

# Estimation de la régularité d'un modèle de diffusion à VS

Marc Hoffmann et Mathieu Rosenbaum

Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR 8050  
et Université de Marne-la-Vallée

# Programme

- Un modèle de diffusion à VS intégrant de la “persistence”
- Le problème (mal posé) de l’estimation de la régularité d’un signal dans une échelle diffusive : analyse en ondelettes et niveaux d’énergie
- Données simulées, données réelles

# Un modèle de diffusion à temps continu intégrant de la “persistence” dans la volatilité

On part de la forme classique d'un modèle de diffusion VS :

$$dY_t = \mu_t dt + \sigma_t dB_t,$$

avec  $Y_t = S_t$  ou  $\log S_t$ , et

$$\sigma_t = \Phi(W_{\tau_t}),$$

$W \perp B$  et  $\tau_t$  changement de temps (régulier et non constant) déterministe.

## Cas particulier

Posons

$$\tau_t := t, \quad f := (\Phi^2)' \circ \Phi^{-1}, \quad g := \frac{1}{2}(\Phi^2)'' \circ \Phi^{-1};$$

de manière équivalente

$$d\sigma_t^2 = g(\sigma_t^2)dt + f(\sigma_t^2)dW_t.$$

On retrouve (partiellement) le contexte standard des diffusions SV, sous des hypothèses (fortes) sur  $\Phi$ .

# Régularité de la fonction de volatilité

On introduit une classe plus large de processus de volatilité :

$$\sigma_t^H := \Phi\left(\int_0^t a(u) dW_u^H\right), \quad 1/2 \leq H \leq 1,$$

$W^H$  mouvement brownien fractionnaire (indépendant de  $B$ ):

$$\mathbb{E} [W_t^H W_s^H] = c(H)(t^{2H} + s^{2H} - |t - s|^{2H}).$$

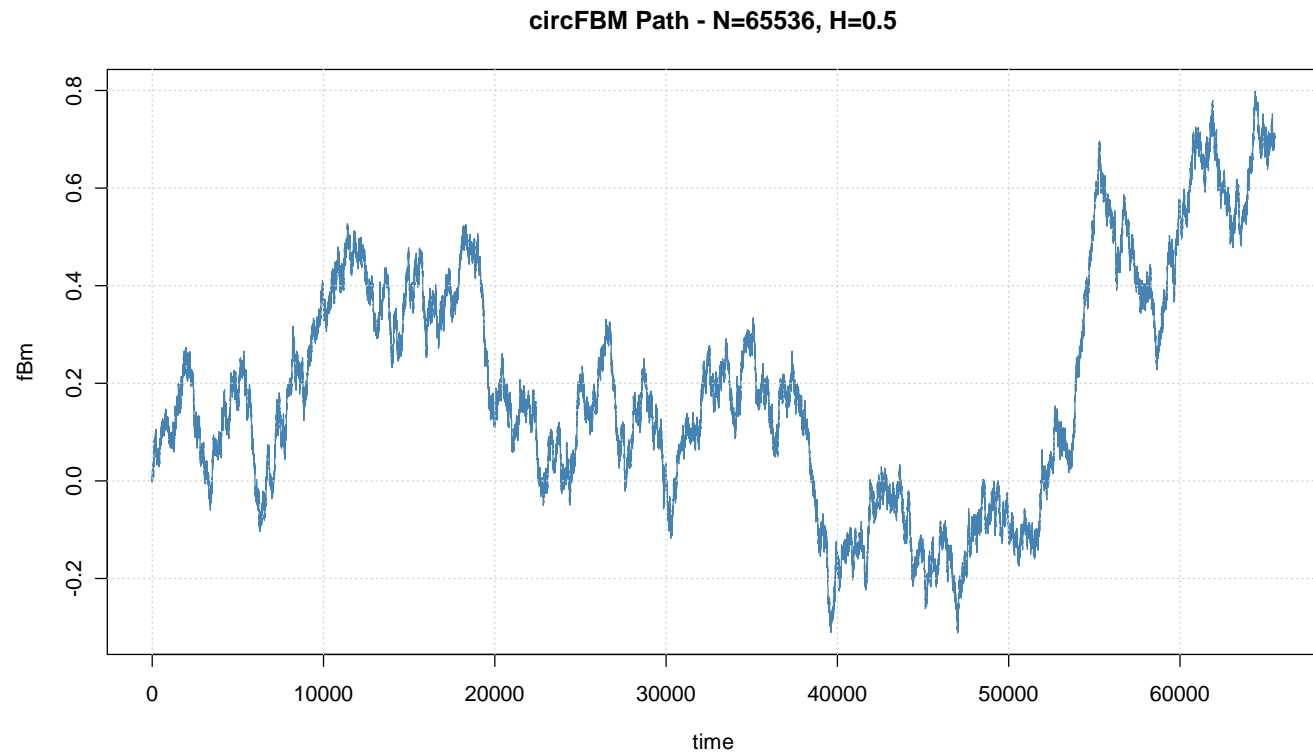


Figure 1: Trajectoire simulée d'un mouvement brownien ( $H = 0.5$ )

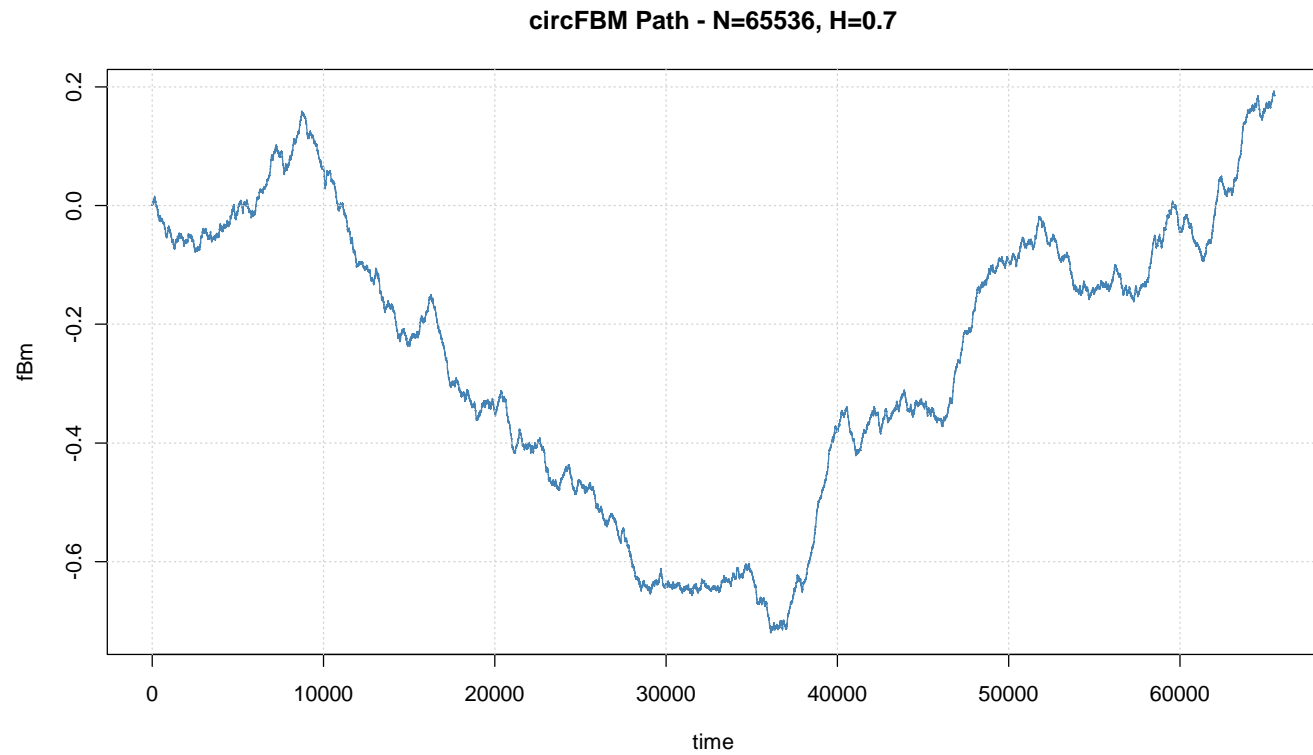


Figure 2: Trajectoire simulée d'un fBM avec  $H = 0.7$

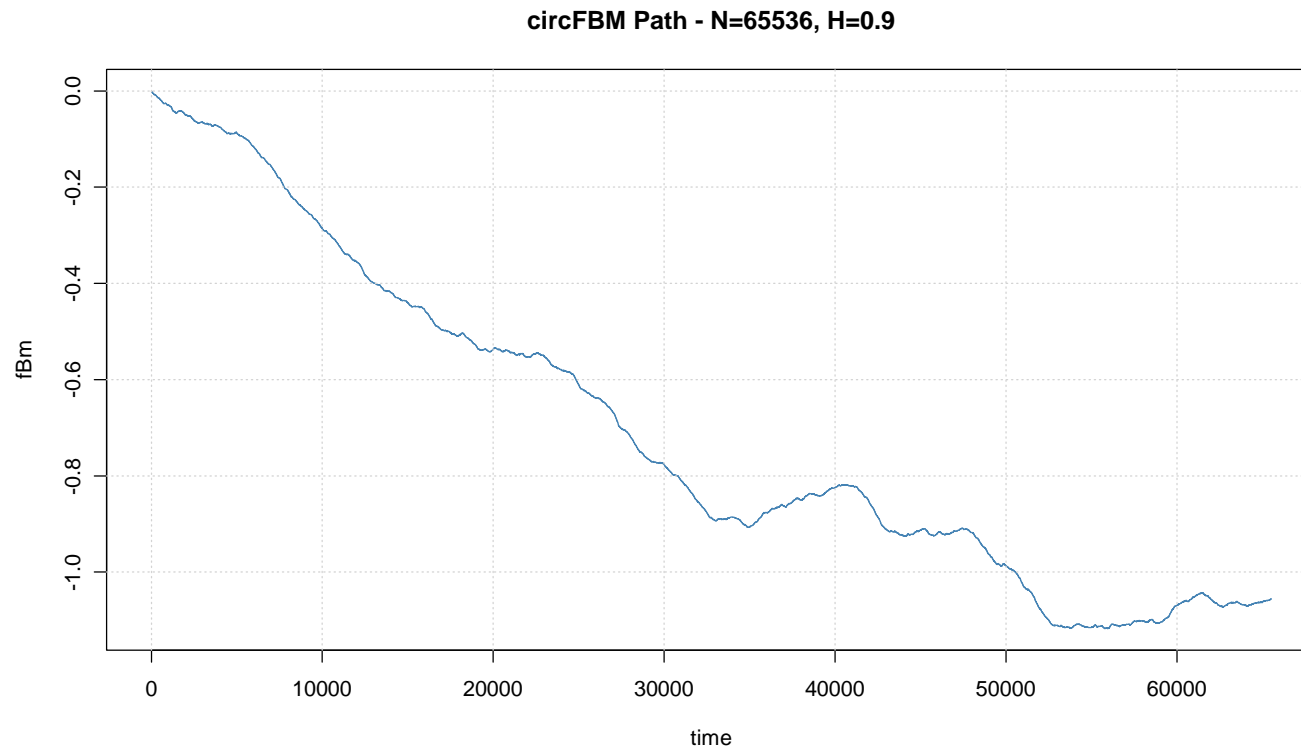


Figure 3: Trajectoire simulée d'un fBM avec  $H = 0.9$

## Régularité des trajectoires de $\sigma_t^H$

$H$  est un index d'oscillation locale, autorisant de la "persistence".

Propriétés de régularité (dans un sens p.s.): Si  $a$  et  $\Phi$  sont suffisamment régulières

$$t \rightsquigarrow \sigma_t^H \in \mathcal{B}_{2,\infty}^H \quad \text{mais pas dans } \mathcal{B}_{2,q}^H \text{ pour } q < \infty,$$

pour tout  $s < H$  :  $t \rightsquigarrow \sigma_t^H \in \mathcal{B}_{\infty,\infty}^s$  mais pas dans  $\mathcal{B}_{\infty,\infty}^H$ .

$\mathcal{B}_{p,q}^s$ : espace de Besov de régularité  $s$  mesurée dans  $L^p$ .

# Programme statistique

$\mathcal{S}$ : expérience statistique engendrée par des données historiques  $(Y_{i\Delta}$  sur  $[0, T]$ ) dans le modèle

$$dY_t = \sigma_t dB_t, \quad \sigma_t = \Phi\left(\int_0^t a(u) dW_u^H\right).$$

- Formellement  $\mathcal{S} = (\mathbb{P}_H^\Delta, H \in [1/2, 1))$ , avec asymptotique  $\Delta$  petit ( $T$  est fixé).
- But: estimer  $H$  ou tester  $H = 1/2$  contre  $H \neq 1/2$  dans  $\mathcal{S}$ .

## Approche analyse du signal (non-paramétrique)

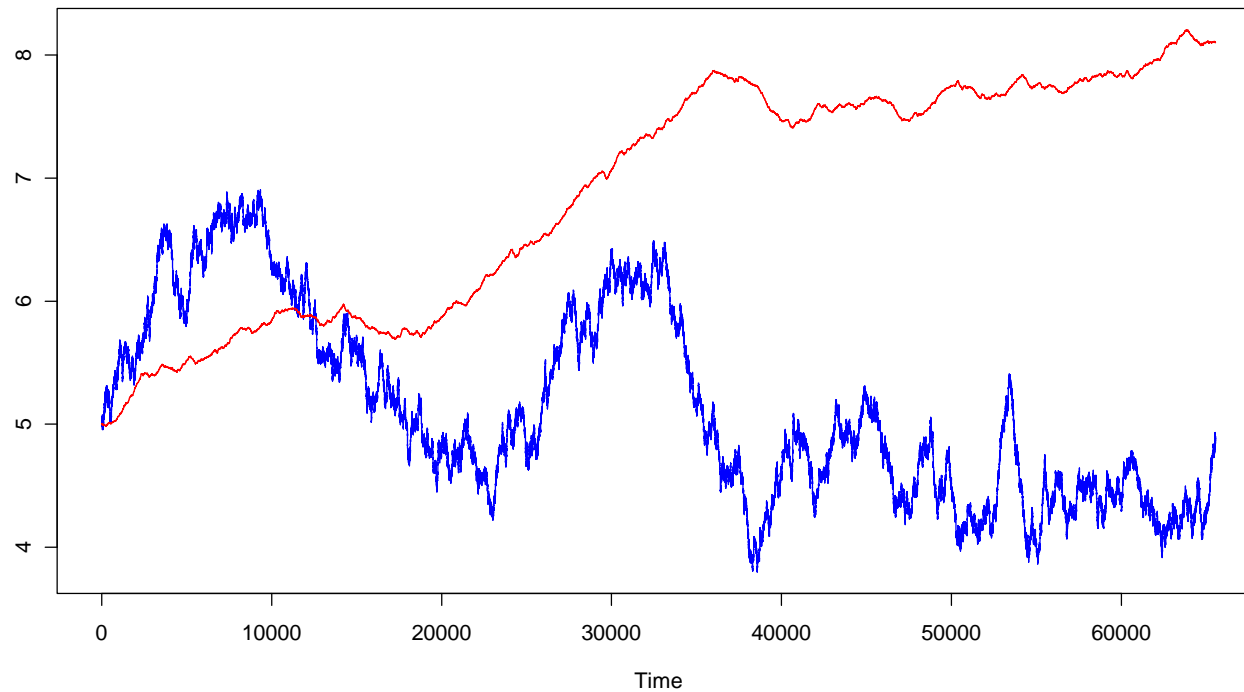
- Régression martingale :

$$Z_i^\Delta := \Delta^{-1} (Y_{i\Delta} - Y_{(i-1)\Delta})^2 \approx \Delta^{-1} \int_{(i-1)\Delta}^{i\Delta} \sigma_s^2 ds + \xi_i^\Delta.$$

Les  $\xi_i^\Delta$  sont (conditionnellement) des accroissements de martingale d'ordre 1:

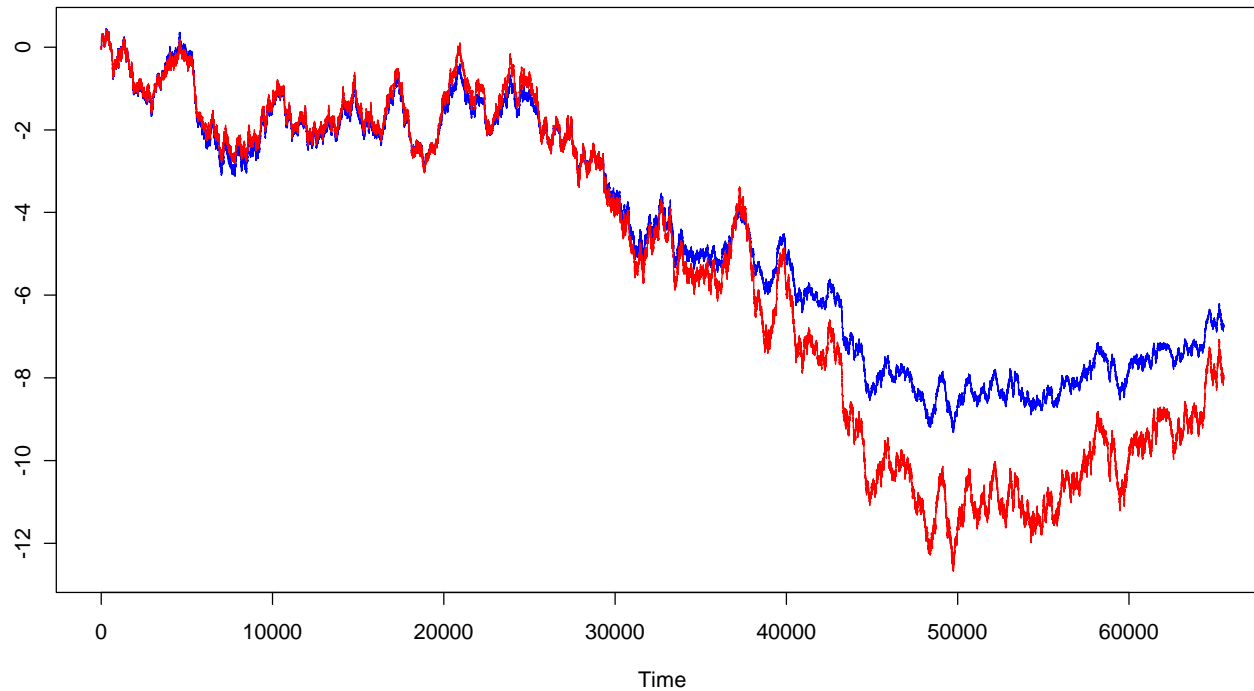
$$Z_i^\Delta \approx \sigma_{i\Delta}^2 + \xi_i^\Delta, \quad i = 1, \dots, c(T)\Delta^{-1}.$$

- Problème (mal posé) : Estimer la régularité du signal  $t \rightsquigarrow \sigma_t^2$  à partir de données bruitées.



11

Figure 4: Simulation de  $\sigma_t$  avec  $\Phi(W_t^H) = 5 + 3W_t^H$ ,  $H = 0.5$  (bleu) et  $H = 0.8$  (rouge),  $\Delta^{-1} = 2^{16}$ .



12

Figure 5: Simulation de  $dY_t = \sigma_t dB_t$ , même trajectoire pour  $B$  (bleu correspond à régularité sur vol. 0.5 et rouge à régularité sur vol. 0.8).

# Le problème mal posé de l'estimation de la régularité d'un signal

- Modèle signal + bruit : on observe

$f(i\Delta) + \xi_i, \quad i = 0, \dots, \Delta^{-1}, \quad \xi_i$  centrées, sous-gaussiennes,

$$f = f_0 \text{ ou } f = f_1.$$

$H(f) := \sup\{\gamma > 0 \text{ t.q. } \forall x, y : |f(x) - f(y)| \lesssim |x - y|^\gamma\}.$

- But : Tester  $H = H(f_0)$  contre  $H = H(f_1)$  avec erreurs de type 1 et 2 prescrites.

Supposons

$$\|f_0 - f_1\|_{L^2} \leq \lambda\sqrt{\Delta}, \lambda > 0.$$

Alors, on ne **peut pas construire de tel test**, même si  $H(f_0)$  et  $H(f_1)$  sont séparés.

Preuve :  $T = 1$  si  $H = H(f_0)$  est rejetée et 0 sinon.

- On impose  $\mathbb{E}_{f_0}^\Delta[T] \leq \alpha$  pour un  $\alpha \in (0, 1)$ .
- On a

$$\mathbb{E}_{f_1}^\Delta[T] \leq \mathbb{E}_{f_0}^\Delta[T] + \|\mathbb{P}_{f_0}^\Delta - \mathbb{P}_{f_1}^\Delta\|_{TV} \leq \alpha + c\Delta^{-1/2}\|f_0 - f_1\|_{L^2}.$$

- Prendre  $f_1 = f_0 + \lambda\Delta^{1/2}\psi$ ,  $\|\psi\|_{L^2} = 1$  et jouer sur  $H(\psi)$ .

# Analyse du mouvement brownien fractionnaire

$$W^H = \sum_j \sum_k d_{jk}(W^H) \psi_{jk}, \quad d_{jk}(W^H) = \langle W^H, \psi_{jk} \rangle$$

$$W^H \in \mathcal{B}_{2,\infty}^H \quad \text{implique} \quad \boxed{\sup_j 2^{2jH} \sum_k d_{jk}^2 < \infty.}$$

De plus

$$\boxed{\inf_j 2^{2jH} \sum_k d_{jk}^2 > 0 \quad \text{en probabilité.}}$$

# Loi d'échelle pour les niveaux d'énergie du fBM

$$Q_j(W^H) := \sum_k d_{jk}(W^H)^2.$$

Loi d'échelle :

$$\boxed{\frac{Q_{j+1}}{Q_j} = 2^{-2H} + \mathcal{O}(2^{-j/2})} \longrightarrow \hat{H} := -\frac{1}{2} \log_2 \frac{Q_{j+1}}{Q_j}.$$

- Comment **passer** de  $W^H$  à  $\sigma^2$ ?
- Comment **estimer**  $d_{jk}(\sigma^2)$  à partir de données bruitées ?
- Comment **choisir un niveau de résolution**  $j$  optimal ?

## Une loi d'échelle pour $\sigma_t^2$

On a

$$\mathbb{E} \left[ \left| 2^{2jH} Q_j(\sigma^2) - c(\psi) \int_0^T a(u)^2 \left[ \frac{d}{dx} \Phi^2 \left( \int_0^u a(v) dW_v^H \right) \right]^2 \right| \right] \lesssim 2^{-j/2}$$

- Ceci permet de déduire les mêmes propriétés (asymptotiques) que pour le cas du fBM.
- La limite non-dégénérée peut être source d'**instabilité** supplémentaire lors des calculs effectifs.

## Construction de $\widehat{H}$

$$\widehat{H}_{\Delta,j} := -\frac{1}{2} \log \frac{\widehat{Q}_{j+1}^{\Delta}}{\widehat{Q}_j^{\Delta}}, \quad j = 1, \dots, J_{\Delta}.$$

$$\widehat{Q}_j^{\Delta} = \sum_k (\widehat{d}_{jk}^{\Delta})^2.$$

- $(\widehat{d}_{jk}^{\Delta})^2$  = carré des coefficients d'ondelettes empiriques – compensateur approprié.
- Choix du niveau de résolution optimal  $j$  dicté par une règle de seuillage des coefficients par bloc.

## Estimation d'erreurs

- Borne supérieure : L'erreur normalisée

$$\Delta^{-1/(4H+2)} (\widehat{H}_{\Delta,j} - H)$$

est tendue, uniformément sur les compacts de  $[1/2, 1)$ .

- Borne inférieure : Pour un  $c > 0$ , on a

$$\liminf_{\Delta \rightarrow 0} \inf_{\hat{H}} \sup_H \mathbb{P} [\Delta^{-1/(4H+2)} |\hat{H} - H| \geq c] > 0.$$

- La “vitesse paramétrique” est obtenue en laissant  $H \rightarrow 0$ .

# Tentative d'illustration numérique sur des données simulées et réelles

- Données simulées : Une trajectoire simulées avec soit  $\Delta^{-1} = 2^{16}$  “Haute Fréquence” – ou  $\Delta^{-1} = 2^{12}$  “Basse Fréquence” .
- Données de marché :
  - **données HF** : contrat futur FGBL (Bund) et EuroDollar (change) échantillonnés avec  $\Delta = 1$  seconde (tick)
    - \* Bund : 01 Mar. 2005 au 30 Nov. 2005.
    - \* EuroDollar : 15 Aug. 2005 au 02 Oct. 2005.

- **données BF** : Bund, EuroDollar (change) et US DollarYen (change) échantillonnés avec  $\Delta = 1$  jour depuis
  - \* Bund : 04 Apr. 1999 au 06 Dec. 2005.
  - \* EuroDollar : 01 Jan. 1999 au 06 Dec. 2005.
  - \* US Dollar-Yen : 02 Jan. 1995 au 06 Dec. 2005.

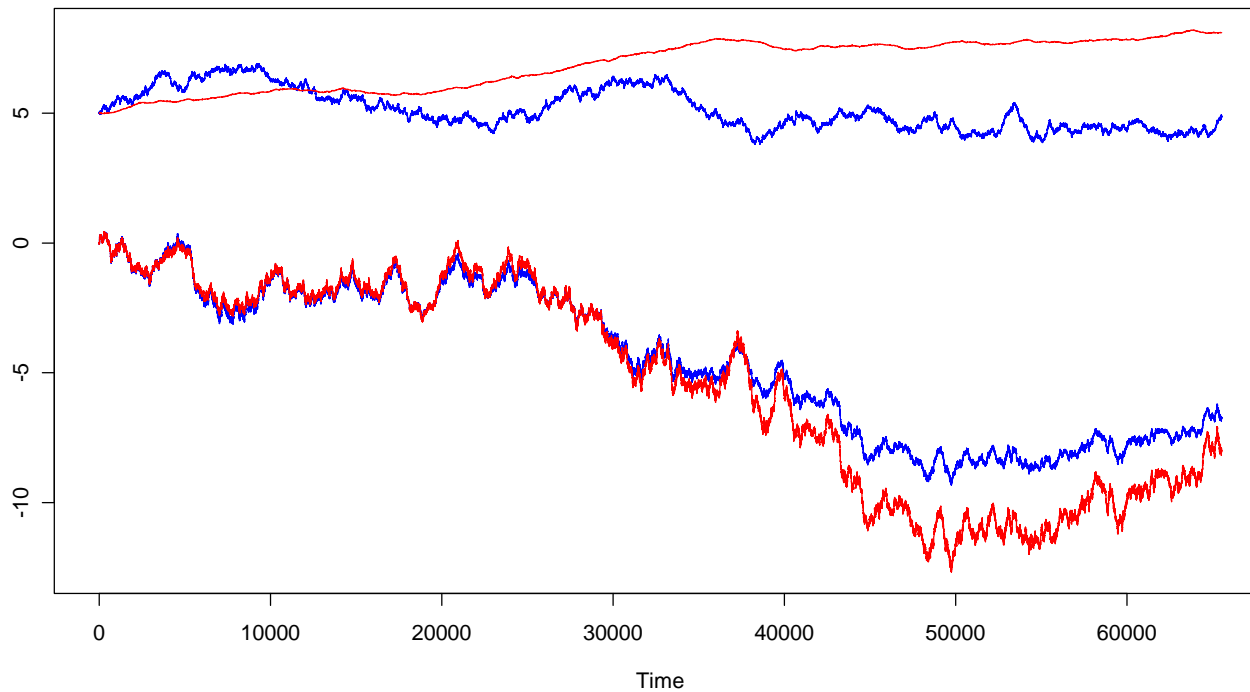


Figure 6: Données simulées :  $\Phi(W_t^H) = 5 + 3W_t^H$ ,  $H = 0.5$  (bleu) et  $H = 0.8$  (rouge),  $\Delta^{-1} = 2^{16}$ .

## Estimation sur données simulées

	$H = 0.5$ (sur volatilité)	$H = 0.8$ (sur volatilité)
régime HF	0.500	0.501
régime BF	0.496	0.500

Expérience de contrôle : recherche d'une échelle diffusive.

	$H = 0.5$	$H = 0.8$
régime HF	0.50, 0.51, 0.58	0.66, 0.76, 0.84, 0.52
régime BF	0.46, 0.57	1.01

Estimation de  $H$  : indice de régularité de la volatilité.

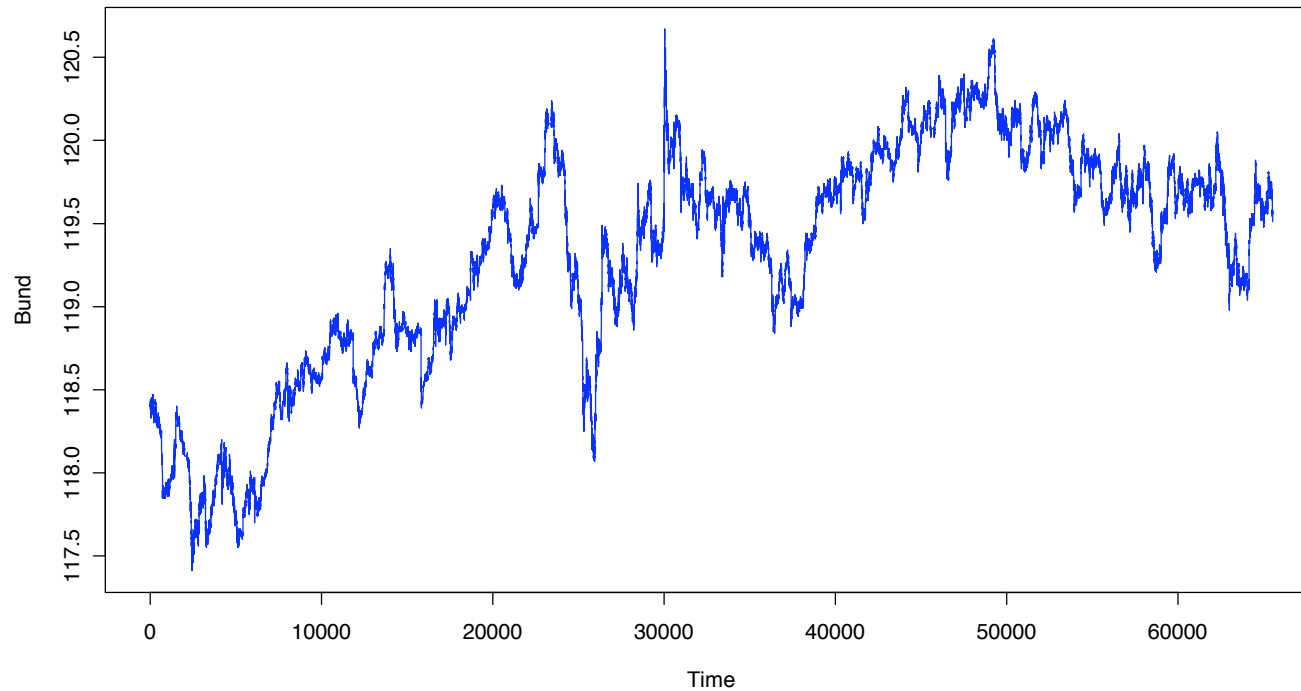


Figure 7: FGBLM5, FGBLU5 and FGBLZ5 (Bund) échantillonné avec  $\Delta = 1$  minute (Bid quote, effet overnight) du 01 Mar. 2005 au 30 Nov. 2005.

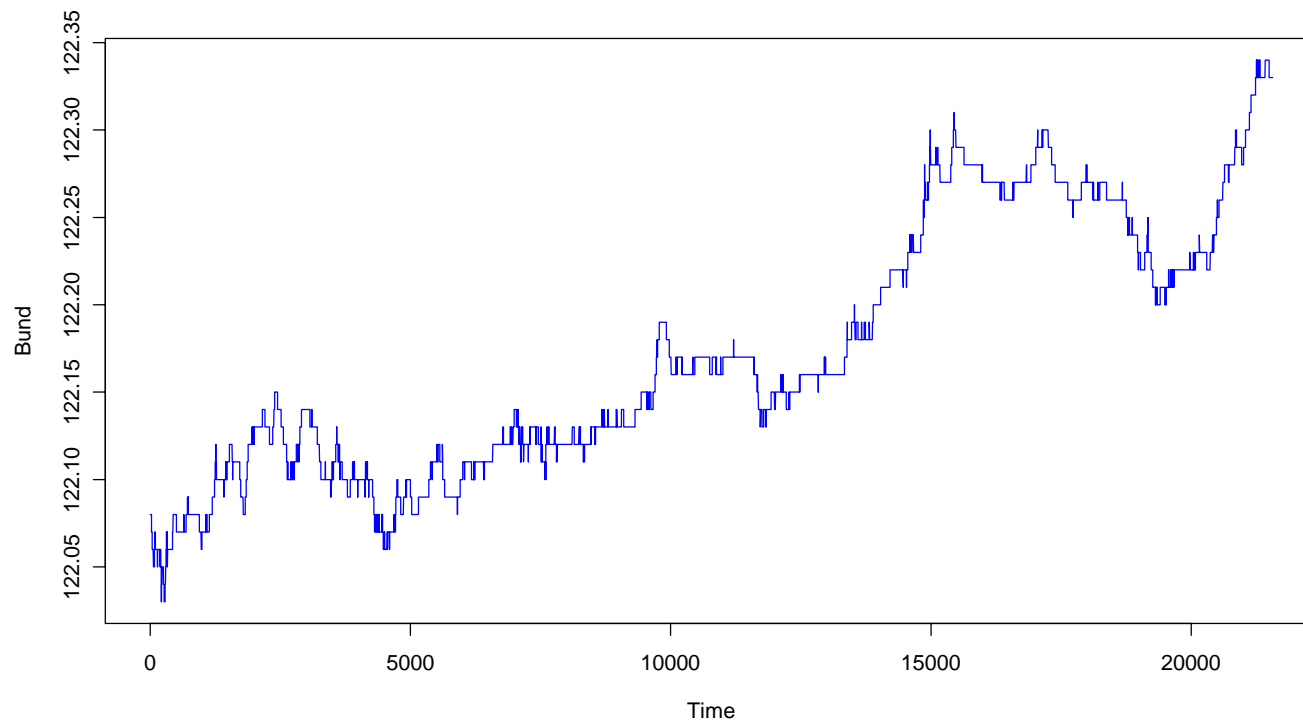


Figure 8: FGBLZ5 (Bund) échantillonné avec  $\Delta = 1$  seconde (Bid quote)  
le 04 Oct. 2005.

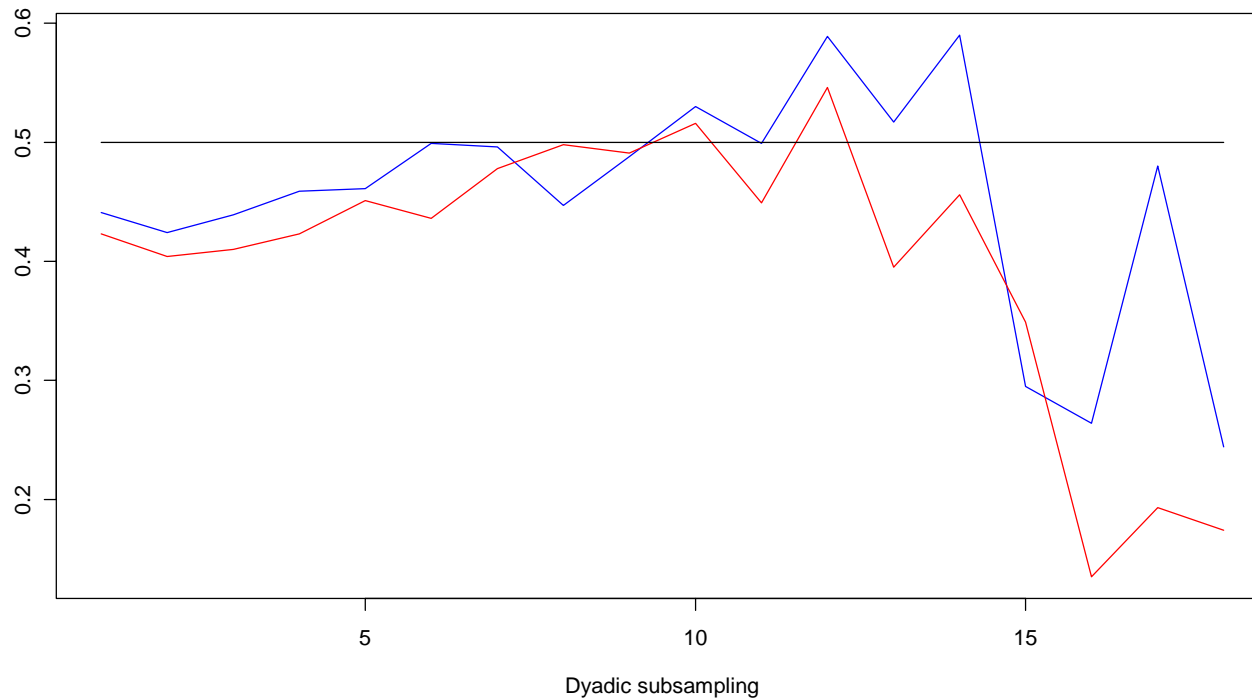


Figure 9: Recherche d'échelles diffusives, FGBL (Bund) du 01 Mar. 2005 au 30 Nov. 2005, (Rouge=1ère période, Bleu=2nd période)

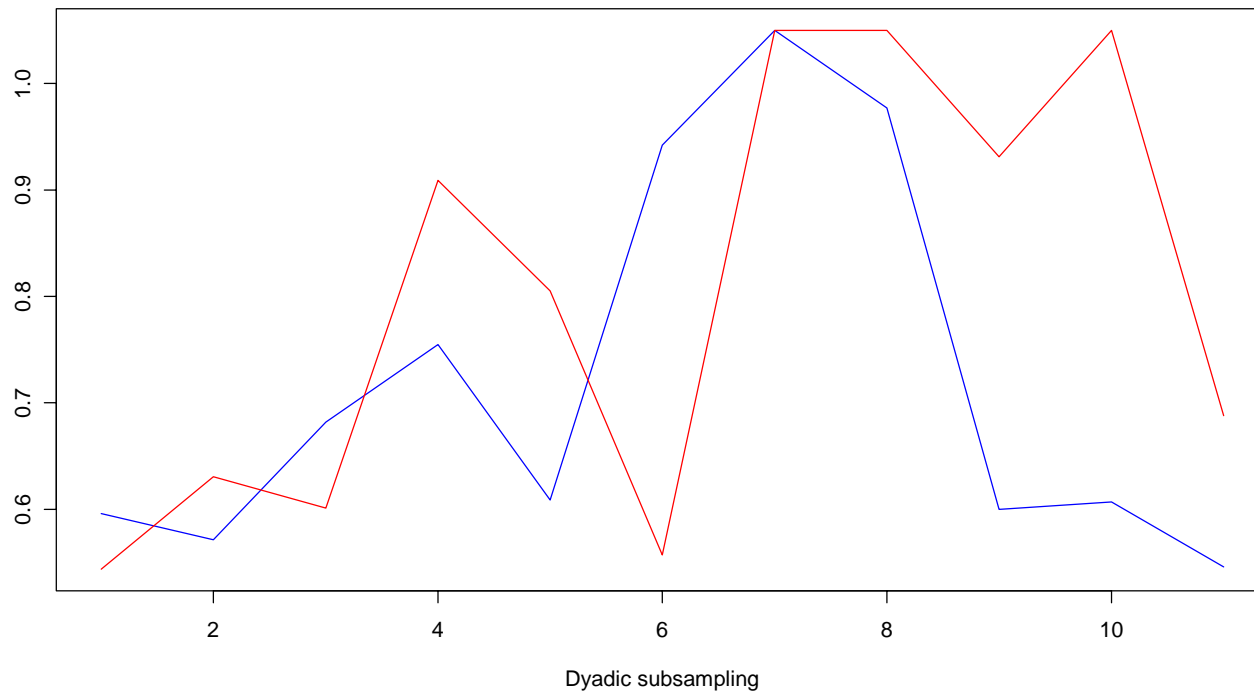


Figure 10: Tentative d'estimation de  $H$  sur le Bund (Rouge=1ère période, Bleue=2nd période)

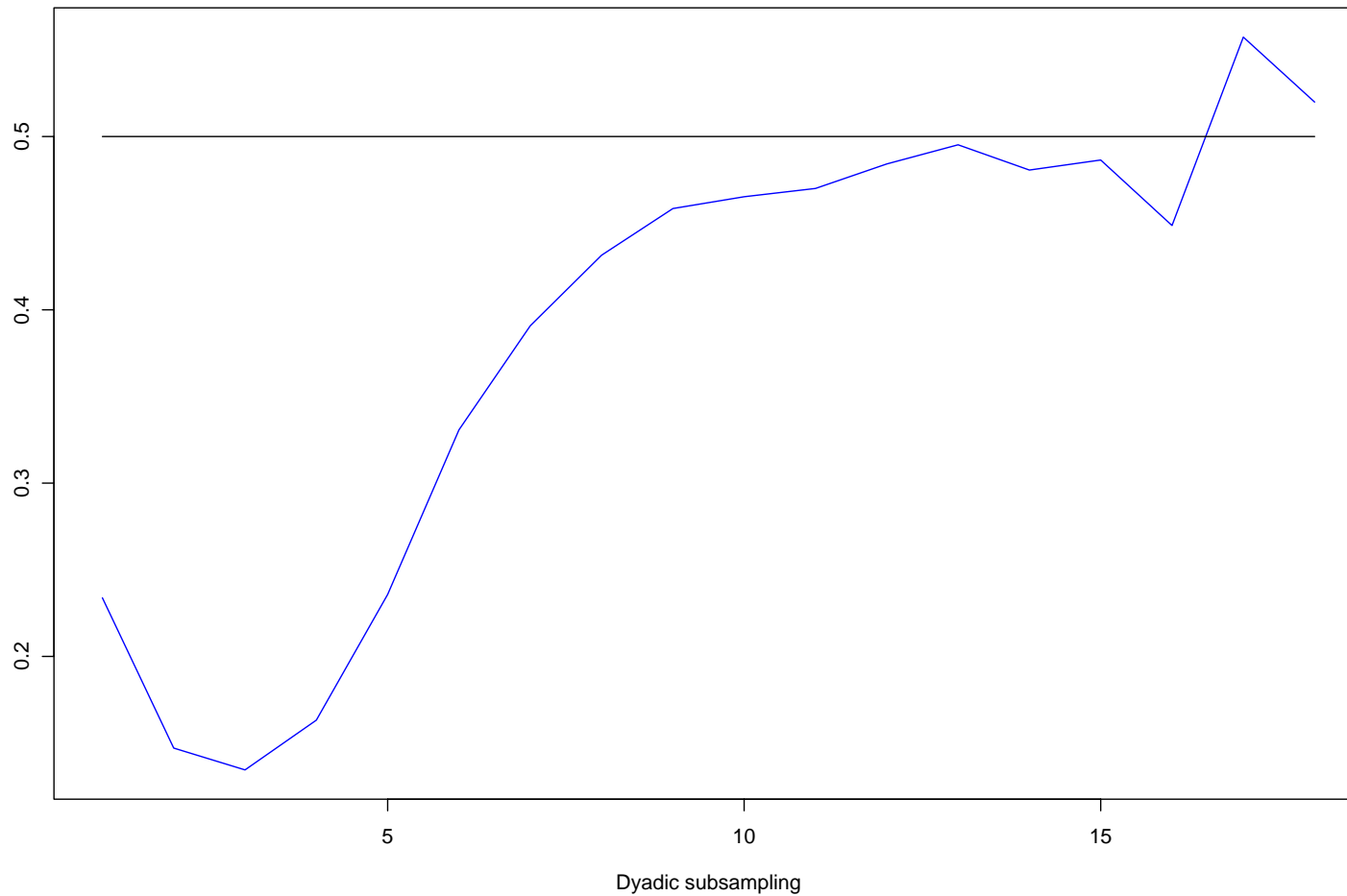


Figure 11: Recherche d'échelles diffusives, EuroDollar (change) du 15 Aug. 2005 au 02 Oct. 2005.

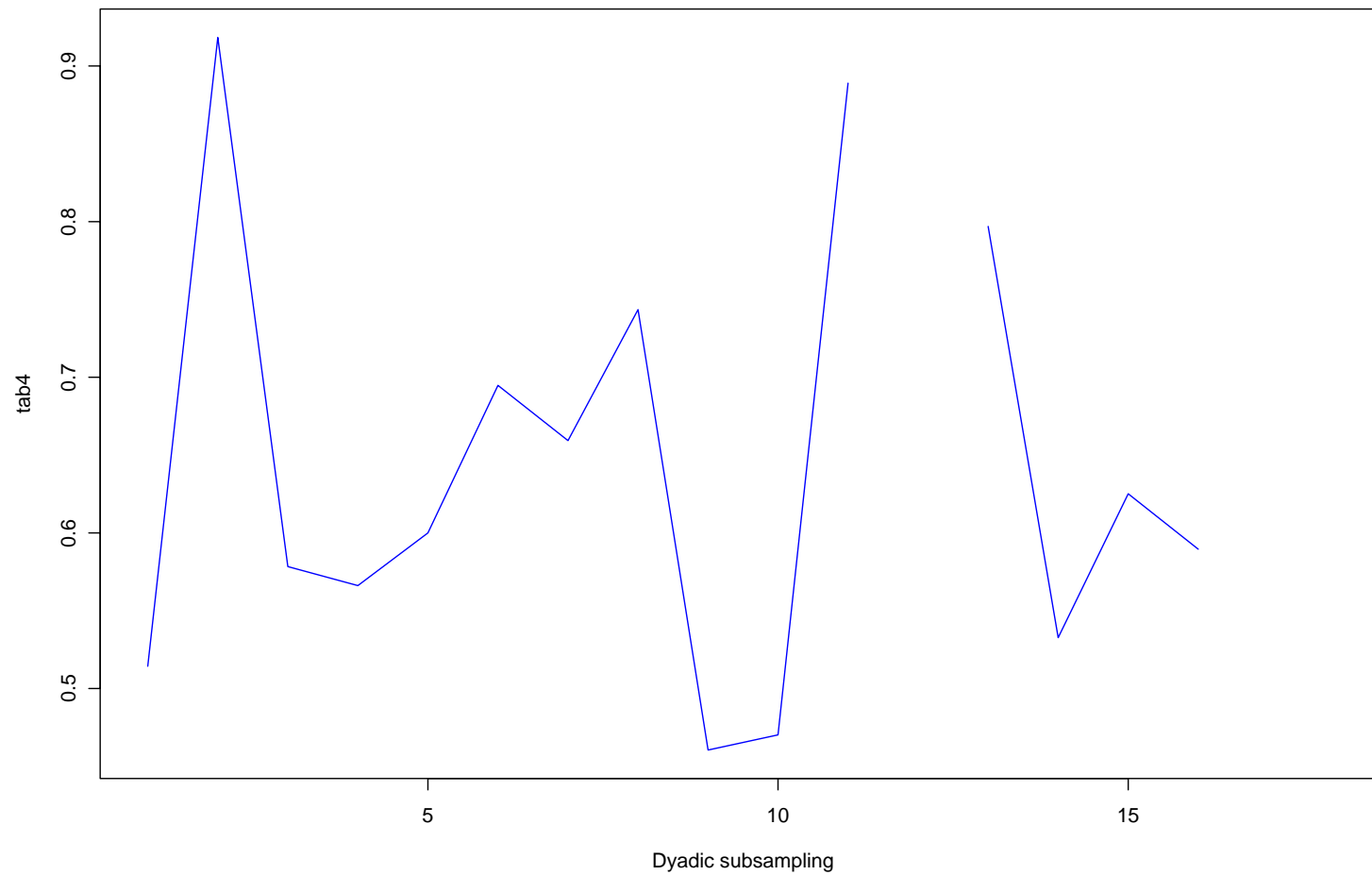


Figure 12: Tentative d'estimation de  $H$  sur EuroDollar (change)

## Données basse fréquence: $\Delta = 1$ jour

données LF : Bund, EuroDollar (change) et US DollarYen (change) échantillonnés avec  $\Delta = 1$  jour

- Bund: 04 Apr. 1999 au 06 Dec. 2005:  $(T_2 - T_1)\Delta^{-1} = 1764$ .
- EuroDollar: 01 Jan. 1999 au 06 Dec. 2005:  $(T_2 - T_1)\Delta^{-1} = 1808$ .
- US Dollar-Yen: 02 Jan. 1995 au 06 Dec. 2005:  $(T_2 - T_1)\Delta^{-1} = 2852$ .

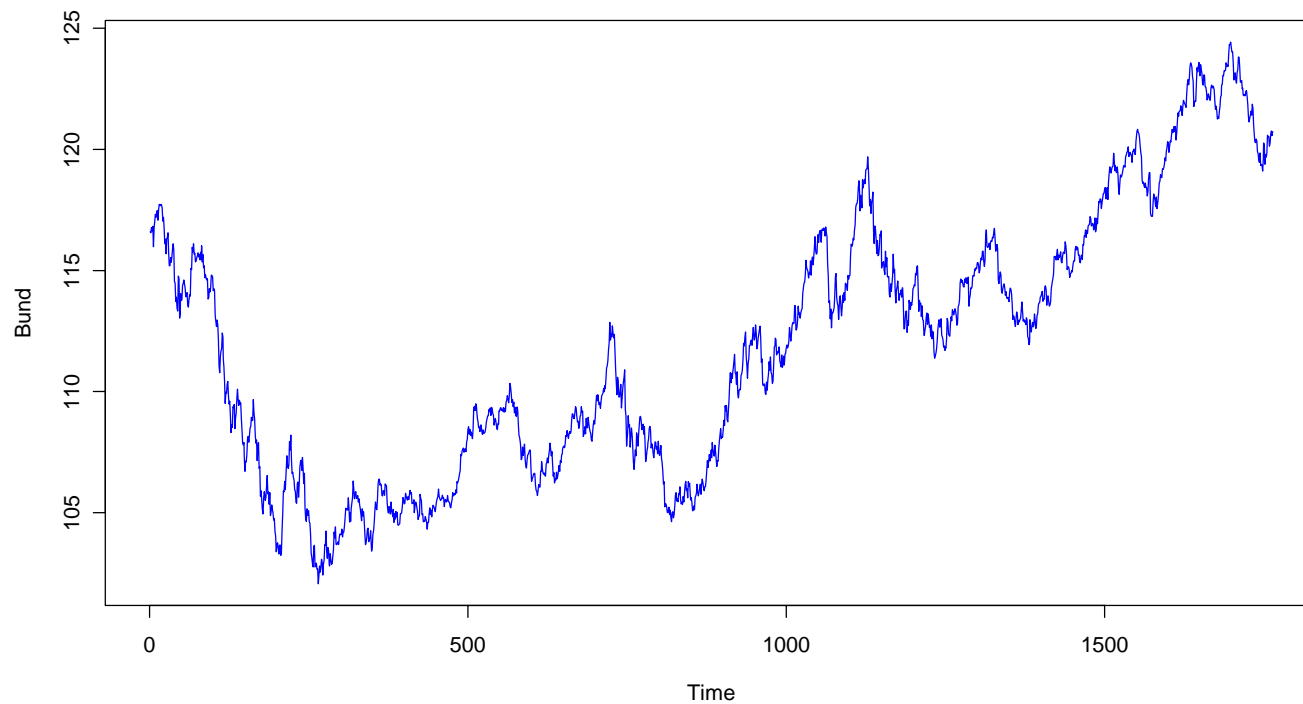


Figure 13: FGBL (Bund) avec  $\Delta = 1$  jour (Last quote), du 04 Apr. 1999 au 06 Dec. 2005.

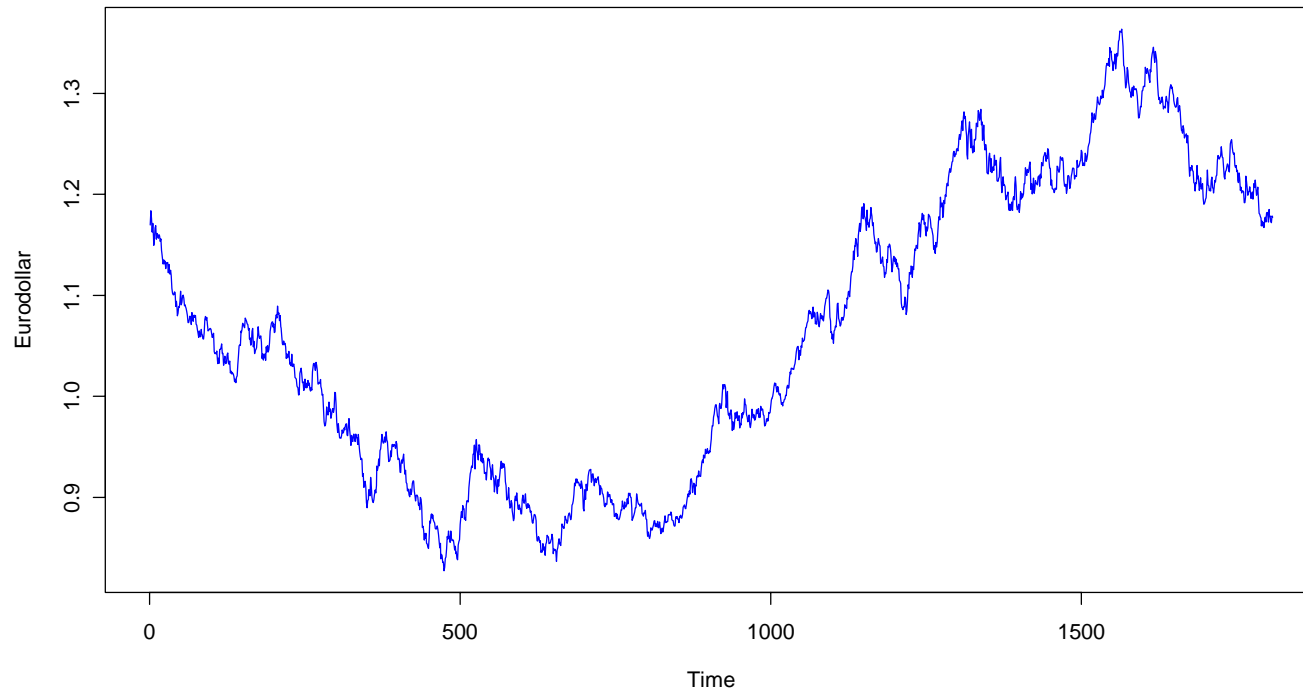


Figure 14: EuroDollar (change) avec  $\Delta = 1$  jours (Last quote), du 01 Jan. 1999 au 06 Dec. 2005.

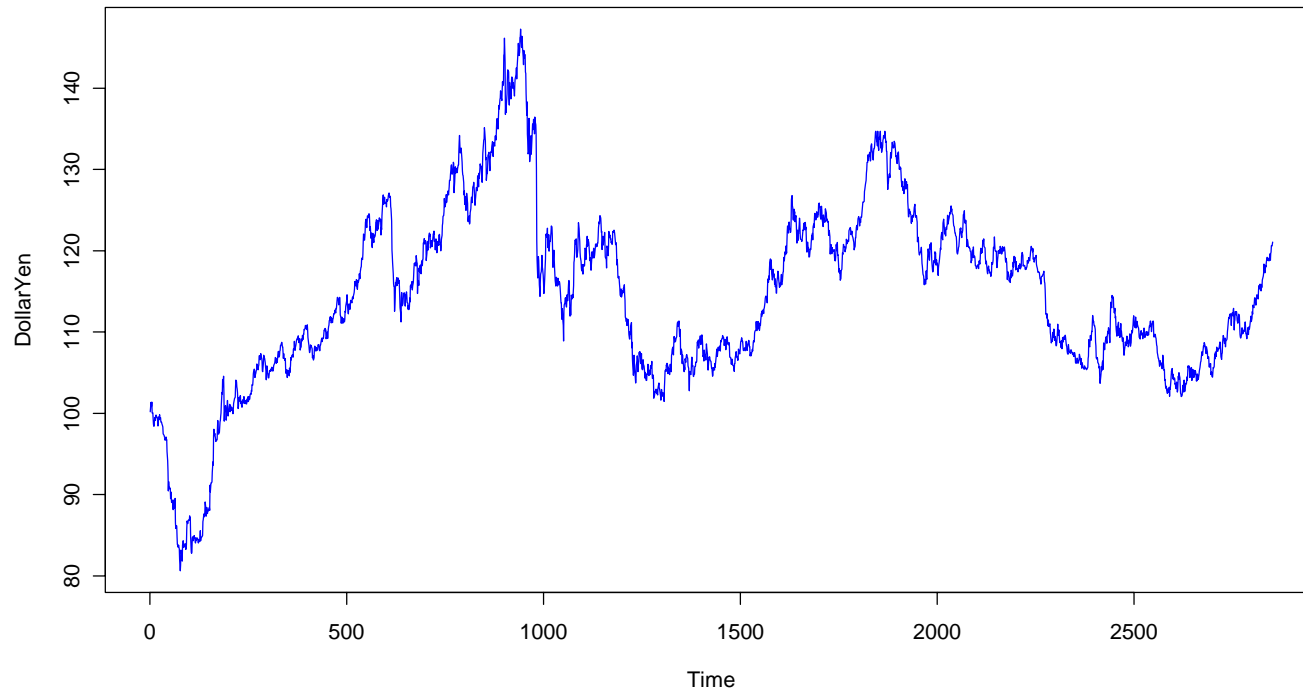


Figure 15: US Dollar-Yen (change) avec  $\Delta = 1$  jour (Last quote), du 02 Jan. 1995 au 06 Dec. 2005.

# Estimation

Echelle diffusive	Bund	EuroDollar	US Dollar-Yen
sous-jacent	0.481	0.453	0.510
log(sous-jacent)	0.480	0.453	0.510

Expérience de contrôle: recherche d'une échelle diffusive.

	Bund	EuroDollar	US Dollar-Yen
sous-jacent	0.49	NA	0.64
log(sous-jacent)	0.66	0.48	NA

Estimation : régularité de la volatilité.

# Conclusion

- Une méthode générique pour estimer la “régularité” à partir de données bruitées dans des modèles de diffusion.
- L'**analogie non-paramétrique** et le cadre **ondelettes-Besov** dicte le choix des compensateurs et les règles d'arrêt des paramètres de résolution.
- **Analogie forte** (équivalence dans certains cas) avec les méthodes de variations quadratiques.

## Conclusion (suite)

- Bien que **théoriquement faisable**, l'analyse statistique met en garde sur la précision d'estimation atteignable.
- Modèle jouet **pas adapté** au bruit de microstructure.
- **Travail numérique préliminaire** sur données simulées encourageant mais insuffisant.
- **Travail sur données réelles** compatible mais pas concluant à ce niveau d'étude.

# Remerciements

Marek Musiela et l'équipe **FIRST/FPG** de BNP-Paribas.