1 Dynamique des prix de 02 actifs observés et impact de la corrélation	20
4.2.2 Estimateur $U_n$ / valeur réelle	21
4.2.3 Distribution simulée de $U_n$	21
4.3 Discussions	22
4.4 Conclusion 4	22



INSTITUT  
POLYTECHNIQUE  
DE PARIS

<b>Conclusion</b>	<b>23</b>
<b>Références</b>	<b>24</b>

## INTRODUCTION

Dans le domaine de la finance à haute fréquence, la modélisation et l'estimation des dynamiques de prix transactionnels sont des problématiques essentielles pour optimiser les stratégies de trading et comprendre les comportements des marchés. Ce rapport s'appuie sur l'étude de plusieurs contributions scientifiques majeures qui abordent des questions clés telles que la volatilité réalisée, les impacts de marché, et les interactions entre les prix latents et les prix observés.

L'article d'Almgren et Chriss propose un modèle détaillé pour la liquidation optimale d'actifs, mettant en avant l'impact des paramètres de marché sur les coûts de transaction et les risques de volatilité.

L'article de Zhang, Mykland et Ait-Sahalia, aborde l'estimation de la volatilité intégrée dans des données à haute fréquence bruyantes. Les auteurs introduisent une méthode innovante qui décompose la volatilité observée en deux échelles de temps différentes, permettant d'extraire des informations fiables même en présence de bruit microstructurel. Cette approche améliore l'estimation de la volatilité réalisée, tout en corrigeant les biais liés au bruit de mesure qui affectent les méthodes classiques.

Ensuite l'article de Robert et Rosenbaum introduit un mécanisme innovant basé sur les zones d'incertitude pour relier les prix transactionnels discrets aux processus latents sous-jacents, tenant compte des effets micro-structurels.

Enfin, les travaux de Hayashi et Yoshida se concentrent sur l'estimation robuste de la covariance entre deux processus de diffusion dans un contexte d'observations non synchrones. En réponse aux limites des estimateurs classiques, tels que l'estimateur de covariance réalisée, les auteurs proposent un nouvel estimateur sans biais et indépendant des choix de synchronisation, offrant ainsi une solution robuste et cohérente pour l'analyse des données financières.

À travers ces études, ce rapport explore les fondements théoriques et pratiques de la modélisation des prix transactionnels et de l'estimation de la volatilité, tout en examinant les limites des approches classiques et les solutions proposées pour les surmonter. Les simulations numériques illustrent les concepts étudiés et permettent de valider empiriquement les modèles développés. Ces travaux ouvrent ainsi la voie à une meilleure compréhension des mécanismes sous-jacents aux marchés financiers modernes.

# 1

## Almgren et Chriss, 2000

### 1.1 PRÉSENTATION DE L'ARTICLE

#### 1.1.1 • RÉSUMÉ

Le but de l'article est d'établir une stratégie de liquidation optimale d'une quantité donnée d'options avant une maturité  $T$ . Pour ce faire, les auteurs commencent par définir un modèle de trading qui est une discrétisation du temps de liquidation et un nombre de parts à liquider à chaque pas de temps. La liste de ces nombres de parts forme la « trajectoire de liquidation ». Ainsi, une trajectoire de liquidation sera utilisée selon qu'elle satisfait un critère d'optimalité. Le critère d'optimalité utilisé ici est la « fonction d'utilité »  $f(x) = \mathbf{E}(x) + \lambda \mathbf{V}(x)$  où  $x$  est le shortfall, c'est-à-dire : le coût de transaction de la stratégie. Le modèle d'Almgren Chriss prend en compte le market impact (temporaire et permanent) et peut être étendu à un modèle avec drift.

#### 1.1.2 • DESCRIPTION DU MODÈLE DE TRADING

Nous détenons un stock de  $X$  unités d'un titre que nous voulons liquider complètement avant le temps  $T$ . Nous divisons  $T$  en  $N$  intervalles de temps de longueur  $\tau = \frac{T}{N}$  et définissons les temps discrets  $t_k = k\tau$ . Nous définissons une trajectoire de négociation comme une liste  $(x_0, \dots, x_N)$  où  $x_k$  est le nombre d'unités que nous prévoyons de détenir au temps  $t_k$ ; ainsi  $x_0 = X$ ,  $x_N = 0$  et  $n_k = x_{k-1} - x_k$  étant le nombre d'actifs vendus entre  $t_{k-1}$  et  $t_k$ .

#### 1.1.3 • DYNAMIQUE DES PRIX

Supposons que le prix initial du titre soit  $S_0$  de sorte que la valeur de marché initiale de notre position soit  $X S_0$ . Le prix du titre évolue en fonction de deux facteurs exogènes : la volatilité et le drift, et d'un facteur endogène : l'impact du marché. La volatilité et le drift sont supposés être le résultat des forces du marché qui se produisent de manière aléatoire et indépendamment de notre activité de trading. Ainsi, le prix évolue de la manière suivante :

$$S_k = S_{k-1} + \sigma \tau^{1/2} \xi_k - \tau g(n_k)$$

$\sigma$  : la volatilité de l'actif

$\xi_k$  : une variable aléatoire avec une moyenne nulle et une variance unitaire

$g(v)$  : caractérise l'impact permanent du marché. Cet impact est le changement du prix d'équilibre du marché dû à notre trading.

#### 1.1.4 • IMPACTS DE MARCHÉ TEMPORAIRE ET PERMANENT

Lorsqu'un trader achète ou vend de nombreux actifs, par le principe de l'offre et la demande, le marché réagit naturellement à ce comportement. Nous distinguons deux types d'impact sur le marché.

D'abord l'impact permanent qui lui est très simple à comprendre. En effet, il est la conséquence de notre trading sur le marché. Ce dernier détectant nos activités sur le marché, ajuste donc naturellement son prix.

Quant à l'impact temporaire, il survient lorsqu'un trader qui décide de vendre entre  $t_{k-1}$  et  $t_k$  une quantité  $n_k$  de l'actif, mais décide de le diviser en plusieurs tranches plus petites dans le but de déterminer les points de liquidité optimaux. Si le nombre total d'unités  $n_k$  est suffisamment important, le prix d'exécution peut diminuer régulièrement entre  $t_{k-1}$  et  $t_k$ , en partie en raison de l'épuisement de l'offre de liquidité à chaque niveau de prix successif.

Nous supposons que cet effet est de courte durée et le modélisons en introduisant une fonction d'impact temporaire sur les prix notée  $h(v)$ . Cet effet agit si le prix, en faisant baisser le prix moyen par action pendant intervalle cours. Ainsi le prix réel par action reçue lors de la vente  $k$  :

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau}\right)$$

mais il faut noter que l'effet de  $h(v)$  n'apparaît pas dans le prix "du marché"  $S_k$

Ainsi après avoir défini tout cela, nous pouvons définir le revenu du trading à la fin de toutes les transactions et l'on appelle cela une Capture de trajectoire. Il s'agit donc de la somme du produit du nombre d'unités  $n_k$  que nous vendons à chaque intervalle de temps multiplié par le prix effectif par action  $\tilde{S}_k$  reçu lors de cette vente. Nous calculons facilement par :

$$\sum_{k=1}^N n_k \tilde{S}_k = XS_0 + \sum_{k=1}^N \left( \sigma \tau^{1/2} \xi_k - \tau g\left(\frac{n_k}{\tau}\right) \right) x_k - \sum_{k=1}^N n_k h\left(\frac{n_k}{\tau}\right)$$

Avec :

- $XS_0$  : Valeur initiale de notre position
- $\sum_{k=1}^N \sigma \tau^{1/2} \xi_k x_k$  : représentant l'effet total de la volatilité
- $-\sum_{k=1}^N x_k \tau g\left(\frac{n_k}{\tau}\right)$  : représente la perte de valeur de notre position totale, causée par la baisse permanente des prix associés à la vente d'une petite partie de la position
- $\sum_{k=1}^N n_k h\left(\frac{n_k}{\tau}\right)$  : est la baisse de prix due à notre vente, en agissant uniquement sur les unités que nous vendons pendant la période  $k$ th

On définit donc le coût total de trading comme la différence  $XS_0 - \sum_{k=1}^N n_k \tilde{S}_k$  la valeur de notre position initial et notre capture.

Ce coût total de trading est étroitement lié à la notion de **shortfall** (ou **implémentation shortfall**), qui représente la perte subie par rapport à la valeur théorique attendue. Le shortfall inclut non seulement les **coûts de transaction directs**, mais aussi l'**impact de marché** et le **risque lié à la volatilité**. Ainsi, nous pouvons écrire :

$$\text{Shortfall} = \text{Coûts de transaction} + \text{Impact de marché} + \text{Effet de la volatilité}$$

Dans notre contexte :

- Les termes  $\sum_{k=1}^N \tau x_k g\left(\frac{n_k}{\tau}\right)$  et  $\sum_{k=1}^N n_k h\left(\frac{n_k}{\tau}\right)$  représentent les **coûts de transaction implicites**, dus à l'impact de marché permanent et temporaire.
- Le terme  $\sum_{k=1}^N \sigma \tau^{1/2} \xi_k x_k$  représente le **risque lié à la volatilité**.

L'objectif est donc de minimiser ce shortfall total, en trouvant un équilibre optimal entre la réduction des coûts de transaction et la gestion du risque de volatilité. Cela constitue la base de la construction d'une **stratégie d'exécution optimale**.

### 1.1.5 • TRAJECTOIRES OPTIMALES - FORMULES D'ALMGREN-CHRISS

Dans le cadre de l'exécution optimale d'un portefeuille, Almgren et Chriss (2000) proposent un modèle permettant de minimiser le **shortfall** total en équilibrant les coûts de transaction et les risques liés à la volatilité du marché. Leur approche repose sur la construction de **trajectoires optimales** pour la liquidation d'une position, en tenant compte de l'impact de marché permanent et temporaire.

**Définition de la trajectoire optimale :** L'objectif est de déterminer la séquence de ventes  $n_k$  qui minimise la fonction suivante :

$$\min_x (E(x) + \lambda V(x)),$$

où :

- $E(x)$  est l'espérance du **coût total de transaction**, incluant l'impact de marché permanent et temporaire.
- $V(x)$  est la **variance** associée au coût de transaction, représentant le risque de volatilité.
- $\lambda$  est un paramètre d'aversion au risque, reflétant l'importance accordée à la minimisation de la variance par rapport au coût moyen.

**Formules de la trajectoire optimale :** Pour un impact de marché linéaire, Almgren et Chriss montrent que la trajectoire optimale  $x_k$  (le nombre d'unités restantes à l'instant  $t_k$ ) suit une fonction exponentielle décroissante donnée par :

$$x_k = \frac{\sinh(\kappa(T - t_k))}{\sinh(\kappa T)} X,$$

où :

- $X$  est la taille initiale de la position.
- $T$  est le temps total alloué pour la liquidation.
- $\kappa$  est un paramètre clé qui dépend de la volatilité du marché ( $\sigma$ ), du coût temporaire ( $\eta$ ) et de l'aversion au risque ( $\lambda$ ) :

$$\kappa = \sqrt{\frac{\lambda \sigma^2}{\eta}}.$$

**Interprétation de la demi-vie :** Un concept important introduit par Almgren et Chriss est celui de la **demi-vie de liquidation** ( $\theta$ ), définie par :

$$\theta = \frac{1}{\kappa}.$$

La demi-vie représente le temps nécessaire pour liquider environ la moitié de la position. Elle est indépendante du temps total  $T$  et ne dépend que des paramètres du marché et du niveau de risque accepté. Une liquidation plus rapide (demi-vie courte) signifie un coût temporaire élevé mais un risque de volatilité réduit, tandis qu'une liquidation plus lente (demi-vie longue) réduit l'impact de marché mais augmente l'exposition à la volatilité.

## 1.2 SIMULATIONS NUMÉRIQUES

L'objectif des simulations est de comprendre l'impact des paramètres du modèle, tels que l'aversion au risque ( $\lambda$ ) et la volatilité ( $\sigma$ ), sur les stratégies de liquidation et les trajectoires des prix. Nous utilisons un cadre discret pour modéliser la liquidation d'un portefeuille sur une période donnée, tout en prenant en compte les coûts liés aux impacts permanents et temporaires des transactions. À l'aide de ces simulations, nous analysons :

- Les trajectoires de prix en fonction de différentes stratégies (impact minimum et variance minimum).
- L'effet de l'aversion au risque sur les trajectoires optimales de liquidation.
- L'impact des erreurs dans l'estimation de la volatilité sur l'utilité réalisée par le trader.

Afin d'obtenir les résultats que nous allons présenter, nous avons utilisé les paramètres suivants :

- $T = 5$  maturité
- $X = 100$  quantité d'actifs à liquider
- $S_0 = 50$  prix initial
- $N = 100$  nombres de pas de temps

- $\sigma = 1$  volatilité
- $\epsilon = 0.0625$  écart acheteur-vendeur
- $\gamma = 2.5e-7$
- $\eta = 2.5e-6$
- $m = 100000$  nombre de simulations

### 1.2.1 • TRAJECTOIRES DES PRIX

Nous avons tracé 10 trajectoires de prix dans chaque cas sur  $T=5$  jours avec  $N=100$  périodes de temps. Ci-dessous, les résultats que nous avons obtenus pour les 02 cas.

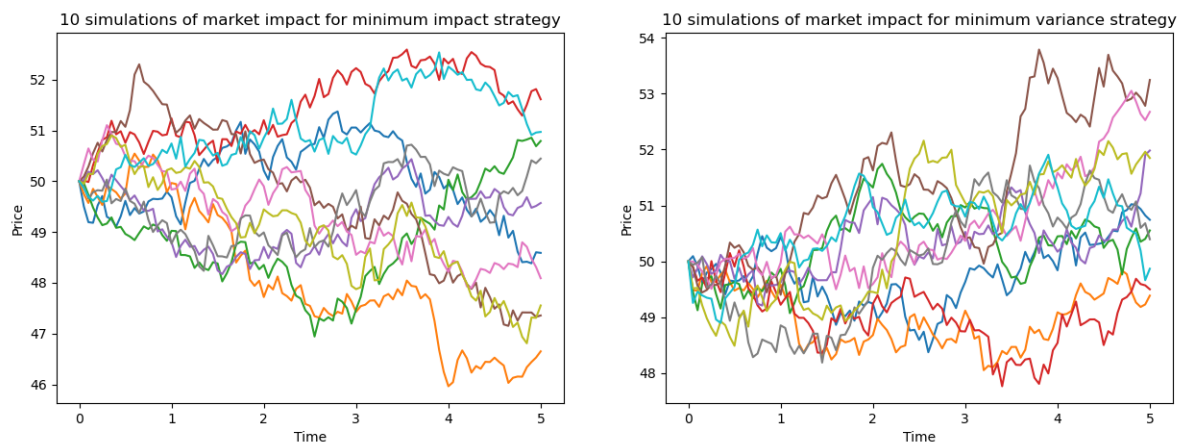


FIGURE 1 – trajectoires des prix

Les résultats montrent que :

- La stratégie à impact minimum distribue les transactions uniformément sur le temps, entraînant une moindre variation des prix.
- La stratégie à variance minimum effectue les transactions de manière plus concentrée, optimisant l'exposition au risque, ce qui engendre une plus grande volatilité des prix.

### 1.2.2 • EFFET DE L'AVERSION SUR LES TRAJECTOIRES OPTIMALES

Cette simulation met en évidence comment le paramètre d'aversion au risque  $\lambda$  influence les trajectoires de liquidation optimale.

En ajustant  $\lambda$ , deux comportements principaux émergent :

- Faible aversion au risque ( $\lambda$  faible) : Le trader privilégie une exécution plus lente pour minimiser les coûts liés à l'impact de marché, acceptant ainsi de s'exposer davantage aux fluctuations du prix.
- Forte aversion au risque ( $\lambda$  élevé) : Le trader cherche à minimiser le risque de prix. Cela conduit à une liquidation rapide, car le trader préfère se débarrasser des actifs rapidement pour éviter l'incertitude liée aux variations de prix, même au prix d'un impact de marché plus important.

Les trajectoires simulées pour différentes valeurs de  $\lambda$  valident ces prédictions théoriques : la stratégie optimale résulte d'un compromis entre l'impact de marché et le risque de prix, selon la tolérance au risque du trader.



Trajectoires optimales de liquidation pour différentes valeurs de lambda

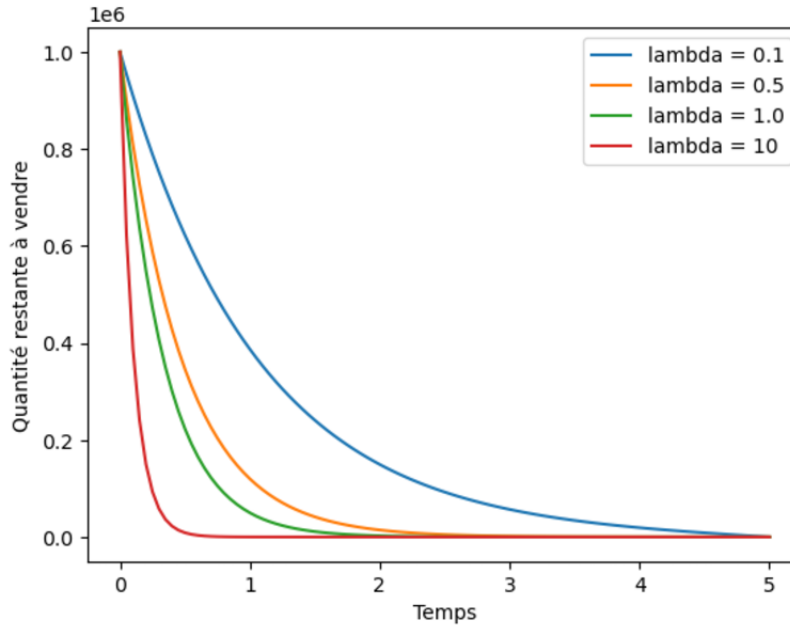


FIGURE 2 – effet de l’aversion sur les trajectoires optimales

### 1.2.3 • EFFET DE LA VOLATILITÉ ESTIMÉE SUR L’UTILITÉ

Dans cette section, nous analysons l’impact des erreurs d’estimation de la volatilité  $\sigma$  sur l’utilité réalisée par le trader.

Nous avons simulé des trajectoires en supposant que le trader applique la stratégie optimale pour une valeur donnée de  $\sigma$ , mais que la volatilité réelle diffère de celle estimée.

Les résultats montrent que :

- Lorsque la volatilité estimée est correcte ( $\sigma_{estimée} = \sigma_{réelle}$ ), l’utilité est minimale.
- Toute erreur sur  $\sigma_{estimée}$  entraîne une augmentation de l’utilité, principalement à cause de la mauvaise calibration de la stratégie.

Cette analyse souligne l’importance d’une estimation précise des paramètres du marché pour optimiser la liquidation d’un portefeuille.

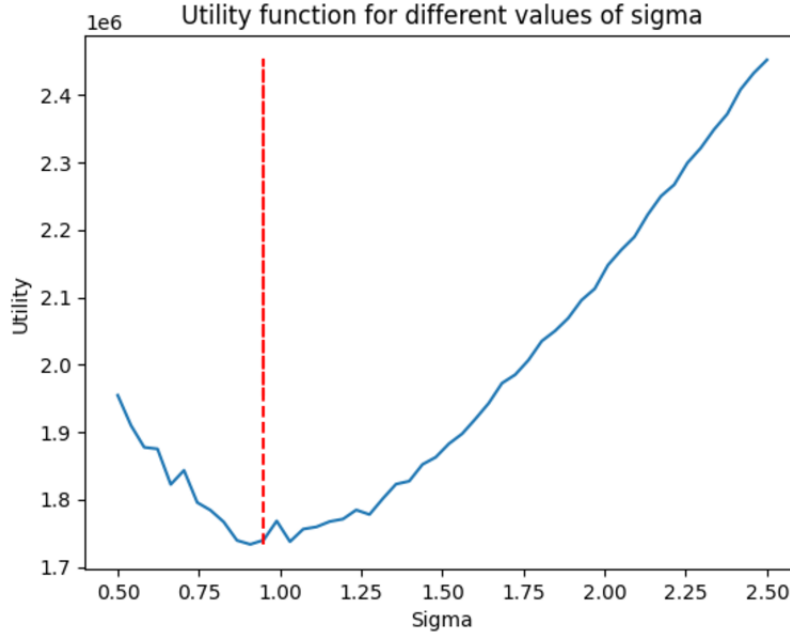


FIGURE 3 – effet de la volatilité estimée sur l'utilité

### 1.3 DISCUSSIONS

Rappelons que la fonction objective choisie dans le cadre de cet article est la fonction d'utilité  $f(x) = E(x) + \lambda V(x)$ . Cette fonction résulte du problème de minimisation :  $\min_{x: V(x) \leq V_*} E(x)$ , elle a l'avantage d'avoir une interprétation financière directe : En effet, le paramètre  $\lambda$  peut être vu comme l'aversion au risque du trader, ainsi :

- Pour un trader très averse au risque, comme on a pu le voir sur la figure 2, la trajectoire de liquidation sera convexe, du fait que le trader préfère se débarrasser le plus vite possible de ses titres afin d'éviter les possibles augmentations de prix liées à l'impact de marché (permanent et temporaire).
- De l'autre côté, un trader peu averse prendra du temps pour liquider ses titres, en espérant pouvoir tirer profit d'une baisse de prix à un moment donné.

En outre, les expériences menées dans le cadre de la "volatilité estimée" nous a permis de voir que pour tirer plein profit d'une liquidation tout en maintenant une variance basse (utilité du shortfall), le trader doit estimer autant correct que se peut, la volatilité du prix du sous-jacent. Ceci nous illustre l'importance de cette estimation.

### 1.4 CONCLUSION 1

Cet article nous a introduit à notre premier problème relevant du domaine des mathématiques financières (la liquidation d'une certaine quantité de titres dans une période prédéterminée) et la modélisation de diverses stratégies de trading pour ce problème. Nous avons également découvert plusieurs notions : l'impact de marché, l'aversion au risque, la tendance (drift) et d'autres notions importantes du monde financier.

Comme relevé par les simulations numériques et la discussion, nous avons vu que le choix de la volatilité influe considérablement les résultats d'une stratégie. De ce fait, notre seconde étude se concentrera sur l'estimation de la volatilité avec le papier : A Tale of Two Time Scales (Un conte de deux échelles de temps).

## 2

## Zhang et al., 2005

## 2.1 PRÉSENTATION DE L'ARTICLE

## 2.1.1 • DESCRIPTION DU MODÈLE

L'article aborde la modélisation du prix d'un actif financier dans un contexte de trading à haute fréquence. Le processus sous-jacent  $S_t$  est défini comme le prix de l'actif, tandis que  $X_t = \log S_t$  suit un processus stochastique d'Itô représenté par  $dX_t = \mu_t dt + \sigma_t dB_t$ , où  $\mu_t$  est le coefficient de drift,  $\sigma_t$  est la variance instantanée, et  $B_t$  est un mouvement brownien standard. Ce modèle permet de caractériser la volatilité intégrée ( $\int_0^T \sigma_t^2 dt$ ) sur une période donnée, une mesure essentielle pour les analyses financières. Traditionnellement, la volatilité est estimée par la somme des carrés des rendements à haute fréquence, une approche appelée "volatilité réalisée". Cependant, dans le contexte de la microstructure des marchés, cette méthode devient biaisée à des fréquences d'échantillonnage élevées en raison du bruit de mesure et d'autres phénomènes comme le spread bid-ask.

## 2.1.2 • DYNAMIQUE DES PRIX

Pour remédier aux problèmes liés au bruit de microstructure, le modèle considère les observations comme  $Y_{t_i} = X_{t_i} + \epsilon_{t_i}$ , où  $X_t$  est le processus de retour latent (véritable) et  $\epsilon_t$  est un bruit indépendant. Ce bruit représente les distorsions induites par des facteurs comme les contraintes de liquidité ou les effets des ordres de marché. Une méthode classique pour éviter le biais est de réduire la fréquence d'échantillonnage, par exemple en utilisant des intervalles de 5 à 30 minutes entre les observations, mais cela implique une perte importante de données. L'article propose une solution innovante en intégrant le bruit directement dans la procédure d'estimation, permettant une correction systématique des effets du bruit. Cette approche améliore considérablement l'estimation de la volatilité tout en exploitant pleinement les données à haute fréquence.

## 2.1.3 • ESTIMATEURS DE LA VOLATILITÉ RÉALISÉE

Afin de résoudre la problématique de l'estimation de la volatilité réalisée durant un intervalle de temps précis, les auteurs de l'article nous propose 5 (4 en réalité) estimateurs rangés par ordre croissant de performance :

1. Estimateur naïf (on ignore complètement le bruit)
2. (2 et 3) Estimateur avec sampling sélectif des données à une fréquence donnée (et détermination de la fréquence optimale)
3. Estimateur avec sampling sélectif plus moyennage
4. Estimateur avec sampling sélectif plus moyennage et correction de biais : l'estimateur "à deux temporalités"

**1. Estimateur naïf**

En effet, cet estimateur se calcule comme suit :  $E_4 = \sum_{t_i, t_{i+1} \in [0, T]} (Y_{t_{i+1}} - Y_{t_i})^2 := [Y, Y]_T^{all}$ . À l'idéal, cet estimateur devrait être proche de la volatilité réalisée  $\sigma^2 T$ . Or, à cause de la microstructure du marché, celui ci est **totalemtent noyé par le bruit**, menant à des résultats non-fiables :

$$\begin{aligned} E_4 &= \sum_{t_i, t_{i+1} \in [0, T]} (Y_{t_{i+1}} - Y_{t_i})^2 \\ &= 2n\mathbb{E}(\epsilon^2) + O_p(n^{\frac{1}{2}}) \end{aligned}$$

Plus encore, on se rend compte qu'il est plutôt un estimateur fiable du bruit. Cela sera confirmé par les simulations numériques.

### 2. Estimateur avec sampling sélectif des données à la fréquence optimale

Cet estimateur calcule la volatilité réalisée de la même manière que précédemment, mais sur un échantillon de données réduit.

En effet, une fréquence d'échantillonnage est donc choisie (elle peut être arbitraire ou optimisée), ce qui résulte en une perte de données :  $E_3 = [Y, Y]_T^{sparse}$ . On trouve que :

$$E_3 = {}^{\mathcal{L}} \langle X, X \rangle_T + 2n_{sparse}^* \mathbb{E}(\epsilon^2) + [4n_{sparse}^* \mathbb{E}(\epsilon^4) + \frac{2T}{n_{sparse}^*} \int_0^T \sigma_t^4 dt]^{\frac{1}{2}} Z_{total},$$

$$\text{where } n_{sparse}^* = \left( \frac{T}{4(\mathbb{E}(\epsilon^2))^2} \int_0^T \sigma_t^4 dt \right)^{\frac{1}{3}} \text{ nombre optimal de points et } Z_{total} \sim \mathcal{N}(0, 1)$$

Ce resultat sera également vérifié par des simulations numériques.

### 3. Estimateur avec sampling sélectif plus moyennage

Cet estimateur entend répondre au problème de perte de données qui se pose avec l'estimateur précédent. Pour ce faire, la fréquence optimale est maintenu, mais désormais,  $E_3$  est calculé sur toutes les sous-grilles de taille  $\bar{n}^*$  et  $E_2$  est la moyenne de ces  $K^* = \lfloor \frac{n}{\bar{n}^*} \rfloor$  estimateurs :  $E_2 := [Y, Y]_T^{(avg)}$ .

$$E_2 = {}^{\mathcal{L}} \langle X, X \rangle_T + 2\bar{n} \mathbb{E}(\epsilon^2) + [4\frac{\bar{n}}{K^*} \mathbb{E}(\epsilon^4) + \frac{4T}{3\bar{n}} \int_0^T \sigma_t^4 dt]^{\frac{1}{2}} Z_{total},$$

$$\text{where } \bar{n}^* = \left( \frac{T}{6(\mathbb{E}(\epsilon^2))^2} \int_0^T \sigma_t^4 dt \right)^{\frac{1}{3}} \text{ nombre optimal de points et } Z_{total} \sim \mathcal{N}(0, 1)$$

On se rend compte que cet estimateur reste biaisé, même s'il y a une réduction du biais comparé au précédent.

### 4. Estimateur avec sampling sélectif plus moyennage et correction de biais

Le dernier estimateur - l'estimateur à deux temporalités - utilise l'estimateur précédent en lui ajoutant un terme de correction de biais. Ainsi,  $E_1 = [Y, Y]_T^{(avg)} - [Y, Y]_T^{all}$ . Comme on peut le voir, cet estimateur fait intervenir deux quantités à des échelles de temps différentes, d'où son nom. On a, en fait :

$$E_5 = {}^{\mathcal{L}} \langle X, X \rangle_T + \frac{1}{n^{\frac{1}{6}}} \left[ \frac{8}{c^2} \mathbb{E}(\epsilon^2)^2 + c \frac{4T}{3} \int_0^T \sigma_t^4 dt \right]^{\frac{1}{2}} Z_{total}$$

$$\text{ou } c^* = \left( \frac{T}{12(\mathbb{E}(\epsilon^2))^2} \int_0^T \sigma_t^4 dt \right)^{-\frac{1}{3}}$$

Comme on peut le voir, cet estimateur est asymptotiquement sans biais. L'article suggère son utilisation car il fait usage de toutes les données disponibles et corrige le bruit et les effets de discrétisations.

## 2.2 SIMULATIONS NUMÉRIQUES

Les simulations numériques effectuées nous ont permis d'atteindre deux objectifs :

- Vérifier les résultats théoriques présentés dans l'article
- Mettre en exergue des limitations du modèle que nous élaborerons en discussions.

Afin d'obtenir les résultats que nous allons présenter, nous avons utilisé les paramètres suivants :

- $\mu = 0$  et  $\sigma = 0.2$  pour le modèle de Black-Scholes
- $S_0 = 100$ ,  $T = 2$

- Bruit :  $0.05\mathcal{N}(0, 1)$
- $n = 1000$  pas de temps
- $m = 100000$  échantillons

En ce qui concerne les échantillonnages, nous nous sommes limités au cas des fréquences d'échantillonnages équidistants pour leur simplicité de mise en œuvre.

### 2.2.1 • VÉRIFICATIONS EMPIRIQUES DES RÉSULTATS THÉORIQUES

Il était question ici d'effectuer des vérifications des formules obtenues dans l'article. Nous avons fait cela en affichant les histogrammes des échantillons obtenus, normalisés judicieusement afin d'obtenir une forme gaussienne standard (Théorème Central Limite).

Sur la figure 4, les résultats que nous avons obtenus pour les 4 différents estimateurs.

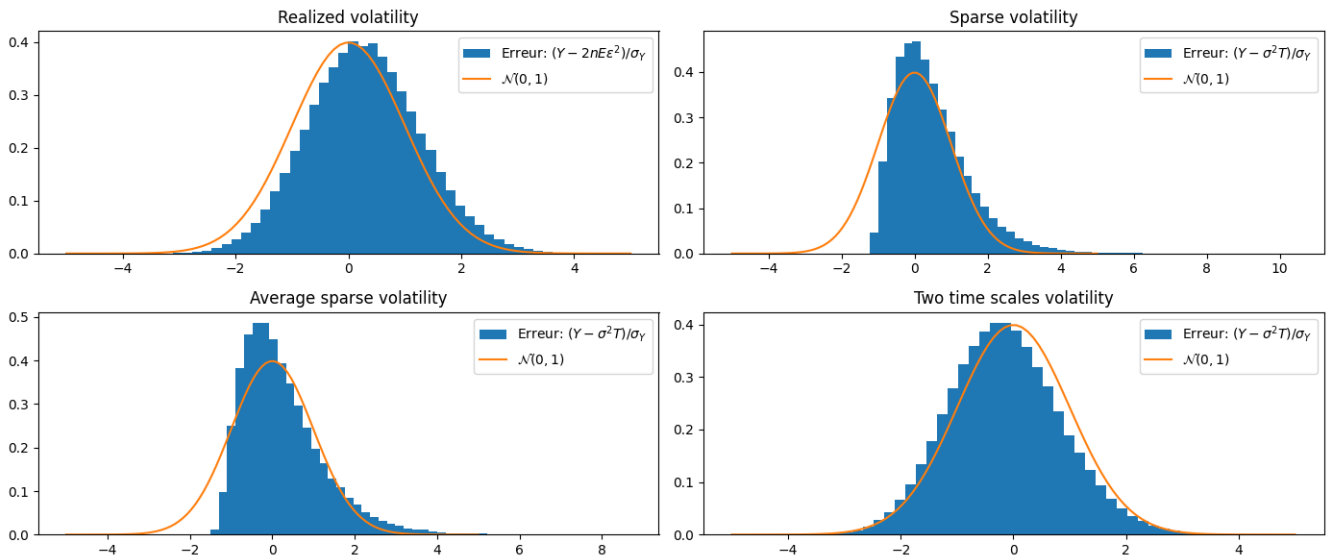


FIGURE 4 – Histogrammes des 4 estimateurs normalisés

#### Focus : La volatilité réalisée - estimateur du bruit ?

Comme on peut le voir, l'estimateur naïf (la volatilité réalisée) est en concordance avec le TCL lorsqu'il est centré par la variance totale du bruit comme établi dans l'article.

En essayant de la centrer par la volatilité totale  $\sigma^2 T$ , voici ce qu'on obtient : la figure 5.

### 2.2.2 • TRAJECTOIRES DE PRIX

La deuxième expérience menée était d'analyser les trajectoires des prix observés en faisant varier le bruit ajouté. (Figure 5)

Comme nous pouvons le voir, malgré les différentes intensités du bruit ajouté, les prix se comportent toujours de manière similaire à des Browniens (continuité).

## 2.3 DISCUSSIONS

- Les différentes simulations numériques menées nous ont permis de nous vérifier les résultats théoriques énoncés. Dans le cas des estimateurs 2 et 3, le décalage avec la gaussienne standard est sûrement dû au

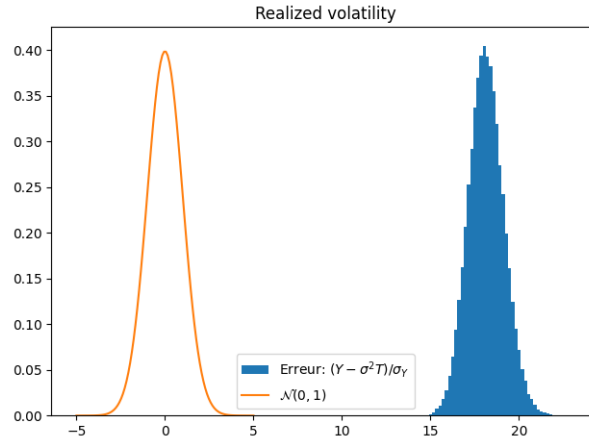


FIGURE 5 – Volatilité réalisée centrée par  $\sigma^2 T$  : complètement noyée par le bruit

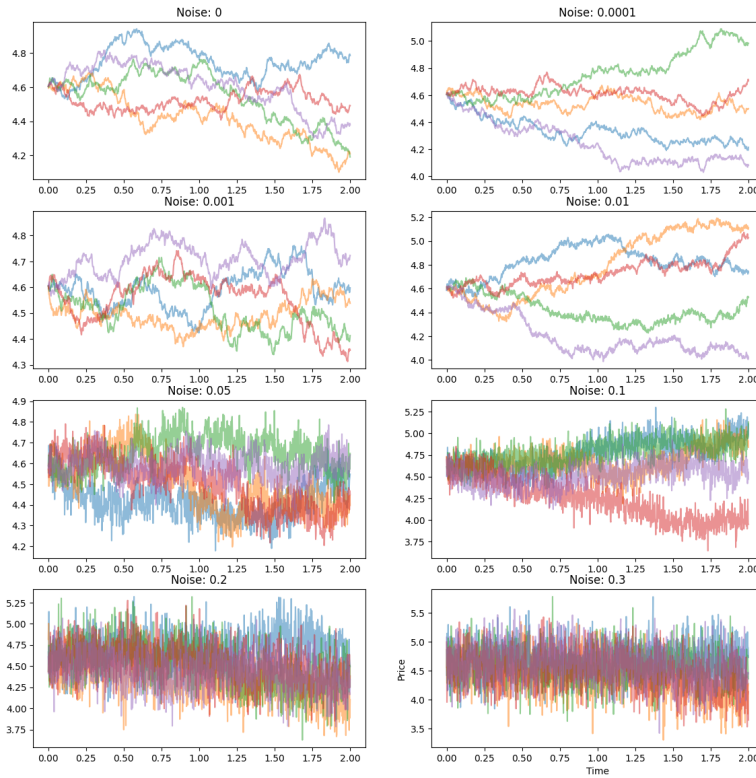


FIGURE 6 – 5 trajectoires des prix par intensité du bruit

fait que l'échantillonnage équidistant n'est pas optimal.

- Cependant, nous avons pu obtenir ces résultats uniquement parce que nous étions dans une situation théorique où nous avons la valeur réelle de la volatilité. En effet, l'article présente un **souci majeur** : Pour estimer la volatilité, les différents estimateurs (formules présentées en 2.1.3) font appel à une intégrale de la volatilité qui est en principe... inconnue. Ainsi, ces formules reposent sur un estimateur oracle.
- Comme on a pu le voir sur la figure 6, les trajectoires des prix obtenus en ajoutant du bruit gaussien indépendant, ne permettent pas d'obtenir la caractéristique discrète typique des prix intraday, comme

dans l'image sur la figure 7.



FIGURE 7 – Exemple de trajectoire intraday. Source : Lynx broker

## 2.4 CONCLUSION 2

Parvenu au terme de ce second article, nous avons étudié 4 estimateurs de la volatilité totale pendant une période donnée. Les étudier successivement nous a permis de comprendre les avantages et inconvénients de chacun et comment et pourquoi l'estimateur "two-time scales" est défini tel quel.

Comme relevé dans les discussions, les prix obtenus par ajout de bruits gaussiens indépendants à un mouvement brownien (et même n'importe quel autre type de bruits plus généralement) ne nous permettent pas d'obtenir des prix semblables à la réalité des prix intraday.

Ainsi, il sera question pour nous dans la suite d'étudier des modèles répliquant cette caractéristique particulière du trading haute fréquence.

## 3

## Robert et Rosenbaum, 2010

## 3.1 PRÉSENTATION DE L'ARTICLE

L'article ROBERT et ROSENBAUM, 2010, répond à une question centrale en finance à haute fréquence : **comment modéliser les prix transactionnels observés tout en tenant compte de leurs caractéristiques spécifiques (discrétion, volatilité intra-jour, et espacement irrégulier) tout en les reliant à un prix latent supposé continu ?**

Les modèles classiques s'attaquant à cette problématique supposent souvent que les prix observés sont des approximations bruitées d'un prix latent continu (un processus semi-martingale). Cependant, ces modèles ne tiennent pas compte de l'aversion des participants au marché pour les changements de prix, ni des mécanismes microstructurels qui influencent les transactions à des fréquences élevées.

Ainsi, le coeur de l'article est l'hypothèse suivante : pour qu'un changement de prix transactionnel se produise, le prix latent doit sortir d'une **zone d'incertitude**. Cette zone d'incertitude reflète les conditions dans lesquelles les participants au marché considèrent un changement de prix comme justifié. Une description plus précise sera faite en 3.1.2.

## 3.1.1 • DESCRIPTION DU MODÈLE

Le modèle repose sur l'idée que les prix transactionnels sont déterminés par un mécanisme stochastique basé sur un prix latent efficace supposé être un semi-martingale d'Itô.

En effet,  $(X_t)_{t \geq 0}$  est le processus des prix latents efficaces ; sur un espace de probabilité filtré bien choisi  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ,  $(Y_t)_{t \geq 0}$  log-prix est une semi-martingale d'Itô :

$$Y_t = \log X_t = \log X_0 + \int_0^t a_u du + \int_0^t \sigma_{-u} dW_u$$

avec  $(W_t)_{t \geq 0}$  un  $\mathcal{F}$ -Mouvement Brownien.

Les prix transactionnels doivent appartenir à une grille de ticks  $\{k\alpha, k \in \mathbb{N}\}$ ,  $\alpha$  est la taille d'un tick. Pour chaque  $k \in \mathbb{N}$  et  $\eta \in ]0, 1[$ ,  $U_k$ , la zone d'incertitude est définie par  $U_k = [0, \infty) \times (d_k, u_k)$  avec

$$d_k = (k + \frac{1}{2} - \eta)\alpha \text{ et } u_k = (k + \frac{1}{2} + \eta)\alpha$$

La largeur de ces zones d'incertitude est contrôlée par  $\eta$ , qui reflète la sensibilité du marché.

## 3.1.2 • ALGORITHME DE CONSTRUCTION DES PRIX

Le but ici est de construire une suite de **temps de sortie** des zones d'incertitude  $(\tau_i)_{i \geq 0}$  capturant les comportements essentiels des participants du marché, à savoir :

- Les participants ne réagissent pas immédiatement aux petites fluctuations du prix latent. Ils considèrent que ces mouvements sont bruités et non significatifs
- Un changement de prix transactionnel se produit uniquement lorsque le prix latent s'éloigne suffisamment du dernier prix enregistré, au point où les participants estiment que ce dernier ne reflète plus la valeur réelle de l'actif.

*Algorithme*

1. Initialisation :  $\tau_0 = 0, \tau_1$  : temps de sortie de  $(d_{k_0-1}, u_{k_0})$ , avec  $k_0 = \frac{X_0^{(\alpha)}}{\alpha}$  où  $X_0^{(\alpha)}$ , la valeur approchée de  $X_0$  à un multiple de  $\alpha$ .



2. Induction : On suppose que  $(L_i)_{i \geq 1} \in \mathcal{F}_{\tau_i}, \tau_i$  temps de sortie de  $(d_{k_i-L_i}, u_{k_i+L_i-1})$  avec  $k_i = \frac{X_{\tau_i}^{(\alpha)}}{\alpha}$  et

$$\tau_{i+1} = \inf\{t : t > \tau_i, X_t \in \{X_{\tau_i}^{(\alpha)} - \alpha(L_i + \frac{1}{2} + \eta), X_{\tau_i}^{(\alpha)} + \alpha(L_i + \frac{1}{2} + \eta)\}\}$$

En ce qui concerne la suite  $(t_i, P_{t_i})_{i \geq 0}$  des couples de temps de transaction et des prix transactionnels, on les obtient comme suit :

1. Initialisation :  $t_0 = 0$  et  $P_0 = X_0^{(\alpha)}$
2. Induction : On suppose qu'entre  $\tau_i$  et  $\tau_{i+1}$ , au moins une transaction a eu lieu :  $\tau_i \leq t_i < \tau_{i+1}, t_i$  : Premier temps de transaction et  $P_{t_i} = X_{\tau_i}^{(\alpha)}$ .  $[t_i, t_{i+1})$  : Intervalle de "flat trading".

### 3.1.3 • ESTIMATION DE LA VOLATILITÉ RÉALISÉE

L'article montre comment le modèle peut être utilisé pour estimer la volatilité intégrée du prix latent. Une méthode simple consiste à :

1. : Estimer le paramètre  $\eta$  :

$$\hat{\eta}_{\alpha,t} = \frac{N_{\alpha,t}^{(c)}}{2N_{\alpha,t}^{(a)}}$$

avec  $N_{\alpha,t}^{(c)}$  (resp.  $N_{\alpha,t}^{(a)}$ ) : Nombre de continuations (resp. alternations) sur  $[0, t]$ . Une continuation (resp. alternation) d'un tick est un saut de prix dont la direction est similaire (opposée) à celle du prix précédent.

2. Reconstruire une approximation des prix latents :

$$X_{\tau_i} = P_{t_i} - \text{sign}(P_{t_i} - P_{t_{i-1}})(1/2 - \hat{\eta}_{\alpha,t})\alpha$$

3. Appliquer la formule connue de l'espérance réalisée (des log-prix)

$$\int_0^t \sigma_s^2 ds = \sum_{\tau_i \leq t} \left( \frac{\hat{X}_{\tau_i} - \hat{X}_{\tau_{i-1}}}{\hat{X}_{\tau_{i-1}}} \right)^2$$

## 3.2 SIMULATIONS NUMÉRIQUES

Pour mener les simulations numériques, nous nous sommes basés sur des cas simplificateurs :  $a_u = 0$  et  $\sigma_{-u} = cte = 0.2$ . Nous avons utilisé les paramètres suivants :

- $X_0 = 100$
- Horizon  $t = 1$  (1 jour de trading)
- Taille de tick  $\alpha = 0.5$ , aversion  $\eta = 0.6$
- Accuracy de simulation  $acc = 0.001$
- Fréquence d'échantillonnage du Brownien  $n = 8 \times 3600 \times acc^{-1} = 288 \times 10^5$
- $L_i = 1, \forall i \in \mathbb{N}$

### 3.2.1 • TRAJECTOIRES DE PRIX

Comme on peut le constater sur la figure 8, les caractéristiques des prix intraday sont bien présentes et les prix construits suivent bien la trajectoire des prix latents.

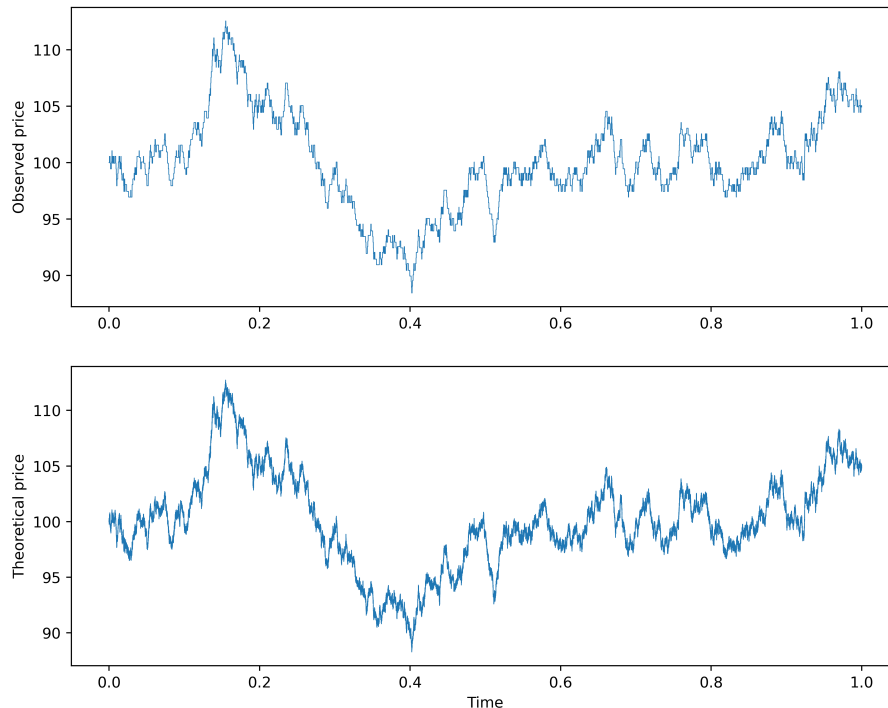


FIGURE 8 – Prix observés Vs prix efficients

### 3.2.2 • ESTIMATION DE LA VOLATILITÉ

Afin d'estimer le paramètre de volatilité, nous avons procédé comme expliqué en 3.1.3.

Les résultats obtenus témoignent de l'efficacité de la méthode.

- $\eta$  estimée : 0.622 (erreur : 3.74 % )
- Volatilité estimée : 0.267 (erreur 33.85%)
- Estimateur Two Time Scales (ZHANG et al., 2005) : 0.211 (5.81%)

La performance de l'estimateur Two Time Scales varie grandement au travers des réalisations. Comme on peut le voir dans la figure 9, le TCL n'est pas vérifié : L'histogramme n'est pas centré et a un écart type considérable par rapport à l'estimateur de volatilité étudié qui a en plus un écart-type de l'ordre de  $\sigma$ .

## 3.3 DISCUSSIONS

Nous avons constaté après lecture de cet article qu'il dépasse les approches traditionnelles en introduisant un mécanisme stochastique réaliste qui explique :

1. **La dynamique des prix transactionnels** : Les prix transactionnels observés sont reliés à un processus latent continu tout en intégrant les effets de microstructure.
2. **La formation des prix observés** : Les prix transactionnels émergent d'interactions entre le prix latent et l'aversion au changement des participants.
3. **Des applications pratiques** : En modélisant précisément les zones d'incertitude, l'approche peut être utilisée pour estimer des grandeurs clés comme la volatilité intégrée ou l'aversion des participants.

Nous avons également constaté les limitations de l'estimateur Two scales : Vu que les prix sont liés les uns aux autres, de par l'algorithme de formation, ils ne peuvent être considérés comme des observations indépendantes et identiquement distribuées (iid), ce qui est l'hypothèse de base de Two scales.

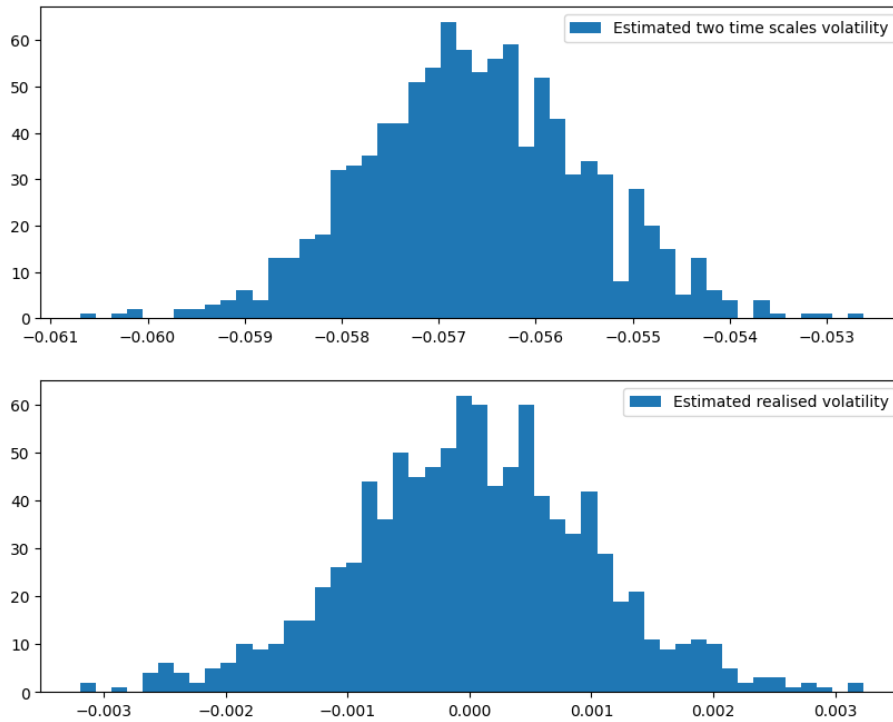


FIGURE 9 – Estimateur Two Time Scales sur les prix du modèle de zones d'incertitude

### 3.4 CONCLUSION 3

Parvenu au terme de cet article, nous avons étudié une approche clé pour la modélisation des prix transactionnels qui tient compte des caractéristiques spécifiques de "l'intraday" et de la microstructure des marchés. Grâce à cette méthode de modélisation, nous pouvons estimer des grandeurs clés du marché avec une très bonne précision.

## 4

### Hayashi et Yoshida, 2005

#### 4.1 PRÉSENTATION DE L'ARTICLE

L'article de Hayashi et Yoshida aborde le problème de l'estimation de la covariance de deux processus de diffusion dans le contexte d'observations non synchrones. Ils analysent les limites des estimateurs classiques, tels que l'estimateur de covariance réalisée, et proposent une méthode alternative basée sur les observations originales, sans nécessiter de synchronisation préalable des données.

##### 4.1.1 • DESCRIPTION DU MODÈLE

Le modèle repose sur des processus de diffusion définis par des équations différentielles stochastiques. Soient  $P_1(t)$  et  $P_2(t)$  deux processus de diffusion décrits par :

$$dP_1(t) = \mu_1(t)dt + \sigma_1(t)dW_1(t), \quad dP_2(t) = \mu_2(t)dt + \sigma_2(t)dW_2(t),$$

où  $W_1(t)$  et  $W_2(t)$  sont des mouvements browniens corrélés avec  $\langle W_1, W_2 \rangle_t = \rho(t)dt$ .

Les observations des prix sont effectuées à des instants aléatoires et non synchrones, ce qui complique l'estimation de la covariance.

L'objectif est d'estimer la covariation cumulative  $\langle P_1, P_2 \rangle_T$ , définie comme :

$$\langle P_1, P_2 \rangle_T = \int_0^T \sigma_1(t)\sigma_2(t)\rho(t)dt.$$

##### 4.1.2 • DYNAMIQUE DES PRIX ET EFFET DE EPPS

L'effet de Epps, observé dans des données haute fréquence, montre que les corrélations entre actifs financiers tendent à diminuer à mesure que la fréquence d'échantillonnage augmente. Cet effet est lié à l'asynchronie des observations, où les prix ne sont pas enregistrés au même instant pour différents actifs.

Dans cet article, l'effet est explicité en montrant que l'estimateur de covariance réalisée :

$$V_h = \sum_{i=1}^m (P_1(t_i) - P_1(t_{i-1}))(P_2(t_i) - P_2(t_{i-1})),$$

présente un biais important lorsque  $h$  tend vers 0, l'intervalle d'échantillonnage, est petit par rapport aux intervalles réels d'observation des transactions. Cela résulte de la perte d'informations sur les mouvements non synchrones des prix.

##### 4.1.3 • ESTIMATEURS DE HY DE LA COVARIANCE

Hayashi et Yoshida introduisent un nouvel estimateur, basé sur les observations originales, qui ne nécessite aucune synchronisation des données :

$$U_n = \sum_{i,j} \Delta P_1(I_i) \Delta P_2(J_j) \mathbf{1}_{I_i \cap J_j \neq \emptyset},$$

où  $\Delta P_1(I_i)$  et  $\Delta P_2(J_j)$  sont les incréments des processus sur les intervalles d'observation  $I_i$  et  $J_j$ , respectivement.

Cet estimateur est démontré comme étant :

- **Consistant** :  $U_n \rightarrow \langle P_1, P_2 \rangle_T$  en probabilité lorsque le nombre d'observations augmente.
- **Sans biais** : Il ne dépend pas des choix arbitraires de synchronisation ou d'imputation.

## 4.2 SIMULATIONS NUMÉRIQUES

Pour mener les simulations numériques, Nous avons utilisé les paramètres suivants :

- Horizon  $T = 1$  (1 jour de trading)
- Nombre de pas dans le temps continu  $n_{steps} = 1000$
- Corrélation entre les deux actifs  $\rho = 0.6$
- Volatilités des actifs  $vol_1, vol_2 = 0.2, 0.3$

### 4.2.1 • DYNAMIQUE DES PRIX DE 02 ACTIFS OBSERVÉS ET IMPACT DE LA CORRÉLATION

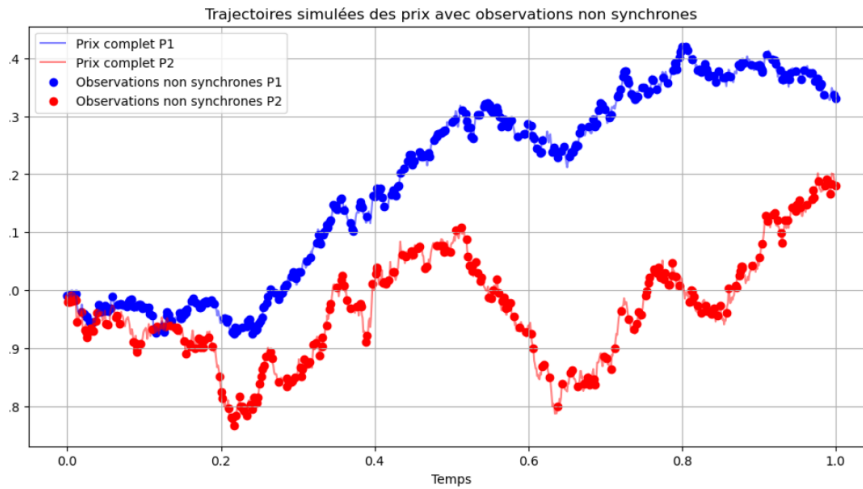


FIGURE 10 – Prix observés-prix efficients de 02 actifs

Les trajectoires des deux actifs corrélés générées dans la simulation montrent clairement l'effet de la corrélation  $= 0.6$ . Les prix des deux actifs suivent des dynamiques similaires, bien qu'ils soient soumis à des variations individuelles influencées par leurs propres volatilités ( $vol_1 = 0.2$ ,  $vol_2 = 0.3$ ). Cette corrélation partielle est bien capturée dans les données simulées, reflétant la dépendance des mouvements des deux actifs.

De plus, la distinction entre les prix observés et les prix efficients met en évidence l'importance de prendre en compte les observations asynchrones pour une estimation précise. Cela illustre également l'effet de la granularité des données haute fréquence et son rôle dans la précision des calculs.

#### 4.2.2 • ESTIMATEUR $U_n$ / VALEUR RÉELLE

```
In [3]: # Fonction pour calculer les variations observées
def calculate_estimator(t1, P1, t2, P2):
    U_n = 0
    for i in range(len(t1) - 1):
        for j in range(len(t2) - 1):
            # Identifier les intervalles qui se chevauchent
            if t1[i+1] > t2[j] and t2[j+1] > t1[i]:
                delta_P1 = P1[i+1] - P1[i]
                delta_P2 = P2[j+1] - P2[j]
                U_n += delta_P1 * delta_P2
    return U_n

# Calcul de l'estimateur
U_rho*vol_1*vol_2
U_n = calculate_estimator(t1_obs, P1_obs, t2_obs, P2_obs)
print(f"Estimateur U_n : {U_n:.4f}")
print(f"valeur réelle U : {U:.4f}")
#print(t1_obs, t2_obs)

Estimateur U_n : 0.0380
valeur réelle U : 0.0360
```

FIGURE 11 – Estimateur  $U_n$  / valeur réelle

La comparaison entre les résultats de l'estimateur  $U_n$  et la covariation réelle intégrée

$$\langle P_1, P_2 \rangle_T$$

confirme la pertinence de l'estimateur proposé par Hayashi et Yoshida. Les simulations montrent que  $U_n$  converge vers la valeur réelle lorsque le nombre d'observations augmente, validant sa consistance.

#### 4.2.3 • DISTRIBUTION SIMULÉE DE $U_n$

Statistique du test de Shapiro-Wilk : 0.9974  
p-value : 1.0522e-01

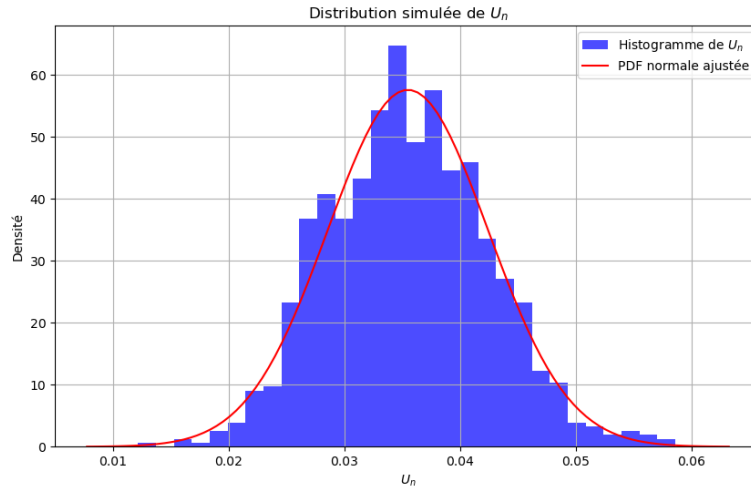


FIGURE 12 – Distribution simulée de  $U_n$

La distribution empirique simulée de  $U_n$  présente une concentration autour de la covariation réelle, confirmant que cet estimateur est non biaisé. Cela est cohérent avec les résultats théoriques selon lesquels  $U_n$  ne dépend pas des choix arbitraires de synchronisation ou d'imputation des données. Cependant, la distribution révèle également une variabilité dépendant des paramètres simulés, notamment la corrélation et les volatilités des actifs. Cette variabilité souligne que la précision de  $U_n$  peut être influencée par les caractéristiques propres aux actifs analysés.

## 4.3 DISCUSSIONS

---

Les simulations réalisées dans le cadre de cet article apportent une perspective innovante en dépassant les approches traditionnelles grâce à une méthode robuste d'estimation en contexte d'observations non synchrones. Elles démontrent :

- La prise en compte de l'asynchronie des données : Contrairement aux estimateurs classiques, l'approche de Hayashi-Yoshida s'affranchit de la nécessité de synchroniser les observations, permettant une modélisation réaliste des données transactionnelles.
- L'efficacité de l'estimation : En utilisant directement les incréments des processus sur les intervalles d'observation, l'estimateur  $Un$  est cohérent et non biaisé, même en présence d'effets complexes comme l'effet de Epps.
- Des applications prometteuses : La méthode proposée ouvre des possibilités d'utilisation sur des données haute fréquence pour estimer des covariations en préservant les caractéristiques intrinsèques des données.

Cependant, les simulations montrent également les limites des méthodes traditionnelles, telles que les estimateurs réalisés, qui subissent des biais significatifs lorsque les données ne respectent pas l'hypothèse d'observations indépendantes et identiquement distribuées. En cela, l'approche de Hayashi-Yoshida se distingue par sa robustesse et son adaptabilité.

## 4.4 CONCLUSION 4

---

L'article met en évidence les limitations des méthodes classiques d'estimation de la covariance dans des données haute fréquence non synchrones et propose une alternative robuste et sans biais. L'estimateur de Hayashi-Yoshida constitue une avancée importante, permettant d'utiliser efficacement les données brutes tout en évitant les biais dus à la synchronisation.

Cette méthode ouvre des perspectives intéressantes pour l'analyse des données financières et pourrait être étendue à des cas plus complexes, notamment des processus avec bruit microstructural ou des observations dépendant des prix, comme cela a été fait dans l'article de Robert et Rosenbaum, 2010.

## CONCLUSION

---

Parvenu au terme de cet enseignement où nous avons étudié 4 articles abordant des problématiques clés présentes dans le domaine du Trading haute fréquence, nous avons tout au long de celui-ci, acquis des connaissances sur des modèles financiers plus anciens comme ALMGREN et CHRISS, 2000 ou plus récents comme ROBERT et ROSENBAUM, 2010.

Cet EA aura donc été l'occasion pour nous de plonger dans le domaine de la recherche en finance haute fréquence : de la lecture des articles à l'implémentation des approches utilisées, ainsi que leur critique constructive.



## RÉFÉRENCES

---

- ALMGREN, R., & CHRISS, N. (2000). Optimal Execution of Portfolio transactions. *Journal of Risk*.
- HAYASHI, T., & YOSHIDA, N. (2005). On Covariance Estimation of Non-synchronously Observed Diffusion Processes. *Bernoulli*.
- ROBERT, C., & ROSENBAUM, M. (2010). A New Approach for the Dynamics of Ultra-High-Frequency Data : The Model with Uncertainty Zones. *Oxford University Press*.
- ZHANG, L., MYKLAND, P., & AIT-SAHALIA, Y. (2005). A Tale of Two Time Scales : Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*.