Quantitative Finance and Economics

https://www.aimspress.com/journal/QFE

QFE, 9(1): 1–39.
DOI: 10.3934/QFE.2025001
Received: 01 October 2024
Revised: 12 December 2024
Accepted: 06 January 2025
Published: 09 January 2025

Research article

Some empirical studies for the applications of fractional G-Brownian motion in finance

Changhong Guo[1], **Shaomei Fang**[2,*], **Yong He**[1] **and Yong Zhang**[1]

[1] School of Management, Guangdong University of Technology, Yinglong Road 161, Guangzhou, Guangdong 510520, P. R. China

[2] Department of Mathematics, South China Agricultural University, Wushan Street 483, Guangzhou, Guangdong 510640, P. R. China

* **Correspondence:** Email: fangsm90@163.com, dz90@scau.edu.cn; Tel: +86-20-87344806; Fax: +86-20-85282366.

Abstract: Since the fractional G-Brownian motion (fGBm) generalizes the concepts of the standard Brownian motion, fractional Brownian motion, and G-Brownian motion, while it can exhibit long-range dependence or antipersistence and feature the volatility uncertainty simultaneously, it can be a better alternative stochastic process in the financial applications. Thus, in this paper, some empirical studies for the financial applications of the fGBm were carried out, where the recent high-frequency data for some selected assets in the financial market are from the Oxford-Man Institute of Quantitative Finance Realized Library. There are two main empirical findings. One was that the H-G-normal distributions associated with the fGBm are more suitable in describing the dynamics of daily returns and increments of log-volatility for these assets than the usual distributions, since they not only characterize the properties of skewness, excess kurtosis, and long-range dependence $\left(\frac{1}{2} < H < 1\right)$ or antipersistence $\left(0 < H < \frac{1}{2}\right)$, but also feature the volatility uncertainty. The other one was that the daily return and log-volatility both behave essentially as fGBm with different $\underline{\sigma}^2$ and $\overline{\sigma}^2$, but Hurst parameters $H < \frac{1}{2}$, at any reasonable time scale. Then a generalized stochastic model for the dynamics of the assets called rough fractional stochastic volatility model driven by fGBm (RFSV-fGBm) was developed. Finally, some parameter estimates and numerical experiments for the RFSV-fGBm model were investigated and carried out.

Keywords: fractional G-Brownian motion; H-G-normal distribution; volatility smoothness; RFSV-fGBm; parameter estimates

JEL Codes: G15, G10

1. Introduction

In the last few decades, some stochastic processes have been widely investigated and applied in the area of mathematical finance. The earliest and most important one is the standard Brownian motion (Bm), which is a continuous Gaussian process with self-similar and independent increments (Klebaner, 2012; Privault, 2013). However, there are two main incompatibilities, as one employs the Bm to the real market as carrying out some empirical studies. The first one is about the Gaussian law, since the asset return distributions in the financial markets always exhibit excess kurtosis and heavy tails (Fama, 1965). The other one is the independence of the increments, which is incompatible with the property of long-range dependence or persistence $\left(\frac{1}{2} < H < 1\right)$ (Lo and MacKinlay, 1988), which is the classical stylized fact for most of the return distributions in the financial markets. In order to characterize the long-range dependence, another stochastic process called fractional Brownian motion (fBm) has been introduced and then widely applied in mathematical finance (Kolmogorov, 1940; Mandelbrot and Van Ness, 1968; Biagini et al. , 2008). However, it always causes some contradictions as one takes the fBm as the driven source to model the asset pricing. If one adopts the stochastic calculus defined for the Bm to fBm directly, the asset pricing models based on fBm will admit arbitrage (Rogers, 1997; Sottinen, 2001); that is because the fBm is neither a Markov process nor a semimartingale, except for the case $H = 1/2$ (Rogers, 1997). One way to exclude the arbitrage opportunity was taking another different definition for the stochastic calculus with respect to fBm, such as the fractional Itô integral based on the Wick product (Hu and Øksendal, 2003; Elliott and Hoek, 2003; Biagini et al. , 2008). However, it will not have a reasonable economic interpretation for the definition of the self-financing trading strategies used here (Björk and Hult, 2005). Anyway, the fBm not only has better-behaved tails but also exhibits long-range dependence $\left(\frac{1}{2} < H < 1\right)$ or antipersistence $\left(0 < H < \frac{1}{2}\right)$, which makes it an important role in financial applications.

The former two stochastic processes focus on characterizing the essential randomness of the financial markets. Besides this kind of randomness, there is another model uncertainty for the markets that needed to be considered, which is always called model ambiguity. Two representative examples are the drift uncertainty (Coquet et al., 2002) and volatility uncertainty (Avellaneda et al., 1995; Lyons, 1995). The latter one means that the volatility is not known precisely but is assumed to lie between two extreme values, as referred to by a certainty band (Avellaneda et al., 1995; Lyons, 1995). Some attempts have been made to handle this type of model uncertainty, such as the robust statistics (Denis and Martini, 2006; Muhle-Karbe and Nutz, 2018), some imprecise continuous-time Markov chains (Krak et al., 2017), non-linear affine processes (Fadina et al., 2019), and so on. Most importantly, Peng (2005, 2007a,b, 2008, 2011, 2019) proposed a formal mathematical approach under the framework of nonlinear expectation and related G-Brownian motion (GBm) on some sublinear space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ to model the volatility uncertainty. The concepts of nonlinear expectation and related GBm have a very rich and interesting new structure and non-trivially generalize the classical one. The applications of nonlinear expectation and related GBm in mathematical finance have been widely studied (Chen and Epstein, 2002; Denis et al., 2011; Epstein and Ji, 2013; Vorbrink, 2014; Soumana-Hima, 2017), and the GBm process has also been extended to some more general cases, as being called the non-linear Lévy processes; we refer readers to Hu and Peng (2009); Neufeld and Nutz (2017); Kühn (2019); Denk et al. (2020) and references therein.

Most recently, in order to take full advantage of fBm in exhibiting the long-range dependence or antipersistence and GBm in featuring the volatility uncertainty, a generalized stochastic process called fractional G-Brownian motion (fGBm) was developed (Guo et al., 2023a,b). Since the fGBm generalizes the concepts of the standard Bm, fBm, and GBm in the framework of nonlinear expectation, it is much more reasonable to consider it to be a better alternative stochastic process used for financial applications. Thus subsequent to these works (Guo et al., 2023a,b), in this paper we are further going to make some empirical studies for the applications of fGBm in finance, as a starting point. There are two main empirical findings. One was that the H-G-normal distributions associated with the fGBm are more suitable in describing the dynamics of daily returns and increments of log-volatility for these assets than the usual distributions, since they not only characterize the properties of skewness, excess kurtosis, and long-range dependence $\left(\frac{1}{2} < H < 1\right)$ or antipersistence $\left(0 < H < \frac{1}{2}\right)$, but also feature the volatility uncertainty. The other one was that the daily return and log-volatility both behave essentially as fGBm with different $\underline{\sigma}^2$ and $\overline{\sigma}^2$, but Hurst parameters $H < \frac{1}{2}$. Based on these empirical findings, we develop a generalized stochastic model for the dynamics of the assets and call it rough fractional stochastic volatility model driven by fGBm (RFSV-fGBm). Furthermore, some parameter estimates and numerical experiments for the RFSV-fGBm model are also investigated and carried out.

The rest of the paper is organized as follows: In Section 2, we provide some notations and preliminaries for the sublinear expectation and fGBm. The definition of the H-G-normal distribution associated with the fGBm is also given out in a similar way. In Section 3, we will make some empirical studies to show that the H-G-normal distributions are more suitable in describing the dynamics of some selected assets, both for the daily returns and increments of log-volatility, than the usual normal distributions and Cauchy distributions, where the recent high-frequency data for these assets are from the Oxford-Man Institute of Quantitative Finance Realized Library. In Section 4, we estimate the smoothness of the daily return and log-volatility processes for these assets under the sublinear expectation, and show that both of them behave essentially as fGBm with different $\underline{\sigma}^2$ and $\overline{\sigma}^2$, but Hurst parameters $H < \frac{1}{2}$, at any reasonable time scale. Based on the empirical results, some stochastic model called rough fractional stochastic volatility model driven by fGBm will be developed for the dynamics of the assets in Section 5. Some parameter estimates for the RFSV-fGBm model were investigated, and some experiments were also carried out in this section. Section 6 concludes. Some additional preliminaries about the theory of sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ and other results about the related GBm are collected in Appendix A. Some parameter estimates and fitted pdfs for the increments of the other NASDAQ index, DAX, and HANG SENG index are also present in Appendix B.

2. Notations and preliminaries

In this section, we give out some necessary notations and preliminaries and summarize some results for fGBm from Guo et al. (2023a,b). For more notations, preliminaries, and other results about the theory of sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ and related GBm, we list them in Appendix A and refer readers to Peng (2005, 2007a,b, 2008, 2011, 2019) and references therein. First, we recall the basic definition of one-dimensional fGBm (Guo et al., 2023a,b), which is stated as follows.

Definition 1. *(Fractional G-Brownian motion) (Guo et al., 2023a,b) Let $H \in (0, 1)$. Then a continuous stochastic process $B_{FG}(t)_{t\in\mathbb{R}^+}$ on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ is called a fractional G-Brownian motion (fGBm) with Hurst parameter H if*

(1) $B_{FG}(0) = 0$*, and for all* $t \geq 0$

$$-\widehat{\mathbb{E}}\left[-B_{FG}(t)\right] = \widehat{\mathbb{E}}\left[B_{FG}(t)\right] = 0. \tag{1}$$

(2) For all $s, t \geq 0$*, there holds*

$$\begin{aligned} \widehat{\mathbb{E}}\left[B_{FG}(t)B_{FG}(s)\right] &= \tfrac{1}{2}\overline{\sigma}^2\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \\ -\widehat{\mathbb{E}}\left[-B_{FG}(t)B_{FG}(s)\right] &= \tfrac{1}{2}\underline{\sigma}^2\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \end{aligned} \tag{2}$$

where $\underline{\sigma}^2 = -\widehat{\mathbb{E}}\left[-B_{FG}^2(1)\right]$ *and* $\overline{\sigma}^2 = \widehat{\mathbb{E}}\left[B_{FG}^2(1)\right]$.

(3) For each $t, s \geq 0$*,* $B_{FG}(t+s) - B_{FG}(s)$ *and* $B_{FG}(t)$ *are identically distributed.*

(4) $\lim\limits_{t\to 0}\widehat{\mathbb{E}}\left[|B_{FG}(t)|^m\right]t^{-2H} = 0$ *for each* $m \in \mathbb{N}$ *and* $m \geq 3$.

Remark 1. *The fGBm* $B_{FG}(t)_{t\in\mathbb{R}^+}$ *generalizes the concepts of the standard Bm* $B(t)_{t\in\mathbb{R}^+}$ *(Privault, 2013), fBm* $B_H(t)_{t\in\mathbb{R}^+}$ *(Biagini et al. , 2008), and GBm* $B_G(t)_{t\in\mathbb{R}^+}$ *(Peng, 2019) in the framework of nonlinear expectation, and it can exhibit the long-range dependence* $\left(\frac{1}{2} < H < 1\right)$ *or antipersistence* $\left(0 < H < \frac{1}{2}\right)$ *property and feature the volatility uncertainty* $\left(\underline{\sigma}^2, \overline{\sigma}^2\right)$ *simultaneously.*

Lemma 1. *(Guo et al., 2023b) Let* $\Delta B_{FG}(t) = B_{FG}(t+\Delta t) - B_{FG}(t)$*. Then*

$$\Delta B_{FG}(t) \overset{d}{=} \xi(t)(\Delta t)^H, \ \ as \ \ \Delta t \to 0, \tag{3}$$

or in continuous form

$$\mathrm{d}B_{FG}(t) \overset{d}{=} \xi(t)(\mathrm{d}t)^H, \ \ as \ \ \mathrm{d}t \to 0, \tag{4}$$

where $\xi(t)$ *is a G-normally distributed random variable* $\xi(t) \sim \mathcal{N}\left(\{0\}, [\underline{\sigma}^2, \overline{\sigma}^2]\right)$*, and* $X \overset{d}{=} Y$ *means* X *and* Y *are identically distributed, as defined in Definition A.3.*

Definition 2. *Let* $X(t)_{t\in\mathbb{R}^+}$ *be a stochastic process on a sublinear expectation space* $(\Omega, \mathcal{H}, \widehat{\mathbb{E}})$*. We define the upper and lower autocovariance functions, upper and lower autocorrelation functions, and upper and lower autocorrelation coefficients of* $X(t)$ *at* t_1 *and* t_2 *as*

(1) Upper and lower autocovariance functions:

$$\begin{aligned} \overline{C}_X(t_1,t_2) &:= \mathbb{C}ov\left(X(t_1), X(t_2)\right) = \widehat{\mathbb{E}}\left[\left(X(t_1) - \widehat{\mathbb{E}}[X(t_1)]\right)\left(X(t_2) - \widehat{\mathbb{E}}[X(t_2)]\right)\right], \\ \underline{C}_X(t_1,t_2) &:= -\mathbb{C}ov\left(-X(t_1), X(t_2)\right) = -\widehat{\mathbb{E}}\left[\left(-X(t_1) - \widehat{\mathbb{E}}[-X(t_1)]\right)\left(X(t_2) - \widehat{\mathbb{E}}[X(t_2)]\right)\right]. \end{aligned}$$

(2) Upper and lower autocorrelation functions:

$$\overline{R}_X(t_1,t_2) := \widehat{\mathbb{E}}\left[X(t_1)X(t_2)\right], \quad \underline{R}_X(t_1,t_2) := -\widehat{\mathbb{E}}\left[-X(t_1)X(t_2)\right].$$

(3) Upper and lower autocorrelation coefficients:

$$\overline{\rho}_X(t_1,t_2) := \frac{\overline{C}_X(t_1,t_2)}{\sqrt{\overline{C}_X(t_1,t_1)}\sqrt{\overline{C}_X(t_2,t_2)}}, \quad \underline{\rho}_X(t_1,t_2) := \frac{\underline{C}_X(t_1,t_2)}{\sqrt{\underline{C}_X(t_1,t_1)}\sqrt{\underline{C}_X(t_2,t_2)}}.$$

Remark 2. *Sometimes for simplicity, we denote*

$$\begin{aligned} &\overline{C}_X(\tau) = \overline{C}_X(t,t+\tau), \quad \underline{C}_X(\tau) = \underline{C}_X(t,t+\tau), \quad \overline{R}_X(\tau) = \overline{R}_X(t,t+\tau), \quad \underline{R}_X(\tau) = \underline{R}_X(t,t+\tau), \\ &\overline{\rho}_X(\tau) = \overline{\rho}_X(t,t+\tau), \quad \underline{\rho}_X(\tau) = \underline{\rho}_X(t,t+\tau). \end{aligned} \tag{5}$$

Definition 3. *Let $X(t)_{t\in\mathbb{R}}$ and $Y(t)_{t\in\mathbb{R}}$ be two stochastic processes on a sublinear expectation space $(\Omega,\mathcal{H},\widehat{\mathbb{E}})$. Then the upper and lower cross-covariance functions, upper and lower cross-correlation functions, and upper and lower cross-correlation coefficients between $X(t)$ at time t_1 and $Y(t)$ at time t_2 are defined by*

(1) Upper and lower cross-covariance functions:

$$\overline{C}_{XY}(t_1,t_2) := \mathbb{C}ov(X(t_1),Y(t_2)) = \widehat{\mathbb{E}}\left[\left(X(t_1)-\widehat{\mathbb{E}}[X(t_1)]\right)\left(Y(t_2)-\widehat{\mathbb{E}}[Y(t_2)]\right)\right],$$
$$\underline{C}_{XY}(t_1,t_2) := -\mathbb{C}ov(-X(t_1),Y(t_2)) = -\widehat{\mathbb{E}}\left[\left(-X(t_1)-\widehat{\mathbb{E}}[-X(t_1)]\right)\left(Y(t_2)-\widehat{\mathbb{E}}[Y(t_2)]\right)\right].$$

(2) Upper and lower cross-correlation functions:

$$\overline{R}_{XY}(t_1,t_2) := \widehat{\mathbb{E}}[X(t_1)Y(t_2)], \quad \underline{R}_{XY}(t_1,t_2) := -\widehat{\mathbb{E}}[-X(t_1)Y(t_2)].$$

(3) Upper and lower cross-correlation coefficients:

$$\overline{\rho}_{XY}(t_1,t_2) := \frac{\overline{C}_{XY}(t_1,t_2)}{\sqrt{\overline{C}_{XX}(t_1,t_1)}\sqrt{\overline{C}_{YY}(t_2,t_2)}}, \quad \underline{\rho}_{XY}(t_1,t_2) := \frac{\underline{C}_{XY}(t_1,t_2)}{\sqrt{\underline{C}_{XX}(t_1,t_1)}\sqrt{\underline{C}_{YY}(t_2,t_2)}}.$$

Definition 4. *(Upper and lower correlation coefficients) If the upper and lower cross-correlation coefficients of $X(t)$ and $Y(t)$ are independent of t. Then we call $X(t)$ and $Y(t)$ to have the constant upper correlation coefficient and lower correlation coefficient at all times, and denote*

$$\overline{\rho} := \overline{\rho}(X,Y) = \overline{\rho}_{XY}(t,t), \quad \underline{\rho} := \underline{\rho}(X,Y) = \underline{\rho}_{XY}(t,t).$$

Definition 5. *(Long-range dependence) (Guo et al., 2023a) A sequence $(X_n)_{n\in\mathbb{N}}$ on a sublinear expectation space $(\Omega,\mathcal{H},\widehat{\mathbb{E}})$ exhibits long-range dependence if the upper autocovariance functions*

$$\overline{C}_X(n) := \overline{C}_X(k,k+n) = \mathbb{C}ov(X_k,X_{k+n}),$$

and lower autocovariance functions

$$\underline{C}_X(n) := \underline{C}_X(k,k+n) = -\mathbb{C}ov(-X_k,X_{k+n}),$$

both satisfy

$$\lim_{n\to\infty}\frac{\overline{C}(n)}{c_1 n^{-\alpha}} = 1, \quad \lim_{n\to\infty}\frac{\underline{C}(n)}{c_2 n^{-\beta}} = 1,$$

for some constants c_1 and c_2, and $\alpha,\beta\in(0,1)$.

Theorem 2. *(Guo et al., 2023a) For a fGBm $B_{FG}(t)_{t\in\mathbb{R}^+}$ with Hurst parameter $H\in(0,1)$ defined on the sublinear expectation space $(\Omega,\mathcal{H},\widehat{\mathbb{E}})$. Then*

(1) G-normally distributed: $B_{FG}(t)\sim\mathcal{N}\left(\{0\},\left[t^{2H}\underline{\sigma}^2,t^{2H}\overline{\sigma}^2\right]\right)$, i.e.,

$$-\widehat{\mathbb{E}}[-B_{FG}(t)] = \widehat{\mathbb{E}}[B_{FG}(t)] = 0, \quad -\widehat{\mathbb{E}}\left[-B_{FG}^2(t)\right] = \underline{\sigma}^2 t^{2H}, \quad \widehat{\mathbb{E}}\left[B_{FG}^2(t)\right] = \overline{\sigma}^2 t^{2H}. \tag{6}$$

(2) Self-similarity:

$$a^{-H}B_{FG}(at) \overset{d}{=} B_{FG}(t), \text{ for any } a>0. \tag{7}$$

(3) Long-range dependence: if $H\in\left(\frac{1}{2},1\right)$.

Lemma 3. *Let $B^1_{FG}(t)_{t\in\mathbb{R}^+}$ and $B^2_{FG}(t)_{t\in\mathbb{R}^+}$ be two fractional G-Brownian motions on a sublinear expectation space $\left(\Omega,\mathcal{H},\widehat{\mathbb{E}}\right)$. If $B^2_{FG}(t)$ is an independent copy of $B^1_{FG}(t)$, and $B^1_{FG}(t)$ and $B^2_{FG}(t)$ have the constant upper correlation coefficient $\overline{\rho}$, then we have*

$$\widehat{\mathbb{E}}\left[\Delta B^1_{FG}(t)\Delta B^2_{FG}(t)\right] = \overline{\rho}\,(\Delta t)^{2H}\,, \tag{8}$$

where $\Delta B^k_{FG}(t) = B^k_{FG}\,(t+\Delta t) - B^k_{FG}(t)$, $k = 1, 2$.

Proof. First, we claim that $B^2_{FG}(t)$ can be expressed as

$$B^2_{FG}(t) = \overline{\rho}B^1_{FG}(t) + \sqrt{1-\overline{\rho}^2}B^*_{FG}(t), \tag{9}$$

where $B^*_{FG}(t)$ is another fGBm on the sublinear expectation space $\left(\Omega,\mathcal{H},\widehat{\mathbb{E}}\right)$ and is an independent copy of $B^1_{FG}(t)$. It is easy to verify that $B^2_{FG}(t)$ defined in (9) is indeed an fGBm on the sublinear expectation space $\left(\Omega,\mathcal{H},\widehat{\mathbb{E}}\right)$. We just prove they are correlated with the upper correlation coefficient $\overline{\rho}$. By Theorem 2 and Lemma A.1, it is easy to find that

$$\widehat{\mathbb{E}}\left[B^2_{FG}(t)\right] = \widehat{\mathbb{E}}\left[\overline{\rho}B^1_{FG}(t) + \sqrt{1-\overline{\rho}^2}B^*_{FG}(t)\right] = 0. \tag{10}$$

Using Lemma A.1 again, we have

$$\begin{aligned}
\overline{C}_{B^2_{FG}B^2_{FG}}(t,t) &= \mathbb{C}\mathrm{ov}\left(B^2_{FG}(t), B^2_{FG}(t)\right)\\
&= \widehat{\mathbb{E}}\left[\left(B^2_{FG}(t) - \widehat{\mathbb{E}}[B^2_{FG}(t)]\right)^2\right]\\
&= \widehat{\mathbb{E}}\left[\left(\overline{\rho}B^1_{FG}(t) + \sqrt{1-\overline{\rho}^2}B^*_{FG}(t)\right)^2\right]\\
&= \widehat{\mathbb{E}}\left[\overline{\rho}^2\left(B^1_{FG}(t)\right)^2 + 2\overline{\rho}\sqrt{1-\overline{\rho}^2}B^1_{FG}(t)B^*_{FG}(t) + \left(1-\overline{\rho}^2\right)\left(B^*_{FG}(t)\right)^2\right]\\
&= \overline{\rho}^2\widehat{\mathbb{E}}\left[\left(B^1_{FG}(t)\right)^2\right] + \left(1-\overline{\rho}^2\right)\widehat{\mathbb{E}}\left[\left(B^*_{FG}(t)\right)^2\right]\\
&= t^{2H},
\end{aligned} \tag{11}$$

where the results of $\widehat{\mathbb{E}}\left[B^1_{FG}(t)B^*_{FG}(t)\right] = 0$ and $\widehat{\mathbb{E}}\left[X^2+Y^2\right] = \widehat{\mathbb{E}}\left[X^2\right] + \widehat{\mathbb{E}}\left[Y^2\right]$ were used, since $B^*_{FG}(t)$ is an independent copy of $B^1_{FG}(t)$. Similarly, we also have

$$\begin{aligned}
\overline{C}_{B^1_{FG}B^2_{FG}}(t,t) &= \mathbb{C}\mathrm{ov}\left(B^1_{FG}(t), B^2_{FG}(t)\right)\\
&= \widehat{\mathbb{E}}\left[\left(B^1_{FG}(t) - \widehat{\mathbb{E}}[B^1_{FG}(t)]\right)\left(B^2_{FG}(t) - \widehat{\mathbb{E}}[B^1_{FG}(t)]\right)\right]\\
&= \widehat{\mathbb{E}}\left[B^1_{FG}(t)\left(\overline{\rho}B^1_{FG}(t) + \sqrt{1-\overline{\rho}^2}B^*_{FG}(t)\right)\right]\\
&= \widehat{\mathbb{E}}\left[\overline{\rho}\left(B^1_{FG}(t)\right)^2 + \sqrt{1-\overline{\rho}^2}B^1_{FG}(t)B^*_{FG}(t)\right]\\
&= \overline{\rho}\widehat{\mathbb{E}}\left[\left(B^1_{FG}(t)\right)^2\right] + \sqrt{1-\overline{\rho}^2}\times 0\\
&= \overline{\rho}t^{2H}.
\end{aligned} \tag{12}$$

Thus, there holds

$$\overline{\rho}\left(B^1_{FG}, B^2_{FG}\right) = \overline{\rho}_{B^1_{FG}B^2_{FG}}(t,t) = \frac{\overline{C}_{B^1_{FG}B^2_{FG}}(t,t)}{\sqrt{\overline{C}_{B^1_{FG}B^1_{FG}}(t,t)}\sqrt{\overline{C}_{B^2_{FG}B^2_{FG}}(t,t)}} = \frac{\overline{\rho}t^{2H}}{\sqrt{t^{2H}}\sqrt{t^{2H}}} = \overline{\rho}. \tag{13}$$

So the claim is verified. Then we have

$$\widehat{\mathbb{E}}\left[\Delta B^1_{FG}(t)\Delta B^2_{FG}(t)\right] = \widehat{\mathbb{E}}\left[\Delta B^1_{FG}(t)\left(\overline{\rho}\Delta B^1_{FG}(t) + \sqrt{1-\overline{\rho}^2}\Delta B^*_{FG}(t)\right)\right] = \overline{\rho}\widehat{\mathbb{E}}\left[\left(\Delta B^1_{FG}(t)\right)^2\right] = \overline{\rho}(\Delta t)^{2H}, \tag{14}$$

where the independence of $B^*_{FG}(t)$ and $B^1_{FG}(t)$ and Lemma A.1 were also applied. This completes the proof of Lemma 3.

Remark 3. *If $B^1_{FG}(t)$ and $B^2_{FG}(t)$ have the constant lower correlation coefficient $\underline{\rho}$, and $B^2_{FG}(t)$ is an independent copy of $B^1_{FG}(t)$. Then we can also set*

$$B^2_{FG}(t) = \underline{\rho}B^1_{FG}(t) + \sqrt{1-\underline{\rho}^2}B^*_{FG}(t), \tag{15}$$

and there will be

$$-\widehat{\mathbb{E}}\left[-\Delta B^1_{FG}(t)\Delta B^2_{FG}(t)\right] = \underline{\rho}(\Delta t)^{2H}. \tag{16}$$

Remark 4. *Sometimes, we also write* (8) *and* (16) *in the continuous form as*

$$\widehat{\mathbb{E}}\left[\mathrm{d}B^1_{FG}(t)\mathrm{d}B^2_{FG}(t)\right] = \overline{\rho}(\mathrm{d}t)^{2H}, \quad \textit{and} \quad -\widehat{\mathbb{E}}\left[-\mathrm{d}B^1_{FG}(t)\mathrm{d}B^2_{FG}(t)\right] = \underline{\rho}(\mathrm{d}t)^{2H}. \tag{17}$$

As it is well known that the anomalous diffusion can be understood from the probability perspective, there is indeed a deep relationship between the stochastic processes and anomalous diffusions. To be more specific, the evolution of the probability distribution of the stochastic processes and anomalous diffusions can be represented as the solutions of some partial differential equations (PDEs). For the standard Bm and fBm, their probability density functions (pdfs) and related PDEs can be seen in Theorem A.3. As a result, for any $t, s \geq 0$, there holds that

$$B(t+s) - B(t) \overset{d}{=} \xi, \quad B_H(t+s) - B_H(t) \overset{d}{=} \eta, \tag{18}$$

where $\xi \sim \mathcal{N}(0, s)$ and $\eta \sim \mathcal{N}_H(0, s)$, which are defined by Definition 6, respectively.

Definition 6. *(Jacod and Protter, 2004) We say that X is a normal variable with parameters μ and σ^2, if its probability density function (pdf) is given by*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}, \tag{19}$$

and we denote $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. We say that X is a H-normal variable with parameters μ, σ^2 and H $(0 < H < 1)$, if its pdf (H-Normal pdf) is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^{4H}}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^{4H}}\right), \quad x \in \mathbb{R}, \tag{20}$$

and we denote $X \sim \mathcal{N}_H\left(\mu, \sigma^2\right)$. Obviously, $X \sim \mathcal{N}_H\left(\mu, \sigma^2\right)$ is equivalent to $X \sim \mathcal{N}\left(\mu, \sigma^{4H}\right)$.

However, for the GBm in the framework of nonlinear expectation, the associated G-normal distribution $\mathcal{N}\left(\{0\},\left[\underline{\sigma}^2,\overline{\sigma}^2\right]\right)$ on the sublinear expectation space $\left(\Omega,\mathcal{H},\widehat{\mathbb{E}}\right)$ is no longer a classical probability distribution but an infinite family of distributions $\{F_\theta\}_{\theta\in\Theta}$, where Θ is the set of all of the intervals $\left[\underline{\sigma}^2,\overline{\sigma}^2\right]$, which also means that the interval $\left[\underline{\sigma}^2,\overline{\sigma}^2\right]$ is used to characterize the unknown family of distributions $\{F_\theta\}_{\theta\in\Theta}$. For any $\varphi(x)\in C_{l,Lip}(\mathbb{R}^n)$, by the representation theorem of sublinear expectation (Peng, 2019), there holds

$$\widehat{\mathbb{E}}\left[\varphi(X)\right]=\sup_{\theta\in\Theta}\mathbb{E}_\theta\left[\varphi(X)\right]=\sup_{\theta\in\Theta}\int_{\mathbb{R}}\varphi(x)\mathrm{d}F_\theta(x). \tag{21}$$

Thus, the distribution of $B_G(t)$ is now characterized by Theorem A.4 and Theorem A.5. In similar to Theorem A.5, we have

Theorem 4. *(fGBm with PDEs) (Guo et al., 2023a) Let $B_{FG}(t)_{t\in\mathbb{R}^+}$ be one-dimensional fGBm with Hurst parameter $H\in(0,1)$ on a sublinear expectation space $\left(\Omega,\mathcal{H},\widehat{\mathbb{E}}\right)$. Then the distribution of $B_{FG}(t)$ is characterized by*

$$\widehat{\mathbb{E}}\left[\varphi\left(B_{FG}(t)\right)\right]=v\left(0,t^{2H}\right),\quad \varphi\in C_{l,Lip}(\mathbb{R}), \tag{22}$$

or

$$\widehat{\mathbb{E}}\left[\varphi\left(B_{FG}(t)\right)\right]=v_H(0,t),\quad \varphi\in C_{l,Lip}(\mathbb{R}), \tag{23}$$

where $v(x,t)$ is the viscosity solution of the G-heat equation (A.15) *with initial condition* (A.16)*, where $G(x)$ is defined by*

$$G(x)=\frac{1}{2}\left(\overline{\sigma}^2x^+-\underline{\sigma}^2x^-\right),\quad \underline{\sigma}^2=-\widehat{\mathbb{E}}\left[-B_{FG}^2(1)\right],\ \textit{and}\ \ \overline{\sigma}^2=\widehat{\mathbb{E}}\left[B_{FG}^2(1)\right], \tag{24}$$

and $v_H(x,t)$ is the viscosity solution of the following nonlinear initial-value problem:

$$\frac{\partial}{\partial t}v_H(x,t)=2Ht^{2H-1}G\left(\frac{\partial^2}{\partial x^2}v_H(x,t)\right),\quad v_H(x,0)=\varphi(x), \tag{25}$$

with the same $G(x)$ defined in (24)*.*

Being parallel to Proposition A.6, we also have

Proposition 5. *If the initial condition $v_H(x,0)$ in the nonlinear initial-value problem* (25) *takes as*

$$v_H(x,0)=\varphi(x)=I_{(0,\infty)}(x), \tag{26}$$

where $I_A(x)$ is the indicator function of a set A. Then the solution of the nonlinear initial-value problem (25) *can be expressed as*

$$v_H(x,t)=\int_{-\infty}^{x}\rho_H(y,t)\,\mathrm{d}y, \tag{27}$$

where $\rho(x,t)$ is a function on $\mathbb{R}\times\mathbb{R}^+$ defined by

$$\rho_H(y,t)=\frac{\sqrt{2}}{\left(\underline{\sigma}+\overline{\sigma}\right)\sqrt{\pi t^{2H}}}\left[\exp\left(-\frac{y^2}{2\overline{\sigma}^2t^{2H}}\right)I_{(-\infty,0]}(y)+\exp\left(-\frac{y^2}{2\underline{\sigma}^2t^{2H}}\right)I_{(0,\infty)}(y)\right]. \tag{28}$$

Proof. This proposition can be proved in a similar way as proving Proposition A.6 in Peng et al. (2023).

Inspired by (A.20) in Proposition A.6 and (28) in Proposition 5, we give out the following definition, being similar to Definition 6. The reader is reminded that this definition is not the same as the G-normal distribution as in Definition A.6. We just take the similar notations in order to be in accordance with the terminologies in Definition 6. For some fixed $\theta = \left[\underline{\sigma}^2, \overline{\sigma}^2\right]$ and H, we define

Definition 7. *We say that X is a G-normal random variable with parameters $\left(\underline{\sigma}, \overline{\sigma}\right)$, if its pdf (G-normal pdf) is given by*

$$f_\theta(x) := f(x) = \frac{\sqrt{2}}{\left(\underline{\sigma} + \overline{\sigma}\right)\sqrt{\pi}}\left[\exp\left(-\frac{(x-\mu)^2}{2\overline{\sigma}^2}\right) I_{(-\infty,\mu]}(x) + \exp\left(-\frac{(x-\mu)^2}{2\underline{\sigma}^2}\right) I_{(\mu,\infty)}(x)\right], \quad x \in \mathbb{R}, \tag{29}$$

and we denote $X \sim \mathcal{GN}\left(\{\mu\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$. We say that X is an H-G-normal random variable, or simply that X is H-G-normally distributed with parameters $\left(\underline{\sigma}, \overline{\sigma}, H\right)$ $(0 < H < 1)$, if its pdf (H-G-normal pdf) is given by

$$f_{\theta,H}(x) := f(x) = \frac{\sqrt{2}}{\left(\underline{\sigma}^{2H} + \overline{\sigma}^{2H}\right)\sqrt{\pi}}\left[\exp\left(-\frac{(x-\mu)^2}{2\overline{\sigma}^{4H}}\right) I_{(-\infty,\mu]}(x) + \exp\left(-\frac{(x-\mu)^2}{2\underline{\sigma}^{4H}}\right) I_{(\mu,\infty)}(x)\right], x \in \mathbb{R}, \tag{30}$$

and we also denote $X \sim \mathcal{GN}_H\left(\{\mu\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$. Clearly, $X \sim \mathcal{GN}_H\left(\{\mu\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$ is equivalent to $X \sim \mathcal{GN}\left(\{\mu\}, \left[\underline{\sigma}^{4H}, \overline{\sigma}^{4H}\right]\right)$.

According to Theorem 4 and Proposition 5, we have the following proposition.

Proposition 6. *Let $B_{FG}(t)_{t\in\mathbb{R}^+}$ be one-dimensional fGBm with Hurst parameter $H \in (0,1)$ on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$. Then for any $\varphi(x) \in C_{l,Lip}(\mathbb{R})$ and $t, s \geq 0$, there holds that*

$$\widehat{\mathbb{E}}\left[\varphi\left(B_{FG}(t+s) - B_{FG}(t)\right)\right] = \widehat{\mathbb{E}}\left[\varphi(\zeta)\right], \tag{31}$$

where

$$\widehat{\mathbb{E}}\left[\varphi(\zeta)\right] = \sup_{\theta\in\Theta} \mathbb{E}_\theta\left[\varphi(\zeta_\theta)\right] = \sup_{\theta\in\Theta} \int_{\mathbb{R}} \varphi(x) f_{s\theta,H}(x)\mathrm{d}x. \tag{32}$$

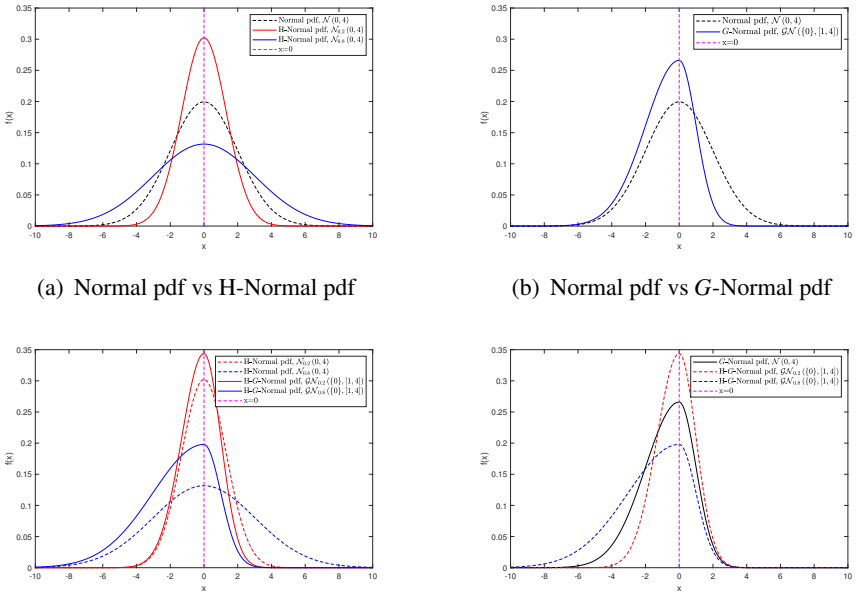
and $\zeta_\theta \sim \mathcal{GN}_H\left(\{0\}, \left[\underline{\sigma}^2 s, \overline{\sigma}^2 s\right]\right)$.

3. Empirical studies of the H-G-Normal distribution

Based on the Proposition 6 in the previous section, we can find that there is indeed a deep relationship between the fGBm and H-G-Normal distribution. So in order to make preparations for the applications of fGBm in mathematical finance, we first make some empirical studies of the H-G-normal distribution in the financial market.

First by (30), we can draw the corresponding plots for some fixed $\theta = \left[\underline{\sigma}^2, \overline{\sigma}^2\right]$ and H. Figure 1 shows some comparisons among the classical normal pdf $\mathcal{N}(0,4)$, H-Normal pdf $\mathcal{N}_H(0,4)$ $(H = 0.2, 0.8)$, G-Normal pdf $\mathcal{GN}(\{0\},[1,4])$ and H-G-Normal pdf $\mathcal{GN}_H(\{0\},[1,4])$ $(H = 0.2, 0.8)$. As we

can see, the classical normal pdf and H-Normal pdf are symmetric about the central axis $x = 0$, while the G-Normal pdf and H-G-Normal pdf are asymmetric. Furthermore, when $\underline{\sigma} < \overline{\sigma}$, they are both left-skewed, and they are right-skewed if $\underline{\sigma} > \overline{\sigma}$. On the other hand, the value of H in the H-Normal pdf and H-G-Normal pdf determines the steepness of the curve. The larger the value of H, the flatter the curve and the higher the degree of dispersion, which stands for a platykurtic distribution. And the smaller the value of H, the steeper the curve and the lower the degree of dispersion, which stands for a leptokurtic distribution compared to a mesokurtic distribution. Thus, compared to the normal pdf, H-Normal pdf, and G-Normal pdf, the H-G-Normal pdf can characterize not only skewness but also excess kurtosis simultaneously, which makes it a better alternative in describing the dynamics of some assets in the financial markets.



(a) Normal pdf vs H-Normal pdf

(b) Normal pdf vs G-Normal pdf

(c) H-Normal pdf vs H-G-Normal pdf

(d) G-Normal pdf vs H-G-Normal pdf

Figure 1. Some comparisons among the classical normal pdf, H-Normal pdf, G-Normal pdf and H-G-Normal pdf.

In reverse, for the dynamics of some assets in the financial markets, we can use the H-G-normal distribution to fit the distribution of evolution laws. As it is well known that the asset returns distributions observed in financial markets do not follow the Gaussian law, but other laws (Fama, 1965). Since the H-G-normal distribution (30) has an advantage in characterizing skewness and excess kurtosis simultaneously, we will make some empirical studies to show that the H-G-normal distribution is indeed a better alternative for describing these dynamics. For convenience, we choose the assets as some well-known indices, which are from the Oxford-Man Institute of Quantitative Finance Realized Library*, since some basic data for these indices have been collected and precomputed. Here we focus only on the S&P500 index, NASDAQ index, DAX, and HANG SENG index for example, where the closing price and the precomputed 5-minute realized variance estimates for the days from January 2, 2004 to December 31,

*http://realized.oxford-man.ox.ac.uk/data/download. The Oxford-Man Institute's Realized Library contains a selection of daily non-parametric estimates of volatility of financial assets, including realized variance (rv) and realized kernel (rk) estimates.

2021, were taken into consideration, unless specified otherwise. For each index, we mainly consider the distributions of the asset returns and the distributions of increments of the indices' log-volatility.

For the first empirical study, we apply the H-G-Normal distribution (30) to fit the distributions of the asset returns. At first, for each index, the daily returns can be computed, and some empirical distribution estimates $\widetilde{f}(x)$ can be established with the smoothing techniques (Bowman and Azzalini, 1997). Here we used the command *ksdensity* in the software *MATLAB* to obtain these priori distributions for simplicity. Second, in order to apply the H-G-Normal distribution (30) $f_{\theta,H}(x) = f(x) = f(x; \underline{\sigma}, \overline{\sigma}, H)$, where we rewrite this notation to emphasize the dependence on the parameters $\underline{\sigma}, \overline{\sigma}$ and H and denote $\Sigma = \left(\underline{\sigma}, \overline{\sigma}, H\right)$, then the three parameters Σ need to be identified. Here we are mainly going to use the Levenberg-Marquardt (LM) iterative algorithm (Madsen et al., 2004; Sun and Yuan, 2005; Fu et al., 2019) to give the estimates $\widehat{\Sigma} = \left(\widehat{\underline{\sigma}}, \widehat{\overline{\sigma}}, \widehat{H}\right)^T$, and the corresponding procedure based on the LM method is summarized in Algorithm 1. Sometimes for simplicity, we can first estimate the parameter H by the so-called rescaled analysis (R/S) method (Mandelbrot, 1972), and estimate the other two left parameters $\underline{\sigma}$ and $\overline{\sigma}$ by the same LM iterative algorithm, which is summarized in Algorithm 2. This simpler algorithm is basically the same as Algorithm 1, and the only difference is the first few steps.

As we know, there are many other alternative models that can be used to capture the distributions of the assets, such as the Cauchy distribution, Student t-distribution, and other stable distributions (Fallahgoul et al., 2017; Nolan, 2020). However, the Student t-distribution is always symmetric, but the skewness for most assets is usually negative, as seen in Table 1. So, the Student t-distribution is not considered to describe the left-skewed data. And thus, we just make some simple comparisons with the common Cauchy distribution. Although the Cauchy distribution is also symmetric, it is more simple and is a member of the Student t-family of distributions with one degree of freedom. The pdf of the Cauchy distribution is given by

$$f(x; a, b) = \frac{b}{\pi\left[(x-a)^2 + b^2\right]}, \tag{33}$$

where real scalar a is the statistical median and positive real b is the half width at half maximum.

In order to assess the quality of the similarity for two series, $f(x_i)$ and $\widetilde{f}(x_i)(i = 1, 2, \cdots, N)$, we mainly calculate some loss functions, which are root mean square error (RMSE), mean absolute error (MAE), and Theil's U statistic (TIC). These loss functions are defined respectively as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[f(x_i) - \widetilde{f}(x_i)\right]^2},\ MAE = \frac{1}{N}\sum_{i=1}^{N}\left|f(x_i) - \widetilde{f}(x_i)\right|,\ TIC = \frac{\sqrt{\sum_{i=1}^{N}\left[f(x_i) - \widetilde{f}(x_i)\right]^2}}{\sqrt{\sum_{i=1}^{N} f^2(x_i)} + \sqrt{\sum_{i=1}^{N} \widetilde{f}^2(x_i)}}. \tag{34}$$

Table 1 shows some descriptive statistics for the daily returns of the S&P500 index, NASDAQ index, DAX, and HANG SENG index, including the h and p-value of the Jarque-Bera normality test. As we can see, the returned values of $h = 1$ and the returned p-values are all below the default significance level of 5% for the four indices, which shows they do not follow the usual normal distributions. From the values of skewness and kurtosis, it is easy to find that they are all left-skewed and leptokurtic distributed. Moreover, we find that all the values of the Hurst parameter H are greater than 0.5, which indicates the long-range dependence. Table 2 lists the parameter estimates for the normal distribution, H-G-normal distribution, and Cauchy distribution, which are used to fit the empirical distributions of the daily mean-centralized returns for the above four indices. The parameter estimates for the

Algorithm 1 Parameter estimates of H-G-Normal pdf by LM method

Input:

Consider the observed priori pdf $\widetilde{f}(x)$, for some $x = \{x_i\}, i = 1, 2, \cdots, N$.

Chose small positive numbers $\varepsilon_1, \varepsilon_2, k_{\max}, \epsilon$ and τ.

Set the initial $\Sigma_0 = \left(\underline{\sigma}_0, \overline{\sigma}_0, H_0\right)^T$.

Output:

$k := 0$; $\nu := 2$; $\Sigma := \Sigma_0$;

Use (30) to compute $f(x; \Sigma)$; $\mathbf{y}(\Sigma) := f(x; \Sigma) - \widetilde{f}(x)$;

Compute

$$\mathbf{J}(\Sigma) := \begin{pmatrix} \frac{\partial f(x_1;\Sigma)}{\partial \underline{\sigma}}, \frac{\partial f(x_1;\Sigma)}{\partial \overline{\sigma}}, \frac{\partial f(x_1;\Sigma)}{\partial H} \\ \frac{\partial f(x_2;\Sigma)}{\partial \underline{\sigma}}, \frac{\partial f(x_2;\Sigma)}{\partial \overline{\sigma}}, \frac{\partial f(x_2;\Sigma)}{\partial H} \\ \vdots \quad \vdots \quad \vdots \\ \frac{\partial f(x_N;\Sigma)}{\partial \underline{\sigma}}, \frac{\partial f(x_N;\Sigma)}{\partial \overline{\sigma}}, \frac{\partial f(x_N;\Sigma)}{\partial H} \end{pmatrix};$$

$\mathbf{A} = \{a_{ij}\} := \mathbf{J}(\Sigma)^T * \mathbf{J}(\Sigma)$; $\mathbf{g} := \mathbf{J}(\Sigma)^T * \mathbf{y}(\Sigma)$;

found $:= (\|g\|_\infty) \le \varepsilon_1$; $\mu := \tau * \max\{a_{ii}\}$;

while not (*found* **and** $(k < k_{\max}$ **do**

  $k := k + 1$; solve $(\mathbf{A} + \mu \mathbf{I} h_{lm} = -\mathbf{g}$;

  if $\|h_{lm}\| \le \varepsilon_2(\|\Sigma\| + \varepsilon_2)$ **then**

    found := **true**

  else

    $\Sigma_{\text{new}} := \Sigma + h_{lm}$;

    $F(\Sigma) := \frac{1}{2}\|\mathbf{y}(\Sigma)\|^2$; $F(\Sigma_{\text{new}} := \frac{1}{2}\|f(x; \Sigma_{\text{new}} - \widetilde{f}(x)\|^2$;

    $\varrho := (F(\Sigma) - F(\Sigma_{\text{new}})/(-h_{lm}^T \mathbf{J}(\Sigma)^T \mathbf{y}(\Sigma) - \frac{1}{2} h_{lm}^T \mathbf{J}(\Sigma)^T \mathbf{J}(\Sigma) h_{lm}$;

    if $\varrho > 0$ **then**

      $\Sigma := \Sigma_{\text{new}}$;

      $\mathbf{A} = \{a_{ij}\} := \mathbf{J}(\Sigma)^T * \mathbf{J}(\Sigma)$;

      $\mathbf{g} := \mathbf{J}(\Sigma)^T * \mathbf{y}(\Sigma)$;

      found $:= (\|g\|_\infty) \le \varepsilon_1$;

      $\mu := \mu * \max\left\{\frac{1}{3}, 1 - (2\varrho - 1)^3\right\}$; $\nu := 2$;

    else

      $\mu := \mu * \nu$; $\nu := 2 * \nu$

    end if

  end if

end while

Algorithm 2 Parameter estimates of $\left(\underline{\sigma}, \overline{\sigma}^T\right)$ by LM method

Input:

1: Consider the observed priori pdf $\widetilde{f}(x)$, for some $x = \{x_i\}, i = 1, 2, \cdots, N$.

2: Chose small positive numbers $\varepsilon_1, \varepsilon_2, k_{\max}, \epsilon$ and τ.

3: Compute the value of H by the so-called rescaled analysis (R/S) method (Mandelbrot, 1972; Lo, 1991); Set the initial $\widehat{\sigma}_0 = \left(\underline{\sigma}_0, \overline{\sigma}_0\right)^T$.

Output:

4: $k := 0$; $\nu := 2$; $\sigma := \widehat{\sigma}_0$;

5: Use (30) to compute $f(x; \sigma, H)$; $\mathbf{y}(\sigma) := f(x; \sigma, H) - \widetilde{f}(x)$;

6: Compute

$$\mathbf{J}(\sigma) := \begin{pmatrix} \frac{\partial f(x_1;\sigma,H)}{\partial \underline{\sigma}}, \frac{\partial f(x_1;\sigma,H)}{\partial \overline{\sigma}} \\ \frac{\partial f(x_2;\sigma,H)}{\partial \underline{\sigma}}, \frac{\partial f(x_2;\sigma,H)}{\partial \overline{\sigma}} \\ \vdots \quad\quad \vdots \\ \frac{\partial f(x_N;\sigma,H)}{\partial \underline{\sigma}}, \frac{\partial f(x_N;\sigma,H)}{\partial \overline{\sigma}} \end{pmatrix};$$

7: $\cdots$

H-normal distribution and G-normal distribution are not given out since $N_H\left(\mu, \sigma^2\right)$ is equivalent to $N\left(\mu, \sigma^{4H}\right)$ and $\mathcal{GN}\left(\{\mu\}, \left[\underline{\sigma}^{4H}, \overline{\sigma}^{4H}\right]\right)$ is equivalent to $\mathcal{GN}_H\left(\{\mu\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$. When applying Algorithm 2 to estimate the parameters of H-G-Normal pdf, we set $\varepsilon_1 = 10^{-16}, \varepsilon_2 = 10^{-16}, k_{\max} = 200, \epsilon = 10^{-3}$ and $\tau = 10^{-16}$. Figure 2 also shows the corresponding pdfs for the daily returns of the four indices. The corresponding loss functions are also listed in Table 2. As we can see, the H-G-normal distributions are obviously superior to the usual normal distributions for all indices, and they also fit better than the Cauchy distributions except for the S&P500 index. Thus, it is reasonable to believe that the H-G-normal distributions have much better behavior in describing the characteristics of the daily returns than the usual normal distributions and Cauchy distributions, especially about the skewness and excess kurtosis.

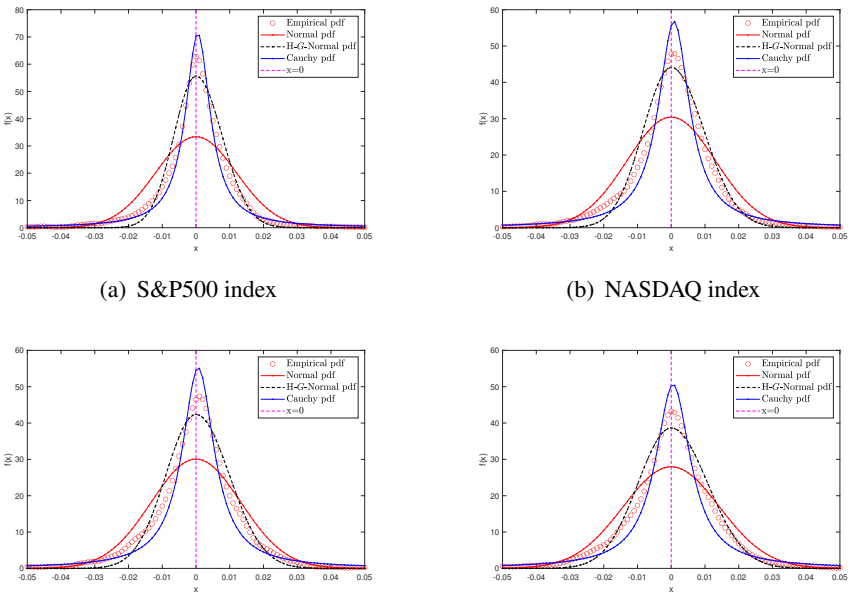
Table 1. Descriptive statistics for the daily returns of the S&P500 index, NASDAQ index, DAX, and HANG SENG index.

Index	Mean	Std.Dev.	Skewness	Kurtosis	$[h, p]$	H(R/S)
S&P500	0.000323	0.011947	−0.547380	17.188369	[1, 1e-3]	0.560995
NASDAQ	0.000454	0.013106	−0.478758	11.926907	[1, 1e-3]	0.520510
DAX	0.000301	0.013259	−0.149373	11.781643	[1, 1e-3]	0.567498
HANG SENG	0.000136	0.014262	−0.052401	11.995654	[1, 1e-3]	0.546451

For the second empirical study, we consider the distributions of increments of the indices' log-volatility, similar to the work done by Gatheral et al. (2018). We also proxy daily spot variances by the precomputed 5-minute realized variance estimates from the Oxford-Man dataset, and set the time grid with mesh

Table 2. Parameter estimates for kinds of distributions and corresponding loss functions.

Index	Loss functions Parameter	Normal σ	H-G-normal $\left(\underline{\sigma}, \overline{\sigma}, H\right)$	Cauchy (a, b)
S&P500	Estimates	0.011947	(0.011367, 0.013145, 0.560995)	(0.000537, 0.004462)
	RMSE	7.274866	3.312260	**2.647351**
	MAE	4.086505	2.171631	**1.647506**
	TIC	0.214513	0.086336	**0.070758**
NASDAQ	Estimates	0.013106	(0.010086, 0.011677, 0.520510)	(0.000811, 0.005603)
	RMSE	5.143355	**2.564910**	2.803242
	MAE	3.171417	**1.866597**	1.870834
	TIC	0.164228	**0.074787**	0.083717
DAX	Estimates	0.013259	(0.015662, 0.017130, 0.567498)	(0.000654, 0.005763)
	RMSE	4.820034	**2.586372**	2.688916
	MAE	2.901935	1.888938	**1.881877**
	TIC	0.155589	**0.076770**	0.081491
HANG SENG	Estimates	0.014262	(0.014569, 0.015874, 0.546451)	(0.000541, 0.006289)
	RMSE	4.399718	**2.323598**	2.685732
	MAE	2.738483	**1.694848**	1.871101
	TIC	0.147835	**0.072120**	0.085018



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

Figure 2. Some fitted pdfs for the daily returns of the S&P500 index, NASDAQ index, DAX, and HANG SENG index.

$\Delta = 1, 5, 25$, and 125 days. Table 3 lists the corresponding results of the descriptive statistics. As we can see, on one hand, being contrary to the skewnesses for the daily returns of these indices as listed in Table 1, all the values of skewness are now positive, which means that they are now right-skewed. On the other hand, they are still leptokurtic distributed since the values of kurtosis are still bigger than 3, although they are more platykurtic than the kurtosis of the daily returns. Similarly, the values of h and p from the Jarque–Bera normality test also indicate that the increments of these indices' log-volatility do not follow the usual normal distributions. Moreover, it is interesting to find that the values of the Hurst parameter H increase with the time grid with mesh Δ, and they are smaller than 0.5 as $\Delta = 1$ and 5 days and bigger than 0.5 in the other two cases. Table 4 shows the parameter estimates for the increments of S&P500 log-volatility with the LM Algorithm 2 and the corresponding loss functions. As we can see, the H-G-normal distributions are better than the usual normal distributions and Cauchy distributions, for all kinds of the time grid Δ. The fitted pdfs with the empirical distributions, usual normal distributions, H-G-normal distributions and Cauchy distributions for various lags Δ of the increments of S&P500 log-volatility are also shown in Figure 3. The parameter estimates and fitted pdfs for the increments of the other NASDAQ index, DAX, and HANG SENG index are also presented in Tables B.1–B.3 and Figures B.1–B.3 in Appendix B.

From the above two empirical results, we find that the H-G-normal distributions have much better behavior in describing the characteristics of the assets in the financial markets, whether for the daily returns or the increments of the indices' log-volatility, especially at skewness and excess kurtosis. On the other hand, the H-G-normal distributions have another three important parameters as $\left(\underline{\sigma}, \overline{\sigma}, H\right)$, which can feature the volatility uncertainty and exhibit long-range dependence $\left(\frac{1}{2} < H < 1\right)$ or antipersistence $\left(0 < H < \frac{1}{2}\right)$. This will make the associated fGBm be a better alternative stochastic process in the financial applications.

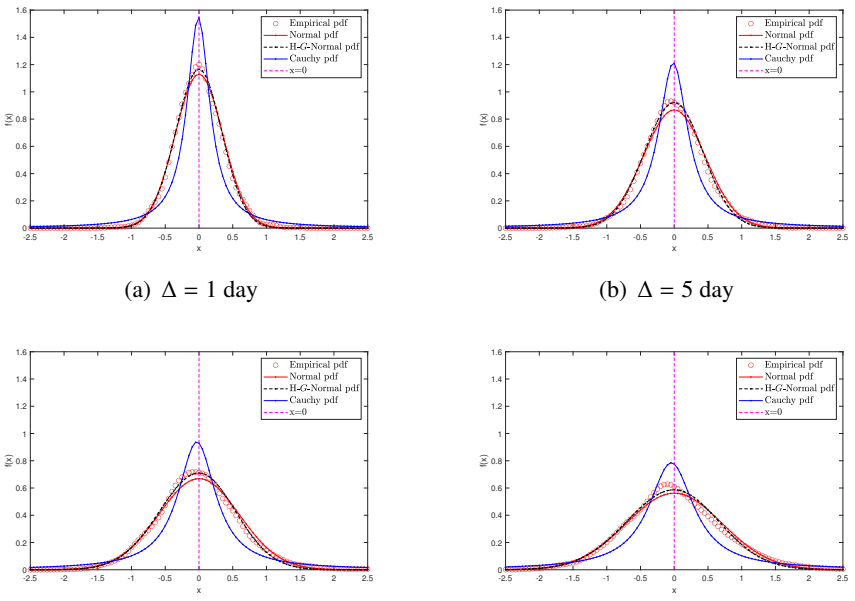
4. Empirical estimates of smoothness

In the previous section, we mainly reported that the H-G-normal distributions can be a better alternative in describing the characteristics of the daily returns and increments of the log-volatility for the four selected assets in the financial markets, where the parameter H was mainly estimated by the R/S method for simplicity. Actually, it can be estimated in many other kinds of way. One of the main results obtained by Gatheral et al. (2018) was that the log-volatility behaves essentially as a fractional Brownian motion with Hurst parameter $H < 1/2$ at any reasonable time scale. We expect that some similar results can also be obtained in the sublinear expectation frame, and this will make some preparations for the proposal of the RFSV-fGBm model in the next section. Thus, we continue to investigate the estimates of the smoothness of the volatility process for the referred assets in this section.

First, for convenience, we denote both the daily price process S_t and the volatility process σ_t as y_t, and suppose that the discrete observations of these processes have been collected, which are still the closing price and precomputed 5-minute realized variance estimates for the days from January 2, 2004, to December 31, 2021, from the Oxford-Man dataset. For a time grid with mesh Δ on $[0, T]$, the data are denoted as $y_0, y_\Delta, \cdots, y_{k\Delta}, \cdots, y_{N\Delta}$, where $k = 0, 1, \cdots, N$ and $N = \lfloor T/\Delta \rfloor$. Now for any $q > 0$, we

Table 3. Descriptive statistics for the increments of log-volatility of the S&P500 index, NASDAQ index, DAX, and HANG SENG index with mesh $\Delta = 1, 5, 25$ and 125 days.

Index	Δ(day)	Mean	Std.Dev.	Skewness	Kurtosis	$[h, p]$	H(R/S)
S&P500	1	-3.236e-05	0.353021	0.108505	3.558133	[1, 1e-3]	0.252654
	5	-1.897e-04	0.459691	0.156357	3.914937	[1, 1e-3]	0.353724
	25	9.521e-04	0.596537	0.392137	3.976984	[1, 1e-3]	0.517795
	125	-5.955e-03	0.710198	0.252874	3.443637	[1, 1e-3]	0.799864
NASDAQ	1	-1.308E-06	0.308525	0.174817	3.708583	[1, 1e-3]	0.254573
	5	-3.984e-05	0.411032	0.169375	3.931580	[1, 1e-3]	0.345253
	25	1.073e-03	0.524908	0.404608	4.202368	[1, 1e-3]	0.511626
	125	-4.675e-03	0.617956	0.175892	3.586046	[1, 1e-3]	0.772418
DAX	1	-2.098e-04	0.310102	0.168699	3.850313	[1, 1e-3]	0.243269
	5	-7.721e-04	0.377672	0.232704	3.882425	[1, 1e-3]	0.367160
	25	-1.241e-03	0.479206	0.475160	4.582901	[1, 1e-3]	0.537616
	125	-1.272e-02	0.591943	0.346047	3.678122	[1, 1e-3]	0.793942
HANG SENG	1	-2.365e-05	0.300528	0.265168	4.393874	[1, 1e-3]	0.247352
	5	-3.326e-04	0.346797	0.175259	4.530594	[1, 1e-3]	0.379670
	25	-3.730e-04	0.405120	0.195580	4.196129	[1, 1e-3]	0.561732
	125	-9.372e-05	0.486879	0.384526	4.109052	[1, 1e-3]	0.839362



(a) Δ = 1 day (b) Δ = 5 day

(c) Δ = 25 day (d) Δ = 125 day

Figure 3. Some pdfs with the empirical distributions, usual normal distributions, H-G-normal distributions, and Cauchy distributions for various lags Δ of the increments of S&P500 log-volatility.

Table 4. Parameter estimates for kinds of distributions for the increments of S&P500 log-volatility and corresponding loss functions.

Index	Loss functions Parameter	Normal σ	H-G-normal $\left(\underline{\sigma}, \overline{\sigma}, H\right)$	Cauchy (a, b)
$\Delta = 1$	Estimates	0.353021	(0.122614, 0.116936, 0.252654)	(-0.008826, 0.205824)
	RMSE	0.019295	**0.015688**	0.107565
	MAE	0.010280	**0.008912**	0.065264
	TIC	0.024167	**0.019497**	0.135834
$\Delta = 5$	Estimates	0.459691	(0.313987, 0.298544, 0.353724)	(-0.014049, 0.262424)
	RMSE	0.024313	**0.015924**	0.092635
	MAE	0.014337	**0.011148**	0.059728
	TIC	0.034580	**0.022311**	0.132155
$\Delta = 25$	Estimates	0.596537	(0.593985, 0.552339, 0.517795)	(-0.032164, 0.339228)
	RMSE	0.025381	**0.018028**	0.079362
	MAE	0.016381	**0.012853**	0.056306
	TIC	0.041112	**0.028769**	0.128834
$\Delta = 125$	Estimates	0.710198	(0.800988, 0.770131, 0.799864)	(-0.043645, 0.405512)
	RMSE	0.024739	**0.021496**	0.069734
	MAE	0.015460	**0.015914**	0.051980
	TIC	0.043867	**0.037702**	0.124073

define

$$m(q,\Delta)=\frac{1}{N}\sum_{k=1}^{N}\left|\log(y_{k\Delta}-\log(y_{(k-1)\Delta}\right|^{q}. \tag{35}$$

In order to compute $\widehat{\mathbb{E}}[\varphi(X)]$ numerically on a sublinear expectation space $\left(\Omega,\mathcal{H},\widehat{\mathbb{E}}\right)$ for some functions $\varphi(x)$ based on a sequence of independent identically distributed samples $\{X_i\}_{i=1}^{N}$, we mainly adopt the basic algorithm named "φ-max-mean" (Peng, 2017, 2019; Jin and Peng, 2021), where the moving window method is applied. We set the window width to be M, where $M\leq N$, and define the statistic

$$m_j(\varphi(x))=\frac{1}{M}\sum_{i=j}^{M+j-1}\varphi(x_i),\ \ j=1,\cdots,N-M+1, \tag{36}$$

and set

$$\underline{m}[\varphi]=\min_j m_j(\varphi(x)),\quad \overline{m}[\varphi]=\max_j m_j(\varphi(x)). \tag{37}$$

Then according to the results by Jin and Peng (2021), as N is sufficient enough, the estimators

$$\underline{\mathbb{T}}[\varphi](X)=\underline{m}[\varphi](X),\quad \overline{\mathbb{T}}[\varphi](X)=\overline{m}[\varphi](X), \tag{38}$$

provide the asymptotically unbiased estimators of the sublinear distribution of X, $\underline{\mu}_{\varphi(X)}=-\widehat{\mathbb{E}}[-\varphi(X)]$ and $\overline{\mu}_{\varphi(X)}=\widehat{\mathbb{E}}[\varphi(X)]$, respectively. That is also to say, $\underline{\mathbb{T}}[\varphi](X)$ is the smallest unbiased estimator for the lower mean $\underline{\mu}_{\varphi(X)}$ and $\overline{\mathbb{T}}[\varphi](X)$ is the largest unbiased estimator for the upper mean $\overline{\mu}_{\varphi(X)}$.

Now based on (36)-(38) and for the process y_t, we redefine

$$\underline{m}(q,\Delta)=\min_j m_j(q,\Delta),\quad \overline{m}(q,\Delta)=\max_j m_j(q,\Delta), \tag{39}$$

where

$$m_j(q,\Delta)=\frac{1}{M}\sum_{i=j}^{M+j-1}\left|\log(y_{j\Delta}-\log(y_{(j-1)\Delta}\right|^{q},\ \ j=1,\cdots,N-M+1. \tag{40}$$

Thus $\underline{m}(q,\Delta)$ and $\overline{m}(q,\Delta)$ can be seen as the empirical counterparts of

$$-\widehat{\mathbb{E}}\left[-\left|\log(y_\Delta-\log(y_0)\right|^{q}\right], \tag{41}$$

and

$$\widehat{\mathbb{E}}\left[\left|\log(y_\Delta-\log(y_0)\right|^{q}\right]. \tag{42}$$

As we can see, when the window width $M=N$ and N is sufficient enough, we have

$$\underline{m}(q,\Delta)=\overline{m}(q,\Delta)=m(q,\Delta), \tag{43}$$

and thus $m(q,\Delta)$ is the empirical counterpart of $\mathbb{E}\left[\left|\log(y_\Delta-\log(y_0)\right|^{q}\right]$ (Gatheral et al., 2018).

Similar to the assumption by Gatheral et al. (2018), we also assume that for some $\underline{s}_q>0,\underline{b}_q>0$ and $\overline{s}_q>0,\overline{b}_q>0$, as Δ tends to zero, there holds

$$M^{q\underline{s}_q}\underline{m}(q,\Delta)\to\underline{b}_q,\quad M^{q\overline{s}_q}\overline{m}(q,\Delta)\to\overline{b}_q. \tag{44}$$

Now we estimate the smoothness parameters $\underline{s}_q$ and $\overline{s}_q$ for each q by computing $\underline{m}(q,\Delta)$ and $\overline{m}(q,\Delta)$ for different values of Δ and regressing $\log \underline{m}(q,\Delta)$ and $\log \overline{m}(q,\Delta)$ against $\log \Delta$. For a given Δ, we fix the window width M and take the averages of several values of $\underline{m}(q,\Delta)$ and $\overline{m}(q,\Delta)$, since they are computed depending on the starting point.

For the daily price processes $y_t = S_t$ of the S&P500 index, NASDAQ index, DAX, and HANG SENG index, the plots of $\log \underline{m}(q,\Delta)$, $\log m(q,\Delta)$, and $\log \overline{m}(q,\Delta)$ vs $\log \Delta$ for different values of q with the fixed window width $M = N/2$ are displayed in Figure 4. For the volatility processes $y_t = \sigma_t$ of these four indices, the corresponding plots are also presented in Figure 5. As we can see, for all the indices and for a given q, the points essentially lie on a straight line, which implies that both the daily returns (log-price increments) and the log-volatility increments follow the scaling property in the nonlinear expectation:

$$-\widehat{\mathbb{E}}\left[-\left|\log(y_\Delta - \log(y_0)\right|^q\right] = \underline{b}_q \Delta^{\underline{\zeta}_q}, \tag{45}$$

and

$$\widehat{\mathbb{E}}\left[\left|\log(y_\Delta - \log(y_0)\right|^q\right] = \overline{b}_q \Delta^{\overline{\zeta}_q}, \tag{46}$$

where $\underline{\zeta}_q = q\underline{s}_q$ and $\overline{\zeta}_q = q\overline{s}_q$ are the slopes of the line associated to q. Furthermore, the smoothness parameters $\underline{s}_q$ and $\overline{s}_q$ do not seem to depend on q. Actually, for the daily price process $y_t = S_t$ with the fixed window width $M = N/2$, plotting $\underline{\zeta}_q$ and $\overline{\zeta}_q$ against q, we obtain that $\underline{\zeta}_q \sim q\underline{H}$ and $\overline{\zeta}_q \sim q\overline{H}$ with $\underline{H} = 0.353988, \overline{H} = 0.461693$ for the S&P500 index; $\underline{H} = 0.378514, \overline{H} = 0.484807$ for the NASDAQ index; $\underline{H} = 0.452087, \overline{H} = 0.475811$ for the DAX, and $\underline{H} = 0.432790, \overline{H} = 0.474951$ for the HANG SENG index, see Figure 6. Figure 7 also shows the corresponding values of $\underline{H}$ and $\overline{H}$ for the volatility process $y_t = \sigma_t$ of these four indices. Here we find that for both the daily price and volatility processes, when the window width was fixed with $M = N/2$, the values of $\overline{H}$ are all smaller than $1/2$. Indeed, for any window width $M \le N$, the results still hold true. Figure 8 and Figure 9 show the values of $\underline{H}$ and $\overline{H}$ under the different window widths for the daily price and volatility processes of the four indices. Here we can also find that when $M = N$, the values of $\underline{H}$ and $\overline{H}$ are equal to H, which is computed from (35) as the same result as obtained by Gatheral et al. (2018).

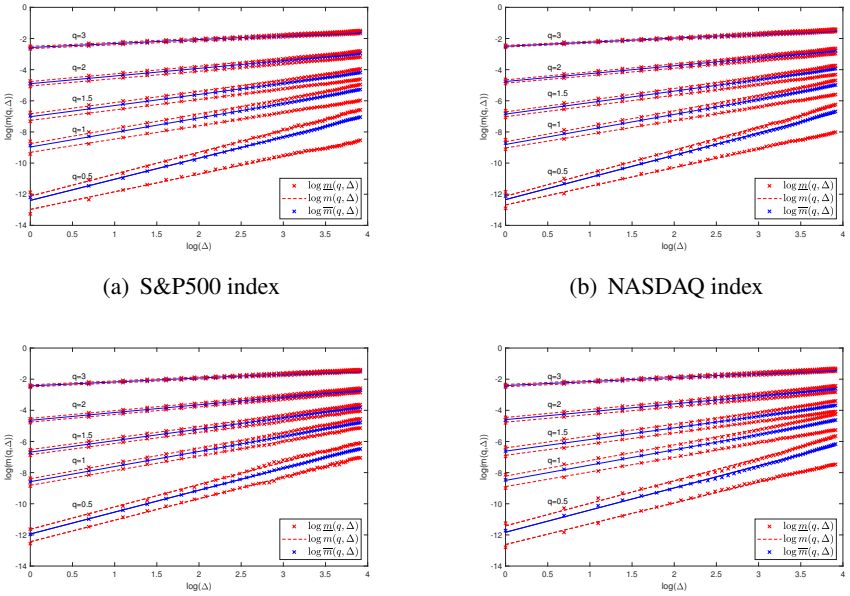
5. Some dynamic model compatible with the empirical studies

5.1. Specification of RFSV-fGBm model

In the previous sections, we have shown that, empirically, both the increments of the daily price (returns) and the increments of log-volatility of various assets enjoy some scaling properties with bounded and constant smoothness parameters and that their distributions are close to H-G-normal distributions. This naturally suggests the following dynamic model for the assets:

$$\begin{cases} \frac{S_{t+\Delta t} - S_t}{S_t} = \mu \Delta t + \sigma_t \left(B^1_{FG}(t+\Delta t) - B^1_{FG}(t)\right), \\ \log \sigma_{t+\Delta t} - \log \sigma_t = \nu \left(B^2_{FG}(t+\Delta t) - B^2_{FG}(t)\right), \\ \widehat{\mathbb{E}}\left[\Delta B^1_{FG}(t) \Delta B^2_{FG}(t)\right] = \overline{\rho}\,(\Delta t)^{2H}, \\ -\widehat{\mathbb{E}}\left[-\Delta B^1_{FG}(t) \Delta B^2_{FG}(t)\right] = \underline{\rho}\,(\Delta t)^{2H}, \end{cases} \tag{47}$$

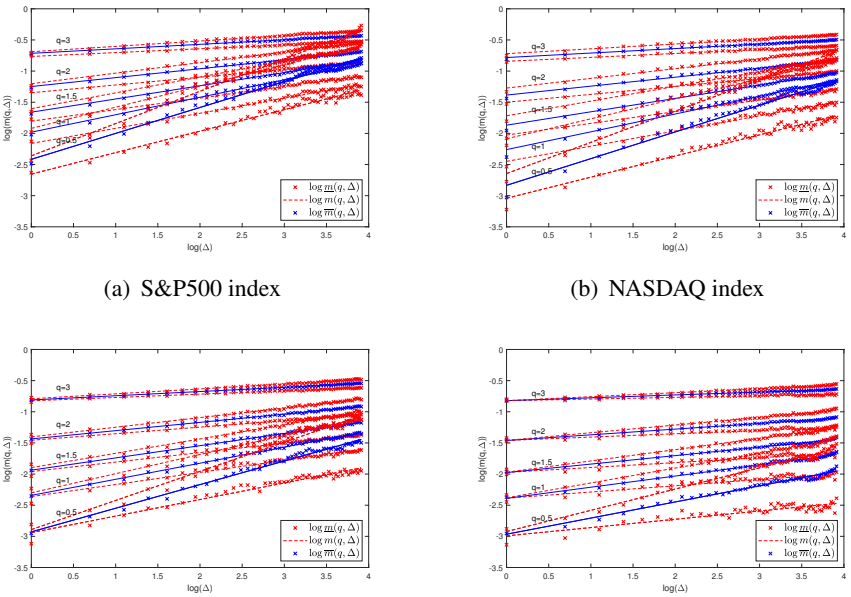
where $\Delta B^k_{FG}(t) = B^k_{FG}(t+\Delta t) - B^k FG(t)$, $k = 1, 2$, $B^1_{FG}(t)$ and $B^2_{FG}(t)$ are the proposed fGBm with a constant upper correlation coefficient $\overline{\rho}$ and constant lower correlation coefficient $\underline{\rho}$, and ν is a positive



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

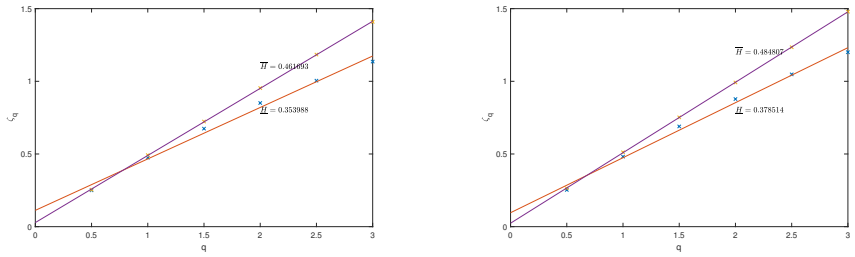
Figure 4. $\log \underline{m}(q,\Delta)$, $\log m(q,\Delta)$, and $\log \overline{m}(q,\Delta)$ vs $\log \Delta$ for the daily price process of S&P500 index, NASDAQ index, DAX, and HANG SENG index.



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

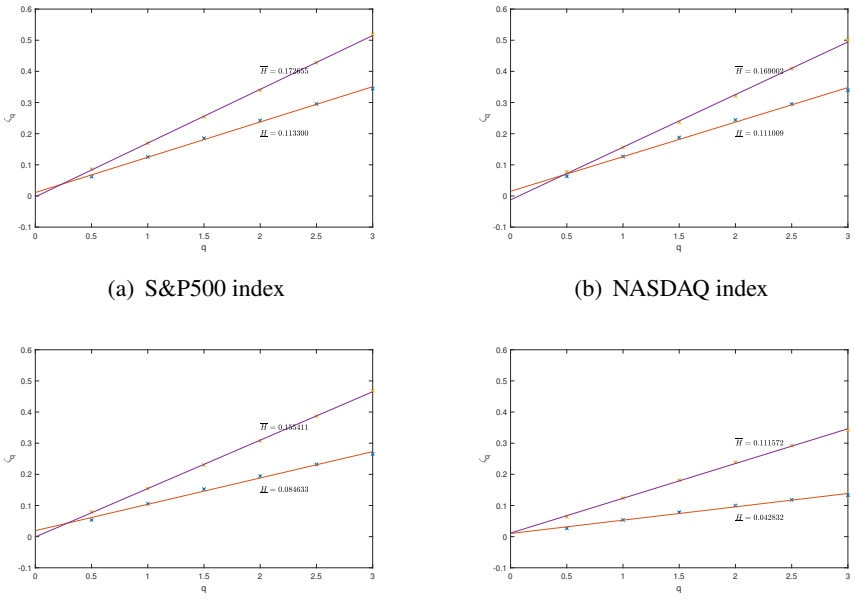
Figure 5. $\log \underline{m}(q,\Delta)$, $\log m(q,\Delta)$, and $\log \overline{m}(q,\Delta)$ vs $\log \Delta$ for the volatility process of S&P500 index, NASDAQ index, DAX, and HANG SENG index.



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

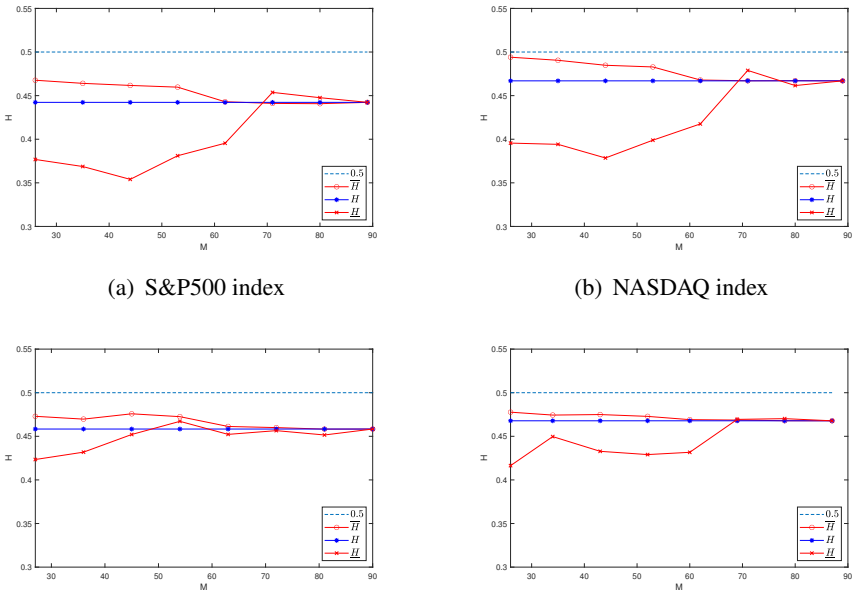
Figure 6. $\underline{\zeta}_q$ and $\overline{\zeta}_q$ against q for the daily price process of S&P500 index, NASDAQ index, DAX, and HANG SENG index.



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

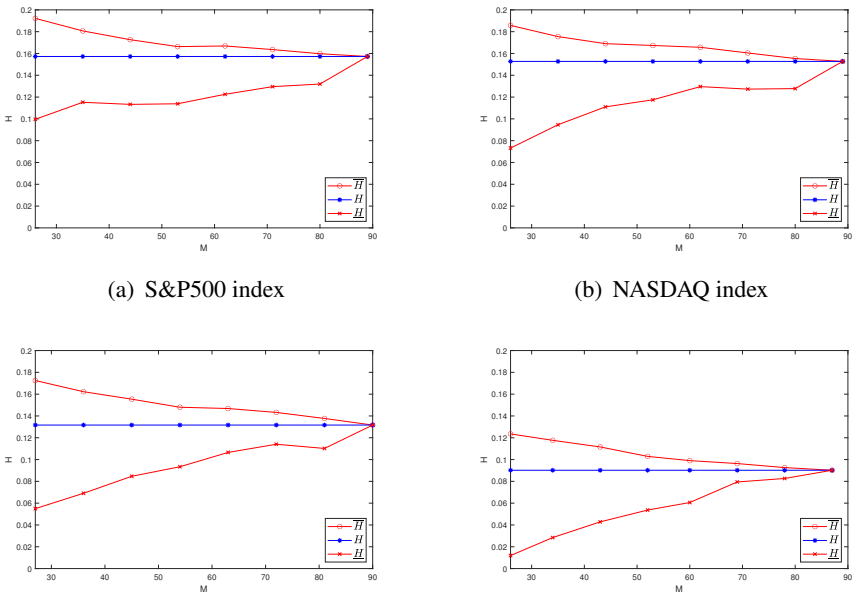
Figure 7. $\underline{\zeta}_q$ and $\overline{\zeta}_q$ against q for the volatility process of S&P500 index, NASDAQ index, DAX, and HANG SENG index.



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

Figure 8. The values of $\underline{H}$ and $\overline{H}$ under the different window widths for the daily price process of the S&P500 index, NASDAQ index, DAX, and HANG SENG index.



(a) S&P500 index (b) NASDAQ index

(c) DAX (d) HANG SENG index

Figure 9. The values of $\underline{H}$ and $\overline{H}$ under the different window widths for the volatility process of the S&P500 index, NASDAQ index, DAX, and HANG SENG index.

constant. By Remark 4, we rewrite (47) into the following continuous form as

$$\begin{cases} \mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma_t S_t \,\mathrm{d}B^1_{FG}(t), \\ \sigma_t = \sigma_0 \exp(\nu B^2_{FG}(t)), \\ \widehat{\mathbb{E}}\left[\mathrm{d}B^1_{FG}(t)\,\mathrm{d}B^2_{FG}(t)\right] = \overline{\rho}\,(\mathrm{d}t)^{2H}, \\ -\widehat{\mathbb{E}}\left[-\mathrm{d}B^1_{FG}(t)\,\mathrm{d}B^2_{FG}(t)\right] = \underline{\rho}\,(\mathrm{d}t)^{2H}, \end{cases} \tag{48}$$

where σ_0 is another positive constant. However, the log-volatility model in (48) is not stationary, while stationarity is desirable both for mathematical tractability and practical application. Thus, similar to the assumption supposed by Gatheral et al. (2018), we also impose the stationarity for the log-volatility model with the following fractional Ornstein–Uhlenbeck process (X_t) (Cheridito et al., 2003):

$$\begin{cases} \sigma_t = \exp(X_t), \\ \mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \mathrm{d}B^2_{FG}(t), \end{cases} \tag{49}$$

where κ and θ are some constants. One should notice that the model of the volatility process does not contain the volatility parameter ν any more. This is due to the characteristics of the fGBm, that the volatility component is endogenously contained in $B_{FG}(t)$. Replacing the log-volatility model in (48) with (49), we arrive at the final model for the dynamics of the assets, and we call it rough fractional stochastic volatility model driven by the fGBm (RFSV-fGBm):

$$\begin{cases} \mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma_t S_t \,\mathrm{d}B^1_{FG}(t), \\ \sigma_t = \exp(X_t), \\ \mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \mathrm{d}B^2_{FG}(t), \\ \widehat{\mathbb{E}}\left[\mathrm{d}B^1_{FG}(t)\,\mathrm{d}B^2_{FG}(t)\right] = \overline{\rho}\,(\mathrm{d}t)^{2H}, \\ -\widehat{\mathbb{E}}\left[-\mathrm{d}B^1_{FG}(t)\,\mathrm{d}B^2_{FG}(t)\right] = \underline{\rho}\,(\mathrm{d}t)^{2H}. \end{cases} \tag{50}$$

Here we need to note that, (50) are the commonly informal shorthand notations for some integral expressions, which involve the stochastic integrals with respect to fGBm. This can be understood in the sense of pathwise Riemann–Stieltjes integral for the purpose of having reasonable economic interpretation, although the stochastic calculus for the fGBm has not been studied thoroughly. Since in the present paper, we will not deal with the integrals with respect to the fGBm, we will leave this topic for another work. As we can see, when the sublinear expectation $\widehat{\mathbb{E}}[\cdot]$ becomes the classical linear expectation $\mathbb{E}[\cdot]$ and $\underline{\sigma}^2 = \overline{\sigma}^2 = \nu$, the fGBm $B_{FG}(t)$ reduces to the fBm, and thus the RFSV-fGBm model (50) will be equivalent to the RFSV model (3.3) and (3.4) in Gatheral et al. (2018):

$$\begin{cases} \sigma_t = \exp(X_t), \\ \mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \nu\,\mathrm{d}W^H_t, \end{cases} \tag{51}$$

where W^H_t is a fBm with the Hurst parameter H.

Remark 5. *Similar to the classical Heston model (Heston, 1993),*

$$\begin{cases} \mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sqrt{V_t} S_t \,\mathrm{d}B^1(t), \\ \mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}B^2(t), \\ \mathrm{d}B^1(t)\,\mathrm{d}B^2(t) = \rho \mathrm{d}t, \end{cases} \tag{52}$$

where $B^1(t)$ and $B^2(t)$ are the standard Bm, we can also generalize the following Heston-type model driven by the fGBm

$$\begin{cases} \mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sqrt{V_t}S_t\,\mathrm{d}B^1_{FG}(t), \\ \mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}B^2_{FG}(t), \\ \widehat{\mathbb{E}}\left[\mathrm{d}B^1_{FG}(t)\,\mathrm{d}B^2_{FG}(t)\right] = \overline{\rho}\,(\mathrm{d}t)^{2H} \\ -\widehat{\mathbb{E}}\left[-\mathrm{d}B^1_{FG}(t)\,\mathrm{d}B^2_{FG}(t)\right] = \underline{\rho}\,(\mathrm{d}t)^{2H}, \end{cases} \tag{53}$$

where $B^1_{FG}(t)$ and $B^2_{FG}(t)$ are the proposed fGBm, and we thus call (53) *the Heston–fGBm model.*

5.2. Parameter estimates of RFSV-fGBm model

In this subsection, we explain how to estimate the parameters of the RFSV-fGBm model (50) from a data set of asset prices. For simplicity, we assume $B^2_{FG}(t)$ is an independent copy of $B^1_{FG}(t)$ and $\overline{\rho} = \underline{\rho} = 0$. Then we have

$$\begin{cases} \mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma_t S_t\,\mathrm{d}B^1_{FG}(t), \\ \sigma_t = \exp(X_t), \\ \mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \mathrm{d}B^2_{FG}(t). \end{cases} \tag{54}$$

The next step is to discretize the RFSV-fGBm model (54). Here we mainly employ the Euler's discretization method for simplicity, and based on Lemma 1, the Euler's discretization of the RFSV-fGBm model (54) can be expressed as

$$\begin{cases} R_{t+\Delta t} = \mu\Delta t + \sigma_t\xi_1(t)(\Delta t)^{H_1}, \\ \sigma_t = \exp(X_t), \\ X_{t+\Delta t} = X_t + \kappa(\theta - X_t)\Delta t + \xi_2(t)(\Delta t)^{H_2}, \end{cases} \tag{55}$$

where $R_{t+\Delta t} = \frac{S_{t+\Delta t} - S_t}{S_t}$ denotes the asset returns, $\xi_i(t)(i = 1, 2)$ are two G-normally distributed random variables as $\xi_i(t) \sim \mathcal{N}\left(\{0\}, \left[\underline{\sigma_i}^2(t), \overline{\sigma_i}^2(t)\right]\right)$. Thus we need to estimate the parameters $\mu, \kappa, \theta, \underline{\sigma_i}^2, \overline{\sigma_i}^2$, and $H_i(i = 1, 2)$. As being supposed that the data for the daily prices and volatilities were observed and collected in the previous section, we still follow this assumption in this section for convenience. That is to say, the discrete data for S_t and X_t have been observed and collected.

Conventionally, there are two main methods for the parameter estimates, which are usually called the method of moments and maximum likelihood estimation. Here for brevity, we mainly adopt the method of moments. First for $\xi_i(t)(i = 1, 2)$, we have

$$-\widehat{\mathbb{E}}\left[-\xi_i(t)\right] = -\widehat{\mathbb{E}}\left[-\xi_i(t)|\,I_t\right] = 0, \quad \widehat{\mathbb{E}}\left[\xi_i(t)\right] = \widehat{\mathbb{E}}\left[\xi_i(t)|\,I_t\right] = 0, \tag{56}$$

and

$$-\widehat{\mathbb{E}}\left[-\xi_i^2(t)\Big|\,I_t\right] = \underline{\sigma_i}^2(t), \quad \widehat{\mathbb{E}}\left[\xi_i^2(t)\Big|\,I_t\right] = \overline{\sigma_i}^2(t), \tag{57}$$

where I_t denotes all the information up to the time t. Then by (55), we have

$$\mu = \widehat{\mathbb{E}}\left[\frac{R_{t+\Delta t}}{\Delta t}\right], \tag{58}$$

and

$$\widehat{\mathbb{E}}\left[X_{t+\Delta t} - (1 - \kappa\Delta t)\,X_t - \kappa\theta\Delta t\right] = 0. \tag{59}$$

Meanwhile, we also notice that

$$\widehat{\mathbb{E}}\left[X_t\xi_2(t)\right] = \widehat{\mathbb{E}}\left[\widehat{\mathbb{E}}\left[X_t\ \xi_2(t)|\ I_t\right]\right] = \widehat{\mathbb{E}}\left[X_t\widehat{\mathbb{E}}\left[\xi_2(t)|\ I_t\right]\right] = 0, \tag{60}$$

and thus

$$\widehat{\mathbb{E}}\left[X_t\left(X_{t+\Delta t} - (1-\kappa\Delta t)\,X_t - \kappa\theta\Delta t\right)\right] = 0. \tag{61}$$

Combining (59) (61) and employing the "φ-max-mean" method (Peng, 2017, 2019; Jin and Peng, 2021) again, we have the parameter estimates for κ and θ. Approximatively for daily data, there is

$$\begin{cases} \widehat{\kappa} = \dfrac{(N-1)\sum\limits_{t=1}^{N-1}\left(X_{t+1}X_t - X_t^2\right) - \sum\limits_{t=1}^{N-1}\left(X_{t+1}-X_t\right)\sum\limits_{t=1}^{N-1}X_t}{\left(\sum\limits_{t=1}^{N-1}X_t\right)^2 - (N-1)\sum\limits_{t=1}^{N-1}X_t^2}, \\ \widehat{\theta} = \dfrac{\sum\limits_{t=1}^{N-1}\left(X_{t+1}-X_t\right) + \widehat{\kappa}\sum\limits_{t=1}^{N-1}X_t}{(N-1)\widehat{\kappa}}, \end{cases} \tag{62}$$

where N is the total number of the observations.

Once the estimates for μ, κ, and θ are obtained, the remaining unestimated parameters $\underline{\sigma_i}^2, \overline{\sigma_i}^2$, and $H_i (i = 1, 2)$ can be easily obtained by the LM method, which can be seen in Algorithm 1 or Algorithm 2 in Section 3.

For the numerical experiment, we still take the S&P500 index, NASDAQ index, DAX and HANG SENG index for examples. As the prices and volatilities for these assets are observed daily, we set $\Delta t = 1/252$ in the numerical computations. By virtue of the *MATLAB* software, the parameter estimates for the four indices are shown in Table 5. As a comparison in contrast, the parameter estimates for the RFSV model (51) (Gatheral et al., 2018) are also present in Table 5, where the estimates are carried out by the method in Brouste and Iacus (2013). As we can see, the parameters μ, κ, and θ estimated by (58) and (62) are the same as the RFSV model (51) by Gatheral et al. (2018) There are some differences between the values of H_i, since they are affected by the variances $\underline{\sigma_i}^2$ and $\overline{\sigma_i}^2$ simultaneously. However, they are all bigger than 0.5 for the price processes but smaller than 0.5 for the volatility processes.

Here we must notice that, on one hand, the parameter estimates by the method of moments maybe not outperform the maximum likelihood estimation, since the method of moments is relatively simple and convenient, while maximum likelihood estimation is nuch more complicated, requesting the expression for the likehood function, as there is no explicit formulation associated with the fGBm. On the other hand, being analogous to the classical calibration problem (Bakshi et al., 1997; Wang et al., 2018), the parameter estimates of the RFSV-fGBm model (50) cannot be completely obtained by the empirical estimates from the historical prices $S(t)$, but need to combine the prices of other financial derivatives, such as the option. For the above two concerns, we leave them in our other research.

Remark 6. *If we take the model* (55) *as forecasting the log-volatility, the first thought that comes to our mind is adopting the Euler–Maruyama method. Once the parameters* $\left(\kappa, \theta, \nu, \underline{\sigma}^2, \overline{\sigma}^2, H\right)$ *estimated, some G-normally distributed random variables* ξ_t *can be generated by the Algorithm 1 in Guo et al. (2023a) based on the known* $\underline{\sigma}^2, \overline{\sigma}^2$*, and* H*. And one can apply*

$$X_{t+\Delta t} = X_t + \kappa\left(\theta - X_t\right)\Delta t + \xi_t\left(\Delta t\right)^H, \tag{63}$$

to complete the prediction. However, this discretization method may not be great, since the parameter estimates are not precise and ξ_t *is random. Meanwhile, there is a lack of results for the well-posedness. So some more accurate forecasting methods need to be investigated, and we will focus on this prediction in another work.*

Table 5. Parameter estimates of RFSV-fGBm model (50) and RFSV model (Gatheral et al., 2018) for the S&P500 index, NASDAQ index, DAX, and HANG SENG index.

Index	Parameter	RFSV-fGBm model (50)	RFSV model (Gatheral et al., 2018)
S&P500	μ	0.081437	0.081437
	$\left[\underline{\sigma_1}, \overline{\sigma_1}\right], H_1$	[0.011263, 0.013032], 0.559841	[1, 1], 0.560995
	(κ, θ, ν)	(0.186665, −2.268302, −−)	(0.186665, −2.268302, 1.307050)
	$\left[\underline{\sigma_2}, \overline{\sigma_2}\right], H_2$	[0.119144, 0.124922], 0.254980	[1, 1], 0.138179
NASDAQ	μ	0.114513	0.114513
	$\left[\underline{\sigma_1}, \overline{\sigma_1}\right], H_1$	[0.009990, 0.011575], 0.519465	[1, 1], 0.520510
	(κ, θ, ν)	(0.197913, −2.205740, −−)	(0.197913, −2.205983, 1.116781)
	$\left[\underline{\sigma_2}, \overline{\sigma_2}\right], H_2$	[0.090725, 0.096604], 0.258132	[1, 1], 0.204523
DAX	μ	0.075888	0.075888
	$\left[\underline{\sigma_1}, \overline{\sigma_1}\right], H_1$	[0.015449, 0.016900], 0.565586	[1, 1], 0.567498
	(κ, θ, ν)	(0.212701, −2.040554, −−)	(0.212701, −2.040554, 1.091078)
	$\left[\underline{\sigma_2}, \overline{\sigma_2}\right], H_2$	[0.083620, 0.091042], 0.253651	[1, 1], 0.132149
HANG SENG	μ	0.034337	0.034337
	$\left[\underline{\sigma_1}, \overline{\sigma_1}\right], H_1$	[0.014380, 0.015663], 0.544761	[1, 1], 0.546451
	(κ, θ, ν)	(0.266193, −2.131269, −−)	(0.266193, −2.131269, 0.974328)
	$\left[\underline{\sigma_2}, \overline{\sigma_2}\right], H_2$	[0.078129, 0.088226], 0.257342	[1, 1], 0.053634

6. Conclusions

In this paper, we carried out some empirical studies for the financial applications of fGBm, which generalizes the concepts of the standard Brownian motion, fBm, and GBm. Since the fGBm can exhibit long-range dependence and feature volatility uncertainty simultaneously, it is a better alternative to capture the intrinsic characteristics of the financial markets. In the present paper, as a starting point, the corresponding H-G-normal distributions associated with the fGBm are used to describe the dynamics of daily returns and increments of log-volatility for some selected assets, and the corresponding parameters of the H-G-normal distributions are estimated by the Levenberg–Marquardt iterative algorithm. The empirical studies show that the H-G-normal distributions are more suitable than the usual normal distributions, as they keep the properties of skewness and excess kurtosis. On the other hand, the H-G-normal distributions contains another three important parameters as $\left(\underline{\sigma}, \overline{\sigma}, H\right)$, which are used to feature the volatility uncertainty $\left(\underline{\sigma}, \overline{\sigma}\right)$ and long-range dependence $\left(\frac{1}{2} < H < 1\right)$ or antipersistence $\left(0 < H < \frac{1}{2}\right)$. By estimating the smoothness of the daily price and volatility processes with the "φ-max-mean" method under the sublinear expectation frame, we also found that both the daily return and log-volatility behave essentially as some fGBm with different $\underline{\sigma}^2$ and $\overline{\sigma}^2$, but Hurst parameters $H < \frac{1}{2}$, at any reasonable time scale. Based on these empirical studies, we developed another generalized model (50) called rough fractional stochastic volatility model driven by the fGBm (RFSV-fGBm) to describe the dynamics of the assets. Finally, some parameter estimates for the RFSV-fGBm model were carried out by the method of moments, and the other three parameters, $\underline{\sigma}^2, \overline{\sigma}^2$, and H were estimated by the LM algorithm for simplicity.

However, we have not carried out the elaborate theoretical research for the financial applications of the proposed fGBm in the present paper. Nevertheless, we believe that the fGBm is indeed a good candidate to model the financial markets from the various empirical evidences, and there are great prospects for the applications of fGBm in the field of mathematical finance. For the theoretical studies, there are many topics that need to be concerned, such as various stochastic calculus with respect to fGBm, the arbitrage problems driven by fGBm, and the corresponding derivative pricing mechanisms. While for the RFSV-fGBm model, it can be further used for financial applications, such as in the volatility modeling, derivative pricing, and risk management, since it is involved with the evolutions of the dynamics for volatility. For some calculation for Value-at-Risk (VaR) under the framework of nonlinear expectation, we refer readers to Peng et al. (2023) and references therein. For the issues referred to above, we leave them for our further works.

Appendices

Appendix A. Sublinear expectation and GBm

Let Ω be a given set and $\mathcal{H}$ be a vector lattice of real-valued functions defined on Ω containing 1, namely, $\mathcal{H}$ is a linear space such that $1 \in \mathcal{H}$ and that $X \in \mathcal{H}$ implies $|X| \in \mathcal{H}$. The space $\mathcal{H}$ is also called a space of random variables. We also denote $C_{l,Lip}(\mathbb{R}^n)(n \in \mathbb{N})$ as the linear space of functions $\varphi(x)$ satisfying the following local Lipschitz condition:

$$|\varphi(x) - \varphi(y)| \leq C\left(1 + |x|^m + |y|^m\right)|x - y|, \text{ for } x, y \in \mathbb{R}^n, \tag{A.1}$$

where the constant $C > 0$ and the integer $m \in \mathbb{N}$ depends on φ.

Definition A.1. *(Sublinear expectation) (Peng, 2019) A sublinear expectation $\widehat{\mathbb{E}}$ is a functional $\widehat{\mathbb{E}} : \mathcal{H} \to \mathbb{R}$ satisfying*

(1) Monotonicity: $\widehat{\mathbb{E}}[X] \leq \widehat{\mathbb{E}}[Y]$, if $X \leq Y$.

(2) Constant preserving: $\widehat{\mathbb{E}}[c] = c$, for $c \in \mathbb{R}$.

(3) Sub-additivity: For each $X, Y \in \mathcal{H}$, $\widehat{\mathbb{E}}[X + Y] \leq \widehat{\mathbb{E}}[X] + \widehat{\mathbb{E}}[Y]$.

(4) Positive homogeneity: $\widehat{\mathbb{E}}[\lambda X] = \lambda\widehat{\mathbb{E}}[X]$, for $\lambda \geq 0$.

Then the triplet $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ is called a sublinear expectation space.

Remark A.1. *If equality holds in (3), then $\widehat{\mathbb{E}}[\cdot]$ reduces to the classical linear expectation $\mathbb{E}[\cdot]$.*

Lemma A.1. *(Peng, 2019) Let $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ be a sublinear expectation space, and let X and Y be two random variables such that $\widehat{\mathbb{E}}[Y] = -\widehat{\mathbb{E}}[-Y]$. Then*

$$\widehat{\mathbb{E}}[X + \alpha Y] = \widehat{\mathbb{E}}[X] + \alpha\widehat{\mathbb{E}}[Y], \text{ for all } \alpha \in \mathbb{R}. \tag{A.2}$$

In particular, if $\widehat{\mathbb{E}}[Y] = -\widehat{\mathbb{E}}[-Y] = 0$, then $\widehat{\mathbb{E}}[X + \alpha Y] = \widehat{\mathbb{E}}[X]$.

Definition A.2. *(Distribution) (Peng, 2019) Let $X = (X_1, X_2, \cdots, X_n)$ be a given n-dimensional random vector on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$. Then the distribution of X under $\widehat{\mathbb{E}}$ is defined by the functional on $C_{l,Lip}(\mathbb{R}^n)$ as*

$$\mathbb{F}_X[\varphi] := \widehat{\mathbb{E}}[\varphi(X)] : \ \varphi \in C_{l,Lip}(\mathbb{R}^n) \to \mathbb{R}.$$

Remark A.2. *The triplet $\left(\mathbb{R}^n, C_{l,Lip}(\mathbb{R}^n), \mathbb{F}_X\right)$ also forms a sublinear expectation space. And the distribution of $X \in \mathcal{H}$ has the four typical parameters:*

$$\underline{\mu} = -\widehat{\mathbb{E}}[-X], \quad \overline{\mu} = \widehat{\mathbb{E}}[X], \quad \underline{\sigma}^2 = -\widehat{\mathbb{E}}[-X^2], \quad \overline{\sigma}^2 = \widehat{\mathbb{E}}[X^2], \tag{A.3}$$

where $[\underline{\mu}, \overline{\mu}]$ and $\left[\underline{\sigma}^2, \overline{\sigma}^2\right]$ characterize the mean-uncertainty and the variance-uncertainty of X, respectively. In some real financial markets, the variance uncertainty of the distribution is the source of the volatility uncertainty (Vorbrink, 2014).

Definition A.3. *(Identically distributed) (Peng, 2019) Let X and Y be two n-dimensional random vectors defined on sublinear expectation spaces $\left(\Omega_1, \mathcal{H}_1, \widehat{\mathbb{E}}_1\right)$ and $\left(\Omega_2, \mathcal{H}_2, \widehat{\mathbb{E}}_2\right)$, respectively. They are called identically distributed, denoted by $X \overset{d}{=} Y$, if*

$$\widehat{\mathbb{E}}_1\left[\varphi(X)\right] = \widehat{\mathbb{E}}_2\left[\varphi(Y)\right], \text{ for all } \varphi \in C_{l,Lip}(\mathbb{R}^n).$$

Definition A.4. *(Independent) (Peng, 2019) On a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$, a random variable $Y \in \mathcal{H}^n$ is said to be independent of another random variable $X \in \mathcal{H}^m$ under $\widehat{\mathbb{E}}$ if, for each function $\varphi \in C_{l,Lip}(\mathbb{R}^{m+n}$,*

$$\widehat{\mathbb{E}}\left[\varphi(X, Y)\right] = \widehat{\mathbb{E}}\left[\widehat{\mathbb{E}}\left[\varphi(x, Y)\right]\Big|_{x=X}\right].$$

Definition A.5. *(Independent copy) (Peng, 2019) Let X, Y be two n-dimensional random vectors on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$. Y is called an independent copy of X if $Y \overset{d}{=} X$ and Y is independent of X.*

Definition A.6. *(G-normal distribution) (Peng, 2019) An n-dimensional random vector $X = (X_1, X_2, \cdots, X_n)$ on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ is called (centered) G-normally distributed if for all $a, b \geq 0$, we have*

$$aX + b\overline{X} \overset{d}{=} \sqrt{a^2 + b^2}X, \tag{A.4}$$

where $\overline{X}$ is an independent copy of X. Here the letter G denotes the function

$$G(A) := \frac{1}{2}\widehat{\mathbb{E}}\left[(AX, X)\right] : \mathbb{S}_n \to \mathbb{R}, \tag{A.5}$$

where $\mathbb{S}_n$ denotes the collection of $n \times n$ symmetric matrices. When $n = 1$, X is also called a G-normally distributed random variable (G-NDRV).

Remark A.6. *For a G-normally distributed random vector X, we have*

$$\underline{\mu} = -\widehat{\mathbb{E}}[-X] = \overline{\mu} = \widehat{\mathbb{E}}[X] = 0,$$

which means that X has no mean-uncertainty. Usually, we denote a G-normally distributed random vector X by $X \sim \mathcal{N}\left(\{0\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$.

Lemma A.2. *(Peng, 2007a, 2019) Let X a G-normally distributed random variable on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$ as $X \sim \mathcal{N}\left(\{0\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$. Then for each convex (resp. concave) function φ in $C_{l,Lip}(\mathbb{R})$, we have*

$$\widehat{\mathbb{E}}\left[\varphi(X)\right] = \frac{1}{\sqrt{2\pi\overline{\sigma}^2}}\int_{\mathbb{R}} \varphi(x)\exp\left(-\frac{x^2}{2\overline{\sigma}^2}\right)\mathrm{d}x, \quad \left(resp. \frac{1}{\sqrt{2\pi\underline{\sigma}^2}}\int_{\mathbb{R}} \varphi(x)\exp\left(-\frac{x^2}{2\underline{\sigma}^2}\right)\mathrm{d}x\right). \tag{A.6}$$

In particular, we have

$$-\widehat{\mathbb{E}}\left[-X\right] = \widehat{\mathbb{E}}\left[X\right] = 0, \quad \widehat{\mathbb{E}}\left[X^{2n+1}\right] = \widehat{\mathbb{E}}\left[-X^{2n+1}\right], n = 1, 2, \cdots, \tag{A.7}$$

$$-\widehat{\mathbb{E}}\left[-|X|^m\right] = \begin{cases} \sqrt{\frac{2}{\pi}}(m-1)!!\underline{\sigma}^m, & if\ \ m\ is\ odd, \\ (m-1)!!\underline{\sigma}^m, & if\ \ m\ is\ even, \end{cases} \tag{A.8}$$

and

$$\widehat{\mathbb{E}}\left[|X|^m\right] = \begin{cases} \sqrt{\frac{2}{\pi}}(m-1)!!\overline{\sigma}^m, & if\ \ m\ is\ odd, \\ (m-1)!!\overline{\sigma}^m, & if\ \ m\ is\ even, \end{cases} \tag{A.9}$$

where m!! denotes the double factorial of the positive integer m.

Theorem A.3. *(Wiersema, 2008) The probability density function $p(x, t)$ of the standard Bm $B(t)_{t\in\mathbb{R}^+}$ at the end of time period $[0, t]$*

$$p(x, t) = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{x^2}{2t}\right), \tag{A.10}$$

solves the classical heat equation

$$\frac{\partial}{\partial t}u(x, t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}u(x, t). \tag{A.11}$$

(Meerschaert and Sikorskii, 2012) The probability density function $p(x,t)$ of the fBm $B_H(t)_{t\in\mathbb{R}^+}$ with Hurst parameter $H \in (0,1)$

$$p(x,t) = \frac{1}{\sqrt{2\pi t^{2H}}} \exp\left[-\frac{1}{2}\left(\frac{x}{\sqrt{t^{2H}}}\right)^2\right], \tag{A.12}$$

solves the diffusion equation with variable coefficients

$$\frac{\partial}{\partial t}u(x,t) = Ht^{2H-1}\frac{\partial^2}{\partial x^2}u(x,t). \tag{A.13}$$

Theorem A.4. *(Peng, 2019) Let X be a G-normally distributed random variable on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$, denoted by $X \sim \mathcal{N}\left(\{0\}, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right)$. Then its distribution is characterized by the function*

$$v(x,t) = \widehat{\mathbb{E}}\left[\varphi(x + \sqrt{t}X)\right], \quad \varphi \in C_{l,Lip}(\mathbb{R}^n). \tag{A.14}$$

In particular, $\widehat{\mathbb{E}}[\varphi(X)] = v(0,1)$, where $v(x,t)$ is the viscosity solution of the following G-heat equation:

$$\frac{\partial}{\partial t}v(x,t) = G\left(\frac{\partial^2}{\partial x^2}v(x,t)\right), \tag{A.15}$$

with initial condition

$$v(x,0) = \varphi(x), \tag{A.16}$$

where $G(x) = \frac{1}{2}\left(\overline{\sigma}^2 x^+ - \underline{\sigma}^2 x^-\right)$, $\underline{\sigma}^2 = -\widehat{\mathbb{E}}[-X^2]$, and $\overline{\sigma}^2 = \widehat{\mathbb{E}}[X^2]$.

Theorem A.5. *(GBm with PDEs) (Guo et al., 2023a; Peng, 2019) Let $B_G(t)_{t\in\mathbb{R}^+}$ be a GBm on a sublinear expectation space $\left(\Omega, \mathcal{H}, \widehat{\mathbb{E}}\right)$. Then the distribution of $B_G(t)$ is characterized by*

$$\widehat{\mathbb{E}}\left[\varphi(B_G(t))\right] = v(0,t), \quad \varphi \in C_{l,Lip}(\mathbb{R}), \tag{A.17}$$

where $v(x,t)$ is the viscosity solution of the initial-value problem (A.15) *and* (A.16)*.*

Proposition A.6. *(Peng et al., 2023) If the initial condition $v(x,0)$ in* (A.16) *takes as*

$$v(x,0) = \varphi(x) = I_{(0,\infty)}(x), \tag{A.18}$$

where $I_A(x)$ is the indicator function of a set A. Then the solution of the nonlinear initial-value problem (A.15) (A.16) *can be expressed as*

$$v(x,t) = \int_{-\infty}^{x} \rho(y,t)\,\mathrm{d}y, \tag{A.19}$$

where $\rho(y,t)$ is a function on $\mathbb{R}\times\mathbb{R}^+$ defined by

$$\rho(y,t) = \frac{\sqrt{2}}{\left(\underline{\sigma} + \overline{\sigma}\right)\sqrt{\pi t}}\left[\exp\left(-\frac{y^2}{2\overline{\sigma}^2 t}\right)I_{(-\infty,0]}(y) + \exp\left(-\frac{y^2}{2\underline{\sigma}^2 t}\right)I_{(0,\infty)}(y)\right]. \tag{A.20}$$

Appendix B. Other Parameter Estimates and Fitted pdfs

In this appendix, we mainly list the parameter estimates and plot some fitted pdfs for the increments of the other NASDAQ index, DAX, and HANG SENG index, as being stated in Section 3.

Table B.1. Parameter estimates for kinds of distributions for the increments of NASDAQ index log-volatility and corresponding loss functions.

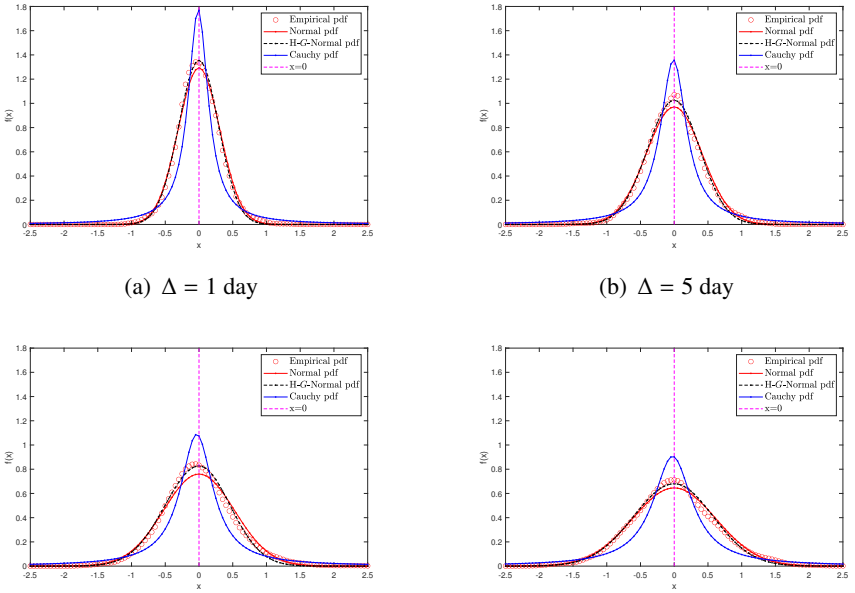
Index	Loss functions / Parameter	Normal σ	H-G-normal $\left(\underline{\sigma}, \overline{\sigma}, H\right)$	Cauchy (a, b)
$\Delta = 1$	Estimates	0.308525	(0.093568, 0.087799, 0.254573)	(-0.011845, 0.179061)
	RMSE	0.026880	**0.020235**	0.117816
	MAE	0.013807	**0.011538**	0.066399
	TIC	0.031383	**0.023352**	0.138524
$\Delta = 5$	Estimates	0.411032	(0.262251, 0.247186, 0.345253)	(-0.011298, 0.233973)
	RMSE	0.026232	**0.018901**	0.096052
	MAE	0.014965	**0.011203**	0.061358
	TIC	0.035308	**0.025090**	0.129577
$\Delta = 25$	Estimates	0.524908	(0.511435, 0.469508, 0.511626)	(-0.033729, 0.292432)
	RMSE	0.034445	**0.023327**	0.083558
	MAE	0.021840	**0.016407**	0.056631
	TIC	0.052119	**0.034561**	0.126059
$\Delta = 125$	Estimates	0.617956	(0.718624, 0.697606, 0.772418)	(-0.023608, 0.351278)
	RMSE	0.024086	**0.019580**	0.074799
	MAE	0.015775	**0.013559**	0.053983
	TIC	0.039793	**0.031933**	0.123844

Table B.2. Parameter estimates for kinds of distributions for the increments of DAX log-volatility and corresponding loss functions.

Index	Loss functions Parameter	Normal σ	H-G-normal $\left(\underline{\sigma}, \overline{\sigma}, H\right)$	Cauchy (a, b)
$\Delta = 1$	Estimates	0.310102	(0.082168, 0.075365, 0.243269)	(-0.005858, 0.175615)
	RMSE	0.032673	**0.021634**	0.109377
	MAE	0.015867	**0.012158**	0.062573
	TIC	0.038152	**0.024850**	0.127842
$\Delta = 5$	Estimates	0.377672	(0.254835, 0.239082, 0.367160)	(-0.017324, 0.216572)
	RMSE	0.027324	**0.019793**	0.103613
	MAE	0.015584	**0.011947**	0.062690
	TIC	0.035251	**0.025198**	0.134224
$\Delta = 25$	Estimates	0.4792067	(0.484797, 0.446528, 0.537616)	(-0.024233 0.265934)
	RMSE	0.034866	**0.021726**	0.088033
	MAE	0.021014	**0.013646**	0.057487
	TIC	0.050357	**0.030713**	0.126646
$\Delta = 125$	Estimates	0.591943	(0.723203, 0.690097, 0.793942)	(-0.047520, 0.345778)
	RMSE	0.026125	**0.023411**	0.085199
	MAE	0.017641	**0.016278**	0.060251
	TIC	0.042325	**0.037676**	0.139263

Table B.3. Parameter estimates for kinds of distributions for the increments of HANG SENG index log-volatility and corresponding loss functions.

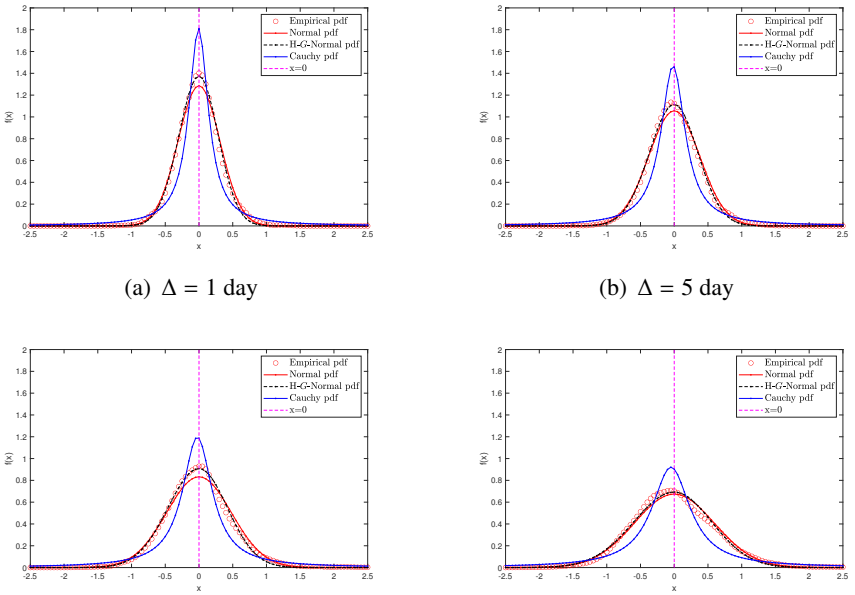
Index	Loss functions Parameter	Normal σ	H-G-normal $\left(\underline{\sigma}, \overline{\sigma}, H\right)$	Cauchy (a, b)
$\Delta = 1$	Estimates	0.300528	(0.080042, 0.070567, 0.247352)	(-0.011521 0.169018)
	RMSE	0.037279	**0.021927**	0.114436
	MAE	0.018719	**0.012151**	0.064298
	TIC	0.042725	**0.024647**	0.131052
$\Delta = 5$	Estimates	0.346797	(0.227484, 0.219210, 0.379670)	(-0.009091, 0.193756)
	RMSE	0.033818	**0.019868**	0.104691
	MAE	0.016900	**0.012471**	0.062321
	TIC	0.041609	**0.023967**	0.128506
$\Delta = 25$	Estimates	0.405120	(0.426562, 0.410713, 0.561732)	(-0.013522, 0.227036)
	RMSE	0.031388	**0.019827**	0.097315
	MAE	0.018102	**0.012321**	0.059415
	TIC	0.041775	**0.025901**	0.129316
$\Delta = 125$	Estimates	0.486879	(0.641497, 0.603030, 0.839362)	(-0.040389, 0.273178)
	RMSE	0.036608	**0.026312**	0.089907
	MAE	0.022861	**0.018406**	0.058577
	TIC	0.053376	**0.037643**	0.130926



(a) Δ = 1 day (b) Δ = 5 day

(c) Δ = 25 day (d) Δ = 125 day

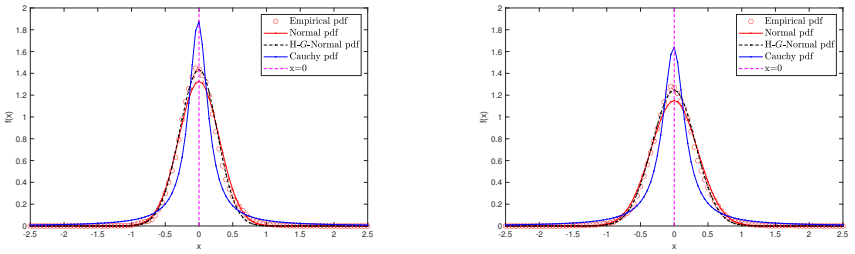
Figure B.1. Some pdfs with the empirical distributions, usual normal distributions, H-*G*-normal distributions, and Cauchy distributions for various lags Δ of the increments of NASDAQ index log-volatility.



(a) Δ = 1 day (b) Δ = 5 day

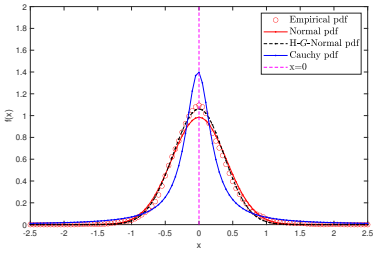
(c) Δ = 25 day (d) Δ = 125 day

Figure B.2. Some pdfs with the empirical distributions, usual normal distributions, H-*G*-normal distributions, and Cauchy distributions for various lags Δ of the increments of DAX log-volatility.

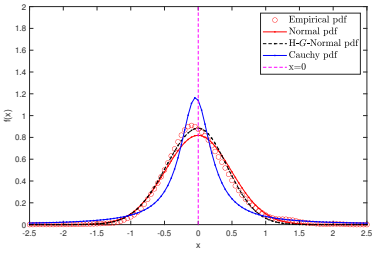


(a) Δ = 1 day

(b) Δ = 5 day



(c) Δ = 25 day



(d) Δ = 125 day

Figure B.3. Some pdfs with the empirical distributions, usual normal distributions, H-*G*-normal distributions, and Cauchy distributions for various lags Δ of the increments of HANG SENG index log-volatility.

Author contributions

Changhong Guo: Writing–original draft, Writing–review & editing, Formal analysis, Visualization, Validation. **Shaomei Fang:** Conceptualization, Formal analysis, Methodology. **Yong He:** Software, Data curation, Writing–review. **Yong Zhang:** Validation, Visualization, Writing–review.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Nos. 72101061 and 72373031), and the Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515012670).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Avellaneda M, Levy A, Parás A (1995) Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl Math Financ* 2: 73–88. https://doi.org/10.1080/13504869500000005

Bakshi G, Cao C, Chen Z (1997) Empirical performance of alternative option pricing models. *J Financ* 52: 2003–2049. https://doi.org/10.1111/j.1540-6261.1997.tb02749.x

Biagini F, Hu Y, Øksendal B, et al. (2008) *Stochastic calculus for fractional Brownian motion and applications*. London: Springer-Verlag.

Björk T, Hult H (2005) A note on Wick products and the fractional Black-Scholes model. *Financ Stoch* 9: 197–209. https://doi.org/10.1007/s00780-004-0144-5

Bowman AW, Azzalini A (1997) *Applied smoothing techniques for data analysis*. New York: Oxford University Press Inc.

Brouste A, Iacus SM (2013) Parameter estimation for the discretely observed fractional Ornstein–Uhlenbeck process and the Yuima R package. *Comput Stat* 28: 1529–1547. https://doi.org/10.1007/s00180-012-0365-6

Chen Z, Epstein L (2002) Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70: 1403–1443. https://doi.org/10.1111/1468-0262.00337

Cheridito P, Kawaguchi H, Maejima M (2003) Fractional Ornstein–Uhlenbeck processes. *Electron J Probab* 8: 1–14. https://doi.org/10.1214/EJP.v8-125

Coquet F, Hu Y, Mémin J, et al. (2002) Filtration-consistent nonlinear expectations and related *g*-expectations. *Probab Theory Relat Fields* 123: 1–27. https://doi.org/10.1007/s004400100172

Denis L, Hu M, Peng S (2011) Function spaces and capacity related to a sublinear expectation: application to *G*-Brownian motion paths. *Potential Anal* 34: 139–161. https://doi.org/10.1007/s11118-010-9185-x

Denis L, Martini C (2006) A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Ann Appl Probab* 16: 827–852. https://doi.org/10.1214/105051606000000169

Denk R, Kupper M, Nendel M (2020) A semigroup approach to nonlinear Lévy processes. *Stochastic Process Appl* 130: 1616–1642. https://doi.org/10.1016/j.spa.2019.05.009

Elliott RJ, Hoek JV (2003) A general fractional white noise theory and applications to finance. *Math Financ* 13: 301–330. https://doi.org/10.1111/1467-9965.00018

Epstein LG, Ji S (2013) Ambiguous volatility and asset pricing in continuous time. *Rev Financ Stud* 26: 1740–1786. https://doi.org/10.1093/rfs/hht018

Fadina T, Neufeld A, Schmidt T (2019) Affine processes under parameter uncertainty. *Probab Uncertain Quant Risk* 4: 1–34. https://doi.org/10.1186/s41546-019-0039-1

Fallahgoul HA, Focardi SM, Fabozzi FJ (2017) *Fractional calculus and fractional processes with applications to financial economics. Theory and application*. London: Elsevier.

Fama E (1965) The behavior of stock market prices. *J Bus* 38: 34–105. https://doi.org/10.1086/294743

Fu H, Liu H, Zheng X (2019) A preconditioned fast finite volume method for distributed-order diffusion equation and applications. *East Asian J Appl Math* 9: 28–44. https://doi.org/10.4208/eajam.160418.190518

Gatheral J, Jaisson T, Rosenbaum M (2018) Volatility is rough. *Quant Financ* 18: 933–949. https://doi.org/10.1080/14697688.2017.1393551

Guo C, Fang S, He Y (2023a) A generalized stochastic process: fractional *G*-Brownian motion. *Methodol Comput Appl Probab* 25: 22. https://doi.org/10.1007/s11009-023-10010-9

Guo C, Fang S, He Y (2023b) Derivation and application of some fractional Black-Scholes equations driven by fractional *G*-Brownian motion. *Comput Econ* 61: 1681–1705. https://doi.org/10.1007/s10614-022-10263-5

Heston SL (1993) A closed-form solution for options with stochastic volatility with applications to band and currency options. *Rev Financ Stud* 6: 327–343. https://doi.org/10.1093/rfs/6.2.327

Hu M, Peng S (2021) G-Lévy processes under sublinear expectations. *Probab Uncertainty Quant Risk* 6: 1–22. https://doi.org/10.3934/puqr.2021001

Hu Y, Øksendal B (2003) Fractional white noise calculus and applications to finance. *Infin Dimens Anal Quantum Probab Relat Top* 6: 1–32. https://doi.org/10.1142/S0219025703001110

Jacod J, Protter P (2004) *Probability essentials*. Berlin:Springer-Verlag.

Jin H, Peng S (2021) Optimal unbiased estimation for maximal distribution. *Probab Uncertain Quant Risk* 6: 189–198. https://doi.org/10.3934/puqr.2021009

Klebaner FC (2012) *Introduction to stochastic calculus with applications*. London: Imperial College Press.

Kolmogorov AN (1940) Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C.R. (doklady) Acad Sci URSS (N.S.)* 26: 115–118. https://api.semanticscholar.org/CorpusID:202489454

Kühn F (2019) Viscosity solutions to Hamilton-Jacobi-Bellman equations associated with sublinear Lévy(-type) processes. *ALEA Lat Am J Probab Math Stat* 16: 531–559. https://doi.org/10.30757/ALEA.v16-20

Krak T, De Bock J, Siebes A (2017) Imprecise continuous-time Markov chains. *Internat J Approx Reason* 88: 452–528. https://doi.org/10.1016/j.ijar.2017.06.012

Lo AW (1991) Long-term memory in stock market prices. *Econometrica* 59: 1279–1313. https://doi.org/10.2307/2938368

Lo AW, MacKinlay AC (1988) Stock market prices do not follow random walks: Evidence from a simple specification test. *Rev Financ Stud* 1: 41–66. https://doi.org/10.1093/rfs/1.1.41

Lyons TJ (1995) Uncertain volatility and the risk-free synthesis of derivatives. *Appl Math Financ* 2: 117–133. https://doi.org/10.1080/13504869500000007

Madsen K, Nielsen HB, Tingleff O (2004) *Methods for non-linear least squares problems, Informatics an Mathematical Modeling*. Copenhagen: Technical University of Denmark.

Mandelbrot BB (1972) Statistical methodology for nonperiodic cycles: from the covariance to R/S analysis. *Ann Econ Soc Meas* 1: 259–290. http://www.nber.org/chapters/c9433

Mandelbrot BB, Van Ness JW (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Rev* 10: 422–437. https://doi.org/10.1137/1010093

Meerschaert MM, Sikorskii A (2012) *Stochastic models for fractional calculus*. Berlin: De Gruyter.

Muhle-Karbe J, Nutz M (2018) A risk-neutral equilibrium leading to uncertain volatility pricing. *Financ Stoch* 22: 281–295. https://doi.org/10.1007/s00780-018-0356-8

Neufeld A, Nutz M (2017) Nonlinear Lévy processes and their characteristics. *Trans Amer Math Soc* 369: 69–95. https://doi.org/10.1090/TRAN/6656

Nolan JP (2020) *Univariate stable distributions. Models for heavy tailed data*. Cham: Springer-Verlag.

Peng S (2005) Nonlinear expectations and nonlinear Markov chains. *Chin Ann Math* 26B: 159–184. https://doi.org/10.1142/S0252959905000154

Peng S (2007a) *G*-expectation, *G*-Brownian motion and related stochastic calculus of Itô's type, *In: Benth FE, Di Nunno G, Lindstrøm T, et al. Stochastic Analysis and Applications*. Berlin: Springer-Verlag, 541–567. https://doi.org/10.1007/978-3-540-70847-6_25

Peng S (2007b) *G*-Brownian motion and dynamic risk measure under volatility uncertainty. preprint, https://doi.org/10.48550/arXiv.0711.2834

Peng S (2008) Multi-dimensional *G*-Brownian motion and related stochastic calculus under *G*-expectation. *Stochastic Process Appl* 118: 2223–2253. https://doi.org/10.1016/j.spa.2007.10.015

Peng S (2017) Theory, methods and meaning of nonlinear expectation theory. *Sci Sin Math* 47: 1223–1254. https://doi.org/10.1360/N012016-00209

Peng S (2019) *Nonlinear expectations and stochastic calculus under uncertainty*. Berlin: Springer-Verlag .

Peng S, Yang S, Yao J (2023) Improving Value-at-Risk prediction under model uncertainty. *J Financ Econ* 21: 228–259. https://doi.org/10.1093/jjfinec/nbaa022

Peng S (2023) *G*-Gaussian processes under sublinear expectations and *q*-Brownian motion in quantum mechanics. *Numerical Algebra, Control and Optimization* 13: 583–603. https://doi.org/10.3934/naco.2022034

Privault N (2013) *Stochastic finance. An introduction with market examples*. Boca Raton: CRC Press.

Rogers LCG (1997) Arbitrage with fractional Brownian motion. *Math Financ* 7: 95–105. https://doi.org/10.1111/1467-9965.00025

Sottinen T (2001) Fractional Brownian motion, random walks and binary market models. *Financ Stoch* 5: 343–355. https://doi.org/10.1007/PL00013536

Soumana-Hima A (2017) *Stochastic differential equations under G-expectation and applications*. Rennes: Université Rennes.

Sun W, Yuan Y (2005) *Optimization theory and methods: nonlinear programming*. New York: Springer-Verlag.

Vorbrink J (2014) Financial markets with volatility uncertainty. *J Math Econ* 53: 64–78. https://doi.org/10.1016/j.jmateco.2014.05.008

Wang X, He X, Bao Y, et al. (2018) Parameter estimates of Heston stochastic volatility model with MLE and consistent EKF algorithm. *Sci Sin* 61: 042202. https://doi.org/10.1007/s11432-017-9215-8

Wiersema UF (2008) *Brownian motion calculus*. Chichester: John Wiley & Sons.

AIMS Press