

Limit Price and LOB

Wentao Wang, Yuhang Guo

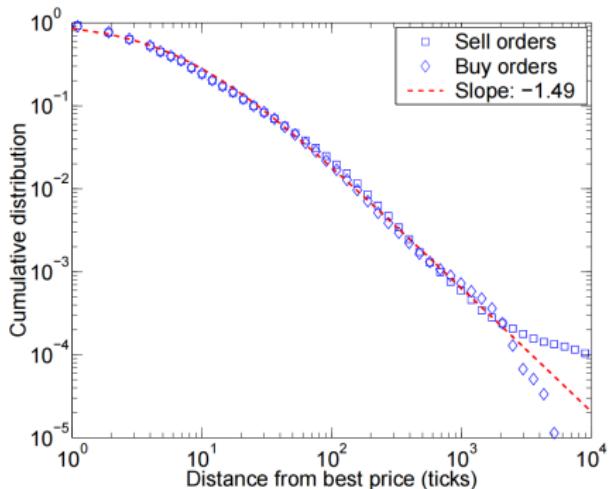
Présentation du cours APM_52112_EP

27 jan 2026

- **Power-Law distribution of limit price:**
 - **Zovko et al.** stat the slow-moving caracteristic of limit price and its coupling with volatility.
 - **Bouchaud et al.** show even if price limits are 'random', order book structure can form automatically.
- **Estimation of relative entropy:** **Sasaki et al.** estimate higher-order derivatives of density without density estimation.
- **Estimation of Hawkes process and asymptotic characteristics:** **Cavaliere et al.** propose a fixed-intensity bootstrap framework to improve finite-sample inference for Hawkes and general point processes.

Power-Law of Relative Limit Price

$$\mathbb{P}(\delta \geq x) \underset{x \rightarrow \infty}{\sim} (x + x_0)^{-\beta}, \quad \beta \approx 1.5.$$



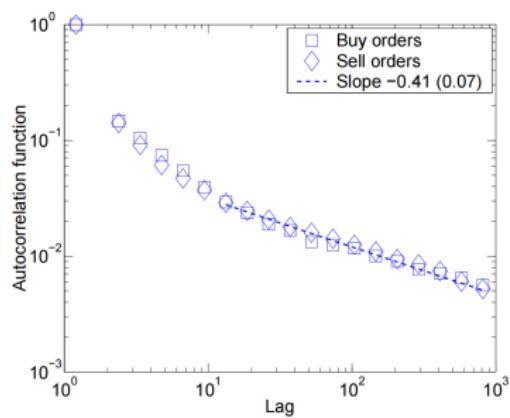
- PDF index 2.5
- Average converge
- Variation diverge

The data exhibits fat tails, indicating the presence of numerous limit orders with extremely low probability of execution yet highly favourable pricing. Traders may adopt a long-term perspective.

Temporal Clustering and Long-Range Dependence

$\delta(t)$ is not i.i.d.; large δ /small δ clusters emerge, vanishing after randomising the temporal order.

$$C(\tau) \underset{\tau \rightarrow \infty}{\sim} \tau^{-0.4}.$$



Delta and volatility link

To ensure the convergence, δ is cutted off at 1000 ticks, and averaged with respect to time.

We define:

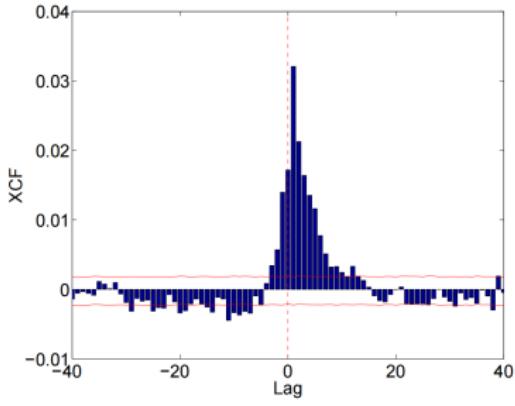
$$\text{XCF}(\tau) = \frac{\langle v(t - \tau) \delta(t) \rangle - \langle v(t) \rangle \langle \delta(t) \rangle}{\sigma_v \sigma_\delta}.$$

A significant positive correlation emerged, with the peak slightly skewed towards $\tau > 0$, indicating that volatility leads limit orders in patience.

Statistical Significance

To eliminate the false correlation caused by long memory of delta and volatility, we construct the null hypothesis as follow:

- Use Theiler method to randomize the phases uniformly in $[0, 2\pi]$, giving same marginal distribution and autocorrelation as $\delta(t)$, but no time alignment with volatility $v(t)$.
- Repeat 300 times and compute the XCF for each surrogate realization, draw 2.5% and 97.5% quantiles.



Mechanism

A positive feedback loop can be observed:

- Volatility $\uparrow \rightarrow$ Traders become more patient ($\delta \uparrow$)
- \rightarrow Near-term order book thins \rightarrow Impact increases
- \rightarrow Volatility further \uparrow

Conclusion

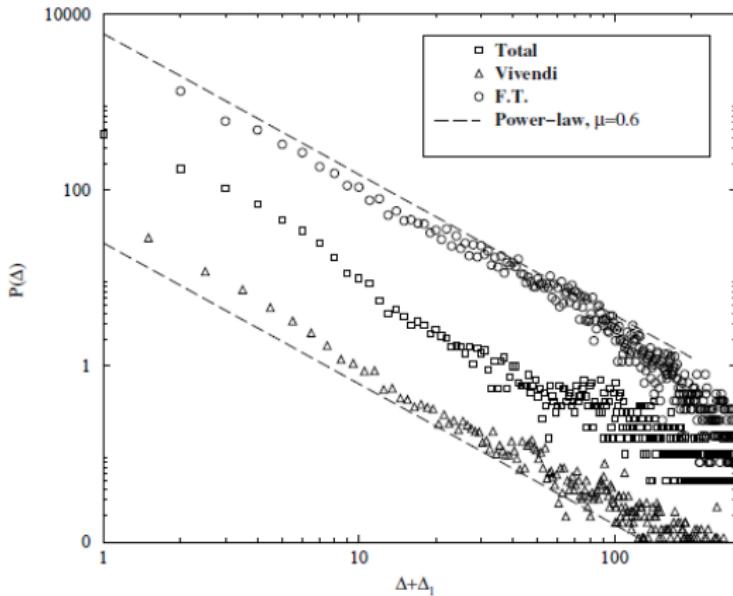
- Relative limit prices exhibit a robust power-law tail, $\mathbb{P}(\delta \geq x) \sim x^{-1.5}$, indicating heavy-tailed patience in order placement.
- The limit price process displays long-range dependence, $C(\tau) \sim \tau^{-0.4}$, implying that order placement is governed by a slow behavioural state.
- Relative limit prices are positively correlated with volatility, with volatility slightly leading, suggesting a feedback mechanism between liquidity and volatility.

Summary

Static properties of three liquid stocks: France-Telecom (F.T.), Vivendi, and Total:

- The distribution of incoming limit orders
- Time - average shape of the order book
- The distribution of volume at best bid/ask

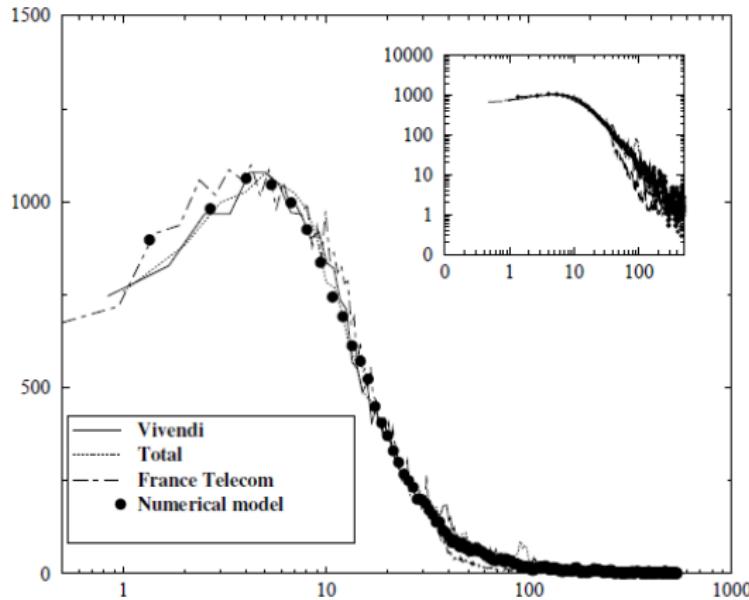
The distribution of incoming limit orders



$$P(\Delta) \propto \frac{\Delta_0^\mu}{(\Delta_0 + \Delta)^{1+\mu}}, \quad \Delta \geq 1. \quad (1)$$

- $\mu \approx 0.6$

Time-average orderbook shape



- has a maximum away from the current bid(ask)

The distribution of volume at bid/ask

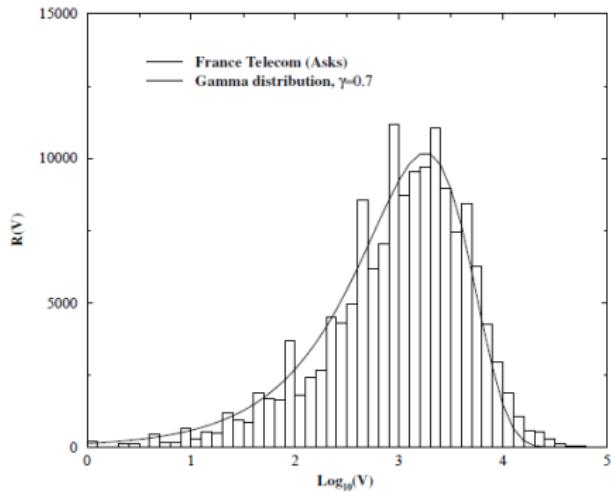


Figure: Histogram of the log-volume at the ask for France Telecom (fit with Gamma distribution of this equation).



$$R(V) \propto V^{\gamma-1} \exp\left(-\frac{V}{V_0}\right) \quad (1)$$

- $\gamma \approx 0.7 - 0.8$ for all three stocks

Zero-Intelligence (ZI) LOB model: motivation and assumptions

- **Motivation:** describe the average shape of the limit order book (LOB) with minimal behavioral assumptions.
- **Zero-intelligence:** agents do not optimize; order flow is generated by simple stochastic rules.
- **Limit order placement:** new limit orders arrive at a distance Δ from the best quote.
- **Market orders:** with probability $p_m(1/6-1/4)$ an incoming order is a market order and removes volume at the best quote.
- **Cancellations:** existing limit orders are removed with constant hazard rate $\Gamma(10^{-3})$ (memoryless cancellations).
- **Price dynamics:** best quotes move when the best queue is depleted; in the theory this is approximated by a diffusion with variance parameter D .

ZI model: core formulae and analytic prediction

- **Empirical placement distribution (power law):**

$$P(\Delta) \propto \frac{\Delta_0^\mu}{(\Delta_0 + \Delta)^{1+\mu}}, \quad \Delta \geq 1. \quad (1)$$

- **Analytic stationary average order book $\rho_{\text{st}}(\Delta)$:** under diffusion approximation for the best quote and exponential cancellations,

$$\rho_{\text{st}}(\Delta) = e^{-\alpha\Delta} \int_0^\Delta du P(u) \sinh(\alpha u) + \sinh(\alpha\Delta) \int_\Delta^\infty du P(u) e^{-\alpha u}, \quad (6)$$

with

$$\alpha^{-1} = \sqrt{\frac{D}{2\Gamma}}.$$

- **Asymptotics (shape):**

$$\rho_{\text{st}}(\Delta) \propto \Delta^{1-\mu} \ (\Delta \rightarrow 0), \quad \rho_{\text{st}}(\Delta) \propto \Delta^{-1-\mu} \ (\Delta \rightarrow \infty).$$

ZI model: results, figures, and takeaway

- **Main result:** despite its simplicity, the ZI model reproduces the *average LOB shape* (a hump away from the best quote and power-law tails).
- **Key scale:** $\alpha^{-1} = \sqrt{D/(2\Gamma)}$ controls how quickly the book relaxes with distance.
- **Limitation:** ignores strategic placement and time-varying order flow; fits average statistics better than short-term dynamics.

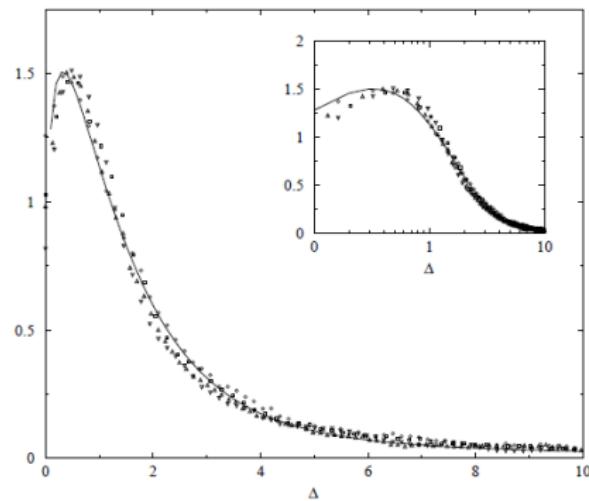
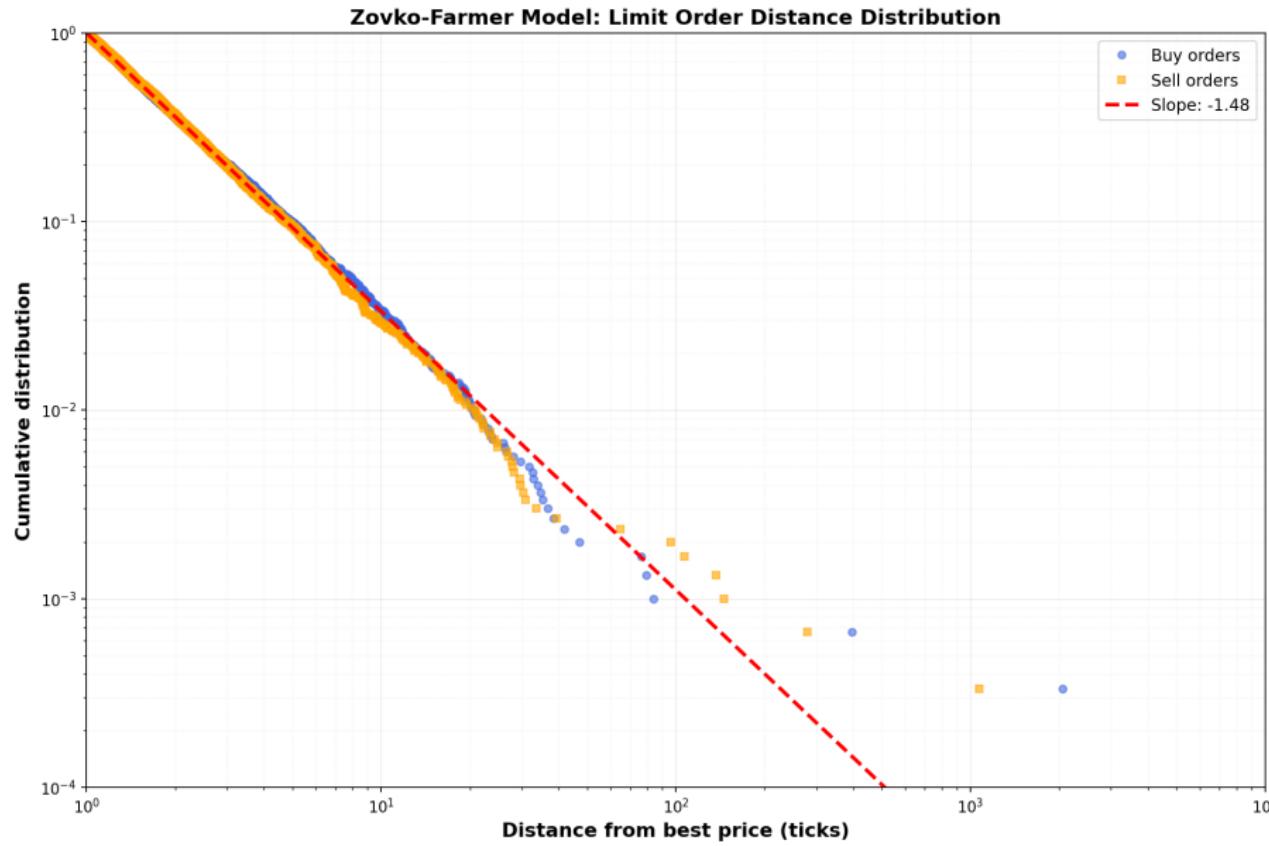


Figure: Average order book: data vs ZI model.

Simulation Sanofi



Introduction

Goal: estimate the higher partial order derivatives of density function:

$$p_{k,\mathbf{j}}(x) = \frac{\partial^k}{\partial x_1^{j_1} \partial x_2^{j_2} \cdots \partial x_d^{j_d}} p(x), \quad \sum_{\ell=1}^d j_\ell = k.$$

- $k = 1$: gradient;
- $k = 2$: hessian.

Traditional KDE method:

$$\hat{p}(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right).$$

With derivative:

$$\hat{p}^{(k)}(x) = \frac{1}{n} \sum_{i=1}^n K^{(k)}\left(\frac{x - x_i}{h}\right).$$

Problem: Bandwidth that maximizes density is not the bandwidth that maximizes derivatives. Higher-order derivatives amplify estimation errors.

MISED Method

Now let $g_{k,j}(x) \approx p_{k,j}(x)$, we define:

$$J(g) = \int (g(x) - p_{k,j}(x))^2 dx.$$

By removing the constant term:

$$J(g) = \int g(x)^2 dx - 2 \int g(x) p_{k,j}(x) dx.$$

We now use integration by parts:

$$\int g(x) p_{k,j}(x) dx. = (-1)^k \int \frac{\partial^k g(x)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}} p(x) dx.$$

Thus we transform the objective to minimize:

$$\tilde{J}(g) = \int g(x)^2 dx - \frac{2(-1)^k}{n} \sum_{i=1}^n \frac{\partial^k g(x_i)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}. \quad (*)$$

Gaussian Kernels

We now apply gaussian kernel to function g:

$$g_{k,j}(x) = \sum_{i=1}^n \theta_{j,i} \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right)$$

Noting:

$$\psi_i(x) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right), \quad \psi(x) = (\psi_1(x), \dots, \psi_n(x))^\top$$

We have:

$$g_{k,j}(x) = \theta_j^\top \psi(x)$$

Since the gaussian kernel is derivable:

$$\frac{\partial^k g_{k,j}(x)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}} = \sum_{i=1}^n \theta_{j,i} \frac{\partial^k}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}} \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) = \sum_{i=1}^n \theta_{j,i} \varphi_{j,i}(x)$$

Analytic Solution

Noting:

$$\varphi_{\mathbf{j}}(x) = (\varphi_{\mathbf{j},1}(x), \dots, \varphi_{\mathbf{j},n}(x))^{\top}$$

Applying the result to formula (*) with L_2 normalization term:

$$\tilde{J}(\theta_{\mathbf{j}}) = \theta_{\mathbf{j}}^{\top} G \theta_{\mathbf{j}} - 2(-1)^k \theta_{\mathbf{j}}^{\top} \mathbf{h}_{\mathbf{j}} + \lambda \theta_{\mathbf{j}}^{\top} \theta_{\mathbf{j}},$$

with : $\mathbf{h}_{\mathbf{j}} = \frac{1}{n} \sum_{i=1}^n \varphi_{\mathbf{j}}(x_i),$

$$G_{i\ell} = \int \psi_i(x) \psi_{\ell}(x) dx = (\pi \sigma^2)^{d/2} \exp\left(-\frac{\|x_i - x_{\ell}\|^2}{4\sigma^2}\right).$$

Objective function is strictly convex, quadratic form, thus possesses closed-form solution:

$$\hat{\theta}_{\mathbf{j}} = (-1)^k (G + \lambda I)^{-1} \mathbf{h}_{\mathbf{j}}.$$

Thus the density derivative estimator is given as:

$$\hat{g}_{k,\mathbf{j}}(x) = \hat{\theta}_{\mathbf{j}}^{\top} \psi(x).$$

CV Parameters Selection

Now we need to choose gaussian kernel width σ and regularization parameter λ .

- Let $\{X^{(t)}\}_{t=1}^T$ be a partition of the dataset X into T disjoint subsets;
- For the t -th fold, the estimator $\hat{g}_{k,j}^{(t)}$ is trained using $X \setminus X^{(t)}$;
- The validation score for the t -th fold is defined as:

$$CV^{(t)}(\sigma, \lambda) = \int (\hat{g}_{k,j}^{(t)}(x))^2 dx - \frac{2(-1)^k}{|X^{(t)}|} \sum_{x \in X^{(t)}} \frac{\partial^k \hat{g}_{k,j}^{(t)}(x)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}};$$

- The overall cross-validation score is given by:

$$CV(\sigma, \lambda) = \frac{1}{T} \sum_{t=1}^T CV^{(t)}(\sigma, \lambda).$$

Then the parameters are chosen as:

$$(\hat{\sigma}, \hat{\lambda}) = \arg \min_{\sigma, \lambda} CV(\sigma, \lambda).$$

NN-KL-Divergence Approximation

Given samples as:

$$X_1 = \{x_i\}_{i=1}^{n_1} \sim p_1, \quad X_2 = \{x_i\}_{i=n_1+1}^{n_1+n_2} \sim p_2$$

Recall the KL-divergence:

$$\text{KL}(p_1 \| p_2) = \int p_1(x) \log \frac{p_1(x)}{p_2(x)} dx$$

We now focus on NNKL-estimator is given as:

$$\widehat{\text{KL}}_{\text{NN}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \log \frac{(n_1 - 1) \text{dist}_1(x_i)^d}{n_2 \text{dist}_2(x_i)^d}$$

Studies show that under large sample limit, the dominant bias term of the NN-KL estimator is given:

$$\boxed{\text{Bias}(x) \propto \frac{\text{tr}(\nabla^2 p_1(x))}{((n_1 - 1)p_1(x))^{2/d} p_1(x)} - \frac{\text{tr}(\nabla^2 p_2(x))}{(n_2 p_2(x))^{2/d} p_2(x)}}$$

Optimal Local Metric for Hessian

Bias is controlled by Hessian, thus minimize local measure to minimize bias.
We now consider the Mahalanobis distance:

$$\|x - x_0\|_A^2 = (x - x_0)^\top A(x - x_0), \quad A \succ 0.$$

It is proven that the optimal matrix A^* is given as:

$$A^* = \arg \min_A \text{tr}(A^{-1}B), \quad \text{s.t. } A^\top = A, |A| = 1, A \succ 0.$$

With:

$$B = \frac{1}{((n_1 - 1)p_1)^{2/d}} \frac{\nabla^2 p_1}{p_1} - \frac{1}{(n_2 p_2)^{2/d}} \frac{\nabla^2 p_2}{p_2}.$$

MISED-based Metric Learning

We rescale the matrix B :

$$\tilde{B} = \frac{1}{(n_1 - 1)^{2/d}} \left(\frac{p_2}{p_1} \right)^{2/d+1} \nabla^2 p_1 - \frac{1}{n_2^{2/d}} \nabla^2 p_2.$$

Now we can estimate B by estimating Hessians and density ratio $\frac{p_2}{p_1}$:

- MISED: estimate $\nabla^2 p_1$, $\nabla^2 p_2$;
- Least-Squares Density Ratio Estimation: estimate $\frac{p_2}{p_1}$.

The decomposition of \tilde{B} is given as:

$$\tilde{B} = [U_+ \ U_-] \begin{pmatrix} \Lambda_+ & 0 \\ 0 & \Lambda_- \end{pmatrix} [U_+ \ U_-]^\top.$$

Thus:

$$A^* \propto [U_+ \ U_-] \begin{pmatrix} d_+ \Lambda_+ & 0 \\ 0 & -d_- \Lambda_- \end{pmatrix} [U_+ \ U_-]^\top.$$

Application to Regime Change Detection

Consider the following settings:

- Data distributions for adjacent time periods with sliding windows;
- KL divergence as variation metric;
- AUC as evaluation metric.

MISED method outperforms Gaussian-based approaches in change point detection, demonstrating stable performance and suitability for real-world time series applications.

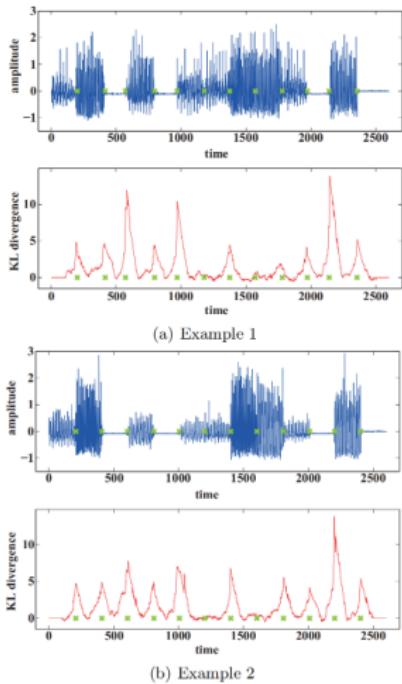


Figure 4: HASC time series data (top) and the KL-divergence estimated by MISED (bottom). Green symbols represent the true change points.

Conclusion

Methodological contributions:

- Proposed MISED, direct density-derivative estimation method, enables stable estimation of density Hessians;
- Reformed bias of nearest-neighbor KL divergence estimation as a metric learning problem;
- Introduced rescaled bias matrix to eliminate unstable density division;
- Obtained optimal local Mahalanobis metric in closed form.

Experimental validations:

- **Synthetic data:** MISED outperforms Gaussian-based methods in non-Gaussian and high-dimensional settings.
- **Change-point detection:** More stable KL estimation leads to improved detection performance on real time-series data.
- **Feature selection:** MISED demonstrates practical effectiveness in high-dimensional, small-sample scenarios.

⇒ MISED provides a robust, non-parametric solution for estimating the KL divergence under complex distributions.

Estimation du processus de Hawkes et propriétés asymptotiques (MLE, normalité asymptotique et bootstrap)

Plan:

- Lecture: Cavaliere et al. (2023) — MLE + théorie asymptotique; bootstrap (FIB/RIB).
- Application: Sanofi — construction d'un processus d'événements, estimation Hawkes sur 2 sous-périodes, et test de rupture des paramètres (théorique vs bootstrap).

Processus ponctuel et intensité conditionnelle

- Un **processus ponctuel** est décrit par le processus de comptage

$$N(t) = \sum_{i \geq 1} \mathbf{1}\{t_i \leq t\}.$$

- **Intensité conditionnelle** (processus régulier / orderly):

$$\lambda(t) := \lim_{\delta \rightarrow 0^+} \frac{1}{\delta} \Pr(N[t, t + \delta] > 0 \mid \mathcal{F}_{t^-}).$$

- Décomposition de Doob–Meyer:

$$N(t) = M(t) + \Lambda(t), \quad \Lambda(t) = \int_0^t \lambda(s) ds,$$

où $M(t)$ est une martingale.

Hawkes : kernel exponentiel

- Hawkes “self-exciting” :

$$\lambda(t) = \mu + \sum_{t_i < t} \gamma(t - t_i).$$

- Kernel exponentiel:

$$\gamma(x) = \alpha e^{-\beta x}, \quad \mu > 0, \alpha \geq 0, \beta > 0.$$

- **Branching ratio** (stabilité):

$$\eta = \int_0^\infty \gamma(x) dx = \frac{\alpha}{\beta} \in (0, 1).$$

- Interprétation: μ = intensité de base; α = amplitude d'excitation; β = vitesse de décroissance.

Vraisemblance et MLE

- Échantillon: événements $0 < t_1 < \dots < t_{n_T} \leq T$.
- Log-vraisemblance (forme générale point process):

$$\ell_T(\theta) = \underbrace{\sum_{i=1}^{n_T} \log \lambda(t_i; \theta)}_{\text{des événements aux instants } t_i} - \underbrace{\int_0^T \lambda(t; \theta) dt}_{\text{pas d'événement sur chaque intervalle } (t_{i-1}, t_i)}$$

- **MLE:**

$$\hat{\theta}_T = \arg \max_{\theta \in \Theta} \ell_T(\theta)$$

$$\theta = (\alpha, \beta, \mu)$$

Hypothèses et résultats asymptotiques (MLE)

- Hypothèses-type (résumé):

- Θ compact; identification; $\lambda(t; \theta)$ prévisible, positive, continue en θ .
- Incréments stationnaires/ergodiques; moments finis.
- Différentiabilité (jusqu'à ordre 3) et information $I(\theta_0) > 0$.

- Consistance (Ogata):

$$\hat{\theta}_T \xrightarrow{P} \theta_0.$$

- Normalité asymptotique:

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} \mathcal{N}(0, I(\theta_0)^{-1}),$$

$$LR_T(\theta_0) := 2(\ell_T(\hat{\theta}_T) - \ell_T(\theta_0)) \xrightarrow{d} \chi_d^2.$$

Pourquoi du bootstrap ?

- En pratique: tailles finies \Rightarrow distributions très asymétriques, approximation gaussienne souvent mauvaise.
- Objectif: approximer la loi finie-échantillon de $\hat{\theta}_T$ et des statistiques de test.
- Cavalier et al. proposent:
 - **FIB (Fixed Intensity Bootstrap)**: intensité bootstrap fixée (non aléatoire conditionnellement aux données).
 - **RIB (Recursive Intensity Bootstrap)**: intensité bootstrap recursive (aléatoire dans le monde bootstrap).

Time-rescaling : idée intuitive

- On observe un processus ponctuel $N(t)$ avec intensité conditionnelle $\lambda(t)$.
- **Intuition :** $\lambda(t)$ joue le rôle d'un *compteur de vitesse* :
 - si $\lambda(t)$ est grande, les événements arrivent plus vite ;
 - si $\lambda(t)$ est petite, ils arrivent plus lentement.
- On définit le **temps transformé** (temps cumulé d'intensité)

$$\Lambda(t) = \int_0^t \lambda(u) du.$$

- Objectif : “**aplatir**” le processus en un Poisson homogène (intensité 1) dans l'échelle de temps $s = \Lambda(t)$.

De l'intensité à la loi d'attente : terme de survie

- Par définition, pour un petit dt ,

$$\Pr(N[t, t + dt) > 0 \mid \mathcal{F}_{t^-}) \approx \lambda(t) dt.$$

- Soit t_{i-1} le dernier événement, et $W_i = t_i - t_{i-1}$ le temps d'attente.
Alors

$$\Pr(W_i > w \mid \mathcal{F}_{t_{i-1}}) = \Pr(N(t_{i-1} + w) - N(t_{i-1}) = 0 \mid \mathcal{F}_{t_{i-1}}).$$

- En multipliant les probabilités de “pas d'événement” sur des petits intervalles et en passant à la limite, on obtient la **fonction de survie** :

$$\boxed{\Pr(W_i > w \mid \mathcal{F}_{t_{i-1}}) = \exp\left(- \int_{t_{i-1}}^{t_{i-1}+w} \lambda(u) du\right)}.$$

Variable transformée : incrément exponentiels et Poisson homogène

- Définissons l'**incrément d'intensité cumulée** entre deux événements

:

$$V_i = \int_{t_{i-1}}^{t_i} \lambda(u) du.$$

- Pour tout $x \geq 0$, il existe un unique $\tau(x)$ tel que $\int_{t_{i-1}}^{\tau(x)} \lambda(u) du = x$ (car Λ est croissante). Alors

$$\Pr(V_i > x \mid \mathcal{F}_{t_{i-1}}) = \Pr(t_i > \tau(x) \mid \mathcal{F}_{t_{i-1}}) = \exp(-x).$$

- Donc

$$V_i \mid \mathcal{F}_{t_{i-1}} \sim \text{Exp}(1).$$

- En posant $S_i = \sum_{j=1}^i V_j$, on a

$$S_i = \int_0^{t_i} \lambda(u) du = \Lambda(t_i),$$

et les S_i sont les instants d'arrivée d'un **Poisson homogène**

FIB : Fixed-Intensity Bootstrap (idée + construction)

- Étape 1: estimer θ_T^* (souvent $\hat{\theta}_T$, ou l'estimateur restreint sous H_0).
- Étape 2: définir l'intensité **fixe** (conditionnellement aux données)

$$\hat{\lambda}(t) = \lambda(t; \theta_T^*), \quad \hat{\Lambda}(t) = \int_0^t \hat{\lambda}(u) du.$$

- Étape 3: générer un processus de Poisson homogène Q^* d'intensité 1 et poser

$$N^*(t) = Q^*(\hat{\Lambda}(t)).$$

- Étape 4: calculer $\hat{\theta}_T^* = \arg \max_{\theta} \ell_T^*(\theta)$ à partir de N^* .

Remarque: conditionnellement aux données, N^* est un Poisson inhomogène d'intensité $\hat{\lambda}(t)$.

Validité asymptotique du FIB (Cavaliere et al.)

- Sous conditions (renforçant légèrement celles de la normalité asymptotique):

$$\sup_{x \in \mathbb{R}} \left| \Pr^* \left(\sqrt{T}(\hat{\theta}_T^* - \theta_T^*) \leq x \right) - \Pr \left(\sqrt{T}(\hat{\theta}_T - \theta_0) \leq x \right) \right| \xrightarrow{P} 0.$$

- Et pour le LR bootstrap:

$$LR_T^*(\theta_T^*) = 2(\ell_T^*(\hat{\theta}_T^*) - \ell_T^*(\theta_T^*)) \xrightarrow{d^*} \chi_d^2.$$

- Intuition: FIB “gèle” la dépendance via $\hat{\lambda}(t)$, ce qui rend l’analyse plus simple et robuste.

RIB : Recursive-Intensity Bootstrap (comparaison)

- Idée: simuler des événements bootstrap *récursivement* comme dans un bootstrap paramétrique classique.
- Construire $\lambda^*(t; \theta)$ en remplaçant l'historique original par l'historique bootstrap $\{t_i^*\}$.
- Génération (schéma):

$$t_i^* = \Lambda^{*-1}(s_i^*), \quad s_i^* = \sum_{j \leq i} v_j^*, \quad v_j^* \sim \text{Exp}(1),$$

puis $\hat{\theta}_T^*$ maximise $\ell_T^*(\theta)$ basé sur $\lambda^*(t; \theta)$.

- **Point clé:** plus coûteux (intégrations / mises à jour séquentielles), mais “paramétrique” naturel.

Application Sanofi : données et définition des événements

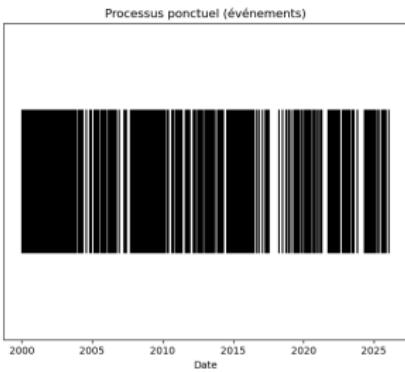
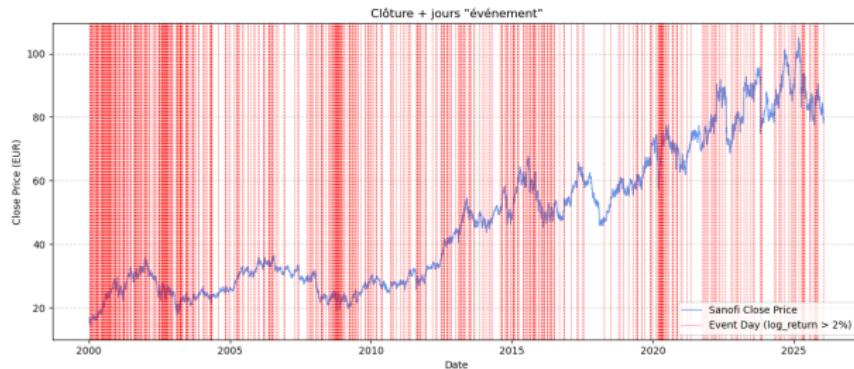
- Série: Sanofi (FR0000120578), **cours de clôture** quotidiens.
- Fenêtre (démonstration): 2000-1-31 → 2026-01-26, 6697 jours de bourse
- Construction d'un **processus d'événements**:

$$\text{événement à } t \iff \ln(r_t) \geq 2\%, \quad r_t = \frac{P_t}{P_{t-1}}.$$

- On travaille en **temps de bourse** (index de jours): $t = 0, 1, \dots, 6697$.
- Découpage: 2 sous-périodes de même longueur: $[0, 3349]$ et $[3349, 6697]$.

Série temporelle et événements

Série temporelle et événements (visualisation)



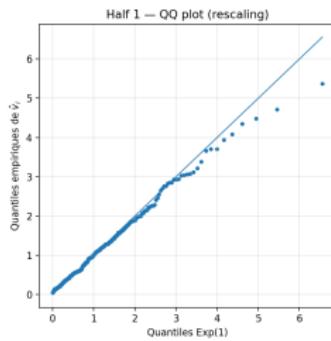
Interprétation: si les événements “se regroupent”, un Hawkes (auto-excitant) peut améliorer un Poisson homogène.

Estimation Hawkes (kernel exponentiel) sur 2 sous-périodes

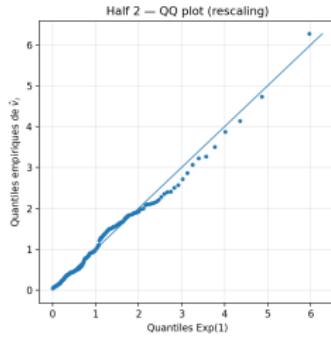
Paramètre	Partie 1	Partie 2	H_0 commun
μ	0.318	0.257	0.240
α	0.014	0.011	0.013
β	0.020	0.019	0.018
$\eta = \alpha/\beta$	0.70	0.56	0.71

- Ici η est modéré (0.56–0.71), ce qui suggère une endogénéité non négligeable (processus loin d'un Poisson).
- En revanche, β est petit (environ 0.02), donc l'excitation a une mémoire longue; le clustering peut être peu visible à l'échelle journalière.

QQ-plot vs Exp(1)



Partie 1



Partie 2

Test de rupture: $H_0 : \theta_1 = \theta_2$

Statistique LR (2 sous-échantillons indépendants)

$$LR = 2 \left[\ell_1(\hat{\theta}_1) + \ell_2(\hat{\theta}_2) - \ell_1(\hat{\theta}_0) - \ell_2(\hat{\theta}_0) \right],$$

où $\hat{\theta}_0 = \arg \max_{\theta} \{ \ell_1(\theta) + \ell_2(\theta) \}$.

Asymptotique (théorie)

$$LR \xrightarrow{d} \chi_d^2, \quad d = 3 (\mu, \alpha, \beta).$$

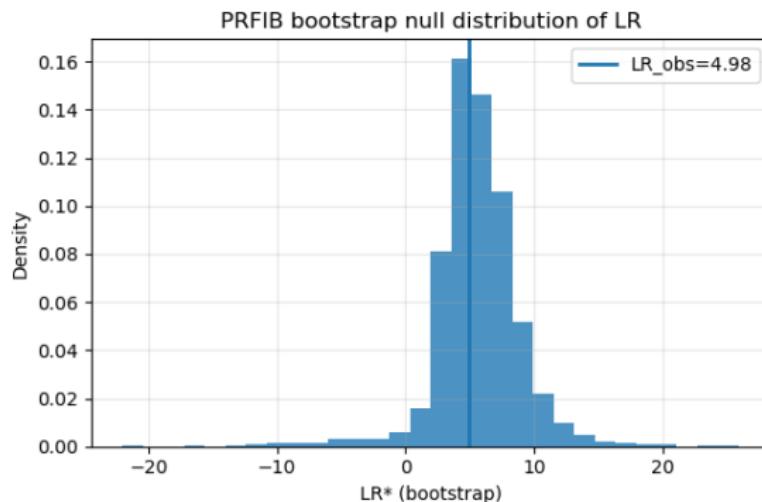
- Observé: $LR = 4.97981 \Rightarrow p_{\chi^2} = 0.173282$.

p-value bootstrap (FIB) pour le test LR

- Sous H_0 , on fixe $\hat{\lambda}(t) = \lambda(t; \hat{\theta}_0)$ dans chaque sous-période et on simule N_1^*, N_2^* via FIB.
- Pour chaque réplication $b = 1, \dots, B$:

$$LR_b^* = 2 \left[\ell_1^*(\hat{\theta}_{1,b}^*) + \ell_2^*(\hat{\theta}_{2,b}^*) - \ell_1^*(\hat{\theta}_0) - \ell_2^*(\hat{\theta}_0) \right].$$

- p-value bootstrap: $\hat{p} = \frac{1}{B} \sum_{b=1}^B \mathbf{1}\{LR_b^* \geq LR_{\text{obs}}\}$.



- $B = 3000$
- $LR_{\text{obs}} = 4.97981$
- $p_{\text{boot}} = 0.5781$