Données haute fréquence

Analyse et modélisation statistique multi-échelle de séries chronologiques financières

Cours de Master - Probabilités et Finances - Sorbonne Université'

Slides de la partie 5

Scale Invariance - Multifractal Models - Rough Volatility - log S-fBm models

Emmanuel Bacry
DR CNRS, CEREMADE, Université Paris-Dauphine, PSL,
CSO, Health Data Hub
emmanuel.bacry@cnrs.fr

Multifractals

A (very) short history

- **Fractionnal Brownian motion** (Mandelbrot, Van Ness, 1968)
- log-normal **Mandelbrot's Random cascades** (1974)
- Gaussian **multiplicative chaos** (Kahane, Peyrière, 1985)
- Multifractal Model for Asset Returns (MMAR) (Calvet, Fischer, Mandelbrot 1997)
- log-infinitely divisible **Multifractal Random Walks/Measures (MRW/MRM)** (Bacry, Muzy, 2001)
- Multifractal products of cylindrical pulses (Barral, Mandelbrot, 2002)
- **Rough volatility models (RVM)** (Jaisson, Gatheral, Rosenbaum, 2014)

$\Longrightarrow$ **log S-fBm models** : a unified framework for RVM models and MRM's (Muzy, Bacry, Wu, 2022)

Scale invariance of the log returns of the price

Some notations

- the log price process $X(t) = \ln(Price(t))$
- the log returns at scale l (supposed to be stationnary)

$$\delta_l X(t) = X(t+l) - X(t)$$

- $M(t) = M([0,t]) =$ integrated volatility on $[0,t]$
- $\delta_l M(t) = M([t,t+l]) =$ integrated volatility on $[t,t+l]$ (whose simplest proxy is $|\delta_l X(t)|$)

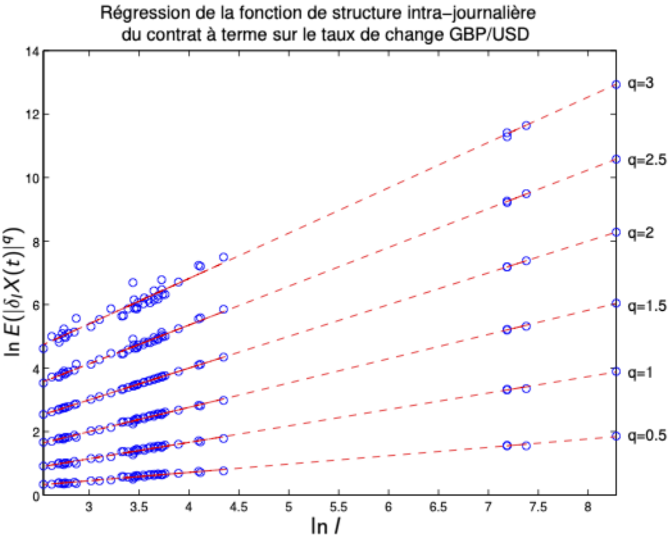
A "definition" of scale invariance

- Scale invariance = Lack of a characteristic scale
  $\Rightarrow$ statistical quantities are power-law as function of time-scale
- The q-order moments statisfy

$$\mathbb{E}(|\delta_l X(t)|^q) \sim l^{\zeta(q)}, \quad \forall q, \text{ when } l \text{ varies}$$

Scale invariance of the log returns of the price

First paper by Mandelbrot et. al. 1997



Thesis A.Kozhemyak, 2007

Self-similar processes : scale invariant processes

$$\mathbb{E}(|\delta_l X(t)|^q) \sim l^{\zeta(q)}$$

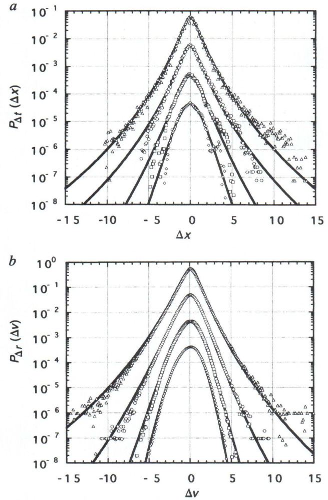
$\zeta(q)$ **linear** $\Longrightarrow$ **Monofractal**

An important example : **Self-similar Processes**

$$\exists H > 0,\ \forall a > 0,\ \ \{X(at)\}_t = \{a^H X(t)\}_t$$

- $\mathbb{E}(|\delta_l X(t)|^q) = C_q l^{qH}$, thus $\zeta(q) = qH$
- H is called the Hurst exponent (regularity exponent)
- e.g. : Brownian motion ($H = 0.5$), fBm ($0 < H < 1$)
  - $H = 0.5 \Rightarrow$ decorrelated (independant) increments
  - $H > 0.5 \Rightarrow$ persistent increments
  - $H < 0.5 \Rightarrow$ contrariant increments
- Shape of distribution of $X(t)$ does not change with t

Log returns are not self-similar



Turbulent cascades in foreign exchange markets, Nature 1998
Ghashghaie, Breymann, Peinke, Talkner, Dodge

Stochastic self-similarity : Multifractal processes

$\zeta(q)$ **non-linear** $\Longrightarrow$ **Multifractal**

The idea :

- Start with self-similarity $\{X(at)\}_t = \{a^H X(t)\}$
- Input stochastic stationnary Hurst exponent $H \to H(t)$
- We set $a^H(t) = W_a$
- $\Rightarrow$ non-linear $\zeta(q)$ and shape of distribution changes with t

Definition : stochastic self-similar process

- T : integral scale
- $\delta_l X(t)$ independant of $\delta_l X(t_1)$ if distance $> T$

$$\{\delta_l X(at)\}_{0 \le t \le T} = \{W_a \delta_l X(t)\}_{0 \le t \le T}$$

where W_a is positive r.v. independant of $\delta_l X(t)$

$$\{X(at)\}_{0\le t\le T} = \{W_a X(t)\}_{0\le t\le T}$$

- Law of $W_a X(t)$ knowing $W_a = w$: $\frac{1}{w} P_X(x/w)$
- Law of $W_a X(t)$ knowing $\ln W_a = u$: $e^{-u} P_X(xe^{-u})$
- If $G_a(u)$ is the law of $\ln W_a$, then

$$P_{X(at)} = \int_{-\infty}^{\infty} G_a(u) e^{-u} P_X(e^{-u} a) du$$

$\Longrightarrow$ **shape of 1-point distribution of $X(t)$ changes with t**

Let fix N, we apply N times $X(at) = W_a X(t)$ with $a^{1/N}$

$$X(at) = \prod_{i=1}^{N} W^{(i)}_{a^{1/N}} \, X(t), \quad \text{with} \quad \{W^{(i)}_{a^{1/N}}\}_i \ \text{ iid}$$

Thus **W_a is log-infinitely divisible**, i.e., $W_a = \prod_{i=1}^{N} W^{(i)}_{a^{1/N}}$

$$\forall q, \quad \mathbb{E}(e^{q \ln W_a}) = \mathbb{E}(W_a^q) = \mathbb{E}(W^q_{a^{1/N}})^N$$

And more generally $\mathbb{E}(W_a^q) = \mathbb{E}(W_{a^{1/r}})^r$

With $r = -\ln a$, we get

$$\mathbb{E}(W_a^q) = a^{-\ln \mathbb{E}(W^q)}, \qquad \text{with} \quad W = W_{e^{-1}}$$

Stochastic self-similarity

Multifractality : "perfect" scaling of the moments + $\zeta(q)$ is non linear

$$\{X(at)\}_{0\leq t\leq T} = \{W_a X(t)\} 0 \leq t \leq T$$

Thus for $t = T$ and $a = l/T$ $(X(0) = 0)$

$$\delta_l X(t) = X(l) = W_{l/T} X(T)$$

Thus

$$\mathbb{E}(|\delta_l X(t)|^q) = \mathbb{E}(|W_{l/T}|^q)\mathbb{E}(|X(T)|^q)$$

Since $\mathbb{E}(|W_a|^q) = a^{-\ln \mathbb{E}(W^q)}$

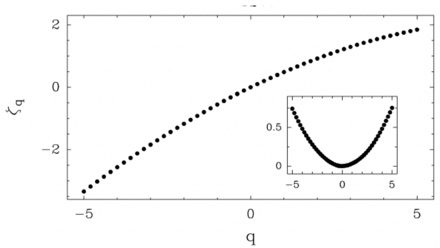
$$\mathbb{E}(|\delta_l X(t)|^q) = C_q (l/T)^{\zeta(q)},$$ **("Perfect scaling")**

with

$\zeta(q) = -\ln \mathbb{E}(W^q)$ **not linear (parabolic if W log-normal)**

Stochastic self-similarity

Multifractality : "perfect" scaling of the moments + $\zeta(q)$ is non linear



S&P500 index 1988-1999

Muzy, Delour Bacry, 1998

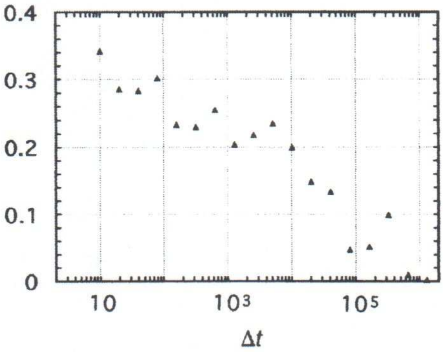
$$\mathbb{E}(W_a^q) = a^{\zeta(q)}$$

- By derivating we get $\mathbb{E}(W_a^q \ln W_a) = \zeta'(q) a^{\zeta(q)} \ln a$
$\longrightarrow q = 0 : \mathbb{E}(\ln W_a) = \zeta'(0) \ln a$
- By derivating again we get
$\mathbb{E}(W_a^q (\ln W_a)^2) = \zeta''(q) a^{\zeta(q)} \ln a + \zeta'(q)^2 a^{\zeta(q)} (\ln a)^2$
$\longrightarrow q = 0 : \mathbb{E}((\ln W_a)^2) = \zeta''(0) \ln a + \zeta'(0)^2 (\ln a)^2$

Thus $\mathrm{Var}(\ln W_a) = -\lambda^2 \ln(a), \quad$ with $\lambda^2 = \zeta''(0)$
Thus

$$\mathrm{Var}(\ln |\delta_l X(t)|) = -\lambda^2 \ln(\frac{l}{T}) + \mathrm{Var}(\ln |X(T)|)$$

where $\lambda^2 = \zeta''(0)$ is called the "intermittency coefficient"



Turbulent cascades in foreign exchange markets, Nature 1998
Ghashghaie, Breymann, Peinke, Talkner, Dodge

$$\delta_l X(0) = W_a \delta_{\frac{l}{a}} X(0) \text{ and } \delta_l X(t) = W_a \delta_{\frac{l}{a}} X\left(\frac{t}{a}\right), \text{ with } \frac{t}{a} + \frac{l}{a} \leq T$$

Thus

$$\mathrm{Cov}\left(\ln|\delta_l X(0)|, \ln|\delta_l X(t)|\right) = \mathrm{Var}\left(\ln W_a\right) + \mathrm{Cov}\left(\ln|\delta_{\frac{l}{a}} X(0)|, \ln|\delta_{\frac{l}{a}} X\left(\frac{t}{a}\right)\right.$$

We take a to separate the two increments as much as possible :

$$t/a + l/a = T \text{ thus } a = (t+l)/T \text{ thus } \frac{l}{a} = \frac{T}{1 + t/l}$$

Thus if $l << t$ then $l/a << 1 \Rightarrow$ the second term is $o(1)$, thus

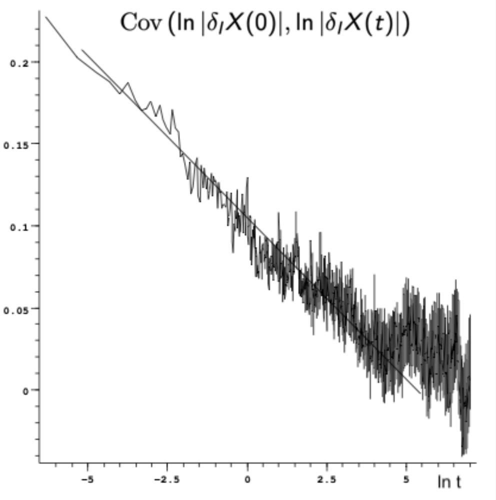
$$\mathrm{Cov}\left(\ln|\delta_l X(0)|, \ln|\delta_l X(t)|\right) = \mathrm{Var}\left(\ln W_{(t+l)/T}\right) + o(1), \quad l << t \leq T$$

Since $\mathrm{Var}\left(\ln W_a\right) = -\lambda^2 \ln(a)$

$$\mathrm{Cov}\left(\ln|\delta_l X(0)|, \ln|\delta_l X(t)|\right) = -\lambda^2 \ln \frac{t}{T} + o(1), \quad l << t \leq T$$

Stochastic self-similarity

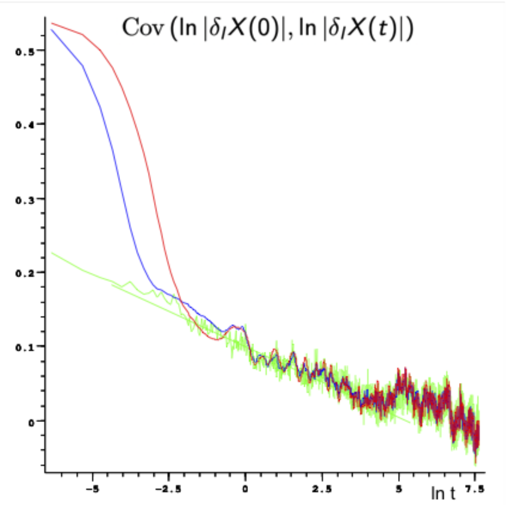
Log-volatility correlation



S&P 500, 5mn log-volatility auto-correlation function, $\lambda^2 \simeq 0.015$
Muzy, Delour, Bacry, 2000

Stochastic self-similarity

Log-volatility correlation



S&P 500, $l =$ 5mn,30mn, 60mn log-vol auto-correlation function

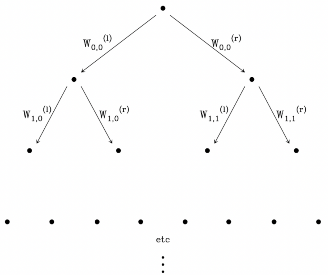
Muzy, Delour, Bacry, 2000

A first multifractal process : Mandelbrot's W-cascade

Mandelbrot, 1974 - Kahane, Peyrière 1985

- A stochastic volatility model (stochastic measure M on [0,T]
- Recursive construction
  - Start with uniform measure on $[0, T]$: M_0
  - Divide $[0, T]$ in two and multiply each part by $W_{0,0}^{(l)}$ and $W_{0,0}^{(r)}$
  - Repeat recursively on each interval (All W's are iid positive)
  - Limit measure satisfies "discrete" stochastic self-similarity property :

$$\{M(t/2)\}_{0 \leq t \leq T} = \{W_{1/2} M(t)\}_{0 \leq t \leq T}, \quad \text{where} \quad M(t) = M([0, t])$$



A stationnary log-normal Multifractal Random Measure

Let

$$M_\ell(t) = \int_0^l e^{\omega_\ell(u)} du$$

be a stochastic measure where $\omega_\ell(u)$ is stationnary log-normal process.

Can we find $\omega_l(u)$ such that we have limiting "perfect" scaling of q-order moments

- $M(t) = \lim_{\ell \to 0} M_\ell(t)$
- $\mathbb{E}(M(t)^q) = C_q t^q \quad \forall 0 \leq t \leq T$

?

The log-normal Multifractal Random Measure (MRM)

Bacry, Delour, Muzy, 2001

Unique solution : ω_l is a gaussian stationnary process with

$$\mathrm{Cov}\left(\omega_\ell(0), \omega_\ell(t)\right) = \begin{cases} -\lambda^2 \ln(t+\ell)/T, & \text{if } t < T. \\ 0, & \text{otherwise.} \end{cases}$$

and $\mathbb{E}(\omega_\ell(t)) = -\mathrm{Var}\left(\omega_\ell(t)\right)/2$

In the limit $\ell \to 0$

- $\mathrm{Cov}\left(\omega_\ell(0), \omega_\ell(t)\right) \to +\infty$
- $\mathbb{E}(\omega_\ell(t)) \to -\infty$
- M_ℓ converges towards a stochastic self-similar measure M such that

$$\mathbb{E}(M(t)^q) = C_q t^q \quad \forall 0 \leq t \leq T$$

This is the so called log-normal MRM measure.

The log-normal MRM construction

How to build corresponding $\omega_l(t)$ process ?

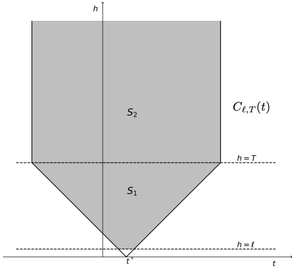
- We consider a non homogeneous 2d gaussian white noise dB on a half plane (t, h) $(h \geq 0)$ with variance

$$\mathbb{E}(dB(t,h)^2)) = \lambda^2 h^{-2} dhdt$$

- We define

$$\omega_\ell(t) = \mu(t) + \int_{C_{\ell,T}(t)} dB(t)$$

with $\mu(t)$ such that $\mathbb{E}(e^{\omega_\ell(t)}) = 1$



Generalization to a log-infinitely divisible MRM

Bacry, Muzy, 2002

- We consider an independently scattered infinitely divisible random measure dB on a half plane (t, h) $(h \geq 0)$ with the measure $\lambda^2 h^{-2} dhdt$
- We define

$$\omega_\ell(t) = \mu(t) + \int_{C_{\ell,T}(t)} dB(t)$$

with $\mu(t)$ such that $\mathbb{E}(e^{\omega_\ell(t)}) = 1$

Then one can prove that M_ℓ converges (when $\ell \to 0$) towards an self-similar log-infinitely divisible stochastic MRM M

Properties of the MRM

The log-normal MRM measure is fully defined by the 2 parameters

- $\lambda^2 = \zeta''(0)$: the intermittency coefficient (generally small)
- T : the integral scale (generally large, "not really meaningful")

Main properties of an MRM

- Stationnary Stochastic Self Similar process
  - T : integral scale
  - $\delta_l X(t)$ independant of $\delta_l X(t_1)$ if distance $> T$ and

$$\{\delta_l X(at)\}_{0 \le t \le T} = \{W_a \delta_l X(t)\}_{0 \le t \le T}$$

  where W_a is log-inf-div. positive r.v. independant of $\delta_l X(t)$
- $\mathbb{E}(M(t)^q) = C_q t^q \quad \forall 0 < t \le T, \quad \forall q$
- $\mathrm{Cov}\left(\ln|\delta_l X(0)|, \ln|\delta_l X(t)|\right) = -\lambda^2 \ln \frac{t}{T} + o(1), \quad l << t \le T$

We define the renormalized magnitude gaussian process as

$$\Omega(t) = \lim_{\ell \to 0} \frac{1}{\lambda} \int_0^t (\omega_\ell(s) - \mathbb{E}(\omega_\ell(s)))ds$$

Then on can prove that for a fixed τ

$$\ln\left(\frac{\delta_\tau M(t)}{\tau}\right) \overset{\lambda}{\simeq} 2\lambda \frac{\delta_\tau \Omega(t)}{\tau},$$

i.e., the process on the right reproduces at the zero and first orders in λ the n-points generalized moments of the process on the left hand-side

$\Longrightarrow$ **allows high performance volatility (or VaR) multi-horizon forecasting**

$$\mathrm{Cov}\left(B_H(s), B_H(s+t)\right) = \sigma^2(s^H + (s+t)^H - 2t^H)$$

We look at it locally around fixed s and we make $H << 1$ and $t << s$

$$\begin{aligned}\mathrm{Cov}\left(B_H(s), B_H(s+t)\right) &\simeq \sigma^2(H\ln s + H\ln(s+t) - 2H\ln t) \\ &\simeq -2\sigma^2 H \ln\frac{t}{s}\end{aligned}$$

This has the same shape as

$$\mathrm{Cov}\left(\omega_\ell(0), \omega_\ell(t)\right) = \begin{cases} -\lambda^2 \ln(t+\ell)/T, & \text{if } t < T. \\ 0, & \text{otherwise.}\end{cases}$$

$$M(t) = \sigma e^{\nu B_H(t)}$$

To have stationnarity, one has to replace $\nu W_H(t)$ by an Orstein-Uhlenbeck version of it $X_H(t)$

$$dX_H(t) = \nu dB_H(t) + \alpha(m - X(t))dt$$

with the reversion time scale $T = 1/\alpha$ is large compared to the observation time scale.
Then

$$\mathrm{Cov}\left(\ln(M(t)), \ln(M(t+l))\right) = \mathrm{Var}\left(\sigma_t\right) - \frac{1}{2}\nu^2 l^{2H} + o(1)$$

The log-S-fBM volatility model

A common framework for MRM and rough volatility, Wu, Muzy, Bacry 2022

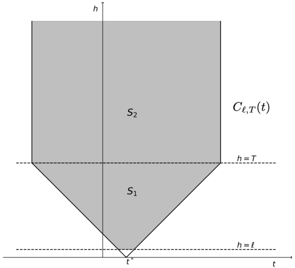
- We consider a non homogeneous 2d gaussian white noise dB on a half plane (t, h) $(h \geq 0)$ with variance

$$\mathbb{E}(dB(t,h)^2)) = \nu^2 H(1-2H)h^{2H-2}dhdt, \quad H > 0$$

- We define the S-fBm process

$$\omega_H(t) = \mu_H(t) + \int_{C_{0,T}(t)} dB(t)$$

with $\mu_H(t)$ such that $\mathbb{E}(e^{\omega_\ell(t)})) = 1$



$$\omega_H(t) = \mu_H(t) + \int_{C_{0,T}(t)} dB(t)$$

One can prove that

$$\mathrm{Cov}\left(\omega_H(0), \omega_H(t)\right) = \begin{cases} \frac{\nu^2}{2}(T^{2H} - t^{2H}) & \text{if } t < T. \\ 0, & \text{otherwise.} \end{cases}$$

and that $\omega_H(t) - \omega_H(0)$ converges towards an fBm when $T \to +\infty$

The log-S-fBm volatility model is then defined by the measure

$$M_H(t) = \int_0^t e^{\omega_H(t)} dt$$

One can prove that

- M_H does correspond to a rough volatility model ($H > 0$) of variance ν^2
- When $H \to 0$ (and $\nu^2 \to +\infty$), M_H converges towards the log-normal MRM measure (noted M_0) with intermittency coefficient $\lambda^2 = \nu^2 H(1 - 2H)$

Estimation of : H and λ^2 (or alternatively ν^2) ?

- Ideally can be made from the scaling property of $\omega_H(t)$, i.e., $\mathbb{E}(|\delta_\tau \omega_H(t)|^q)$ as a function of τ
- But $\omega_H(t)$ is not observable, a proxy is needed (Δ fixed) :

$$\mathbb{E}(|\ln M_{H,\Delta}(t+\tau) - \ln M_{H,\Delta}(t)|^q) \text{ with } M_{H,\Delta}(t) = \int_t^{t+\Delta} e^{\omega_H(t)} dt$$

$\longrightarrow$ **moment scaling based estimation can be highly biased**

GMM approach based either on

- analytical formula for covariance

$$C_M(\Delta, \tau) = \mathbb{E}(M_{H,\Delta}(t) M_{H,\Delta}(t+\Delta))$$

- approximated ($\lambda^2 << 1$) formula for covariance

$$C_{\ln M}(\Delta, \tau) = \mathbb{E}(\ln M_{H,\Delta}(t) \ln M_{H,\Delta}(t+\Delta))$$

$\lambda^2 = 0.02$	$H = 0$	$H = 0.02$	$H = 0.08$	$H = 0.15$
$\widehat{H}$ (GMM_M)	0.010 (0.01)	0.007 (0.015)	0.077 (0.033)	0.146 (0.05)
$\widehat{H}$ ($GMM_{\ln M}$)	0.010 (0.01)	0.018 (0.015)	0.082 (0.02)	0.153 (0.02)
$\widehat{\lambda}^2$ (GMM_M)	0.010 (0.01)	0.010 (0.01)	0.018 (0.006)	0.021 (0.005)
$\widehat{\lambda}^2$ ($GMM_{\ln M}$)	0.019 (0.001)	0.020 (0.001)	0.019 (0.002)	0.020 (0.002)
$\lambda^2 = 0.1$	$H = 0$	$H = 0.02$	$H = 0.08$	$H = 0.15$
$\hat{H}$ (GMM_M)	0.010 (0.02)	0.018 (0.02)	0.11 (0.22)	0.16 (0.26)
$\hat{H}$ ($GMM_{\ln M}$)	0.010 (0.01)	0.02 (0.01)	0.078 (0.02)	0.16 (0.02)
$\widehat{\lambda}^2$ (GMM_M)	0.08 (0.03)	0.08 (0.02)	0.09 (0.045)	0.08 (0.07)
$\widehat{\lambda}^2$ ($GMM_{\ln M}$)	0.095 (0.001)	0.10 (0.005)	0.10 (0.008)	0.10 (0.008)

Table – Mean values and standard deviations estimation errors as obtained from estimations realized on 50 independent samples of length $L = 2^{14}$ of log S-fBM stochastic volatility model.

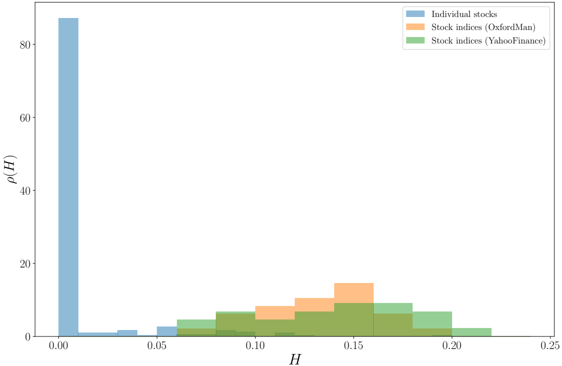


Figure – Probability density distribution of Hurst exponent estimation $\hat{H}$ for the 296 individual stocks (blue bars) of the Yahoo Finance database (OHLC data, 20+ years, green bars) and for the 24 stock indices of the Oxford-Man Institute database (20+ years, orange bars).

The log-S-fBM volatility model

Estimation on financial time-series

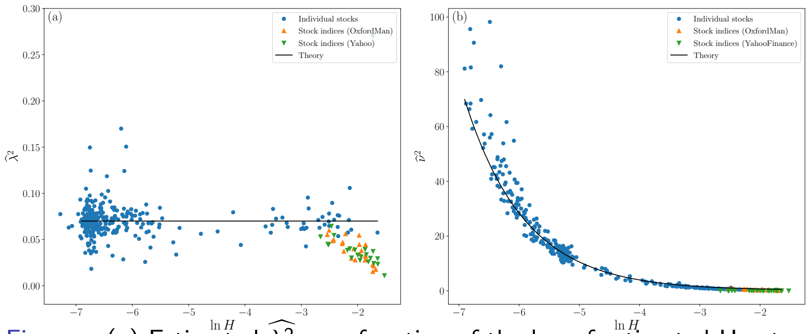


Figure – (a) Estimated $\widehat{\lambda^2}$ as a function of the log of estimated Hurst exponent H. Solid line ($\lambda^2 = 0.07$) corresponds to best fit of individual stock data. (b) Estimated $\widehat{\nu^2}$ as a function of the log of the estimated Hurst exponent H. Solid line log S-fBM expression $\lambda^2 = \nu^2 H(1-2H)$. In (a) and (b) blue dots = stock data from Yahoo Finance, orange up-pointing triangles = index data from Oxford-Man database, down-pointing green triangles = index data from Yahoo Finance database.