Volatility is rough, isn't it?

Dmitry Sotnikov, Kirill Sokolov, Aleksey Kozhevnikov

22 June 2022

Abstract

The roughness of volatility has been discussed for several years up till now. We try to answer the question whether geometric fractional Brownian motion is an adequate model of this phenomenon. The statistical evidence of market volatility roughness led to the growth of popularity of models based on fractional Brownian motion with $H < \frac{1}{2}$. However, different studies show that roughness of the volatility process may be caused by the microstructure noise. We use a statistical test to demonstrate that this noise is not white noise independent from the volatility path. It was shown that different roughness estimators can significantly overestimate it for different reasons. Furthermore, we propose a statistic based on autocorrelation function and roughness estimation and provide simulation results to demonstrate inconsistency of the geometric fractional Brownian motion as an instantaneous volatility model.

1 Introduction

Rough volatility models, i.e. stochastic processes driven by fractional Brownian motion with a small Hurst parameter H, have recently become an extremely popular tool in volatility modeling. It was shown in [?] that the volatility behavior can be well described with the following process:

$$\sigma_t = \exp\{\nu W_t^H\}, \quad H < \frac{1}{2}, \quad \nu > 0, \tag{1}$$

where W^H denotes fractional Brownian motion with the Hurst parameter H. This model will be referred to as geometric fractional Brownian motion (gfBM). The parameter was supposed to be around 0.1. The hypothesis of rough volatility gained popularity and led to appearance of rough extensions of standard pricing models, see, e.g., [?], [?], [?], [?]. However, at the same time the statistical analysis used in [?] to prove the roughness of volatility has been criticized in [?], [?], [?]. It was mentioned that the effect of roughness can be caused by the microstructure noise arising from realized volatility estimation. An adopted Whittle estimator was introduced in [?] to estimate H given the trajectory of realized volatility process. Application of this approach to market data showed that the volatility should be even rougher with the Hurst parameter value of about 0.04.

We consider already given arguments for and against the roughness of volatility and use simulation to test several approaches of the Hurst parameter estimation. Approximating the microstructure noise with the Gaussian white noise we explain why and to what extent the approach proposed in [?] underestimates the actual value of H. We also propose a statistic, which involves both the estimation of H and autocorrelation of squares of volatility log-increments, to demonstrate that the market data can hardly follow the geometric fractional Brownian motion (1).

2 Realized volatility

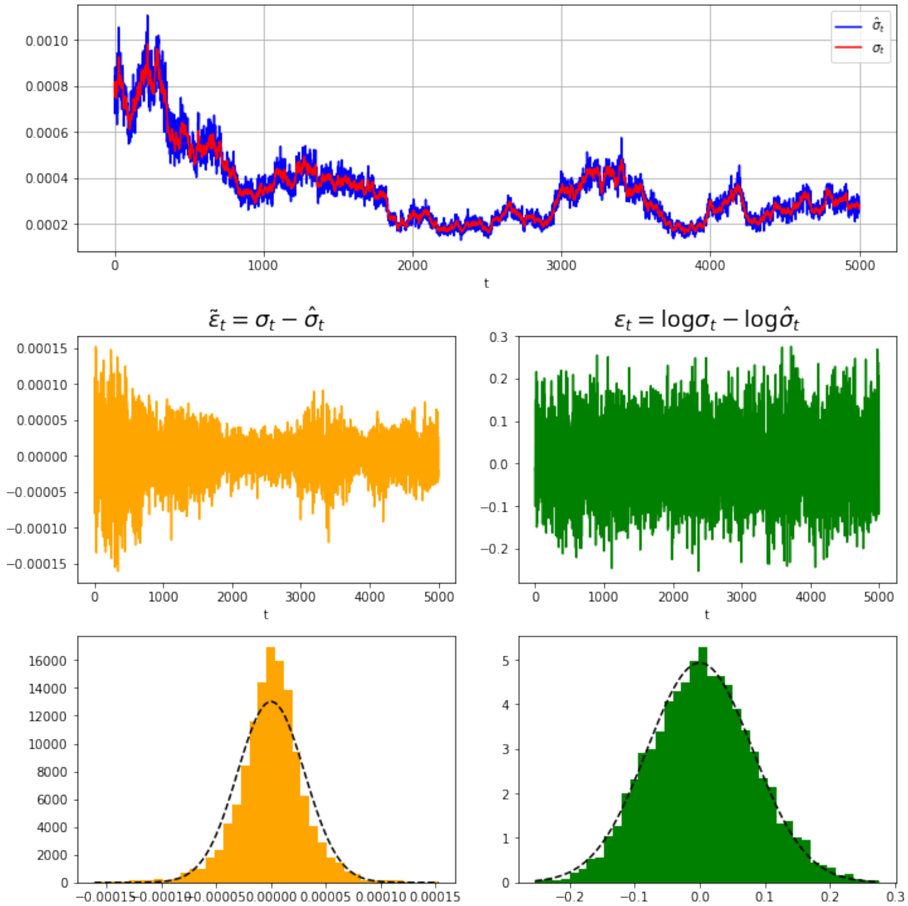

Figure 1: Microstructure noise for geometric Brownian motion (gBM) instantaneous volatility $\sigma_t = \exp\{\nu W_t\}$ with $\nu = 0.02,\ m = 78,\ \sigma_0 = e^{-7.2}$.

Here and below, we consider the following price model:

$$dS_t = \sigma_t S_t \, dB_t, \quad S_0 = s_0, \tag{2}$$

where B_t is a standard Brownian motion and *instantaneous volatility* σ_t is a predictable stochastic process. However, the volatility process can not be observed directly, but should be estimated given the price process. A common approach is to estimate daily[1] *realized volatility* (given m price process observations a day)

$$\hat{\sigma}_t = \sqrt{\sum_{j=1}^{m} \left| \log S_{t+\frac{j}{m}} - \log S_{t+\frac{j-1}{m}} \right|^2} \approx \sigma_t, \quad t = 1, \ldots, N. \tag{3}$$

We also introduce the *microstructure noise*

$$\varepsilon_t = \log \sigma_t - \log \hat{\sigma}_t, \quad t = 1, \ldots N. \tag{4}$$

Note that we subtract log volatility as in this case the noise will be stationary and Gaussian (see fig.1).

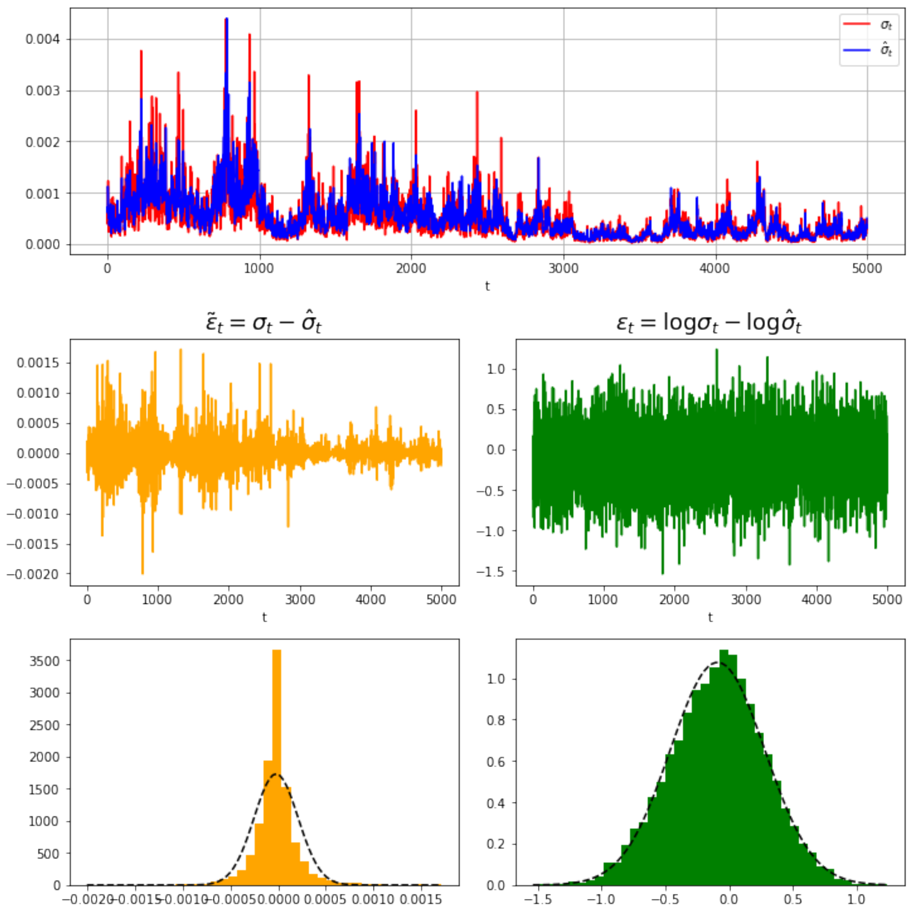

Figure 2: Microstructure noise for geometric fractional Brownian motion (gfBM) instantaneous volatility $\sigma_t = \exp\{\nu W_t^H\}$ with $H = 0.1,\ \nu = 0.5,\ m = 78,\ \sigma_0 = e^{-7.2}$.

This choice is also consistent with the limit theorem proved in [?] claiming that $\frac{\sqrt{m}}{2}\varepsilon_t$ converges to the Gaussian white noise independent from σ_t. How-

[1]Here we assume that time interval $\Delta t = 1$ is equal to one day.

ever, we will show that the noise is far from independent, especially in case of rough volatility model (see fig.2).

2.1 Microstructure noise testing

We implemented the statistical test from [?] to check whether the Gaussian white noise is an appropriate model for microstructure noise. More precisely, we consider the model

$$X_t = \log \hat{\sigma}_t = W_t^H + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \eta) \text{ i.i.d}, \tag{5}$$

and test the hypothesis

$$\mathcal{H}_0 \colon H = H_0, \quad \eta = \eta_0, \tag{6}$$

using the autocovariance $\mathrm{Cov}(\Delta X_t, \Delta X_{t+1})$ as a statistic.

However, the exact value of the parameter η is unknown in realized volatility estimation. Hence, we use sample variance $\eta_0 = \dfrac{1}{N}\sum_{i=t}^{N}(\varepsilon_t - \bar{\varepsilon})^2$ instead. Simulating 1000 trajectories of the fractional Brownian motion $(W_t^H)_{t=1,\dots,256}$ with $H = 0.1$ and $W_0^H = -5$ to obtain volatility process $\sigma_t = \exp\{W_t^H\}$, we tested the hypothesis at significance level $\alpha = 0.05$ for two processes:

1. $X_t = \log \sigma_t + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0, 0.45)$ is a Gaussian white noise. The value of 0.45 was chosen to be close to the variance of the microstructure noise. Note, that in this experiment, although being known exactly, the value of η_0 was statistically estimated to eliminate the differences with the second experiment.

2. $X_t = \log \hat{\sigma}_t$. The value of η_0 was estimated as a sample variance of $\log \sigma_t - \log \hat{\sigma}_t$.

The results shown in table 1 demonstrate that the process in the second experiment can not be represented as a sum of fractional Brownian motion and a Gaussian white noise.

Experiment	Rejected	Total
1. Gaussian white noise	47	1000
2. Microstructure noise	1000	1000

Table 1: Statistical testing results.

Moreover, simulation shows (see fig.3) that in case of rough processes the variance of the realized volatility log increments can be lower than the one for instantaneous volatility, which contradicts the hypothesis that ε_t is independent from σ_t. On the contrary, numerical experiments show strong correlation between the increments $\Delta \log \sigma_t = \log \sigma_t - \log \sigma_{t-1}$ and ε_t:

$$\mathrm{Corr}(\Delta \log \sigma_t, \varepsilon_t) \approx -0.72, \quad \mathrm{Corr}(\Delta \log \sigma_{t+1}, \varepsilon_t) \approx 0.54. \tag{7}$$

We can also note that the effect of microstructure noise differs dramatically from the effect of the white noise, so we can conclude that Gaussian white noise approximation is not suitable for rough volatility processes.

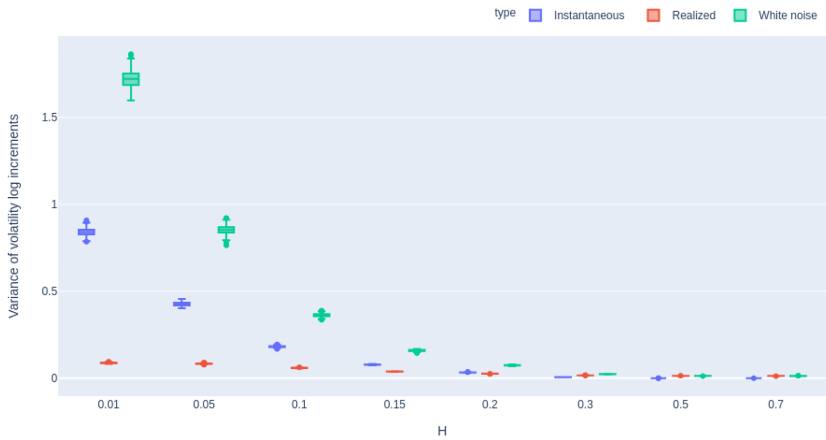

Figure 3: Sample variances of $\Delta \log \sigma_t$ for instantaneous, realized and corrupted by white noise volatility processes.

3 Why volatility may be rough?

As for the market data, we will use a realized volatility process (5593 days from 2000-01-03 to 2022-04-26) of S&P index from the Oxford-Man Institute of Quantitative Finance Realized Library[2]. The 5 minutes realized variance estimator corresponds to the value of $m = 78$ observations a day, that is why it will be used further in numerical experiments.

The trajectories of $\hat{\sigma}_t$ and $\log \hat{\sigma}_t$ are displayed in fig.4. A log increments diagram in fig.5 demonstrates that the increments of $\log \hat{\sigma}_t$ are Gaussian. Hence, it seems reasonable to consider Gaussian processes to model the volatility.

3.1 m-estimator

In this section, we will briefly discuss a statistical argument from [?] in favor of rough volatility models. The proposed approach (which we will call *m-estimator*) is based on the following property of the fractional Brownian motion:

$$\mathbb{E}|W^H_{t+\Delta} - W^H_t|^q = K_q \Delta^{qH}, \tag{8}$$

where K_q denotes q-th absolute moment of standard Gaussian distribution.

Assuming that the volatility follows geometric fractional Brownian motion

$$\sigma_t = \exp\{\nu W^H_t\}, \tag{9}$$

one can calculate empirical estimator of the log increments' q-th moment

$$m(q, \Delta) = \frac{1}{N} \sum_{k=1}^{N} \left| \log(\sigma_{k\Delta}) - \log(\sigma_{(k-1)\Delta}) \right|^q \approx K_q \Delta^{qH} \tag{10}$$

[2]http://realized.oxford-man.ox.ac.uk/data/download

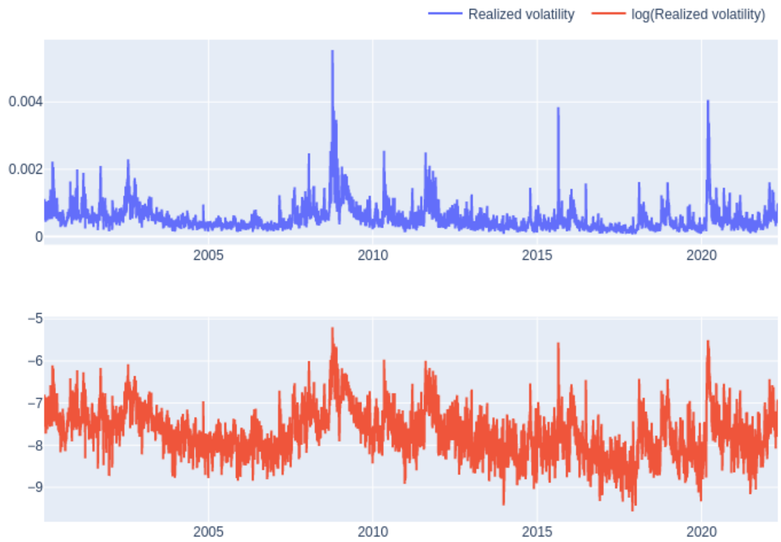

Figure 4: Realized volatility of S&P index.

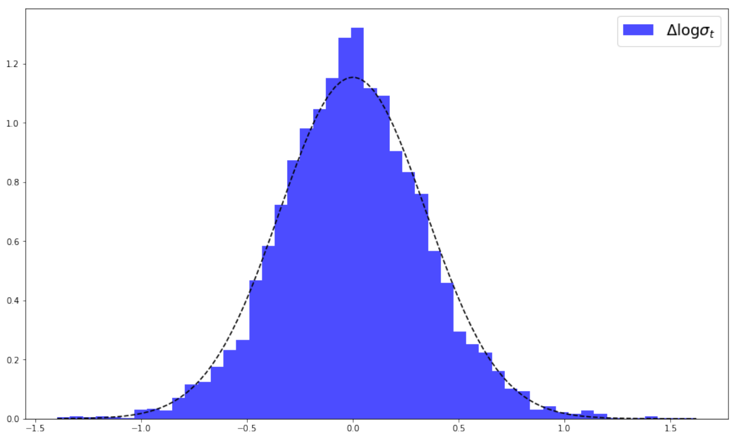

Figure 5: Distribution of the market volatility log increments.

for different values of q and Δ[3] and use linear regression

$$\log m(q, \Delta) = b + \zeta_q \log \Delta \tag{11}$$

to estimate the coefficients $\zeta_q \approx qH$. One more linear regression

$$\zeta_q = \hat{H} q \tag{12}$$

provides an estimate $\hat{H}$ of the Hurst parameter H.

This approach applied to market data gives an estimate $\hat{H} \approx 0.15$, see fig.6, the fact, used by the authors of [?] to claim that the volatility process is a rough gfBM.

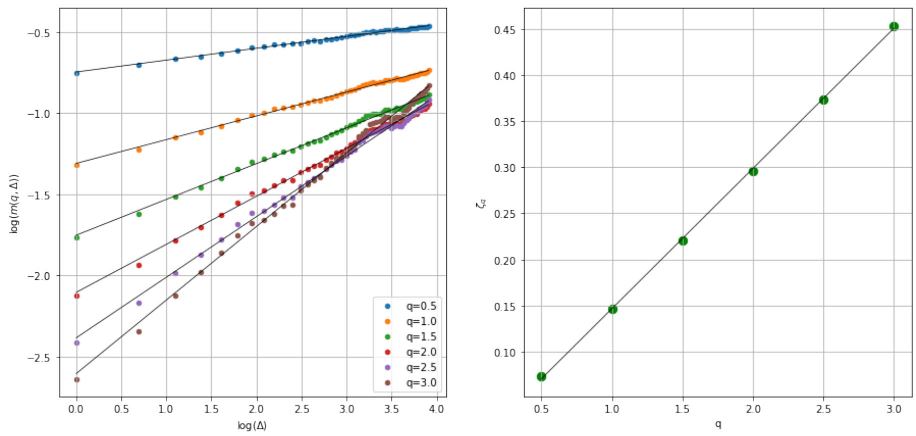

Figure 6: m-estimator for market data. $\hat{H} = 0.15$.

However, m-estimator has two serious drawbacks:

- The method assumes that the analyzed process is gfBM and relies on the property (8). However, for Gaussian processes not satisfying this equation the estimator will not make sense. Further in section 4 we will consider OUOU process which is an Ito process (and, consequently, has $H = 0.5$), but m-estimator gives much smaller value.
- It estimates H of the realized volatility rather than instantaneous volatility, which may be corrupted by microstructure noise, see fig.7. Therefore, we cannot draw any conclusion about the roughness of σ_t based on this estimator. This fact has been already mentioned in [?] and [?]. The authors of [?] proposed a quasi-likelihood estimator for H from realized volatility and claimed that the volatility should be even rougher with $\hat{H} = 0.04$. However, this approach also suffers from the first drawback: it chooses the process with the maximal likelihood, but it chooses only from gfBM processes and exploits the fact that geometric Brownian motion is completely different from the market process. It means that the method does not estimate the roughness of the process but rather fits the parameters of the gfBM.

[3] As in original paper, we take $\Delta = 1, \ldots, 50$.

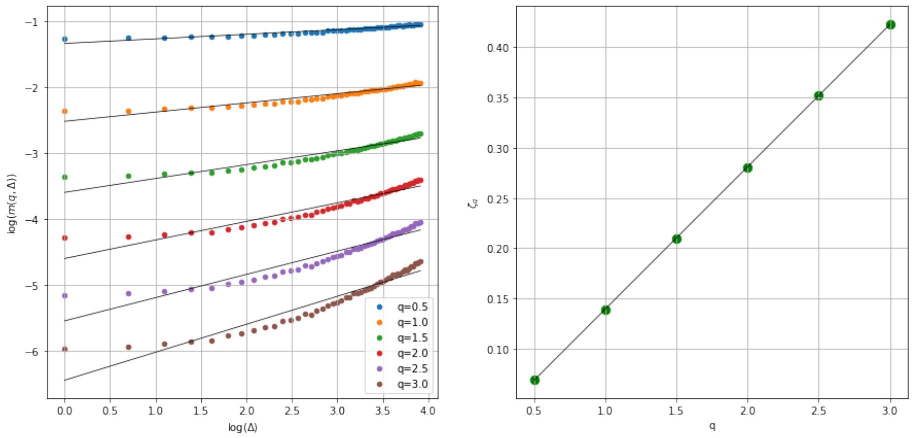

Figure 7: m-estimator for realized volatility $\hat{\sigma}_t$, $m = 78$. Instantaneous volatility is gBM $\sigma_t = \exp\{\nu W_t\}$ with $\nu = 0.02$, $\sigma_0 = e^{-7.2}$. $\hat{H} = 0.14$.

3.2 m-estimator for noisy processes

In this section, we will show how the Gaussian white noise added to log volatility changes the result of m-estimator. As before, we suppose that the volatility follows gfBM

$$\sigma_t = \exp\{\nu W_t^H.\} \tag{13}$$

The observed data is corrupted by a Gaussian white noise in the following way:

$$\tilde{\sigma}_t = \sigma_t e^{\varepsilon_t}, \quad \varepsilon_t \sim \mathcal{N}(0, \eta) \ \text{ i.i.d.} \tag{14}$$

Lemma 1. *Under given assumptions*

$$\mathbb{E}\, m(\Delta, q) = \mathbb{E}|\log(\tilde{\sigma}_{k\Delta}) - \log(\tilde{\sigma}_{(k-1)\Delta})|^q = K_q(\nu^2\Delta^{2H} + 2\eta)^{\frac{q}{2}}. \tag{15}$$

Proof. The left-hand side can be rewritten in a following way:

$$\mathbb{E}\, m(\Delta, q) = \mathbb{E}\left|\nu W_{t+\Delta}^H - \nu W_t^H + \varepsilon_{t+\Delta} - \varepsilon_t\right|^q. \tag{16}$$

Both $(W_{t+\Delta}^H - \nu W_t^H)$ and $(\varepsilon_{t+\Delta} - \varepsilon_t)$ have Gaussian distribution with variance $\nu^2\Delta^{2H}$ and 2η respectively. Since the noise is independent from W^H, the value of $\mathbb{E}\, m(\Delta, q)$ is equal to the q-th absolute moment of normal distribution with variance $(\nu^2\Delta^{2H} + 2\eta)$. ■

Then the regression in m-estimator has the form

$$\log \mathbb{E}\, m(\Delta, q) = \frac{q}{2}\log(\sigma^2\Delta^{2H} + 2\eta) + \log K_q \sim \zeta_q \log \Delta + b. \tag{17}$$

This is equivalent to applying linear regression to the nonlinear function

$$f(x) = \frac{q}{2}\log\left(e^{2Hx} + 2\frac{\eta}{\nu^2}\right) \sim \zeta_q x + b, \tag{18}$$

The ratio $\dfrac{\nu^2}{\eta}$ can be interpreted as signal-to-noise ration, a measure of non-linearity of function f. Applying linear regression we can obtain the result of m-estimator application to the noisy data (if sample moments were equal to theoretical ones). Comparison of theoretical curves in fig.8 with the empirical ones in fig.7 shows that the proposed approach allows us accurately evaluate the error of m-estimator in case of gBM volatility. The dependence of $\hat{H}$ on signal-to-noise ratio $\frac{\nu^2}{\eta}$ for different H is displayed in fig.9. However, we will see that the approximation of microstructure noise with white noise is not appropriate for small values of H.

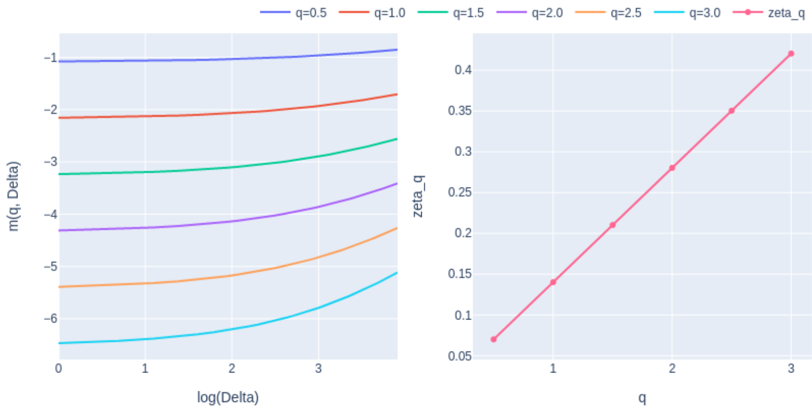

Figure 8: Theoretical value of m-estimator for noisy gBM with $\nu = 0.02$. η is the variance of the corresponding microstructure noise. $\hat{H} = 0.14$.

4 What is roughness and how to estimate it?

We will compare several approaches to estimate the roughness of the trajectory via simulation. However, we haven't given a rigorous definition of roughness yet. Indeed, this concept can be formalized in different ways, for example:

- Roughness of the trajectory x can be defined as $H(x) = \frac{1}{p(x)}$, where

$$p(x) = \inf\{p \geq 1 \colon \text{ x has finite variation of } p\text{-th order}\}. \tag{19}$$

  This approach and an estimator based on it was proposed in [?].

- We can also define roughness as $H(x) = 2 - D(x)$, where $D(x)$ is a fractal dimension (see [?]) of the path x.

- It can be also defined via Hölder regularity:

$$H(x) = \sup\{\alpha \colon \ x \text{ has Hölder regularity } \alpha - \varepsilon \text{ for any } \varepsilon > 0\}. \tag{20}$$

In any case, roughness should be a property of the trajectory, rather than the parameter in definition of fractional Browian motion. However, it can be

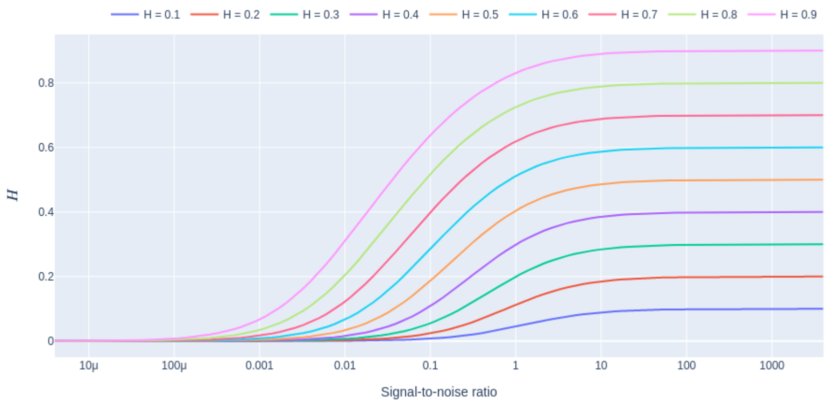

Figure 9: Dependence of m-estimator on signal-to-noise ratio $\frac{\nu^2}{\eta}$.

shown that W^H has roughness H for any proposed definition. We will say that methods relying on the properties of fBM estimate *roughness in a narrow sense*. On the other side, the notion of roughness based on model independent properties of path will be referred to as *roughness in a wide sense*.

We should also emphasize that we try to estimate roughness, being a property of continuous trajectory, based on discrete sample. The analysis of fractional Brownian motion highly involves and benefits from its self-similarity, but in general, processes are not self-similar and roughness estimator may suffer from insufficiently high sampling frequency.

Example 1. *Let us consider a stationary Ornstein–Uhlenbeck (OU) process*

$$dX_t = -\lambda X_t \, dt + \sigma \, dW_t \tag{21}$$

with constant variance $\eta = \frac{\sigma^2}{2\lambda}$. *Let us fix* $\eta = 1$. *For any finite grid* $\{t_1, \ldots, t_n\}$ *there exist large enough* σ *and* λ *s.t.* $\frac{\sigma^2}{2\lambda} = 1$ *and the discretized process* $(X_{t_j})_{j=1,\ldots,n}$ *will be indistinguishable from Gaussian white noise with unit variance.*

This simple example shows that measuring of roughness in a wide sense is not correct without any prior knowledge about the process. Moreover, the fact, that we try to measure the roughness of the discrete process corrupted with a noise of a rather complex nature, requires us to treat and interpret the obtained results very carefully.

Example 2. *Double Ornstein–Uhlenbeck process (OUOU) was proposed in [**?**] as an alternative to rough volatility models. An oscillating mean-reverting behavior of log volatility (fig.4) prompts to consider OU process with slowly varying* *mean also following the Ornstein–Uhlenbeck SDE:*

$$\begin{cases} dY_t = -\lambda_Y (Y_t - \mu)\, dt + \sigma_Y \, dB_t, \\ dX_t = -\lambda_X (X_t - Y_t)\, dt + \sigma_X \, dW_t, \end{cases} \tag{22}$$

where W and B are independent Brownian motions. We choose the parameters

$$\sigma_x = 0.3, \ \sigma_y = 0.05, \ \lambda_X = 0.3, \ \lambda_Y = 0.005, \ \mu = -7.2, \ X_0 = -7.2, \tag{23}$$

for the process to resemble market log-volatility (see fig.10). m-estimator applied to the trajectory of pure OUOU (without microstructure noise) gives $\hat{H} = 0.13$ (see fig.11), although it is clear that the process is not rough in a wide sense.

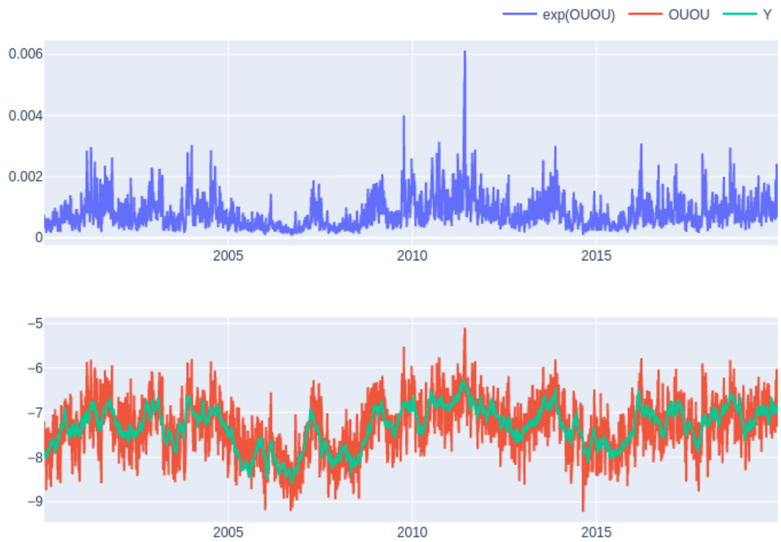

Figure 10: Volatility generated with OUOU process with parameters (23).

We will use OUOU process with fixed parameters (23) to demonstrate that all the considered approaches underestimate the actual roughness $H = \dfrac{1}{2}$ of the process (as it is a bivariate diffusion process).

4.1 Roughness estimators

We will compare H estimators from the following list:

1. **m-estimator** was already discussed in details in previous sections.
2. **w-estimator**. We will call *w-estimator* an approach proposed in [**?**] based on normalized p-th variation $w(x, p)$ and the path variation index $p(x)$,

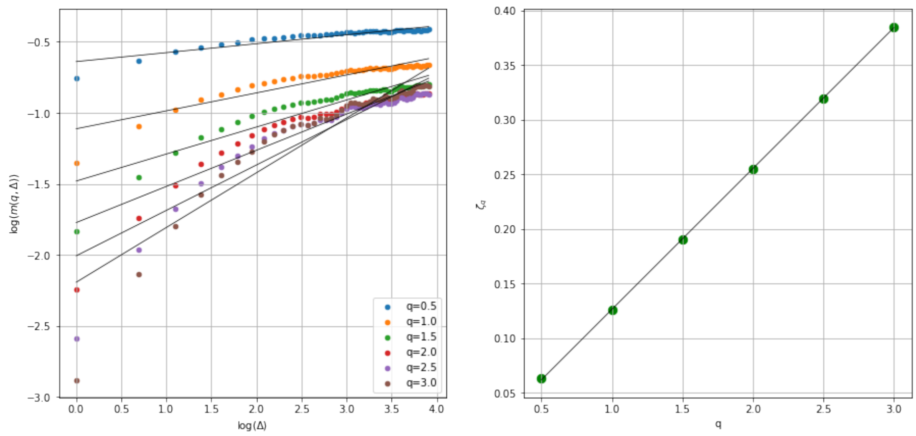

Figure 11: m-estimator for OUOU process with the parameters (23). $\hat{H} = 0.13$.

see the article for details. It was shown that if x has finite non-zero p-th variation for some $p > 1$ and the p-th variation is absolutely continuous then there exists the normalized p-variation and for every $t \in (0, T]$

$$w(x, q)(t) = \begin{cases} \infty, & q > p, \\ t, & q = p, \\ 0, & q < p. \end{cases} \tag{24}$$

It was also proved that stochastic integrals of adapted square integrable processes and fractional Brownian motion have linear normalized variation (of order 2 and $\frac{1}{H}$ respectively) along the dyadic partition sequence almost surely. In order to find an estimator one should calculate sample p-th normalized variation statistic $\hat{w}(x, p)(t)$ and solve the equation

$$\hat{w}(x, \hat{p})(T) = T. \tag{25}$$

Its root provides an estimator for the trajectory roughness $\hat{H} = \frac{1}{\hat{p}}$.

3. **R/S analysis** is described in details in [?]. Let us denote

$$h_n = \log \frac{\hat{\sigma}_n}{\hat{\sigma}_{n-1}}, \quad H_n = \sum_{k=1}^{n} h_k, \quad \bar{h}_n = \frac{H_n}{n} \tag{26}$$

and introduce the following statistics:

$$\mathcal{R}_n = \max_{k \leq n} \left(H_k - \frac{k}{n} H_n \right) - \min_{k \leq n} \left(H_k - \frac{k}{n} H_n \right), \tag{27}$$

$$\mathcal{S}_n^2 = \frac{1}{n} \sum_{k=1}^{n} (h_k - \bar{h}_n)^2, \quad \mathcal{Q}_n = \frac{\mathcal{R}_n}{\mathcal{S}_n}. \tag{28}$$

The statistics enjoys two facts useful in Hurst parameter estimation:

- If h_n are i.i.d., then for $n \gg 1$

$$\log \frac{\mathcal{R}_n}{\mathcal{S}_n} \approx \log \sqrt{\frac{2}{\pi}} + \frac{1}{2} \log n.$$

- If $h = (h_n)$ is fractional Gaussian noise, then for $n \gg 1$

$$\log \frac{\mathcal{R}_n}{\mathcal{S}_n} \approx \log c + H \log n.$$

Thus, in order to estimate H one should do a linear regression of $\log \mathcal{Q}_n$ on $\log n$ to obtain a slope coefficient $\hat{H}$. An application of this approach is displayed on fig.12.

4. **Adopted Whittle estimator** was already discussed in section 3. It assumes that the price and volatility dynamics is described with SDEs

$$dS_t = \sigma_t S_t \, dB_t, \quad d \log \sigma_t^2 = \varkappa_t \, dt + \nu \, dW_t^H \tag{29}$$

and uses a quasi-likelihood function based on the realized volatility sample given the number m of observations per day. The quasi-likelihood estimator of the pair $(\hat{H}, \hat{\nu})$ can be found as a solution of certain optimization problem (see [?]).

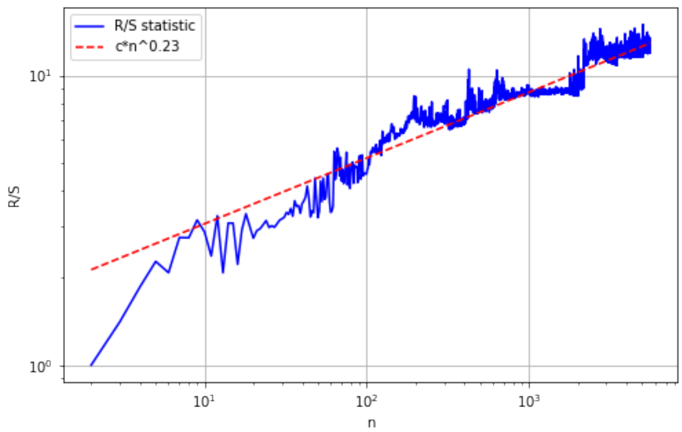

Figure 12: R/S analysis for market data. $\hat{H} = 0.23$.

Note that only w-estimator deals with roughness in a wide sense, as the other estimators highly rely on the properties of fBM or an assumption about dynamics (in case of Adopted Whittle estimator). An application of estimates to market data is displayed in table 2.

4.2 Numerical results

We will compare accuracy of the considered estimations for OUOU (22) with the parameters (23) and gfBM

$$\sigma_t = \exp\{\nu W_t^H\}, \quad t = 1, \dots T, \tag{30}$$

Estimator	$\hat{H}$
m-estimator	0.152
w-estimator	0.130
R/S analysis	0.227
Adopted Whittle estimator	0.075

Table 2: Statistical testing results.

$$T = 5000, \quad H \in \{0.01, 0.05, 0.1, 0.15, 0.2, 0.3, 0.5, 0.7\}. \tag{31}$$

The value of ν is taken equal to T^{-H} in order to keep $\mathrm{Var}\left[\log \sigma_T\right]$ constant and equal to 1. Such choice prevents the price trajectories from rapid convergence to zero and seems consistent with the market data.

Remark 1. *Note that the dependence of log-realized volatility* (3) *is not positively homogeneous with respect to* ν *as it can be rewritten in a following way:*

$$\hat{\sigma}_t^2 = \frac{1}{m}\sum_{k=1}^{m}\left|\int_{t+\frac{k-1}{m}}^{t+\frac{k}{m}} e^{\nu W_s^H}\, dB_s - \frac{1}{2}\int_{t+\frac{k-1}{m}}^{t+\frac{k}{m}} e^{2\nu W_s^H}\, ds\right|^2. \tag{32}$$

Thus the value of ν *might significantly affect the microstructure noise and lead to changes in realized volatility behavior. That is why we believe the estimation of* ν *is an important part in fitting gfBM model.*

The results of simulation of 250 trajectories for each set of parameters are presented in appendix B. We see that roughness estimations for instantaneous and realized volatility differ dramatically. Moreover, microstructure noise has different effect for large values of H (makes the trajectory much rougher) and for small H (makes it more regular). An example of the OUOU process demonstrates that even for a simple diffusion process considered methods can significantly overestimate the roughness of the trajectory. We also notice that the effect of microstructure noise coincides with the effect of Gaussian white noise when $H \geq 0.5$, although it is completely different in other cases.

5 Is volatility actually rough?

We have seen that the conclusions about roughness based only on the estimation of the realized volatility Hurst parameter cannot be trusted. Hence, we propose to consider an autocorrelation function (ACF) of $\Delta \log \hat{\sigma}_t$ and $(\Delta \log \hat{\sigma}_t)^2$. These functions for market volatility process are presented in fig.13. Using ACF function, we can finally distinguish OUOU and gfBM, as the second one has negative correlation of log-increments. We can also see that market ACF looks very similar to the ACF of fgBM with $H = 0.01$. However, their trajectories do not look alike and their estimations of H are significantly different.

We will combine the obtained knowledge about market realized volatility trajectory in order to show that the hypothesis

$$\mathcal{H}_0\colon\ \sigma_t = \exp\{\nu W_t^H\}, \quad \nu > 0, \quad H < \frac{1}{2} \tag{33}$$

should be rejected. More precisely, we will take into account the following facts:

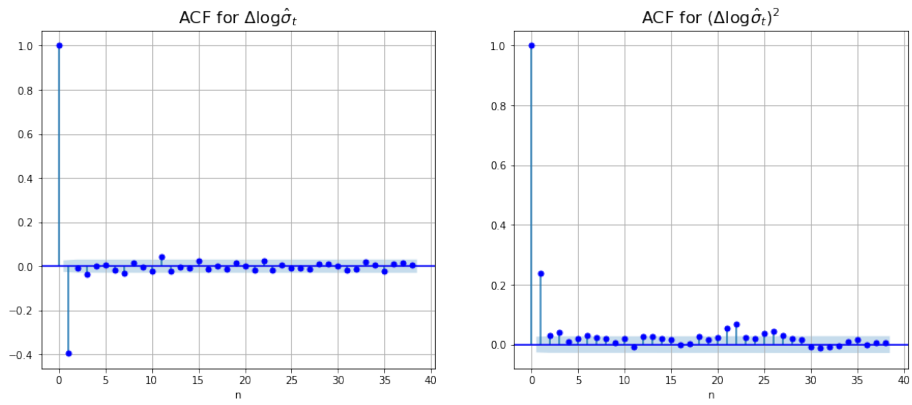

Figure 13: ACF of $\Delta \log \hat{\sigma}_t$ and $(\Delta \log \hat{\sigma}_t)^2$ for market data.

- The observed process is a realized volatility calculated according to (3).
- The estimation of H (via m-estimator) is $\tilde{H} = 0.152$.
- Autocorrelation $\mathrm{Corr}((\Delta \log \hat{\sigma}_t)^2, (\Delta \log \hat{\sigma}_{t+1})^2)$ is equal to $\tilde{\rho} = 0.238$.
- The variance $\mathrm{Var}[\Delta \log \hat{\sigma}_t]$ is equal to $\tilde{v} = 0.119$.

5.1 Bivariate statistic $(\hat{H}, \hat{\rho})$

Suppose for a moment that we know the exact value of H. Then for each trajectory $\sigma_t(\nu)$ we can calculate realized volatility $\hat{\sigma}_t(\nu)$ and estimate the variance of its log-increments. We will use the root of the equation

$$\mathrm{Var}[\Delta \log \hat{\sigma}_t(\hat{\nu})] = \tilde{v} \tag{34}$$

as an estimator $\hat{\nu}$. Doing this, we should realize that the function on the left-hand side, being a sample variance, is not deterministic. However, we have already seen its distribution (fig.3) and can claim that in case of realized volatility the variance is approximately constant.

Then for different values of H we apply the following sampling algorithm to obtain statistic $T = (\hat{H}, \hat{\rho})$:

1. Sample a trajectory of fBM W_t^H (T = 5000, m = 78).
2. Estimate $\hat{\nu}$ according to (34).
3. Calculate the realized volatility $\hat{\sigma}_t$ corresponding to the instantaneous volatility process $\sigma_t = \exp\{\hat{\nu} W_t^H\}$.
4. Use m-estimator to obtain $\hat{H}$.
5. Calculate the autocorrelation $\hat{\rho} = \mathrm{Corr}((\Delta \log \hat{\sigma}_t)^2, (\Delta \log \hat{\sigma}_{t+1})^2)$.

The distribution of the statistic is displayed in fig.14. We see that the point corresponding to market data is a strong outlier from empirical distribution and, hence, we should reject the hypothesis that it was generated by this distribution.

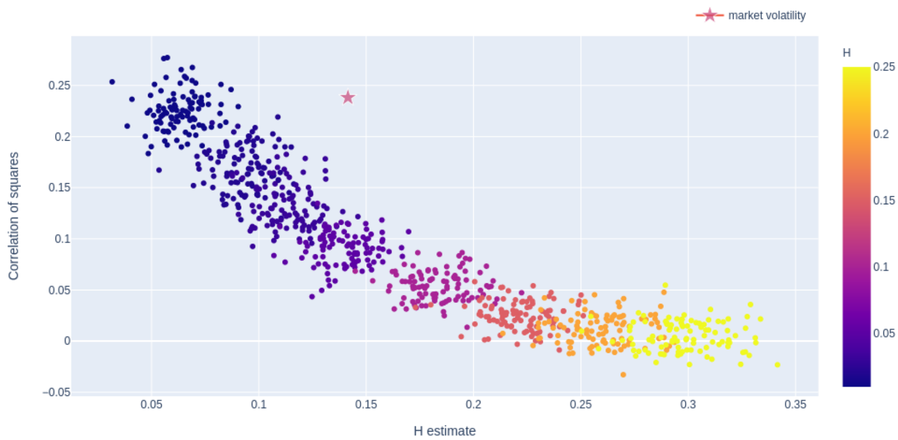

Figure 14: Empirical distribution of $(\hat{H}, \hat{\rho})$ for different H.

Table 3 demonstrates that similar relation between $\hat{H}$ and $\hat{\rho}$ is maintained for other indices. It should be noted that the ACF value of 0.463 looks extremely unrealistic for rough gfBM volatility model, as in this case

$$\mathrm{Corr}((\Delta \log \sigma_t)^2, (\Delta \log \sigma_{t+1})^2) = (2^{2H-1} - 1)^2 \in \left(0, \frac{1}{4}\right). \tag{35}$$

Index	$\hat{H}$	$\hat{\rho}$
S&P	0.141	0.238
FTSE	0.094	0.301
Nikkei 225	0.120	0.216
DAX	0.105	0.364
Russel	0.045	0.463

Table 3: Statistic $(\hat{H}, \hat{\rho})$ for different markets.

6 Conclusion

The provided numerical analysis shows that the problem of market instantaneous volatility roughness estimation is much more delicate than it seemed to be in the beginning. When analyzing the volatility trajectory we should remember the fact that it is corrupted by the microstructure noise whose nature is not completely understood yet. However, we have seen that its effect on fBM with $H \geq \frac{1}{2}$ is similar to the effect of the Gaussian white noise. However, for rough trajectories the effect of noise is much more complex as it is correlated with the volatility process and can diminish the variance of the process increments. Furthermore, the dependence of the noise structure on vol-of-vol parameter ν is not clear too. More detailed study of the microstructure noise is an interesting subject for the future research.

We have also claimed that roughness should be understood wider, as a property of path and, hence, non-parametric criteria should be used to estimate it. All the considered estimators can significantly overestimate the actual roughness of the volatility process for two reasons: the exposure to microstructure noise and low sampling frequency. Thus, the estimators of roughness based on the discrete trajectories of realized volatility can be unreliable. A perfect estimator, if it exists, should take into account microstructure noise, as it was done in [**?**], but it should also be model independent.

We have also proposed a statistic based on the ACF function and the estimation of H to show that market volatility process can hardly be obtained as a realized volatility in gfBM model suggested in [**?**]. Nevertheless, rough volatility models can be a very good proxy for market volatility, as they can reproduce the effect of antipersistency and strong (unfortunately, not enough) correlation of the increments.

A ACF

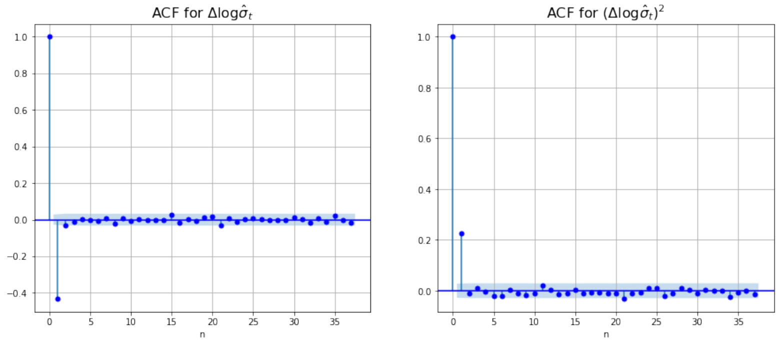

Figure 15: ACF of realized volatility for fgBM with $H = 0.01$.

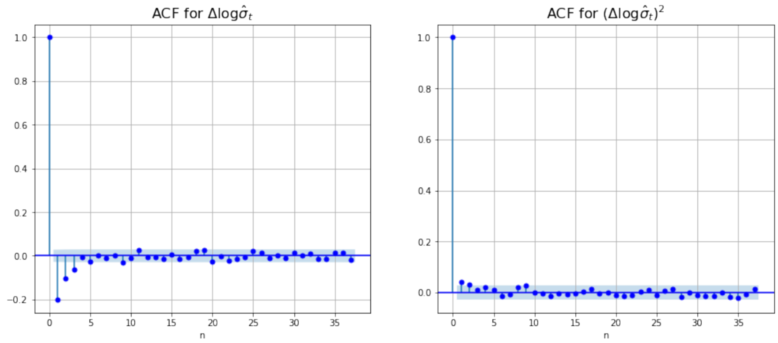

Figure 16: ACF of realized volatility for fgBM with $H = 0.15$.

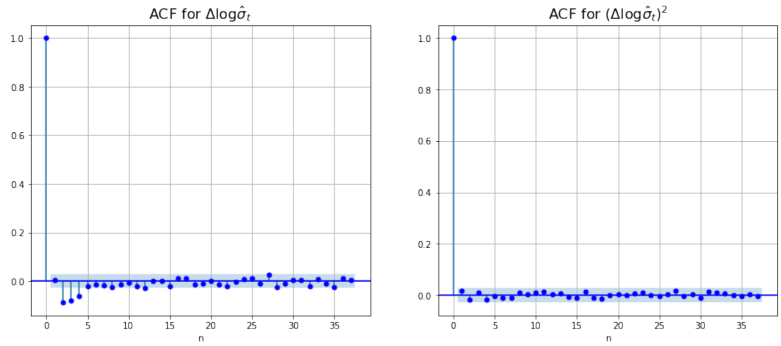

Figure 17: ACF of realized volatility for OUOU.

B Accuracy of estimators

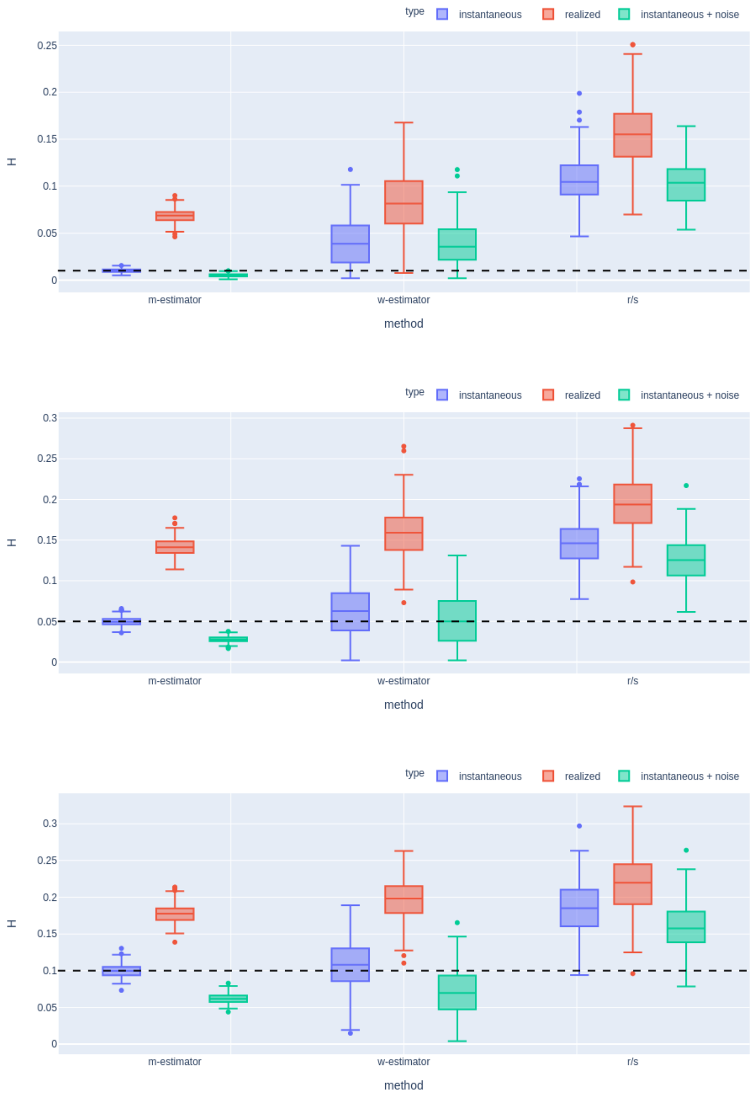

Figure 18: fgBM with $H \in \{0.01, 0.05, 0.1\}$.

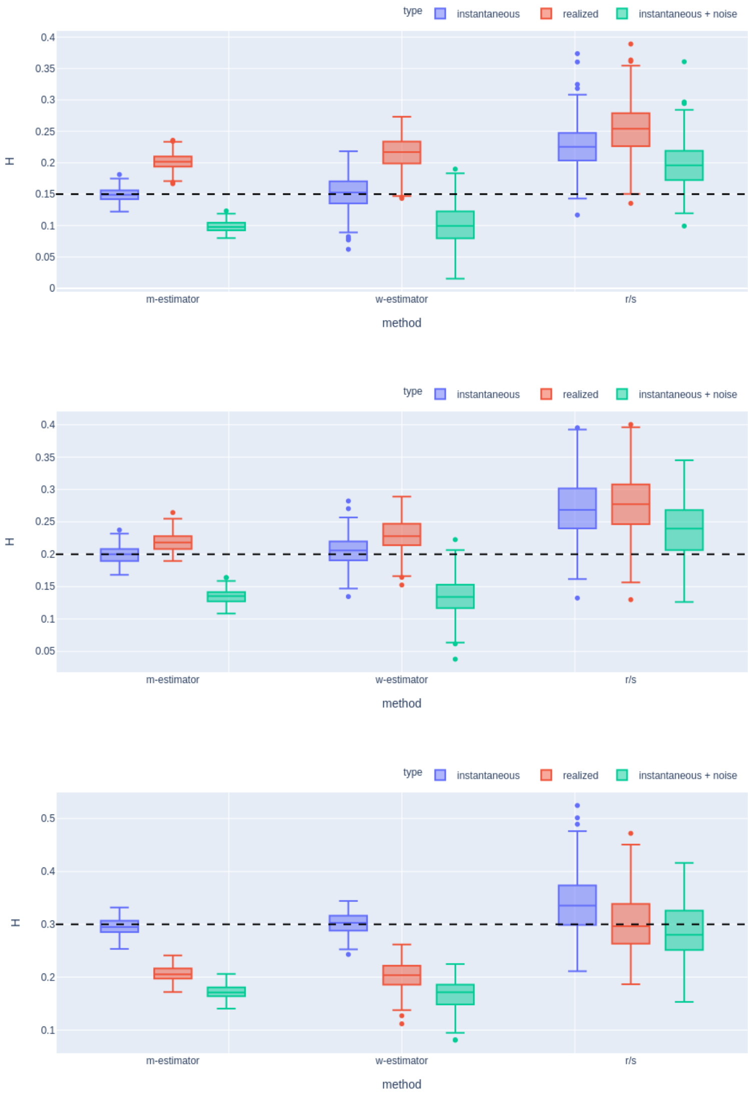

Figure 19: fgBM with $H \in \{0.15, 0.2, 0.3\}$.

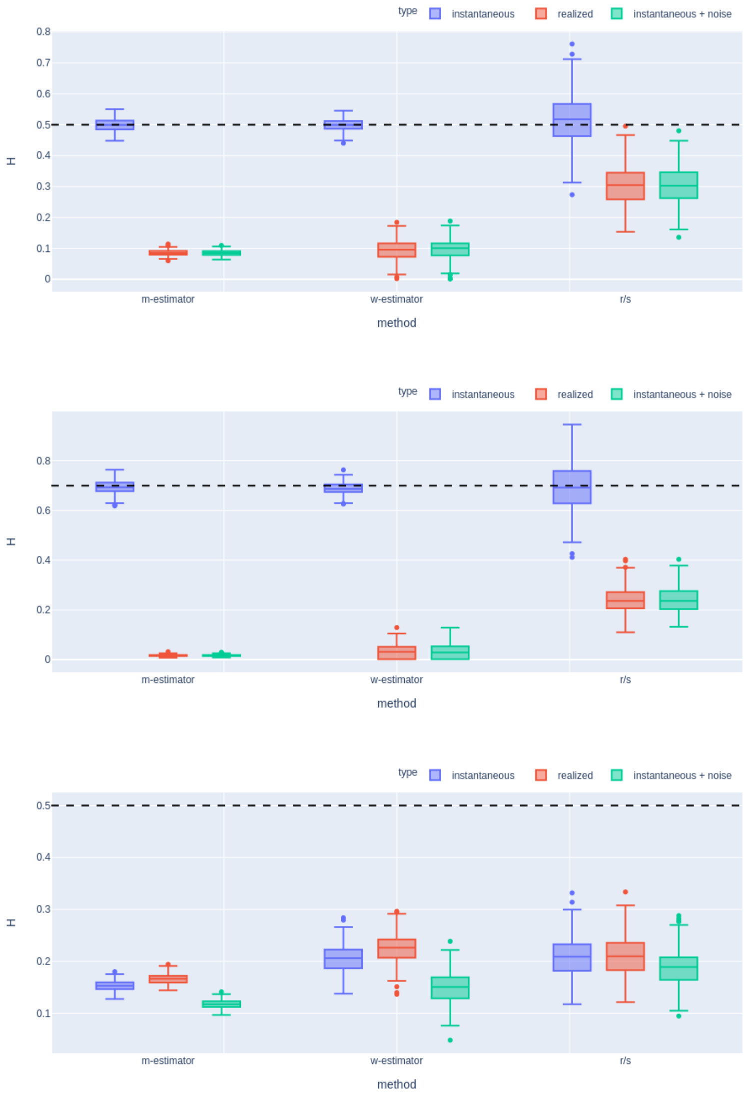

Figure 20: fgBM with $H \in \{0.5, 0.7\}$ and OUOU.

,3,