

ϵ -Strong Simulation of Fractional Brownian Motion and Related Stochastic Differential Equations

Yi Chen

Northwestern University, yichen2016@u.northwestern.edu,

Jing Dong

Columbia University, jing.dong@gsb.columbia.edu,

Hao Ni

University College London, h.ni@ucl.ac.uk,

Consider the fractional Brownian Motion (fBM) $B^H = \{B^H(t) : t \in [0, 1]\}$ with Hurst index $H \in (0, 1)$. We construct a probability space supporting both B^H and a fully simulatable process \hat{B}_ϵ^H such that

$$\sup_{t \in [0, 1]} |B^H(t) - \hat{B}_\epsilon^H(t)| \leq \epsilon$$

with probability one for any user specified error parameter $\epsilon > 0$. When $H > 1/2$, we further enhance our error guarantee to the α -Hölder norm for any $\alpha \in (1/2, H)$. This enables us to extend our algorithm to the simulation of fBM driven stochastic differential equations $Y = \{Y(t) : t \in [0, 1]\}$. Under mild regularity conditions on the drift and diffusion coefficients of Y , we construct a probability space supporting both Y and a fully simulatable process \hat{Y}_ϵ such that

$$\sup_{t \in [0, 1]} |Y(t) - \hat{Y}_\epsilon(t)| \leq \epsilon$$

with probability one. Our algorithms enjoy the tolerance-enforcement feature, i.e., the error bounds can be updated sequentially. Thus, the algorithms can be readily combined with other simulation techniques like multilevel Monte Carlo to estimate expectation of functionals of fBMs efficiently.

Key words: fractional Brownian motion, Stochastic differential equation, Monte Carlo Simulation

1. Introduction The fractional Brownian motion (fBM) of Hurst parameter $H \in (0, 1)$, $\{B^H(t) : t \geq 0\}$, is a centered real-valued Gaussian process with covariance function

$$r(s, t) := E[B^H(s)B^H(t)] = \frac{1}{2} (|s|^{2H} + |t|^{2H} - |s - t|^{2H}). \quad (1)$$

When $H = 1/2$, fBM is a Brownian motion (BM) which has independent increments. When $H < 1/2$, the increments of fBM are negatively correlated. In contrast, when $H > 1/2$, fBM has positively correlated increments and displays long-range dependence. Generally, fBM can be viewed as extension of BM. As it allows the increment to be correlated, it has been applied in communications engineering (Norros, 1995), biology and physics (Höfling and Franosch, 2013). See also Laskin (2000) and Nualart (2006) for applications in finance and turbulence. Due to the correlated increments (lack of Markovian structure), very few closed-form expressions are known for performance measures related to functionals of fBMs. In this context, simulation-based numerical method has been a powerful tool to conduct performance analysis for fBM driven processes.

In this paper, we develop a new class of algorithms to construct paths of fBM and fBM driven stochastic differential equations (SDEs) with strong error guarantees. In particular, the algorithm allows us construct a probability space supporting both a fBM and a fully simulatable paths \hat{B}_ϵ^H , such that

$$\sup_{0 \leq t \leq 1} |B^H(t) - \hat{B}_\epsilon^H(t)| \leq \epsilon \text{ w.p. } 1.$$

Moreover, when $H > 1/2$, for $\alpha \in (1/2, H)$, we can further construct a fully simulatable paths $\hat{B}_{\epsilon(\alpha)}^H$, such that

$$\sup_{0 < s \leq t \leq 1} \frac{|B^H(t) - \hat{B}_{\epsilon(\alpha)}^H(t) - (B^H(s) - \hat{B}_{\epsilon(\alpha)}^H(s))|}{|t - s|^\alpha} \leq \epsilon \text{ w.p. } 1.$$

For $H > 1/2$, the control of the α -Hölder norm allows us to use the rough-path theory [Lyons \(1998\)](#) to construct a probability space supporting both a fBM-driven SDE

$$dY(t) = \mu(Y(t))dt + \sigma(Y(t))dB^H(t)$$

and a sequence of fully simulatable path \hat{Y}_ϵ , such that

$$\sup_{0 \leq t \leq 1} |Y(t) - \hat{Y}_\epsilon(t)| \leq \epsilon \text{ w.p. } 1.$$

In addition to the strong error guarantee, the framework we developed also enjoys the *tolerance-enforcement* feature. Specifically, for any sequence $0 < \epsilon_n < \epsilon_{n-1} < \dots < \epsilon_1$, we can adaptively simulate X_{ϵ_n} conditional on $X_{\epsilon_1}, \dots, X_{\epsilon_{n-1}}$. The tolerance-enforcement allows us to easily combine our procedure with other simulation techniques such as Multilevel Monte Carlo (MLMC) for efficient estimations of expectations [Giles \(2008b\)](#), and various randomization techniques to remove estimation bias (see for example [Rhee and Glynn \(2012\)](#); [Beskos et al. \(2012\)](#)). In this paper, we provide a concrete demonstration as how to combine our algorithm with MLMC. The strong error bound provides the extra benefit of simplifying the rate of convergence analysis and complexity analysis in this case.

In terms of the computational complexity, our algorithm achieves the near optimal complexity. Specifically, for fBM, to achieve an ϵ error bound, the computational cost is $O_p(\epsilon^{-1/(H-\delta)})$ for any $\delta > 0$. When $H > 1/2$, for fBM driven SDEs, under suitable regularity conditions on the drift and diffusion terms, to achieve an ϵ error bound, the computation cost is $O_p(\epsilon^{-1/(2\alpha-1)})$ for any $\alpha \in (1/2, H)$. When combined with MLMC for expectation estimation, to achieve an ϵ^2 mean squared error (MSE), for Lipschitz continuous functionals of fBM, we are able to reduce the computational complexity to the canonical $O_p(\epsilon^{-2} \log(1/\epsilon))$ when $H > 1/2$. For fBM driven SDEs, we are able to reduce the computational complexity to the canonical $O_p(\epsilon^{-2} \log(1/\epsilon))$ when $\alpha > 3/4$.

The simulation framework we developed in this paper is an important extension of the framework developed in [Blanchet et al. \(2017\)](#) for BM driven SDEs, and is the very first of its kind for fBM and fBM driven SDEs (with almost sure error bound and tolerance-enforcement feature). In the process of developing the simulation algorithms, we also extend existing results and prove new results about properties of fBM. These results may be of independent interests to the analysis of fBM. In particular, for midpoint displacement decomposition (wavelet expansion using the Haar wavelets) of fBM, we establish its convergence rate in the uniform norm for $H \in (0, 1)$, and in the α -Hölder norm for $H \in (1/2, 1)$, $\alpha < H$ almost surely. These results rely on detailed analysis of the decay rate of the conditional variance of fBM at different dyadic levels. For fBM driven SDEs, we provide explicit characterization of the constant term for the error induced by Euler scheme. This extends previous results on the convergence rate of Euler scheme in path-by-path construction of fBM driven SDEs.

Throughout the paper, we denote

$$\|u\|_\infty := \sup_{0 \leq t \leq 1} |u(t)| \text{ and } \|u\|_\alpha := \sup_{0 \leq s < t \leq 1} \frac{|u(s) - u(t)|}{|s - t|^\alpha}$$

as the supremum norm and α -Hölder norm of a generic function u on $[0, 1]$ respectively. For \mathbf{v} , which can be a vector, a matrix, or a tensor, we use $\|\mathbf{v}\|$ to denote the maximum of the absolute

value of its entries.

The rest of the paper is organized as follows. We conclude this section with a brief review of the literature to put our work in the right context. We introduce the basic idea of our algorithmic development in Section 2. We then introduce properties of the midpoint displacement decomposition of fBM in Section 3. This provides the theoretical basis for the construction of our algorithm. The details of the actual simulation algorithm for fBM are provided in Section 4. In Section 5, we extend the algorithm to SDEs driven by fBM with Hurst index $H > 1/2$. We explain how our algorithm can be combined with MLMC in Section 6. Lastly, in Section 7 we conduct some numerical experiments as a sanity check of our development and provide some comments about implementations. All of the proofs of the technical lemmas are delayed until the Appendix.

1.1. Literature review Our work is closely related to the line of research on *simulation of fBM*. Existing methods for simulating fBM can be divided into two categories, exact method and approximation method (see Dieker (2004) for a detailed survey).

The exact methods aim to generate the fBM at a fixed finite set of time points from the exact distribution. To carry out this task efficiently is highly nontrivial, due to the correlation structure of fBM. Naive implementation using the Cholesky decomposition has a complexity of $O(n^3)$ for n points in general. More efficient methods have been developed in the literature, mostly of them utilizing the stationarity of the process. For example, the Hosking method (also known as Durbin or Levinson method) (e.g., Levinson (1946), Durbin (1960)) generates sample path recursively, which avoids calculating the inverse of the covariance matrix. Its computational complexity is $O(n^2)$ for a set of size n . The circulant embedding method, which is originally proposed by Davies and Harte (1987) and later generalized by Dietrich and Newsam (1997) and Wood and Chan (1994). The basic idea is to find the square root decomposition of the covariance matrix by embedding the covariance matrix in a so-called circulant covariance matrix. This method can further reduce the complexity to $O(n \log(n))$. Our method aim to recover the whole fBM path, but it builds on being able to generate the fBM at finite set of points, dyadic points in particular, exactly. Thus, some of these techniques developed in the literature will be incorporated into our algorithmic development. However, as our algorithm also relies on sequentially update the set of points at finer and finer scales, we lose some of the stationarity structure.

The approximation methods aim to generate approximations of the fBM sample path. The (conditionalized) random midpoint displacement (RMD) method generates the sample path recursively in a carefully designed order (Lau et al., 1995; Norros et al., 1999). When generating the next sample, the (conditionalized) RMD method speeds up by using only partial samples generated, instead of whole history. It achieves a computational complexity of order $O(n)$ for n points, but the path we constructed may lose certain properties of the original sample path (e.g. long-range dependence) and it is not clear what error guarantee we would be able to achieve here. Some approximation methods build on special representations of fBM. For example, Mandelbrot and Ness (1968) represents the fBM as a stochastic integral with respect to ordinary BM and approximates the integral via Riemann sum. Other representations of the fBM that are used to develop approximation algorithms through truncation include wavelet decompositions (see for example Abry and Sellan (1996), Meyer et al. (1999), and Ayache and Taqqu (2003)), spectral decompositions (see for example Dieker and Mandjes (2003), Paxson (1997) and Dzhaparidze and Zanten (2004)). Our method also builds a suitable infinite series decomposition of the fBM. However, instead of applying a deterministic truncation level, our truncation level is adapted to the sample path, i.e. random. Our work extends this line of literature by achieving a stronger error guarantee. We also note that like a lot of the approximation methods developed in the literature, our algorithm is for pre-specified fixed time horizons.

The simulation framework we developed is also closely related to recent development in ϵ -strong simulation. The ϵ -strong simulation refers to constructing a fully simulatable path whose deviation from the true path is uniformly bounded by ϵ with probability one. Beskos et al. (2012) is among the first to develop the concept ϵ -strong simulation. In Beskos et al. (2012), the authors develop an ϵ -strong simulation algorithm for Brownian motion. Blanchet and Chen (2015) and Blanchet et al. (2017) later extend the framework to reflected Brownian motion and multidimensional stochastic differential equations respectively. One important application of ϵ -strong simulation algorithm is to build unbiased estimators for expectations involving functionals of the sample path (Beskos et al., 2012), or to build exact simulation algorithm for the corresponding stochastic processes (at a finite collection of time points) (Beskos and Roberts, 2005; Chen and Huang, 2013). Blanchet and Zhang (2017) extends the algorithm developed in Blanchet et al. (2017) to construct exact simulation algorithm for multidimensional SDEs. Pollock et al. (2013) considers the SDEs with jumps and provides a comprehensive discussion on ϵ -strong and exact simulation. See also Glynn (2016) for an extension review of recent development in exact simulation and unbiased estimation algorithms. Our work contributes to this line of work by extending the ϵ -strong simulation framework to fBM and fBM driven SDEs.

In terms of the methodology. Our development builds on the idea of *record-breakers*. This idea was first introduced in Blanchet and Sigman (2011) for exact sampling of stochastic perpetuities. Later similar ideas have been applied to exact simulation of queueing models in steady-state (Blanchet and Dong, 2015; Blanchet and Chen, 2015), and max-stable processes and related random fields (Liu et al., 2016). Our algorithm also builds on idea of *Bernoulli factory* (Huber, 2016), but in the actual implementation, we avoid using the Bernoulli factory by applying some properly constructed change-of-measures.

An important tool in developing ϵ -strong simulation algorithm for fBM driven SDEs is the *rough path theory* (Lyons, 1998). The rough path theory provides us with a path-by-path construction of the SDEs. By lifting a driving signal to a rough path, the mapping from a driving rough path to the solution to the SDE is uniformly continuous under the p -variation metric. Most related to our setting, Coutin and Qian (2002) construct a geometric rough path associated with fBM where the Hurst index $H > 1/4$ and develops a Skorohod integral representation of the geometric rough path. More recently, Nualart and Tindel (2011) develop the construction of the rough path above fBM using Volterra's representation for any $H \in (0, 1)$.

There are also works analyzing the discretization error for fBM driven SDEs. For example, Mishura (2008) investigates the rate of convergence of the Euler scheme. Neuenkirch and Nourdin (2007) conduct convergence analysis of a few different discretization schemes. Deya et al. (2012) propose a modified Milstein scheme without using the actual Lévy area. More recently, Hu et al. (2016) introduce a modified Euler scheme that works well when H approaches $1/2$. Most previous works focus on weaker error bound than what is established in our work. Using rough path theory, we are able to study the discretization error in a path-by-path sense. Similar techniques are used in Davie (2008); Blanchet et al. (2017). In this paper, we focus on the Euler scheme to demonstrate the basic idea of the ϵ -strong framework. We view the exploration of more sophisticated discretization scheme as an interesting future research direction.

Our ϵ -strong simulation algorithm can be combined with MLMC. MLMC is first proposed in Giles (2008b) to reduce the computational complexity for the expectation estimation of SDEs (driven by BM) via Euler scheme. MLMC use the multigrids ideas and has $O(\epsilon^{-2}(\log(\epsilon))^2)$ computational complexity to achieve a MSE of $O(\epsilon^2)$, which is a significant improvement from the naive Monte Carlo method. This idea is further enhanced in Giles (2008a) by combing with Milstein scheme. The idea of MLMC are also extended to the estimation of functionals of more general stochastic processes. For example, Dereich and Heidenreich (2011) proposes a MLMC algorithm for Lévy-driven SDEs and Bayer et al. (2016) extends this idea for SDEs driven by general Gaussian noise using the rough path theory.

2. Basic idea We start by introducing the basic idea of our algorithmic development. Recall that fBM B^H is a centered real-valued Gaussian process with the covariance function given in (1). By Kolmogorov continuity theorem, fBM has a continuous modification. Moreover, for any $\alpha \in (0, H)$ and $T > 0$, this modification is α -Hölder continuous on $[0, T]$. In this paper, we refer to the modification as the fBM, and focus on a finite time interval $[0, 1]$ with $B^H(0) = 0$.

The algorithmic development consists of two steps. First, we identify an infinite series expansion of fBM, i.e.

$$B^H(t) = \sum_{k=0}^{\infty} \Lambda_k(t) W_k \quad (2)$$

where Λ_k 's are a sequence of basis functions, and W_k 's are the random coefficients. We then develop an algorithm to truncate the infinite sum up to a finite but random level, K , so that the error induced by the truncated terms is suitably controlled, e.g.

$$\sup_{0 \leq t \leq 1} \left| \sum_{k=0}^K \Lambda_k(t) W_k - B^H(t) \right| \leq \epsilon \text{ w.p. } 1$$

In terms of the infinite series expansion for fBM, several of them are developed in the literature, which are based on wavelet decomposition (multi-resolution framework) (see for example Meyer et al. (1999)) or Karhunen-Loève type of expansion (spectral theory) (see for example Dzhaparidze and Zanten (2004)). Our consideration here is twofold. First, the infinite series expansion needs to converge fast in an almost sure sense under uniform norm or even the α -Hölder norm. Second, the corresponding simulation algorithm can be implemented efficiently.

In this paper, our main development follows Lévy's midpoint displacement technique, which corresponds to the wavelet decomposition using the Haar wavelets. The challenge here is that when $H \neq 1/2$, the coefficients are correlated. We shall provide more analysis about the random coefficient terms in Section 3. The actual midpoint displacement construction goes as follows. Let D_n be the dyadic discretization of order n and Δ_n be the mesh of the discretization. Specifically,

$$D_n = \{t_0^n, t_1^n, \dots, t_{2^n}^n\}, \text{ where } t_i^n = i/2^n, i = 0, 1, \dots, 2^n, \text{ and } \Delta_n = 1/2^n.$$

We use $\mathbf{B}_n^H = (B^H(t_0^n), \dots, B^H(t_{2^n}^n))$ to denote the value of fBM at discretization level n . Given a realization of \mathbf{B}_n^H , we can construct a continuous path B_n^H over time interval $[0, 1]$ via linear interpolation and we call B_n^H a dyadic discretization of fBM of level n . We notice that $B_n^H(t) - B_{n-1}^H(t)$ has zero-value on D_{n-1} . At the augmented points $D_n/D_{n-1} = \{t_{2i+1}^n\}_{i=0, \dots, 2^{n-1}-1}$, where $i = 0, 1, \dots, 2^n - 1$, we have

$$B_n^H(t_{2i+1}^n) - B_{n-1}^H(t_{2i+1}^n) = B^H(t_{2i+1}^n) - \frac{1}{2}(B^H(t_i^{n-1}) + B^H(t_{i+1}^{n-1}))$$

This is what we refer to as the midpoint displacement. In Section 3, we show that the following infinite series representation is valid

$$B^H(t) = \sum_{k=0}^{\infty} [B_k^H(t) - B_{k-1}^H(t)], \quad (3)$$

i.e. B_n^H converges to B^H almost surely under the supremum norm (Theorem 2). Here $B_{-1}^H(t)$ denotes a zero-valued constant function. Note that when we truncate the infinite series at level n , we obtain B_n^H . We also establish the rate of convergence for B_n^H (Theorem 3 and Theorem 4). Specifically,

$$\|B_n^H - B^H\|_{\infty} = O(2^{-(H-\delta)n}) \text{ for any } \delta \in (0, H).$$

Notice that when the W_k 's in the series representation [2](#) are uncorrelated, [Kühn and Linde \(2002\)](#) shows that the optimal rate of convergence, under the L_2 norm, of such series representation of fBM is $O(n^{-H}(1 + \log n)^{1/2})$. The midpoint displacement representation achieves almost the same rate of convergence ¹.

We next introduce the algorithmic development to truncate the infinite sum. Our goal here is to control the error of the infinite truncated terms. To achieve this, we adopt the strategy of “record-breakers” ([Blanchet et al., 2017](#)). The general idea of record-breakers is to define a sequence of events called record-breakers, which satisfies the following two conditions

- C1) The following event happens with probability one: beyond some random but finite level, there will be no more record breakers;
- C2) By knowing that there are no more record breakers, the contribution of the infinite remaining terms are well under control.

In our case, we say a record is broken at level n if

$$\|B_n^H - B_{n-1}^H\|_\infty \geq \rho \cdot 2^{-(H-\delta)n}.$$

Here $\delta \in (0, H)$ and $\rho > 0$ are tuning parameters, which will be specified in [Theorem 1](#) of [Section 3](#). The choices of these parameters will affect the efficiency of the implementation of our algorithm. We also denote N as the level of the last record-breaker, i.e.

$$N = \sup \{n \geq 1 : \|B_n^H - B_{n-1}^H\|_\infty \geq \rho \cdot 2^{-(H-\delta)n}\}.$$

For C1), we show in [Theorem 3](#) that N has a finite moment generating function. For C2), we notice that for $n \geq N$, we have

$$\left\| \sum_{k=n+1}^{\infty} [B_k^H - B_{k-1}^H] \right\|_\infty \leq \rho \cdot \sum_{k=n+1}^{\infty} 2^{-(H-\delta)k}.$$

Thus, once we know the time of the last record-breaker N , to achieve a certain accuracy ϵ , we just need to find $N(\epsilon) > N$, such that

$$\rho \cdot \sum_{k=N(\epsilon)+1}^{\infty} 2^{-(H-\delta)k} < \epsilon,$$

then $\|B_{N(\epsilon)}^H - B^H\|_\infty < \epsilon$. The error bound is achieved in a path-by-path sense. In addition, we show in [Theorem 4](#) that conditional on N , we also have an explicit upper bound for the α -Hölder norm of B^H in a path-by-path sense. This is important to develop the ϵ -strong simulation algorithm for fBM driven SDEs as outlined in [Section 5](#).

The remaining task is to find the last record-breaker N . This is challenging as N is not a stopping time for the filtration generated by the levels. We use techniques from rare-event simulation to overcome the challenge. Our strategy is to find the record-breakers sequentially until we find the last one. Let τ_k to denote the level of the k -th record break, i.e.

$$\begin{aligned} \tau_0 &= 0, \\ \tau_k &= \inf \{n \geq \tau_{k-1} + 1 : \|B_n^H - B_{n-1}^H\|_\infty > \rho \cdot 2^{-(H-\delta)n}\}, \quad k \geq 1. \end{aligned}$$

¹ The n -th dyadic level involves 2^n time points (Gaussian random variables).

Then we have

$$N = \sup\{k \geq 1 : \tau_k < \infty\}.$$

The general idea is as follows. Conditional on $B_{\tau_k}^H$, we first use a change-of-measure to generate $B_{\tau_{k+1}}^H$. We then apply an acceptance-rejection step using a properly defined likelihood ratio. If the path is accepted, we find the τ_{k+1} and $B_{\tau_k}^H$, otherwise, we claim that τ_k is the level of the last record-breaker, i.e. $N = \tau_k$. The details of the algorithmic developments are provided in Section 4.

3. Midpoint displacement of fBM In this section, we analyze the midpoint displacement construction of fBM. This provides the theoretical foundation for our algorithmic development. Specifically, we establish the validity of the infinite series expansion (3), and analyze its rate of convergence under both the uniform norm and the α -Hölder norm. The analysis also provides us a way to construct the record-breakers.

Recall that at the augmented points $D_k/D_{k-1} = \{t_{2^j+1}^k\}_{j=0,1,\dots,2^k-1}$, we have

$$B_k^H(t_{2^j+1}^k) - B_{k-1}^H(t_{2^j+1}^k) = B^H(t_{2^j+1}^k) - \frac{1}{2}(B^H(t_j^{k-1}) + B^H(t_{j+1}^{k-1})).$$

For convenience, we denote by

$$a_j^k := B^H(t_{2^j+1}^k), \quad b_j^k := \frac{1}{2}(B^H(t_j^{k-1}) + B^H(t_{j+1}^{k-1})). \quad (4)$$

Then since $B_k^H(t) - B_{k-1}^H(t)$ is linear over intervals $[t_{2^j}^k, t_{2^j+1}^k]$ and $[t_{2^j+1}^k, t_{2^j+2}^k]$, we have

$$\|B_k^H - B_{k-1}^H\|_\infty = \max_{0 \leq j \leq 2^k-1} |a_j^k - b_j^k|.$$

We first establish the convergence rate of $\|B_k^H - B_{k-1}^H\|_\infty$, which lays the foundation of subsequent results. We define

$$\ell_k := 2^{-(H-\delta)k}, \text{ for any fixed } \delta \in (0, H).$$

For any constant $\nu > 0$, we denote by

$$K(\nu) = \sup\{n \geq 1 : 4\sqrt{n} > \nu \cdot 2^{\delta n}\}. \quad (5)$$

Then we have the following theorem establishing bounds for $\|B_k^H - B_{k-1}^H\|_\infty$.

THEOREM 1. *For any constant $\nu, \nu^* > 0$, for all $k > K(\nu)$, we have*

$$\mathbb{P}(\|B_k^H - B_{k-1}^H\|_\infty \geq \rho \ell_k) = \mathbb{P}\left(\max_{0 \leq j \leq 2^k-1} |a_j^k - b_j^k| \geq \rho \ell_k\right) \leq 2 \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\},$$

where $\rho = 2(\nu + \nu^*)$.

Before we prove Theorem 1, we would like to comment that ν and ρ are parameters characterizing the record-breakers. We have some freedom in choosing these parameters, and there is a tradeoff involved. For larger ρ , the record-breakers are less likely to happen and it is relatively faster to find the last record-breaker. However, as a cost, we need to truncate at a higher level to achieve the desired accuracy. We provide more discussion about the choice of these parameters in practice in Section 7.

Proof of Theorem 1. Note that for fBM with Hurst index $H \neq 1/2$, $a_j^k - b_j^k$ is not a centered Gaussian random variable. In the following, we use

$$c_j^k = \mathbb{E}[B^H(t_{2^{k-1}+1}^k) | \mathbf{B}_{k-1}^H], \quad j = 0, 1, \dots, 2^{k-1} - 1,$$

to denote the conditional expectation of fBM at the augmented points D_k/D_{k-1} given the values of fBM on D_{k-1} . Then we have

$$\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - b_j^k| \leq \max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k| + \max_{0 \leq j \leq 2^{k-1} - 1} |c_j^k - b_j^k|. \quad (6)$$

The two terms in the right-hand side of inequality (6) correspond to the variance and bias. In what follows, we will establish bounds for each of them.

It is easy to see that

$$\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - b_j^k| \geq \rho \ell_k\right) \leq \underbrace{\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k| > \rho \ell_k / 2\right)}_{(V)} + \underbrace{\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1} - 1} |c_j^k - b_j^k| > \rho \ell_k / 2\right)}_{(B)}. \quad (7)$$

The rest of the proof is divided into two parts. We first establish a bound for (V), which corresponds to the variance. We then establish a bound for (B), which corresponds to the bias. In subsequent analysis, we need several auxiliary results that are summarized in lemmas.

For (V), we have

$$\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k| > \rho \ell_k / 2\right) = \mathbb{E}\left[\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k| > \rho \ell_k / 2 \mid \mathbf{B}_{k-1}^H\right)\right]. \quad (8)$$

In the following, we use $\mathbb{P}_{k-1}(\cdot)$ to denote $\mathbb{P}(\cdot | \mathbf{B}_{k-1}^H)$, which is the conditional probability given the values of \mathbf{B}_{k-1}^H . We also use \mathbb{E}_{k-1} and \mathbb{V}_{k-1} to denote corresponding conditional expectation and variance. Then under the probability measure $\mathbb{P}_{k-1}(\cdot)$, a_j^k is a Gaussian random variable with mean c_j^k and variance $\sigma_{kj}^2 := \mathbb{V}_{k-1}(a_j^k)$. The following lemma upper bounds σ_{kj}^2 uniformly for all j 's.

LEMMA 1. *For all $k \geq 1$ and $j = 0, 1, \dots, 2^{k-1} - 1$, we have*

$$\mathbb{V}_{k-1}(a_j^k) = \sigma_{kj}^2 \leq 2 \cdot 2^{-2kH}.$$

Using Lemma 1 and Borell-TIS inequality, for any $u > 0$, we obtain

$$\mathbb{P}_k\left(\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k| - \mathbb{E}_k\left[\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k|\right] > u\right) \leq \exp\{-u^2 \cdot 2^{-2kH-2}\}. \quad (9)$$

In order to get rid of the expectation in inequality (9), we need the following lemma to upper bound the expectation.

LEMMA 2. *Let X_1, X_2, \dots, X_n be a sequence of (not necessarily independent) centered Gaussian random variables whose variances are uniformly bounded by σ^2 . Then we have*

$$\mathbb{E}\left[\max_{1 \leq i \leq n} |X_i|\right] \leq 2\sqrt{2 \log(2n)} \cdot \sigma.$$

By applying Lemma 1 and 2, we obtain that for any $\nu > 0$, there exists a $K(\nu)$ such that for all $k \geq K(\nu)$,

$$\mathbb{E}_{k-1}\left[\max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - c_j^k|\right] \leq 4\sqrt{k} \cdot 2^{-kH} \leq \nu \ell_k. \quad (10)$$

For any constant $\nu^* > 0$, by setting $u = \nu^* \ell_k$ in inequality (9) and using inequality (10), we obtain

$$\mathbb{P}_{k-1} \left(\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - c_j^k| > (\nu + \nu^*) \ell_k \right) \leq \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\}.$$

Then based on (8), for the unconditional probability, we have

$$\mathbb{P} \left(\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - c_j^k| > (\nu + \nu^*) \ell_k \right) \leq \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\}. \quad (11)$$

Now, we turn to (B). In contrast to the previous proof where we deal with the conditional probability $\mathbb{P}_{k-1}(\cdot)$ first, in this part, we consider the unconditional probability $\mathbb{P}(\cdot)$ directly. In the following, to simplify notations, let $\mathbf{b}_k = (b_0^k, \dots, b_{2^{k-1}-1}^k)$ and $\mathbf{c}_k = (c_0^k, \dots, c_{2^{k-1}-1}^k)$. Then by definition, we have

$$\mathbf{b}_k = \mathbf{M}_{k-1} \mathbf{B}_{k-1}^H, \text{ where } \mathbf{M}_{k-1} = \begin{bmatrix} 1/2 & 1/2 & 0 & \cdots & 0 \\ 0 & 1/2 & 1/2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1/2 \end{bmatrix}_{2^{k-1} \times (2^{k-1}+1)}.$$

Based on the conditional distribution of multivariate Gaussian random vector and the covariance function of fBM, we have $\mathbf{c}_k = \mathbf{N}_{k-1} \mathbf{B}_{k-1}^H$, where $\mathbf{N}_{k-1} = \boldsymbol{\Sigma}_{12}^{(k-1)} \cdot [\boldsymbol{\Sigma}_{22}^{(k-1)}]^{-1}$. Here, $\boldsymbol{\Sigma}_{22}^{(k-1)}$ and $\boldsymbol{\Sigma}_{12}^{(k-1)}$ take the form

$$\boldsymbol{\Sigma}_{22}^{(k-1)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots \\ 0 & 2^{2H} & (2^{2H} + 4^{2H} - 2^{2H})/2 & \cdots & \cdots \\ 0 & (2^{2H} + 4^{2H} - 2^{2H})/2 & 4^{2H} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}_{(2^{k-1}+1) \times (2^{k-1}+1)}$$

which is a $(2^{k-1} + 1)$ by $(2^{k-1} + 1)$ matrix with (i, j) -th entry $(|2i - 2|^{2H} + |2j - 2|^{2H} - |2i - 2j|^{2H})/2$;

$$\boldsymbol{\Sigma}_{12}^{(k-1)} = \begin{bmatrix} 1^{2H} & (1^{2H} + 2^{2H} - 1^{2H})/2 & (1^{2H} + 4^{2H} - 3^{2H})/2 & \cdots & \cdots \\ 3^{2H} & (3^{2H} + 2^{2H} - 1^{2H})/2 & (3^{2H} + 4^{2H} - 1^{2H})/2 & \cdots & \cdots \\ 5^{2H} & (5^{2H} + 2^{2H} - 3^{2H})/2 & (5^{2H} + 4^{2H} - 1^{2H})/2 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}_{2^{k-1} \times (2^{k-1}+1)}$$

which is a 2^{k-1} by $(2^{k-1} + 1)$ matrix with (i, j) -th entry $(|2i - 1|^{2H} + |2j - 2|^{2H} - |2i - 2j + 1|^{2H})/2$. We remark that here the inverse $[\cdot]^{-1}$ is interpreted as generalized inverse.

Then we have $\mathbf{c}_k - \mathbf{b}_k = (\mathbf{N}_{k-1} - \mathbf{M}_{k-1}) \mathbf{B}_{k-1}^H$ and its covariance matrix is given by

$$\boldsymbol{\Sigma}^{(k)} = (\mathbf{N}_{k-1} - \mathbf{M}_{k-1}) \boldsymbol{\Sigma}_{22}^{(k-1)} (\mathbf{N}_{k-1} - \mathbf{M}_{k-1})^\top \cdot \Delta_k^{2H}.$$

The following lemma bounds the diagonal entries of $\boldsymbol{\Sigma}^{(k)}$, which correspond to the variances of random variables $c_j^k - b_j^k$.

LEMMA 3. *The diagonal entries of $\boldsymbol{\Sigma}^{(k)}$ are uniformly upper bounded by $2 \cdot 2^{-2kH}$.*

Then by using Borell-TIS inequality again, we have

$$\mathbb{P} \left(\max_{0 \leq j \leq 2^{k-1}-1} |c_j^k - b_j^k| - \mathbb{E} \left[\max_{0 \leq j \leq 2^{k-1}-1} |c_j^k - b_j^k| \right] > \nu^* \ell_k \right) \leq \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\}.$$

Similar to the proof for (V), we can get rid of the expectation and obtain that for all $k > K(\nu)$,

$$\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1}-1} |c_j^k - b_j^k| > (\nu + \nu^*)\ell_k\right) \leq \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\}. \quad (12)$$

Finally, combining (11) (for (V)) and (12) (for (B)), we obtain

$$\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \geq 2(\nu + \nu^*)\ell_k\right) \leq 2 \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\},$$

which concludes the proof of Theorem 1. \square

3.1. Validity of the expansion Let $\mathcal{C}([0, 1])$ be the space of continuous functions over $[0, 1]$ equipped with uniform norm $\|\cdot\|_\infty$. The next theorem establish the validity of (3).

THEOREM 2. *The sample paths of $B_n^H(t)$ converge to a fBM $B^H(t)$ in $\mathcal{C}([0, 1])$ almost surely. In other words,*

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \|B_n^H - B^H\|_\infty = 0\right) = 1.$$

Proof. We first prove that the sequence $\{B_n^H\}_{n \geq 1}$ is a Cauchy sequence in $\mathcal{C}([0, 1])$ almost surely. Then its limiting process exists almost surely due to the completeness of $\mathcal{C}([0, 1])$. Second, we show that the limiting process is a Gaussian process and has the same covariance structure as the fBM.

Since the tail bound $2 \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\}$ established in Theorem 1 is summable, by Borel-Cantelli Lemma, we have

$$\mathbb{P}\left(\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \geq \rho \ell_k, \text{ i.o.}\right) = 0.$$

Hence, there exists a random variable N , which is finite almost surely, such that for all $k \geq N$, $\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \leq \rho \ell_k$. Then for arbitrary $\epsilon > 0$, when n, m large enough, we have

$$\|B_n^H - B_m^H\|_\infty \leq \sum_{k=n+1}^m \|B_k^H - B_{k-1}^H\|_\infty \leq \sum_{k=n+1}^m \max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \leq \sum_{k=n+1}^m \ell_k < \epsilon.$$

Thus, by definition, $\{B_n^H\}_{n \geq 1}$ is a Cauchy sequence in $\mathcal{C}([0, 1])$ almost surely. Since $\mathcal{C}([0, 1])$ is complete, there exists a stochastic process $X(t)$ such that $\|B_n^H - X\|_\infty$ converge to 0 almost surely.

We next show that $\{X(t)\}_{t \geq 0}$ is indeed a fBM. Consider an arbitrary finite collection of time points $(t_1^*, \dots, t_m^*) \in [0, 1]$. For each $1 \leq i \leq m$, there exists a sequence of points $t_n^{n_i} \in D_n$ such that $t_n^{n_i} \rightarrow t_i^*$. Note that

$$\begin{aligned} |X(t^*) - B_n^H(t_n^{n_i})| &\leq |X(t^*) - X(t_n^{n_i})| + |X(t_n^{n_i}) - B_n^H(t_n^{n_i})| \\ &\leq |X(t^*) - X(t_n^{n_i})| + \|X - B_n^H\|_\infty, \end{aligned}$$

and $B_n^H(t_n^{n_i})$ is Gaussian. It implies that $(X(t_1^*), \dots, X(t_m^*))$ is the strong limit of a sequence of Gaussian random vectors, and hence, is itself also Gaussian. Thus, $X(t)$ is a Gaussian process. Moreover, by the construction of $B_n^H(t)$, the covariance matrix of $(B_n^H(t_n^{n_1}), \dots, B_n^H(t_n^{n_m}))$ is $\Sigma^H(t_n^{n_1}, \dots, t_n^{n_m})$, where Σ^H is the covariance matrix function of fBM. Since Σ^H is continuous, we have $\Sigma^H(t_n^{n_1}, \dots, t_n^{n_m}) \rightarrow \Sigma^H(t_1^*, \dots, t_m^*)$, which implies that X has same covariance matrix function as fBM. Thus X is a fBM. \square

Theorem 2 indicates that the representation

$$B^H(t) - B_n^H(t) = \sum_{k=n+1}^{\infty} [B_k^H(t) - B_{k-1}^H(t)] \quad (13)$$

is well defined.

3.2. Convergence analysis in uniform norm and α -Hölder norm In this section, we study the convergence rate of (3). We first investigate the rate of convergence in uniform norm, which provides the basis for the ϵ -strong simulation of fBM. Then, for fBM with Hurst index $H > 1/2$, we strengthen our result under α -Hölder norm, which is necessary for the ϵ -strong simulation of fBM driven SDEs.

Recall our definition of the record-breaker and the last breaking time N . Based on our analysis in Section 3.1, we say that a record breaker happens at level k if

$$\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \geq \rho \ell_k = \rho 2^{-(H-\delta)k}, \quad (14)$$

and $N = \sup\{k \geq 1 : \max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \geq \rho \ell_k\}$. The following theorem shows that conditions C1) and C2) are satisfied for our definition of the record-breaker.

THEOREM 3. *For any fixed $\delta \in (0, H)$ and $t > 0$, $\mathbb{E}[\exp\{tN\}] < \infty$. When $n > N$,*

$$\|B^H - B_n^H\|_\infty \leq \frac{\rho \cdot 2^{-(H-\delta)(n+1)}}{(1 - 2^{-(H-\delta)})}.$$

Proof. The moment generating function of N can be written as

$$\mathbb{E}[\exp\{tN\}] = \mathbb{E}\left[\int_0^\infty 1\{\exp\{tN\} \geq u\} du\right] = \int_0^\infty \mathbb{P}(N \geq \log(u)/t) du.$$

We have

$$\begin{aligned} \mathbb{P}(N \geq \log(u)/t) &\leq \sum_{k=\lceil \log(u)/t \rceil}^{\infty} \mathbb{P}(\text{record broken at level } n) \\ &\leq \sum_{k=\lceil \log(u)/t \rceil}^{\infty} 2 \exp\{-\nu^{*2} \cdot 2^{2k\delta-2}\} \quad (\text{by Theorem 1}) \\ &\leq C \exp\{-\nu^{*2} \cdot u^{2\delta/t}\}, \end{aligned}$$

where C is a constant sufficiently large. Since $\exp\{-\nu^{*2} \cdot u^{2\delta/t}\}$ is integrable, $\mathbb{E}[\exp\{tN\}]$ is finite.

Now for $n > N$, according to the representation (13), we have

$$\|B^H - B_n^H\|_\infty \leq \rho \cdot \sum_{k=n+1}^{\infty} 2^{-(H-\delta)k} = \frac{\rho \cdot 2^{-(H-\delta)(n+1)}}{1 - 2^{-(H-\delta)}}.$$

□

For $0 < \alpha < 1$, the α -Hölder norm of a function $f(\cdot)$ over interval $[0, 1]$ is defined as

$$\|f\|_\alpha = \sup_{0 \leq s < t \leq 1} \frac{|f(s) - f(t)|}{|s - t|^\alpha}.$$

For the α -Hölder norm, we only consider fBM B^H with $H > 1/2$. Then for all $\alpha \in (1/2, H)$, the sample paths of B^H are α -Hölder continuous almost surely. By the representation (13), we have the following upper bound for the α -Hölder norm of discretization error

$$\|B_n^H - B^H\|_\alpha \leq \sum_{k=n+1}^{\infty} \|B_k^H - B_{k-1}^H\|_\alpha.$$

For each discretization level k , the following lemma gives a computable bound of $\|B_k^H(t) - B_{k-1}^H(t)\|_\alpha$.

LEMMA 4. For all $k \geq 1$, we have

$$\|B_k^H - B_{k-1}^H\|_\alpha \leq 2^{\alpha(k-1)+2} \cdot \max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k|.$$

The following theorem establishes convergence rate of (3) in the α -Hölder norm.

THEOREM 4. For any fixed $\alpha \in (1/2, H)$ and $\delta \in (0, H - \alpha)$, when $n > N$

$$\|B^H - B_n^H\|_\alpha \leq \frac{\rho 2^{2-\alpha} \cdot 2^{-(H-\alpha-\delta)(n+1)}}{1 - 2^{-(H-\alpha-\delta)}}.$$

Proof. By the definition of N , for all $k > N$, we have $\max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k| \leq \rho 2^{-(H-\delta)k}$. Then according to Lemma 4, we have

$$\|B_k^H - B_{k-1}^H\|_\alpha \leq 2^{\alpha(k-1)+2} \cdot \rho 2^{-(H-\delta)k} = \rho 2^{2-\alpha} \cdot 2^{-(H-\alpha-\delta)k}.$$

□

As a result of Theorem 4, once we find N , we also have an upper bound for the α -Hölder norm of the fBM sample path. Specifically,

$$\|B^H\|_\alpha \leq \|B_N^H\|_\alpha + \frac{\rho 2^{2-\alpha} \cdot 2^{-(H-\alpha-\delta)(N+1)}}{1 - 2^{-(H-\alpha-\delta)}}.$$

4. Simulation Algorithm In this section, we introduce our ϵ -strong simulation algorithm in details. Based on the theoretical foundation built in the previous section, our simulation algorithm includes two main steps. First, we simulate the fBM up to level N , where N is the level of the last record-breaker. Notice that once we find N , the truncation error at level $n > N$, is controlled by

$$\rho \cdot \sum_{k=n+1}^{\infty} 2^{-(H-\delta)k} = \frac{\rho \cdot 2^{-(H-\delta)(n+1)}}{1 - 2^{-(H-\delta)}}.$$

Second, we find the truncation level $N(\epsilon)$ such that $(1 - 2^{-(H-\delta)})^{-1} \cdot \rho \cdot 2^{-(H-\delta)N(\epsilon)} \leq \epsilon$. In this step, if $N(\epsilon) \leq N$, we have already obtained an ϵ -strong approximated sample path of fBM by simulating the path up to level N . Otherwise, we need to refine the path from level N to level $N(\epsilon)$. We summarize our main simulation algorithm in Algorithm 1. The details of the first step (finding

Algorithm 1 ϵ -Strong Simulation of fBM (SFBM)

- 1: **Input:** Hurst index H , simulation accuracy ϵ , record-breaker parameter ρ, δ .
 - 2: **Find the last record-breaker:**
 - 3: Call Algorithm 2 (SLRB): set $[N, \text{SP}] \leftarrow \text{SLRB}(H, \rho, \delta)$.
 - 4: **Find the truncation level:** set $N(\epsilon) \leftarrow \max\{N, \lceil \log_2(\rho \cdot (\epsilon(1 - 2^{-(H-\delta)}))^{-1}) / (H - \delta) \rceil\}$.
 - 5: **If** $N(\epsilon) > N$:
 - 6: **Simulate until the level $N(\epsilon)$ using acceptance-rejection method:**
 - 7: **Repeat:** sample fBM at $D_{N(\epsilon)}/D_N$, under the nominal measure conditional on the value of the fBM at D_N .
 - 8: **Until:** no record-breakers happen at levels $N + 1, \dots, N(\epsilon)$.
 - 9: Set $\text{SP} \leftarrow \text{Union}(\text{SP}, \text{value of fBM at } D_{N(\epsilon)}/D_N)$.
 - 10: **Output:** $B_{N(\epsilon)}^H(t)$, the piecewise linear interpolation of SP .
-

the last record-breaker) is further outlined in Algorithm 2. The second step (refining the dyadic

approximation up to the truncation level) involves simple acceptance-rejection method which is already detailed in Algorithm 1. We note that sampling the fBM at $D_{N(\epsilon)}$ given its values at D_N can be implemented straightforwardly by Cholesky decomposition given the conditional mean and covariance matrix. However, this implementation has high computational cost. We will provide a more efficient recursive implementation in Section 4.3.

In Algorithm 2, we find the record-breakers sequentially until the last one. Finding the next record-breaker is a challenging task, as the next record-breaker may never happen. We overcome the difficulty here by using techniques from rare-event simulation. Intuitively, breaking the record at level n is a rare event for large values of n . The details of how to find the next record-breaker are summarized in Algorithm 3.

Algorithm 2 Simulation of the Last Record-Breaker (SLRB)

- 1: **Input:** Hurst index H , record-breaker parameter ρ, δ .
 - 2: **Determine the starting level:**
 - 3: Calculate $N = N^*(\rho, \delta)$. Sample \mathbf{B}_N^H and store it in array SP.
 - 4: Set $I \leftarrow 1$.
 - 5: **While** $I = 1$:
 - 6: **Find the next record-breaker:**
 - 7: Call Algorithm 3 (SNRB): set $[I, N, \text{SP}] \leftarrow \text{SNRB}(N, \mathbf{B}_N^H, H, \rho, \delta)$
 - 8: **Output:** $[N, \text{SP}]$.
-

REMARK 1. In Algorithm 2, we need to start the dyadic approximation from a nontrivial starting level $N^*(\rho, \delta)$. $N^*(\rho, \delta)$ is defined later in equation (21) for technical reasons. Specifically, $N^*(\rho, \delta)$ is to ensure a proper bound on the likelihood ratio of the change-of-measure that we will apply in Algorithm 3. In practice, we can twist the record-breaker parameters, $(\rho, \delta, \nu, \nu^*)$, to obtain a reasonable starting level.

In Algorithm 3, we apply a change of measure technique to find the next record-breaker or claim that the record will never be broken again. The basic idea is as follows. Assuming that we have already simulated the fBM until level n . We denote by τ the level of the next record-breaker. Our goal is to find the next record-breaker or claim that $\tau = \infty$, which means the record will never be broken again. In order to determine whether $\tau < \infty$, we essentially want to generate a Bernoulli random variable with success probability $\mathbb{P}_n(\tau < \infty) := \mathbb{P}(\tau < \infty | \mathbf{B}_n^H)$. However, the exact value of $\mathbb{P}_n(\tau < \infty)$ is intractable. To overcome this difficulty, note that we can rewrite the probability as

$$\mathbb{P}_n(\tau < \infty) = \sum_{m=1}^{\infty} \mathbb{P}_n(\tau = n + m) = \sum_{m=1}^{\infty} \frac{\mathbb{P}_n(\tau = n + m)}{g_n(m)} \cdot g_n(m), \quad (15)$$

where $\{g_n(m)\}_{m \geq 1}$ is a carefully designed distribution taking value in \mathbb{Z}^+ such that

$$\mathbb{P}_n(\tau = n + m) / g_n(m) \leq 1,$$

for all $m \geq 1$. Note that $\{g_n(m)\}_{m \geq 1}$ can be interpreted as a potential realization of $\tau - n$.

We now introduce our specific choice of $\{g_n(m)\}_{m \geq 1}$.

$$g_n(m) = Z_n^{-1} \cdot 2^{n+m} \cdot \exp\{-\rho^2/8 \cdot 2^{2(n+m)\delta}\}, \quad (16)$$

where Z_n is the normalizing constant such that the sum of $g_n(m)$ equals to one. If we first generate M from $\{g_n(m)\}_{m \geq 1}$ and then, based on the value of M , generate a Bernoulli random variable with success probability $\mathbb{P}_n(\tau = n + M) / g_n(M)$. Then it is easy to see from equation (15) that we obtain

a Bernoulli random variable whose success probability is $\mathbb{P}_n(\tau < \infty)$. Moreover, if the Bernoulli trail is a success, we also know that the next record-breaker will happen at level M .

We next provide the general idea as how to generate a Bernoulli random variable with success probability $\mathbb{P}_n(\tau = n + M)/g_n(M)$. Given a realization $M = m$, in order to generate the desired Bernoulli random variable, we apply the change-of-measure technique. Note that $\mathbb{P}_n(\tau = n + m) = \mathbb{E}_n[1\{\tau = n + m\}]$. Thus, we first sample τ and the associated fBM sample paths ω from a properly constructed probability measure $\mathbb{Q}_n^{(m)}$. Then, we have

$$\mathbb{E}_{\mathbb{Q}_n^{(m)}} \left[\frac{d\mathbb{P}_n(\omega)}{d\mathbb{Q}_n^{(m)}(\omega)} \cdot \frac{1\{\tau = n + m\}}{g_n(m)} \right] = \frac{\mathbb{P}_n(\tau = n + m)}{g_n(m)}, \quad (17)$$

where $\mathbb{E}_{\mathbb{Q}_n^{(m)}}$ is the expectation with respect to $\mathbb{Q}_n^{(m)}$. If we can upper bound the likelihood ratio $d\mathbb{P}_n/d\mathbb{Q}_n^{(m)}$ by $g_n(m)$, then we generate U , a uniform random variable over $[0, 1]$ independent of everything else. When

$$U < \frac{d\mathbb{P}_n(\omega)}{d\mathbb{Q}_n^{(m)}(\omega)} \cdot \frac{1\{\tau = n + m\}}{g_n(m)},$$

we accept ω as the trajectory leading to the next record-breaker, i.e. we get \mathbf{B}_τ^H ; otherwise, we claim that $\tau = \infty$.

The actual construction of change-of-measure involves more subtleties. For example, we not only need to consider the level of the next record-breaker, $n + m$, but also the actual time index, and whether we break the record due to a large positive deviation or negative deviation. The details of these subtleties are deferred to Section 4.1. It is also in general not easy to bound the likelihood ratio $d\mathbb{P}_n/d\mathbb{Q}_n^{(m)}$. We thus introduce another technical step before we give the actual algorithm. We define the ‘‘bounded conditional expectation’’ condition (BCE). We will later show in Section 4.1 that under this condition, the likelihood ratio is properly bounded. In addition, this condition only gets violated a finite number of times almost surely.

DEFINITION 1. (BCE condition) We say that \mathbf{B}_n^H satisfies the bounded conditional expectation condition, if for all $m \geq 1$ and $1 \leq k \leq 2^{n+m-1}$,

$$|\mu_n(m, k)| = |\boldsymbol{\beta}^\top \mathbb{E}[\boldsymbol{\alpha}_n(m, k) | \mathbf{B}_n^H]| \leq \rho/2 \cdot \ell_{n+m}, \quad (18)$$

where

$$\boldsymbol{\beta} = (1/2, -1, 1/2)^\top, \quad \boldsymbol{\alpha}_n(m, k) = (B^H(t_{2k-2}^{n+m}), B^H(t_{2k-1}^{n+m}), B^H(t_{2k}^{n+m}))^\top. \quad (19)$$

When simulating the next record-breaker, we will first check if \mathbf{B}_n^H satisfies the BCE condition. If not, we will keep generating more refined levels under the nominal measure until the BCE condition is satisfied, before we apply the change-of-measure. The details of how to check whether the BCE condition is met are laid out in Algorithm 5, which we defer to Section 4.1 after we introduce a few more technical results. Similarly, we also defer the details of the change-of-measure to Algorithm 5 in Section 4.1.

4.1. Change of measure In this section, we provide details of our construction of a new measure under which the record-breaker is more likely to happen. Recall that the setting is that we have already generated \mathbf{B}_n^H and a proposed next record breaking level $n + m$, where m is sampled from distribution $\{g_n(m)\}_{m \geq 1}$. Our goal is to find a way to generate a path such that the next record-breaker is more likely to happen at level $n + m$, and the likelihood ratio can be properly bounded.

Algorithm 3 Simulation of the Next Record-Breaker (SNRB)

```

1: Input: Current level  $n$  and values  $\mathbf{B}_n^H$ , Hurst index  $H$ , record-breaker parameter  $\rho, \delta$ .
2: Initialize: Set AP, an array of values of fBM at augmented points, to be NULL, and  $J \leftarrow 0$ .
3: Refine dyadic approximation until BCE condition is satisfied:
4:   While  $J = 0$ :
5:     Checking BCE condition:
6:       Call Algorithm 5 (BCEC): set  $J \leftarrow \text{BCEC}(H, \rho, \delta, n, \mathbf{B}_n^H)$ .
7:       If  $J = 1$ : break.
8:     Refine dyadic approximation to next level:
9:       Sample fBM at  $D_{n+1}/D_n$ , under the nominal measure conditional on  $\mathbf{B}_n^H$ , and
       then store it in AP. Update  $n \leftarrow n + 1$ ,  $\mathbf{B}_n^H \leftarrow \text{Union}(\mathbf{B}_n^H, \text{AP})$ .
10:      If a record-breaker happens at level  $n$ : break.
11: If  $J = 1$ , apply change-of-measure:
12:   Sample  $M$  from distribution  $\{g_n(m)\}_{m \geq 1}$ .
13:   Call Algorithm 4 (ECM): set  $[I, \text{AP}] \leftarrow \text{ECM}(H, \rho, \delta, n, \mathbf{B}_n^H, M)$ .
14: Output:
15:   If  $J = 0$  : return  $[1, n, \mathbf{B}_n^H]$ ,
16:   Else if  $I = 1$  : return  $[1, n + M, \text{Union}(\mathbf{B}_n^H, \text{AP})]$ ,
17:   Else: return:  $[0, n, \mathbf{B}_n^H]$ .

```

Recall the definition of $\alpha_n(m, k)$ and β in equation (19). Then based on the definition of the record-breakers, we say that a record is broken at level $n + m$, position k , if

$$|\beta^\top \alpha_n(m, k)| > \rho \ell_{n+m} = \rho 2^{-(H-\delta)(n+m)}.$$

Furthermore, we say that the record-breaker is up-crossing if

$$\beta^\top \alpha_n(m, k) > \rho \ell_{n+m},$$

and downward-crossing if

$$\beta^\top \alpha_n(m, k) < -\rho \ell_{n+m}.$$

Let

$$\Xi_n^{(m,k)}(\theta) = \mathbb{E}_n[\exp\{\theta \cdot \beta^\top \alpha_n(m, k)\}] = \mathbb{E}[\exp\{\theta \cdot \beta^\top \alpha_n(m, k)\} | \mathbf{B}_n^H],$$

which is the moment generating function of $\beta^\top \alpha_n(m, k)$ conditional on the value of \mathbf{B}_n^H . We also assume that the conditional probability density of $\alpha_n(m, k)$ under measure $\mathbb{P}_n(\cdot)$ is $\psi_n^{(m,k)}$. We also denote the conditional covariance matrix and expectation of $\alpha_n(m, k)$ by $\Sigma_{\alpha_n(m,k)}$ and $\mu_n(m, k)$, respectively.

In what follows, we start by introducing the change-of-measure under the BCE condition. We then show that the BCE condition can only be violated a finite number of times almost surely. Under the BCE condition, we first sample K from the set $\{1, 2, \dots, 2^{n+m-1}\}$ uniformly. The random variable K roughly proposes the position of the next record-breaker. We also sample π from the set $\{+, -\}$ uniformly. The random variable π roughly proposes whether the record-breaker is up-crossing (+) or downward-crossing (-). Second, given $M = m, K = k$, if $\pi = +$, we apply exponential tilting to $\psi_n^{(m,k)}$ with tilting parameter

$$\theta_n^+(m) = \rho/2 \cdot 2^{(m+n)(H+\delta)2}. \quad (20)$$

² The tilting parameter $\theta_n^+(m)$ is carefully chose to make sure that the record-breaking event is more likely to happen under the tilted measure and the likelihood ratio is suitably bounded. See Section 4.2 for more details.

Specifically, we sample $\alpha_n(m, k)$ from the density

$$\tilde{\psi}_n^{(m,k,+)}(x_1, x_2, x_3) = \psi_n^{(m,k)}(x_1, x_2, x_3) \cdot \exp\left\{\theta_n^+(m) \cdot \left(\frac{1}{2}(x_1 + x_3) - x_2\right) - \log(\Xi_n^{(m,k)}(\theta_n^+(m)))\right\},$$

Note that the tilted distribution $\tilde{\psi}_n^{(m,k,+)}$ is still Gaussian. In particular, $\tilde{\psi}_n^{(m,k,+)}$ is the density of the multivariate Gaussian

$$N(\boldsymbol{\mu}_n(m, k) + \theta_n^+(m) \cdot \boldsymbol{\Sigma}_{\alpha_n(m,k)} \boldsymbol{\beta}, \boldsymbol{\Sigma}_{\alpha_n(m,k)}).$$

If $\pi = -$, we apply the exponential tilting to $\psi_n^{(m,k)}$ with tilting parameter $\theta_n^-(m) = -\theta_n^+(m)$.

After we have sampled $\alpha_n(m, k)$ under the tilted measure, given the values of $\alpha_n(m, k)$ and \mathbf{B}_n^H , we sample fBM at the remained dyadic points

$$D_{n+m} / (D_n \cup \{t_{2k-2}^{n+m}, t_{2k-1}^{n+m}, t_{2k}^{n+m}\})$$

under the nominal measure. This step is achieved by calculating the conditional expectation and covariance matrix, and then sampling from the corresponding multivariate Gaussian distribution. We use \mathbb{Q}_n to denote the tilted measure introduced as above. We also denote $\mathbb{Q}_n^{(m,k,+)}$ as the conditional probability measure $\mathbb{Q}_n(\cdot | M = m, K = k, \pi = +)$. Let

$$\begin{aligned} \Theta_n^+(m, k) &:= d\mathbb{P}_n / d\mathbb{Q}_n^{(m,k,+)} \cdot g_n(m)^{-1} \cdot 2^{n+m} \\ &= g_n(m)^{-1} \cdot 2^{n+m} \cdot \exp\left\{-\theta_n^+(m) \cdot \boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) + \log(\Xi_n^{(m,k)}(\theta_n^+(m)))\right\}. \end{aligned}$$

Similarly, we define $\mathbb{Q}_n^{(m,k,-)}$ as the conditional measure $\mathbb{Q}_n(\cdot | M = m, K = k, \pi = -)$ and

$$\begin{aligned} \Theta_n^-(m, k) &:= d\mathbb{P}_n / d\mathbb{Q}_n^{(m,k,-)} \cdot g_n(m)^{-1} \cdot 2^{n+m} \\ &= g_n(m)^{-1} \cdot 2^{n+m} \cdot \exp\left\{-\theta_n^-(m) \cdot \boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) + \log(\Xi_n^{(m,k)}(\theta_n^-(m)))\right\}. \end{aligned}$$

From the definition of $g_n(m)$ in (16), the normalizing constant Z_n is given by

$$Z_n = \sum_{m=1}^{\infty} 2^{n+m} \cdot \exp\{-\rho^2/8 \cdot 2^{2(n+m)\delta}\}.$$

We denote by

$$N^*(\rho, \delta) = 1 + \sup\{n \geq 1 : Z_n > 1\}. \quad (21)$$

Note that for all $n \geq N^*(\rho, \delta)$, we have $Z_n \leq 1$. The next lemma shows that under the BCE condition, $\Theta_n^+(m, k)$'s and $\Theta_n^-(m, k)$'s are suitably bounded. This result is important in constructing our change-of-measure procedure.

LEMMA 5. *For $n \geq N^*(\rho, \delta)$, under the BCE condition, when $M = m$, $K = k$, and $\pi = +$,*

$$\Theta_n^+(m, k) \cdot 1 \left\{ (1/2, -1, 1/2)^\top \boldsymbol{\alpha}_n(m, k) > \rho \ell_{n+m} \right\} \leq 1;$$

when $M = m$, $K = k$, and $\pi = -$,

$$\Theta_n^-(m, k) \cdot 1 \left\{ (1/2, -1, 1/2)^\top \boldsymbol{\alpha}_n(m, k) < -\rho \ell_{n+m} \right\} \leq 1.$$

We're now ready to present our actual algorithm. We denote by

$$R_{n+m} = \sum_{k=1}^{2^{n+m-1}} 1\{|\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k)| > \rho \ell_{n+m}\},$$

the total number of record-breakers at level $n+m$. We also generate a uniformly distributed random variable U over interval $[0, 1]$, independent of everything else. When $\pi = +$, we return 1, if

$$U < \Theta_n^+(m, k) \cdot R_{n+m}^{-1} \cdot 1\left\{\left\{\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) > \rho \ell_{n+m}\right\} \cap \bigcap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\right\},$$

and return 0 otherwise³. Here $\mathcal{C}_n(j)$ denotes the event that there are record-breakers at level $n+j$. When $\pi = -$, the procedure is similar. The details of this simulation procedure are summarized in Algorithm 4 and we prove that it works in Section 4.2.

Algorithm 4 Exponential Change of Measure (ECM)

- 1: **Input:** Hurst index H , record-breaker parameters ρ, δ , current level n and values \mathbf{B}_n^H , level difference m .
 - 2: **Initialize:** set AP, the array of values of fBM at augmented points D_{n+m}/D_n , to be NULL.
 - 3: **Sample position of record-breaker:** sample $K = k$ uniformly from $\{1, 2, \dots, 2^{n+m-1}\}$.
 - 4: **Determine up-crossing or downward-crossing:** sample π uniformly from $\{+, -\}$.
 - 5: **Sample candidate path:**
 - 6: **If** $\pi = +$: sample $\boldsymbol{\alpha}_n(m, k)$ from the exponential tilted measure $\tilde{\psi}_n^{(m, k, +)}$.
 - 7: **If** $\pi = -$: sample $\boldsymbol{\alpha}_n(m, k)$ from the exponential tilted measure $\tilde{\psi}_n^{(m, k, -)}$.
 - 8: Sample fBM at the remained points until the discretization level $n+m$.
 - 9: **Calculate likelihood ratio:**
 - 10: **If** $\pi = +$: calculate $\Theta_n^+(m, k)$.
 - 11: **If** $\pi = -$: calculate $\Theta_n^-(m, k)$.
 - 12: **Calculate total number of record-breakers at level $n+m$:**
 - 13: Set $R_{n+m} \leftarrow \sum_{k=1}^{2^{n+m-1}} 1\{|\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k)| > \rho \ell_{n+m}\}$.
 - 14: **Sample from uniform distribution:** generate $U \sim \mathcal{U}[0, 1]$.
 - 15: **Determine the next record-breaker via acceptance-rejection:**
 - 16: **If** $\pi = +$: set

$$I \leftarrow 1\left\{U < \Theta_n^+(m, k) \cdot R_{n+m}^{-1} \cdot 1\left\{\left\{\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) > \rho \ell_{n+m}\right\} \cap \bigcap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\right\}\right\}.$$
 - 17: **If** $\pi = -$: set

$$I \leftarrow 1\left\{U < \Theta_n^-(m, k) \cdot R_{n+m}^{-1} \cdot 1\left\{\left\{\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) < -\rho \ell_{n+m}\right\} \cap \bigcap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\right\}\right\}.$$
 - 18: **If** $I = 1$: AP=values of fBM at augmented points D_{n+m}/D_n .
 - 19: **Output:** I and AP.
-

It remains to show that the BCE condition is only violated a finite number times. In addition, we also need to have an efficient way to check whether the BCE condition is satisfied at a specific level n .

Let

$$\mathcal{E}_n = \{|\mu_n(m, k)| > \rho/2 \cdot 2^{-(n+m)(H-\delta)}, \text{ for some } m \geq 1 \text{ and } 1 \leq k \leq 2^{n+m-1}\},$$

³ We define $0/0 = 0$ by convention.

Algorithm 5 BCE Checking (BCEC)

-
- 1: **Input:** Hurst index H , record-breaker parameters ρ, δ , current level n and values \mathbf{B}_n^H .
 - 2: **Initialize:** set $J \leftarrow 1$.
 - 3: **Determine maximal checking level:**
 - 4: Set $\gamma_n \leftarrow \max_{1 \leq i \leq 2^{n+1}} |(\boldsymbol{\Sigma}_n^{-1} \mathbf{B}_n^H)_i|$.
 - 5: Set $M_n \leftarrow \max\{1, \lceil (H - \delta)^{-1} \cdot \log_2((2^{n+1} + 2)\gamma_n/\rho) - n \rceil\}$.
 - 6: **For** $1 \leq m \leq M_n$:
 - 7: **For** $1 \leq k \leq 2^{n+m-1}$:
 - 8: Calculate conditional expectation $\mu_n(m, k)$.
 - 9: **If** $|\mu_n(m, k)| \geq \rho/2 \cdot 2^{-(n+m)(H-\delta)}$: set $J \leftarrow 0$, break.
 - 10: **Output:** Return J .
-

where $\mu_n(m, k)$ is the conditional expectation defined in (18). By Definition 1, \mathcal{E}_n denotes the event that the BCE condition is violated at level n . We also define

$$\mathcal{N}_{\mathcal{E}} = \sup\{n \geq 1 : \mathcal{E}_n \text{ happens}\}$$

LEMMA 6. *The BCE condition is violated a finite number times almost surely, i.e.*

$$\mathbb{P}\left(\limsup_{n \rightarrow \infty} \mathcal{E}_n\right) = 0.$$

Moreover, $\mathcal{N}_{\mathcal{E}}$ has a finite moment generating function.

We next introduce the algorithm to check whether the BCE condition is satisfied at level n given the value of \mathbf{B}_n^H . The development of the algorithm is highly nontrivial. In particular, directly checking by definition is impractical, as we need to evaluate $\mu_n(m, k)$ for infinitely many terms. To overcome this difficulty, we need to further explore the structure of $\mu_n(m, k)$. We use γ_n to denote the maximal absolute value of the entries in vector $\boldsymbol{\Sigma}_n^{-1} \mathbf{B}_n^H$, where $\boldsymbol{\Sigma}_n$ is the covariance matrix of \mathbf{B}_n^H . We have the following lemma to characterize the bound for $\mu_n(m, k)$.

LEMMA 7. *For all $m \geq 1$ and $1 \leq k \leq 2^{n+m-1}$, we have*

$$|\mu_n(m, k)| \leq \gamma_n \cdot (2^n + 1) \cdot 2^{-2(n+m)H}.$$

According to Lemma 7, for a fixed level n and values \mathbf{B}_n^H , the decay rate of $|\mu_n(m, k)|$ (with respect to m) is $O(2^{-2mH})$. However, to check the BCE condition, we only need to compare $|\mu_n(m, k)|$ with $\rho/2 \cdot 2^{-(n+m)(H-\delta)}$, whose decay rate (with respect to m) is $O(2^{-m(H-\delta)})$. Hence, when

$$m \geq M_n = \max\left\{1, \lceil (H - \delta)^{-1} \cdot \log_2((2^{n+1} + 2)\gamma_n/\rho) - n \rceil\right\},$$

the following inequality

$$|\mu_n(m, k)| \leq \rho/2 \cdot 2^{-(n+m)(H-\delta)}$$

always holds. This implies that to check whether BCE condition holds at level n , we only need to calculate a finite collection of $\mu_n(m, k)$, i.e. $m = 1, \dots, M_n$. The details is summarized in Algorithm 5, which outputs 1 when the BCE condition is satisfied and 0, otherwise.

4.2. Proof of algorithm In this section, we provide detailed proof to show that our algorithm actually works. The proof is divided into two main steps (two Theorems). We first show that under the BCE condition, for the output of Algorithm 3, $\mathbb{Q}_n(M = m, I = 1) = \mathbb{Q}_n(\tau = n + m)$ and $\mathbb{P}(I = 0) = \mathbb{P}_n(\tau = \infty)$ (Theorem 5), where \mathbb{Q}_n denotes the measure induced by Algorithm 3. We then show that when $M = m$ and $I = 1$, the output path, i.e. the fBM on the augmented points D_{n+m}/D_n , follows from the distribution $\mathbb{P}_n(\cdot|\tau = n + m)$ (Theorem 6).

THEOREM 5. *For the output of Algorithm 3, when $n \geq N^*(\rho, \delta)$ and the BCE condition holds, I , is a Bernoulli random variable with success probability $\mathbb{P}_n(\tau < \infty)$. Moreover,*

$$\mathbb{Q}_n(M = m, I = 1) = \mathbb{P}_n(\tau = n + m).$$

Proof. By definition, we have

$$\mathbb{Q}_n(I = 1) = \sum_{m=1}^{\infty} \mathbb{Q}_n(I = 1, M = m) = \sum_{m=1}^{\infty} \mathbb{Q}_n(I = 1|M = m) \cdot g_n(m).$$

We next show that $\mathbb{Q}_n(I = 1|M = m) = \mathbb{P}_n(\tau = n + m)/g_n(m)$. Note that if X is a random variable taking value in $[0, 1]$ and U is a uniformly distributed random variable on $[0, 1]$ independent of X , then $1\{U < X\}$ is a Bernoulli random variable of success probability $\mathbb{E}[X]$. Recall that in our algorithm, both $\Theta_n^+(m, k)$ and $\Theta_n^-(m, k)$ are bounded by one. Then by the definition of I and Lemma 5, we have the following decomposition

$$\begin{aligned} & \mathbb{Q}_n(I = 1|M = m) \\ &= \mathbb{E}_{\mathbb{Q}_n^{(m)}} \left[1 \{ \{\beta^\top \alpha_n(m, K) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot \Theta_n^+(m, K) \cdot R_{n+m}^{-1} \Big| \pi = + \right] \cdot \frac{1}{2} \\ & \quad + \mathbb{E}_{\mathbb{Q}_n^{(m)}} \left[1 \{ \{\beta^\top \alpha_n(m, K) < -\ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot \Theta_n^-(m, K) \cdot R_{n+m}^{-1} \Big| \pi = - \right] \cdot \frac{1}{2} \\ &= \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m, k, +)}} \left[1 \{ \{\beta^\top \alpha_n(m, k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot \Theta_n^+(m, k) \cdot R_{n+m}^{-1} \right] \cdot \frac{1}{2^{n+m}} \\ & \quad + \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m, k, -)}} \left[1 \{ \{\beta^\top \alpha_n(m, k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot \Theta_n^+(m, k) \cdot R_{n+m}^{-1} \right] \cdot \frac{1}{2^{n+m}}. \end{aligned}$$

By the definition of weighted likelihood ratio $\Theta_n^+(m, k)$ and $\Theta_n^-(m, k)$, we further have

$$\begin{aligned} & \mathbb{Q}_n(I = 1|M = m) \\ &= \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m, k, +)}} \left[1 \{ \{\beta^\top \alpha_n(m, k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot R_{n+m}^{-1} \cdot \frac{d\mathbb{P}_n}{d\mathbb{Q}_n^{(m, k, +)}} \cdot g_n(m)^{-1} \right] \\ & \quad + \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m, k, -)}} \left[1 \{ \{\beta^\top \alpha_n(m, k) < -\ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot R_{n+m}^{-1} \cdot \frac{d\mathbb{P}_n}{d\mathbb{Q}_n^{(m, k, -)}} \cdot g_n(m)^{-1} \right]. \end{aligned}$$

Note that as

$$1 \{ |\beta^\top \alpha_n(m, k)| > \ell_{n+m} \} = 1 \{ \beta^\top \alpha_n(m, k) > \ell_{n+m} \} + 1 \{ \beta^\top \alpha_n(m, k) < -\ell_{n+m} \},$$

we finally have

$$\begin{aligned} \mathbb{Q}_n(I = 1|M = m) &= \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_n \left[1 \{ \{ |\beta^\top \alpha_n(m, k)| > \ell_{n+m} \} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot R_{n+m}^{-1} \right] \cdot g_n(m)^{-1} \\ &= \mathbb{E}_n \left[1 \{ \mathcal{C}_n(m) \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \right] = \mathbb{P}_n(\tau = n + m) / g_n(m). \end{aligned}$$

Above all, for any $m \geq 1$, we have $\mathbb{Q}_n(U = 1, M = m) = \mathbb{P}_n(\tau = n + m)$. \square

Based on Theorem 5, if we obtain an output $I = 0$ from Algorithm 3, we claim that the record-breaker will not happen again after level n . Otherwise, i.e. if we obtain $I = 1$ and associated $M = m$ from Algorithm 3, we claim that the first next record-breaker happens at level $n + m$. In the later case, Algorithm 4 also outputs a path leading to the next record-breaker. We next show that the output of Algorithm 4 when $I = 1$ is a realization of the fBM conditional on that the next record-breaker happens at level $n + m$.

THEOREM 6. *For the output of Algorithm 4, given $M = m$, $I = 1$, the distribution of the values of fBM on the augmented points D_{n+m}/D_n , follows $\mathbb{P}_n(\cdot | \tau = n + m)$.*

Proof. Let $\delta D_n^m = D_{n+m}/D_n$. We consider a sequence of measurable sets H_j where j is chosen such that $t_j \in \delta D_n^m$. We only need to show that

$$\mathbb{P}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m | \tau = n + m) = \mathbb{Q}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m | I = 1, M = m).$$

By definition, we have

$$\begin{aligned} \mathbb{P}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m | \tau = n + m) &= \mathbb{P}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m, \tau = n + m) / \mathbb{P}_n(\tau = n + m) \\ &= \mathbb{E}_n \left[\prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \cdot 1\{\tau = n + m\} \right] / \mathbb{P}_n(\tau = n + m). \end{aligned}$$

Similar to the proof of Theorem 5, we have

$$\begin{aligned} &\mathbb{E}_n \left[\prod_{\substack{t_j \in \delta D_n^m \\ 2^{n+m-1}}} 1\{B^H(t_j) \in H_j\} \cdot 1\{\tau = n + m\} \right] \\ &= \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_n \left[\prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \cdot R_{n+m}^{-1} \cdot 1\{\{\beta^\top \alpha_n(m, k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\} \right] \\ &\quad + \sum_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \cdot R_{n+m}^{-1} \cdot 1\{\{\beta^\top \alpha_n(m, k) < -\ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\}. \end{aligned}$$

The first term in above equation, corresponding to up-crossing record-breaker, and

$$\begin{aligned} &\sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m, k, +)}} \left[\prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \right. \\ &\quad \left. \times 1\{\{\beta^\top \alpha_n(m, k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\} \cdot R_{n+m}^{-1} \cdot d\mathbb{P}_n / d\mathbb{Q}_n^{(m, k, +)} \right] \\ &= \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m, k, +)}} \left[\prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \right. \\ &\quad \left. \cdot 1\{\{\beta^\top \alpha_n(m, k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c\} \cdot R_{n+m}^{-1} \cdot \Theta_n^+(m, k) \right] \cdot g_n(m) / 2^{n+m}. \end{aligned}$$

Similar equation holds for the second term, corresponding to downward-crossing record-breaker.

On the other hand, we have

$$\begin{aligned} &\mathbb{Q}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m | I = 1, M = m) \\ &= \mathbb{Q}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m, I = 1 | M = m) \cdot g_n(m) / \mathbb{Q}_n(I = 1, M = m) \end{aligned}$$

We notice that

$$\begin{aligned} &\mathbb{Q}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m, I = 1 | M = m) = \mathbb{E}_{\mathbb{Q}_n} \left[1\{I = 1\} \cdot \prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} | M = m \right] \\ &= \sum_{k=1}^{2^{n+m-1}} \sum_{\pi=+, -} \mathbb{E}_{\mathbb{Q}_n^{(m, k, \pi)}} \left[1\{I = 1\} \prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \right] \cdot \frac{1}{2^{n+m}}. \end{aligned}$$

Now as

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}_n^{(m,k,+)}} \left[1\{I = 1\} \mid \mathbf{B}_{n+m}^H \right] \\ &= \mathbb{E}_{\mathbb{Q}_n^{(m,k,+)}} \left[1\{ \{\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m,k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot \Theta_n^+(m,k) \cdot R_{n+m}^{-1} \mid \mathbf{B}_{n+m}^H \right], \end{aligned}$$

for the up-crossing part, we have

$$\begin{aligned} & \sum_{k=1}^{2^{n+m}} \mathbb{E}_{\mathbb{Q}_n^{(m,k,+)}} \left[1\{I = 1\} \cdot \prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \right] = \sum_{k=1}^{2^{n+m-1}} \mathbb{E}_{\mathbb{Q}_n^{(m,k,+)}} \left[\prod_{t_j \in \delta D_n^m} 1\{B^H(t_j) \in H_j\} \right. \\ & \quad \left. \times 1\{ \{\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m,k) > \ell_{n+m}\} \cap \cap_{j=1}^{m-1} [\mathcal{C}_n(j)]^c \} \cdot R_{n+m}^{-1} \cdot \Theta_n^+(m,k) \right] \cdot \frac{1}{2^{n+m}}. \end{aligned}$$

Similar result holds for the downward-crossing part. Finally, by Theorem 5, $\mathbb{P}_n(\tau = n+m) = \mathbb{Q}_n(U = 1, M = m)$. Hence, we have

$$\mathbb{P}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m \mid \tau = n+m) = \mathbb{Q}_n(B^H(t_j) \in H_j, t_j \in \delta D_n^m \mid U = 1, M = m).$$

□

Theorem 5, together with Theorem 6, justifies the correctness of our algorithm.

4.3. Computational complexity analysis In this section, we analyze the computational complexity⁴ of our algorithm. Note that the main component in our algorithm is to generate the last record-breaker and the associated path of the fBM. Once we have found the last record-breaker, to achieve an ϵ error bound, the complexity is $O(2^{3N(\epsilon)})$ if we use the naive Cholesky decomposition to sample the fBM at $D_{N(\epsilon)}/D_N$. Using a recursive construction of fractional Brownian bridge, we can reduce the complexity to

$$O(2^{N(\epsilon)} \log(2^{N(\epsilon)})) = O(\epsilon^{-1/(H-\delta)} \log(\epsilon^{-1/(H-\delta)})).$$

We provide more details about this recursive construction towards the end of this section.

We use \bar{N} to denote the last level we need to generate in order to determine the level of the last record-breaker. Note that in our algorithm, in order to apply the change-of-measure technique, we need to refine the dyadic approximation until the BCE condition is satisfied. Thus, $\bar{N} \geq N$. We also denote by $C_{\bar{N}}$ the associated computational complexity. This includes the cost of determining that there will be no more record-breakers. The following theorem establishes that the computational cost is finite in expectation.

THEOREM 7. *For the cost of generating the last record-breaker in Algorithm 2, $C_{\bar{N}}$, we have*

$$\mathbb{E}[C_{\bar{N}}] < \infty.$$

Proof. Recall that in our algorithm, the computational complexity arises from two main procedures. The first is finding the next record-breaker or claiming that there is no record-breakers any more. The second is refining dyadic approximation until the BCE condition is satisfied.

Recall that N is the time of the last record-breaker, which is also an upper bound for the number of while loops in Algorithm 2. We also recall that \mathcal{N}_ϵ is the last level the BCE condition is violated. Then we have

$$\bar{N} \leq \max\{N, \mathcal{N}_\epsilon + 1\}.$$

⁴We refer to the computational complexity as the total number of uniform random variables we need to generate and the number of basic calculations. For example, the Cholesky decomposition of an $n \times n$ covariance matrix has a computational complexity of $O(n^3)$.

At current level n , we first analyze the computational complexity of checking the BCE condition. For now on, for simplicity, we use C to denote a generic constant, which may differ from line to line. Based on Algorithm 5, we need to calculate the conditional expectation until level $n + M_n = O(n + \log_2(\gamma_n))$, whose computational complexity is at most $O(2^{3n}\gamma_n^3)$. If the BCE condition is satisfied, then the computational complexity of applying the change-of-measure is of order

$$\sum_{m=1}^{\infty} g_n(m) \cdot c_n(m) \leq CZ_n^{-1} \cdot \sum_{m=1}^{\infty} 2^{4(n+m)} \cdot \exp\{-\rho^2/8 \cdot 2^{2(n+m)\delta}\} \leq C \cdot 2^{3n},$$

where $c_n(m) = O(2^{3(n+m)})$, which denotes the computational complexity of Algorithm 3 conditional on $M = m$. If the BCE condition is broken, we refine the dyadic approximation until a level $n^* \leq \mathcal{N}_\varepsilon + 1$ where the BCE condition is satisfied. For each refinement level, we need to check the BCE condition. Hence, the corresponding computational complexity is $O(\sum_{j=n+1}^{n^*} 2^{3j}\gamma_j^3)$.

Above all, the total computational complexity for a while loop in Algorithm 5 can be upper bounded by

$$C \cdot \left(\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{3j}\gamma_j^3 + 2^{3\bar{N}} \right),$$

As there are at most \bar{N} loops. We have

$$C_{\bar{N}} \leq C \cdot \left(\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{3j}\gamma_j^3 + 2^{3\bar{N}} \right) \cdot \bar{N} = C \sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{3j}\gamma_j^3 \cdot \bar{N} + C2^{3\bar{N}} \cdot \bar{N}. \quad (22)$$

Since N and \mathcal{N}_ε have finite moment generating function, so it is \bar{N} . Then the second term in (22) is bounded in expectation. We next establish an upper bound for the first term. By Cauchy-Schwartz inequality, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{3j}\gamma_j^3 \cdot \bar{N} \right] &\leq \left(\mathbb{E} \left[\left(\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{6j} \right) \cdot \left(\sum_{j=1}^{\mathcal{N}_\varepsilon+1} \gamma_j^6 \right) \right] \right)^{1/2} \cdot (\mathbb{E}[\bar{N}^2])^{1/2} \\ &\leq C \cdot (\mathbb{E}[\bar{N}^2])^{1/2} \cdot \left(\mathbb{E} \left[\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{6\mathcal{N}_\varepsilon} \cdot \gamma_j^6 \right] \right)^{1/2}. \end{aligned}$$

For the last term in above inequality, recall that γ_n is the maximal absolute value of the entries in the vector $\Sigma_n^{-1} \mathbf{B}_n^H$, which follows multivariate normal distribution $N(0, \Sigma_n^{-1})$. By Fubini's Theorem and Cauchy-Schwartz inequality, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{6\mathcal{N}_\varepsilon} \cdot \gamma_j^6 \right] &= \mathbb{E} \left[\sum_{j=1}^{\infty} 2^{6\mathcal{N}_\varepsilon} \cdot \gamma_j^6 \cdot \mathbf{1}\{j \leq \mathcal{N}_\varepsilon + 1\} \right] \\ &= \sum_{j=1}^{\infty} \mathbb{E} [2^{6\mathcal{N}_\varepsilon} \cdot \gamma_j^6 \cdot \mathbf{1}\{j \leq \mathcal{N}_\varepsilon + 1\}] \leq \left(\mathbb{E} [2^{12\mathcal{N}_\varepsilon}] \right)^{1/2} \cdot \sum_{j=1}^{\infty} \left(\mathbb{E} [\gamma_j^{12} \cdot \mathbf{1}\{j \leq \mathcal{N}_\varepsilon + 1\}] \right)^{1/2} \\ &\leq \left(\mathbb{E} [2^{12\mathcal{N}_\varepsilon}] \right)^{1/2} \cdot \sum_{j=1}^{\infty} (\mathbb{E}[\gamma_j^{24}])^{1/4} \cdot (\mathbb{P}(\mathcal{N}_\varepsilon \geq j-1))^{1/4}. \end{aligned}$$

It is easy to see that $\mathbb{E}[\gamma_j^{24}] \leq C \cdot 2^j \cdot 2^{24Hj} = C \cdot 2^{(1+24H)j}$. Similar to the proof of Theorem 3, using the decay rate of $\mathbb{P}(\mathcal{N}_\varepsilon \geq j)$ proved in Lemma 6, we obtain $\mathbb{E}[\sum_{j=1}^{\mathcal{N}_\varepsilon+1} 2^{6\mathcal{N}_\varepsilon} \cdot \gamma_j^6] < \infty$. Hence, $\mathbb{E}[C_{\bar{N}}] < \infty$, i.e. our algorithm has finite expected computational complexity. \square

In our algorithm, once we have found the level of the last record-breaker N , we only need to refine the dyadic approximation until the desired truncation level $N(\epsilon)$, conditioning on that the record-breakers do not happen beyond N . Note that N does not depend on ϵ . Thus, N can be treated as a constant. Assuming that $N(\epsilon) > N$, then we need to sample the fBM at time points $D_{N(\epsilon)}/D_N$ conditional on \mathbf{B}_N^H , which involves $O(2^{N(\epsilon)})$ correlated Gaussian random variables. If we sample naively, i.e., calculating the conditional distribution first and sample from it, the computational complexity is $O(2^{3N(\epsilon)})$. This complexity can be reduced to $O(2^{N(\epsilon)} \log(2^{N(\epsilon)}))$ using a recursive construction of Gaussian bridge. We next introduce the details of this recursive construction. The algorithm is based on the Gaussian bridge construction developed in [Sottinen and Yazigi \(2014\)](#). We summarize the main idea in the following lemma.

LEMMA 8. *Let $\{X_t\}_{t \geq 0}$ be a stationary Gaussian process with covariance function $r(s, t)$. Then the distribution of X_t conditional on that $X_{t_k} = y_k, k = 1, \dots, n$ is same with that of X_t^n , which is defined recursively as*

$$\begin{aligned} X_t^0 &= X_t \\ X_t^k &= X_t^{k-1} - \frac{r_{k-1}(t, t_k)}{r_{k-1}(t_k, t_k)} \cdot (X_{t_k}^{k-1} - y_k), \quad k = 1, \dots, n \end{aligned}$$

where

$$\begin{aligned} r_0(s, t) &= r(s, t) \\ r_k(s, t) &= r_{k-1}(s, t) - \frac{r_{k-1}(s, t_k) \cdot r_{k-1}(t_k, t)}{r_{k-1}(t_k, t_k)}, \quad k = 1, \dots, n \end{aligned}$$

In our case, to sample the fBM at $D_{N(\epsilon)}/D_N$ conditional on \mathbf{B}_N^H , we first sample without taking condition using the Davies and Harte method ([Davies and Harte \(1987\)](#)), which has complexity $O(2^{N(\epsilon)} \log(2^{N(\epsilon)}))$. Then, for each $t \in D_{N(\epsilon)}/D_N$, we use the recursion in Lemma 8 to calculate the value of fBM at t after taking condition for 2^N steps whose complexity is approximately $O(2^{3N})$. We repeat this procedure $O(2^{N(\epsilon)})$ times. Above all, the total computational complexity of refine the dyadic approximation from level N to $N(\epsilon)$ is $O(2^{N(\epsilon)} \log(2^{N(\epsilon)}))$.

5. ϵ -Strong simulation of stochastic differential equation driven by fBM In this section, we extend the ϵ -strong simulation algorithm for fBM to stochastic differential equations (SDEs) driven by fBM with $H > 1/2$ via rough path theory. Consider a d -dimensional SDE

$$d\mathbf{Y}(t) = \boldsymbol{\mu}(\mathbf{Y}(t)) dt + \boldsymbol{\sigma}(\mathbf{Y}(t)) d\mathbf{B}^H(t), \quad \mathbf{Y}(0) = \mathbf{y}(0), \quad (23)$$

where $\mathbf{B}^H(t)$ is a d' -dimensional fBM (each component is an independent standard one-dimensional fBM), $\boldsymbol{\mu}(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\boldsymbol{\sigma}(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$ are driven vector fields, corresponding to the drift and the volatility, respectively. For any fixed $\epsilon > 0$, our goal is to construct a probability space, supporting both $\mathbf{Y}(t)$ and a fully simulatable path $\hat{\mathbf{Y}}_\epsilon(t)$ such that

$$\sup_{t \in [0, 1]} |\mathbf{Y}(t) - \hat{\mathbf{Y}}_\epsilon(t)| \leq \epsilon, \quad \text{a.s.}$$

The construction of $\hat{\mathbf{Y}}_\epsilon(t)$ builds on our ability to estimate the driving fBM and its corresponding α -Hölder norm. Particularly, for any $1/2 < \alpha < H$, the sample path of fBM is α -Hölder continuous almost surely. In this case, by the rough path theory, the solution of SDE (23) can be defined path by path and the mapping from \mathbf{B}^H to \mathbf{Y} is continuous under the α -Hölder norm ([Lyons, 1998](#))⁵.

⁵ For $H \leq 1/2$, to define the corresponding SDE in a path by path sense, high order iterated integrals of \mathbf{B}^H need to be specified.

Therefore, if we can control the error of the simulated driving signals, by continuous mapping type of argument, we can also control the error of the simulated SDEs.

In what follows, we shall first lay out the main idea of our algorithmic development. We then present the theoretical foundation and derivation in Section 5.1. The construction of the approximated solution is based on simple Euler scheme. For dyadic discretization D_n , we define $\mathbf{Y}_n(t_k^n)$, $k = 1, 2, \dots, 2^n$, via the recursion

$$\begin{aligned} \mathbf{Y}_n(t_0^n) &= 0, \\ \mathbf{Y}_n(t_{k+1}^n) &= \mathbf{Y}^{(n)}(t_k^n) + \boldsymbol{\mu}(\mathbf{Y}_n(t_k^n)) \cdot \Delta_n + \boldsymbol{\sigma}(\mathbf{Y}_n(t_k^n)) \cdot (\mathbf{B}^H(t_{k+1}^n) - \mathbf{B}^H(t_k^n)). \end{aligned}$$

We then construct the whole path $\mathbf{Y}_n(t)$ via piecewise-constant interpolation,

$$\mathbf{Y}_n(t) = \mathbf{Y}_n(t_k^n), \quad t \in [t_k^n, t_{k+1}^n).$$

The challenge here is to choose an appropriate discretization level $N_Y(\epsilon)$, such that

$$\|\mathbf{Y}_{N_Y(\epsilon)} - \mathbf{Y}\|_\infty \leq \epsilon, \quad \text{a.s.}$$

In Theorem 9 below, we establish that

$$\|\mathbf{Y}_n - \mathbf{Y}\|_\infty \leq G \cdot \Delta_n^{2\alpha-1}, \quad (24)$$

where G can be characterized explicitly, and it depends on the α -Hölder norm of \mathbf{B}^H . Therefore, we shall first use the ϵ -strong simulation algorithm we developed in Section 4 to find an upper bound for $\|\mathbf{B}^H\|_\alpha$. We can then upper bound G , and set

$$N_Y(\epsilon) = \lceil \log_2(\epsilon^{-1}G/(2\alpha-1)) \rceil.$$

The actual algorithm is summarized in Algorithm 6.

5.1. Rate of convergence of Euler scheme In this section, we present the details of the rate of convergence of Euler scheme. This result is an extension of Lejay (2010). Particularly, we explicitly characterize the constant in front of $\Delta_n^{2\alpha-1}$ for the Euler scheme at dyadic discretization D_n . This is important for our algorithmic development, as we need to know G in (24) to find the required discretization level N_Y .

We first introduce a few notations to simplify the exposition. Consider the following multidimensional ordinary differential equation (ODE) system y driven by vector-valued signal x

$$\begin{bmatrix} dy^1(t) \\ dy^2(t) \\ \vdots \\ dy^d(t) \end{bmatrix} = \begin{bmatrix} f_{11}(\mathbf{y}(t)) & f_{12}(\mathbf{y}(t)) & \cdots & f_{1h}(\mathbf{y}(t)) \\ f_{21}(\mathbf{y}(t)) & f_{22}(\mathbf{y}(t)) & \cdots & f_{2h}(\mathbf{y}(t)) \\ \vdots & \vdots & \ddots & \vdots \\ f_{d1}(\mathbf{y}(t)) & f_{d2}(\mathbf{y}(t)) & \cdots & f_{dh}(\mathbf{y}(t)) \end{bmatrix} \cdot \begin{bmatrix} dx^1(t) \\ dx^2(t) \\ \vdots \\ dx^h(t) \end{bmatrix}, \quad \begin{bmatrix} y^1(0) \\ y^2(0) \\ \vdots \\ y^d(0) \end{bmatrix} = \begin{bmatrix} y_0^1 \\ y_0^2 \\ \vdots \\ y_0^d \end{bmatrix}, \quad (25)$$

where $\mathbf{y}(t) = [y^1(t), \dots, y^d(t)]^\top$ and $\mathbf{x}(t) = [x^1(t), \dots, x^h(t)]^\top$. It is easy to see that SDE (23) can be written in form of (25). We only need to set

$$\mathbf{f}(\cdot) = \begin{bmatrix} \boldsymbol{\mu}(\cdot) \\ \boldsymbol{\sigma}(\cdot) \end{bmatrix} \quad \text{and} \quad d\mathbf{x}(t) = \begin{bmatrix} dt \\ d\mathbf{B}^H(t) \end{bmatrix},$$

and then dimension $h = d' + 1$. From now on, all of our derivation will be based on the notations in (25). In the following of this section, we use \mathbf{f} to denote the matrix $[f_{ij}]_{d \times h}$. Note that we will use bold letters to denote matrices or vectors. Furthermore, we assume that \mathbf{x} is α -Hölder continuous with $1/2 < \alpha \leq 1$, which is to say, $|\mathbf{x}(t) - \mathbf{x}(s)| \leq C_\alpha |s - t|^\alpha$ for some $C_\alpha \in (0, \infty)$.

The solution to equation (25) is formally defined in terms of the Young integral. In what follows, we shall start by a brief introduction to Young integral and then quantify error of the Euler approximation scheme.

Algorithm 6 ϵ -Strong Simulation of SDE Driven by fBM (SSDE)

- 1: **Input:** Accuracy ϵ , vector fields $\boldsymbol{\mu}, \boldsymbol{\sigma}$, Hölder norm order α , Hurst index H , record-breaker parameters ν, ρ, δ .
- 2: **Estimate α -Hölder norm :**
- 3: **For** i **in** $[d']$:
- 4: Call SLRB (Algorithm 2): set $[\hat{B}^i, N^i] \leftarrow \text{SLRB}(H, \nu, \rho, \delta)$.
- 5: Calculate the upper bound for the α -Hölder norm:

$$\hat{C}_\alpha^i \leftarrow \|\hat{B}^i\|_\alpha + \frac{\rho 2^{2-\alpha} \cdot 2^{-(H-\alpha-\delta)(N^i+1)}}{1 - 2^{-(H-\alpha-\delta)}}.$$

- 6: Set $\hat{C}_\alpha = \max_{i \in [d']} \hat{C}_\alpha^i \wedge 1$.
 - 7: **Determine the truncation level:**
 - 8: Calculate G using \hat{C}_α , then set $N_Y \leftarrow \lceil \log_2(\epsilon^{-1}G/(2\alpha - 1)) \rceil$.
 - 9: **Refine the approximation of fBM:**
 - 10: **For** i **in** $[d']$:
 - 11: **If** $N^i < N_Y$: refine dyadic approximation \hat{B}^i until level N_Y via acceptance-rejection.
 - 12: **Solve the SDE by Euler scheme:**
 - 13: $\mathbf{Y}_{N_Y}(t_0^{N_Y}) \leftarrow 0$
 - 14: **For** k **in** $[2^{N_Y}]$:
 - 15: $\mathbf{Y}_{N_Y}(t_{k+1}^{N_Y}) \leftarrow \mathbf{Y}_{N_Y}(t_k^{N_Y}) + \boldsymbol{\mu}(\mathbf{Y}_{N_Y}(t_k^{N_Y}))\Delta_{N_Y} + \boldsymbol{\sigma}(\mathbf{Y}_{N_Y}(t_k^{N_Y}))(\hat{\mathbf{B}}^H(t_{k+1}^{N_Y}) - \hat{\mathbf{B}}^H(t_k^{N_Y}))$.
 - 16: **Output:** Return $\mathbf{Y}_{N_Y}(t)$, the piecewise constant interpolation of $\{\mathbf{Y}_{N_Y}(t_k^{N_Y})\}_{k=0, \dots, 2^{N_Y}}$.
-

DEFINITION 2. We say that $\mathbf{y}(t)$ is a solution of equation (25) if for all $t \in [0, 1]$,

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(\mathbf{y}(s)) d\mathbf{x}(s),$$

where the integration is interpreted as Young integral.

Young integral is an extension of Riemann-Stieltjes integral for paths with finite p -variation, $1 < p < 2$, but potentially infinite total variation. Recall that we call a continuous path $u(t)$ defined on $[0, 1]$ has finite p -variation if

$$\sup_{\Pi} \sum_{t_i \in \Pi} |u(t_{i+1}) - u(t_i)|^p < \infty,$$

where $\Pi = \{t_i\}_{i \geq 0}$ is a set of finite partitions of $[0, 1]$. Then we have

DEFINITION 3. (Young integral) Let $u(t)$ and $v(t)$ be continuous paths on $[0, 1]$ with finite p -variation and q -variation respectively, such that

$$\frac{1}{p} + \frac{1}{q} > 1.$$

Then the limit of Riemann sum as the mesh of the partition $|\Pi|$ goes to zero

$$\lim_{|\Pi| \rightarrow 0} \sum_{t_i \in \Pi} v(t_i) \cdot (u(t_{i+1}) - u(t_i)),$$

exists and is unique. We use $\int_0^1 v(s) du(s)$, to denote this limit and call it the Young integral of $v(s)$ with respect to $u(s)$.

A special case of finite p -variation path is α -Hölder continuous path. Note that if u is α -Hölder continuous, i.e. $|u(s) - u(t)| \leq \|u\|_\alpha \cdot |s - t|^\alpha$ for some $\|u\|_\alpha \in (0, \infty)$, then u has finite p -variation with $p = 1/\alpha$. From now on, we use $H_{\alpha, [0,1]}(u)$ to denote the α -Hölder norm of u on interval $[0, 1]$. In this case, we have the following Young-Lóeve estimate (Lejay (2010)).

THEOREM 8. (*Young-Lóeve estimate.*) *Assume that the integrator u and the integrand v are Hölder continuous of exponents α and β with $\alpha + \beta > 1$, respectively. Then for any $s, t \in [0, 1]$,*

$$\left| \int_s^t v(r) du(r) - v(s)(u(t) - u(s)) \right| \leq K(\alpha + \beta) \cdot H_{\alpha, [0,1]}(u) \cdot H_{\beta, [0,1]}(v) \cdot |t - s|^{\alpha + \beta},$$

where $K(\alpha + \beta) = 1 + \sum_{n \geq 1} n^{-(\alpha + \beta)}$. Moreover, for any finite partition $\Pi_{s,t}$ of $[s, t]$,

$$\left| \sum_{t_i \in \Pi_{s,t}} v(t_i)(u(t_{i+1}) - u(t_i)) - v(s)(u(t) - u(s)) \right| \leq K(\alpha + \beta) \cdot H_{\alpha, [0,1]}(u) \cdot H_{\beta, [0,1]}(v) \cdot |t - s|^{\alpha + \beta},$$

We are ready to introduce the Euler scheme. For dyadic discretization points $D_n = \{t_k^n\}_{k=0, \dots, 2^n}$, we define

$$\mathbf{y}_n(t_{k+1}^n) = \mathbf{y}_n(t_k^n) + \mathbf{f}(\mathbf{y}_n(t_k^n)) \cdot (\mathbf{x}(t_{k+1}^n) - \mathbf{x}(t_k^n)), \quad k = 0, 1, \dots, 2^n - 1. \quad (26)$$

Based on the values $\{\mathbf{y}_n(t_k^n)\}_{k=0, \dots, 2^n}$, we further construct a continuous path over $[0, 1]$ via piecewise constant interpolation, i.e.

$$\mathbf{y}_n(t) = \mathbf{y}_n(t_k^n), \quad t \in [t_k^n, t_{k+1}^n). \quad (27)$$

Then we call $\mathbf{y}_n(t)$ an approximated solution of level n via Euler scheme. Our goal is to control the uniform norm between the approximated solution $\mathbf{y}_n(t)$ and exact solution $\mathbf{y}(t)$.

To ensure the existence of solution of ODE (25) and control the approximation error, we impose the following smoothness condition on the vector field \mathbf{f} .

ASSUMPTION 1. *We assume that $\mathbf{f}(\cdot) \in C^2(\mathbb{R}^d)$ and $\max\{|\mathbf{f}|, |\nabla \mathbf{f}|, |\nabla^2 \mathbf{f}|\} < \infty$, where $|\mathbf{f}| = \max_{ij} |f_{ij}|$, $|\nabla \mathbf{f}| = \max_{ij} |\nabla f_{ij}|$, and $|\nabla^2 \mathbf{f}| = \max_{ij} |\nabla^2 f_{ij}|$.*

We also define the following constants. Let

$$\begin{aligned} G_1^* &= 2h \cdot \lceil (2dhC_\alpha K(2\alpha) |\nabla \mathbf{f}|)^{1/\alpha} \rceil^{1-\alpha} \cdot |\mathbf{f}| \cdot C_\alpha, & G_2^* &= dh \cdot K(2\alpha) \cdot |\nabla \mathbf{f}| \cdot C_\alpha \cdot G_1^*, \\ L &= 4/(1 - 2^{1-2\alpha}) \cdot (hC_\alpha)^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}|, & \omega &= (h|\mathbf{f}|C_\alpha/L)^{1/\alpha}, \\ G_1 &= (L + h|\mathbf{f}|C_\alpha) \cdot (1 + \omega^{-1}), & G_2 &= \max\left\{ (2\omega^{-\alpha} + \omega^{-1-\alpha}) \cdot (L + h|\mathbf{f}|C_\alpha), L \right\}. \end{aligned}$$

In addition, we define a sequence of constants $\{\Gamma_k\}$ and $\{\Upsilon_k\}$ via recursion

$$\begin{aligned} \Gamma_1 &= 2G_2^*, \quad \Upsilon_1 = (4\zeta)^{-1} \cdot \Gamma_1, \\ \Gamma_k &= 2(G_2^* + v \cdot \Upsilon_{k-1}), \quad \Upsilon_k = (4\zeta)^{-1} \cdot \Gamma_k + \Upsilon_{k-1}, \quad k \geq 2, \end{aligned} \quad (28)$$

where

$$\begin{aligned} \zeta &= h \cdot K(2\alpha) \cdot C_\alpha \cdot (d \cdot |\nabla \mathbf{f}| + d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1)), \\ v &= C_\alpha \cdot (d^2 h \cdot K(2\alpha) \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1) + d \cdot |\nabla \mathbf{f}|). \end{aligned}$$

Under Assumption 1, the general theory of Young integral equation ensures the existence and uniqueness for the solution of equation (25). The following theorem characterizes the rate of convergence of the Euler scheme under the uniform norm and is the main result of this section.

THEOREM 9. *Under Assumption 1,*

$$\|\mathbf{y}_n(t) - \mathbf{y}(t)\|_\infty \leq G \cdot \Delta_n^{2\alpha-1} := (\Upsilon_{\lceil(4\zeta)^{1/\alpha}\rceil} + G_1^*) \cdot \Delta_n^{2\alpha-1}.$$

REMARK 2. By the definition of G , it is easy to see that G is an increasing function with respect to C_α , the α -Hölder norm of \mathbf{x} .

Proof. The proof of Theorem 9 relies on the several lemmas. The first one establishes the existence of the solution to equation (25) and its properties.

LEMMA 9. *Under Assumption 1, the solution of equation (25) $\mathbf{y}(t)$ exists. We also have the following estimates. For all $0 \leq s < t \leq 1$,*

$$\begin{aligned} |\mathbf{y}(s) - \mathbf{y}(t)| &\leq G_1^* \cdot |s - t|^\alpha, \\ |\mathbf{y}(t) - \mathbf{y}(s) - \mathbf{f}(\mathbf{y}(s)) \cdot (\mathbf{x}(t) - \mathbf{x}(s))| &\leq G_2^* \cdot |s - t|^{2\alpha}. \end{aligned}$$

The next lemma is a counterpart of Lemma 9. It establishes similar properties for the solutions obtain by the Euler scheme.

LEMMA 10. *For all the dyadic discretization time points $t_j^n, t_r^n \in [0, 1]$, we have the estimates*

$$\begin{aligned} |\mathbf{y}_n(t_j^n) - \mathbf{y}^{(n)}(t_r^n)| &\leq G_1 \cdot |t_j^n - t_r^n|^\alpha \\ |\mathbf{y}_n(t_j^n) - \mathbf{y}^{(n)}(t_r^n) - \mathbf{f}(\mathbf{y}_n(t_r^n)) \cdot (\mathbf{x}(t_j^n) - \mathbf{x}(t_r^n))| &\leq G_2 \cdot |t_j^n - t_r^n|^{2\alpha}. \end{aligned}$$

We also need to define a restricted α -Hölder norm. Specifically, given the dyadic partition D_n and a path $\mathbf{x}(t)$, the restricted α -Hölder norm on D_n is defined as

$$H_\alpha(\mathbf{x}|D_n) = \sup_{0 \leq i < j \leq 2^n} \frac{|\mathbf{x}(t_i^n) - \mathbf{x}(t_j^n)|}{|t_i^n - t_j^n|^\alpha}.$$

Note that in the restricted α -Hölder norm, we do not require that the path is well-defined on points outside of D_n . Intuitively, the restricted α -Hölder norm measures the α -Hölder continuity of the solution obtained via Euler scheme on D_n . We have the following lemma on the restricted α -Hölder norm of $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)$ on D_n .

LEMMA 11. *For dyadic discretization D_n ,*

$$H_\alpha(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)|D_n) \leq (d \cdot |\nabla \mathbf{f}| + d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1)) \cdot H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) + d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1) \cdot |\mathbf{y}(0) - \mathbf{y}_n(0)|.$$

With the above lemmas, we are ready to prove Theorem 9. Let

$$\mathcal{J}_i \equiv \int_{t_i^n}^{t_{i+1}^n} (\mathbf{f}(\mathbf{y}(s)) - \mathbf{f}(\mathbf{y}(t_i^n))) d\mathbf{x}(s).$$

By definition of Euler scheme and the solution of equation 25, we have

$$\mathbf{y}_n(t_{i+1}^n) - \mathbf{y}(t_{i+1}^n) = \mathbf{y}_n(t_i^n) - \mathbf{y}(t_i^n) + \left(\mathbf{f}(\mathbf{y}_n(t_i^n)) - \mathbf{f}(\mathbf{y}(t_i^n)) \right) \cdot (\mathbf{x}(t_{i+1}^n) - \mathbf{x}(t_i^n)) - \mathcal{J}_i,$$

and furthermore, for all $0 \leq \ell < k \leq 2^n$,

$$(\mathbf{y}_n(t_k^n) - \mathbf{y}(t_k^n)) - (\mathbf{y}_n(t_\ell^n) - \mathbf{y}(t_\ell^n)) = \sum_{i=\ell}^{k-1} \left(\mathbf{f}(\mathbf{y}_n(t_i^n)) - \mathbf{f}(\mathbf{y}(t_i^n)) \right) \cdot (\mathbf{x}(t_{i+1}^n) - \mathbf{x}(t_i^n)) - \sum_{i=\ell}^{k-1} \mathcal{J}_i.$$

Using the Young-Lóeve estimate, we have the following bounds

$$\left| \sum_{i=\ell}^{k-1} \left(\mathbf{f}(\mathbf{y}_n(t_i^n)) - \mathbf{f}(\mathbf{y}(t_i^n)) \right) \cdot (\mathbf{x}(t_{i+1}^n) - \mathbf{x}(t_i^n)) - \left(\mathbf{f}(\mathbf{y}_n(t_\ell^n)) - \mathbf{f}(\mathbf{y}(t_\ell^n)) \right) \cdot (\mathbf{x}(t_k^n) - \mathbf{x}(t_\ell^n)) \right| \leq h \cdot K(2\alpha) \cdot H_\alpha(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)|D_n) \cdot C_\alpha \cdot |t_k^n - t_\ell^n|^{2\alpha},$$

where $K(2\alpha) = 1 + \sum_{n=1}^{\infty} n^{-2\alpha}$ and C_α is the α -Hölder norm of \mathbf{x} . By Lemma 9, we also have

$$|\mathcal{J}_i| \leq G_2^* \cdot |t_{i+1}^n - t_i^n|^{2\alpha}.$$

Note that

$$\sum_{i=\ell}^{k-1} |t_{i+1}^n - t_i^n|^{2\alpha} \leq |t_k^n - t_\ell^n| \cdot \Delta_n^{2\alpha-1} \leq |t_k^n - t_\ell^n|^\alpha \cdot \Delta_n^{2\alpha-1}.$$

Then we have

$$\begin{aligned} & \left| (\mathbf{y}_n(t_k^n) - \mathbf{y}(t_k^n)) - (\mathbf{y}_n(t_\ell^n) - \mathbf{y}(t_\ell^n)) \right| \\ & \leq h \cdot K(2\alpha) \cdot H_\alpha(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)|D_n) \cdot C_\alpha \cdot |t_k^n - t_\ell^n|^{2\alpha} + \\ & \quad \left(\mathbf{f}(\mathbf{y}_n(t_\ell^n)) - \mathbf{f}(\mathbf{y}(t_\ell^n)) \right) \cdot (\mathbf{x}(t_k^n) - \mathbf{x}(t_\ell^n)) + G_2^* \cdot |t_k^n - t_\ell^n|^\alpha \cdot \Delta_n^{2\alpha-1}. \end{aligned}$$

By the definition of restricted α -Hölder norm,

$$|\mathbf{f}(\mathbf{y}_n(t_\ell^n)) - \mathbf{f}(\mathbf{y}(t_\ell^n))| \leq H_\alpha(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)|D_n) \cdot |t_\ell^n|^\alpha + |\mathbf{f}(\mathbf{y}_n(0)) - \mathbf{f}(\mathbf{y}(0))|,$$

combined with Lemma 11, we have

$$\begin{aligned} H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) & \leq h \cdot K(2\alpha) \cdot (|t_k^n - t_\ell^n|^\alpha + |t_\ell^n|^\alpha) \cdot C_\alpha \cdot H_\alpha(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)|D_n) + \\ & \quad G_2^* \cdot \Delta_n^{2\alpha-1} + C_\alpha \cdot |\mathbf{f}(\mathbf{y}_n(0)) - \mathbf{f}(\mathbf{y}(0))| \\ & \leq h \cdot K(2\alpha) \cdot C_\alpha \cdot (d \cdot |\nabla \mathbf{f}| + d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1)) \cdot (|t_k^n - t_\ell^n|^\alpha + |t_\ell^n|^\alpha) \\ & \quad \times H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) + G_2^* \cdot \Delta_n^{2\alpha-1} + C_\alpha \cdot (d^2 h \cdot K(2\alpha) \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1) + d \cdot |\nabla \mathbf{f}|) \\ & \quad \times |\mathbf{y}_n(0) - \mathbf{y}(0)|. \end{aligned}$$

By the definition of ζ and v in (28),

$$H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) \leq \zeta (|t_k^n - t_\ell^n|^\alpha + |t_\ell^n|^\alpha) \cdot H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) + G_2^* \cdot \Delta_n^{2\alpha-1} + v |\mathbf{y}_n(0) - \mathbf{y}(0)|.$$

Let $T_0 = 0$ and $T_1 = \max\{t \in D_n, t \leq (4\zeta)^{-1/\alpha}\}$. Then

$$H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n \cap [T_0, T_1]) \leq 2G_2^* \cdot \Delta_n^{2\alpha-1} + 2v \cdot |\mathbf{y}_n(0) - \mathbf{y}(0)|.$$

Since the Euler scheme and the exact solution have same initial value, $|\mathbf{y}_n(0) - \mathbf{y}(0)| = 0$, then we have

$$H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n \cap [T_0, T_1]) \leq 2G_2^* \cdot \Delta_n^{2\alpha-1},$$

and furthermore,

$$|\mathbf{y}_n(T_1) - \mathbf{y}(T_1)| \leq 2G_2^* \cdot (4\zeta)^{-1} \cdot \Delta_n^{2\alpha-1}.$$

Now we let $T_2 = \max\{t \in D_n, t \leq 2(4\zeta)^{-1/\alpha}\}$. For $T_1 \leq t_\ell^n < t_k^n \leq T_2$, we have

$$H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) \leq \zeta (|t_k^n - t_\ell^n|^\alpha + |t_\ell^n - T_1|^\alpha) \cdot H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) + G_2^* \cdot \Delta_n^{2\alpha-1} + v |\mathbf{y}_n(T_1) - \mathbf{y}(T_2)|.$$

As a result,

$$H_\alpha(\mathbf{y} - \mathbf{y}_n | D_n \cap [T_1, T_2]) \leq 2(G_2^* + v \cdot 2G_2^* \cdot (4\zeta)^{-1}) \cdot \Delta_n^{2\alpha-1},$$

and

$$|\mathbf{y}_n(T_2) - \mathbf{y}(T_2)| \leq \left[2(G_2^* + v \cdot 2G_2^* \cdot (4\zeta)^{-1}) \cdot (4\zeta)^{-1} + 2G_2^* \cdot (4\zeta)^{-1} \right] \cdot \Delta_n^{2\alpha-1}.$$

We need to repeat this procedure at most $k^* = \lceil (4\zeta)^{1/\alpha} \rceil$ times in order to cover the whole interval $[0, 1]$ and we can obtain the a sequence of bounds

$$\begin{aligned} H_\alpha(\mathbf{y} - \mathbf{y}_n | D_n \cap [T_{k-1}, T_k]) &\leq \Gamma_k \cdot \Delta_n^{2\alpha-1}, \\ |\mathbf{y}_n(T_k) - \mathbf{y}(T_k)| &\leq \Upsilon_k \cdot \Delta_n^{2\alpha-1}, \end{aligned}$$

where $\{\Gamma_k\}$ and $\{\Upsilon_k\}$ are defined via recursion (28). Note that $\{\Gamma_k\}$ and $\{\Upsilon_k\}$ are increasing sequences, and for any $t_i^n \in [T_{k-1}, T_k]$, we have $|\mathbf{y}_n(t_i^n) - \mathbf{y}(t_i^n)| \leq \Upsilon_k \cdot \Delta_n^{2\alpha-1}$. Therefore,

$$\sup_{t_i^n \in D_n} |\mathbf{y}_n(t_i^n) - \mathbf{y}(t_i^n)| \leq \Upsilon_{\lceil (4\zeta)^{1/\alpha} \rceil} \cdot \Delta_n^{2\alpha-1}.$$

Finally, for any $t \in [0, 1]$, there exists i such that $t_i^n \leq t < t_{i+1}^n$. Then we have

$$\begin{aligned} |\mathbf{y}_n(t) - \mathbf{y}(t)| &\leq |\mathbf{y}_n(t_i^n) - \mathbf{y}_n(t)| + |\mathbf{y}_n(t) - \mathbf{y}(t)| + |\mathbf{y}(t_i^n) - \mathbf{y}(t)| \\ &\leq \Upsilon_{\lceil (4\zeta)^{1/\alpha} \rceil} \cdot \Delta_n^{2\alpha-1} + G_1^* \cdot \Delta_n^\alpha \leq (\Upsilon_{\lceil (4\zeta)^{1/\alpha} \rceil} + G_1^*) \cdot \Delta_n^{2\alpha-1}. \end{aligned}$$

□

6. Application to multilevel Monte Carlo In this section, we demonstrate how our ϵ -strong simulation algorithm can be easily combined with multilevel Monte Carlo (MLMC). We start with a brief introduction of the MLMC framework (Giles, 2008b). Our objective is to estimate $\alpha = \mathbb{E}[g(B^H)]$, where g is a functional of the fBM path. The MLMC estimator takes the following form

$$\hat{\alpha}_K = \sum_{k=0}^K \frac{1}{r_k} \sum_{i=1}^{r_k} D_k(i) \quad (29)$$

where $D_k(i)$'s are i.i.d. copies of some properly defined level differences. For example, $D_k(i) = g(B_k^H) - g(B_{k-1}^H)$. Assuming $g(B_{-1}^H) = 0$. Then $\mathbb{E}[\hat{\alpha}_K] = g(B_K^H)$, which implies the bias of the estimator only depends on the bias at the highest level K . On the other hand, we have

$$\mathbb{V}(\hat{\alpha}_K) = \sum_{k=0}^K \frac{1}{r_k} \mathbb{V}(D_k),$$

i.e. the variance depends on the variance at different levels. Thus, by using appropriate coupling to create the level differences, D_k 's, and smartly allocate the computational budget, r_k 's, we can achieve substantial computational cost reduction comparing to naive Monte Carlo method⁶. As we shall explain next, the advantage of our ϵ -strong simulation algorithm is that it provides an elegant way to construct D_k 's. It is also straightforward to calculate the variances of D_k 's. In what follows, we denote $\mathcal{C}(D_k)$ as the computational cost of generating one copy of D_k .

We consider two cases for the functional form g .

⁶ By naive Monte Carlo method, we mean generating i.i.d. copies of $g(B_K^H)$.

- I) g is Lipschitz continuous with respect to the supremum norm, and once B^H is given, $g(B^H)$ can be evaluated in closed form. In this case, we can construct $D_k(i)$'s as i.i.d. copies of $g(B_k^H) - g(B_{k-1}^H)$. The coupling is created by using the same fBM path truncated at different levels for B_k^H and B_{k-1}^H .
- II) g maps B^H to the solution to an SDE at a fixed time point. In this case, we can set $D_k(i)$'s as i.i.d. copies of $Y_k - Y_{k-1}$. Similarly to Case I, the coupling is created by constructing Y_k and Y_{k-1} using the same fBM path truncated at different levels.

For Case I, let L denote the Lipschitz constant of g . Then for any fixed $\delta \in (0, H)$,

$$\begin{aligned} \mathbb{V}(D_k) &\leq \mathbb{E}[(g(B_k^H) - g(B_{k-1}^H))^2] \\ &\leq 2 \cdot (\mathbb{E}[(g(B_k^H) - g(B^H))^2] + \mathbb{E}[(g(B_{k-1}^H) - g(B^H))^2]) \\ &\leq \left(\frac{2L\rho}{1 - 2^{-H+\delta}} \right)^2 \cdot \Delta_k^{2(H-\delta)}, \end{aligned}$$

where the last inequality follows from the definition of record-breakers. As for the computational complexity, according to the analysis in Section 4.3, we have $\mathcal{C}(D_k) = O(\Delta_k^{-1} \log(\Delta_k^{-1}))$, where the recursive Gaussian bridge based method is used.

In this case, for a given mean square error (MSE) bound ϵ^2 , we can set $K = C_1 \log(1/\epsilon)$, such that $\mathbb{E}[\hat{\alpha}_K] - \alpha = O(\Delta_K^{H-\delta}) = O(\epsilon)$. We can also set $r_k = C_2 \Delta_k^{2(H-\delta)} \epsilon^{-2} \log(1/\epsilon)$, such that $\mathbb{V}(\hat{\alpha}_K) = O(\epsilon^2)$. With our choice of K and r_k , the total computational cost of $\hat{\alpha}_K$ is

$$\sum_{k=0}^K r_k \mathcal{C}(D_k) = O \left(\epsilon^{-2} \log(1/\epsilon) \sum_{k=1}^K \Delta_k^{2(H-\delta)-1} \log(\Delta_k^{-1}) \right)^7.$$

When $2(H - \delta) > 1$, the cost is $O(\epsilon^{-2} \log(1/\epsilon))$; otherwise the cost is $O(\epsilon^{-1/(H-\delta)} \log(1/\epsilon)^2)$. Note that the total computational cost of naive Monte Carlo estimator is $O(\epsilon^{-2-1/(H-\delta)} \log(1/\epsilon))$.

We believe that the complexity that can be achieved using MLMC in this setting is near optimal. This is based on the fact that mid-point displacement decomposition we employ achieves the optimal rate of convergence in terms of the bias/error of the truncated fBM path.

For Case II, under Assumption 1, we have

$$\begin{aligned} \mathbb{V}(D_k) &\leq \mathbb{E}[(Y_k - Y_{k-1})^2] \\ &\leq \mathbb{E}[(Y_k - Y)^2] + E[(Y_{k-1} - Y)^2] \\ &\leq 4G \cdot \Delta_k^{2(2\alpha-1)} \end{aligned}$$

for any fixed $\alpha \in (1/2, H - \delta)$. We also have $\mathcal{C}(D_k) = O(\Delta_k^{-1} \log(\Delta_k^{-1}))$. Following similar lines of analysis as in Case I, we can show that in Case II, to achieve an MSE of order ϵ^2 , we set $K = C_3 \log(1/\epsilon)$ such that $\mathbb{E}[\hat{\alpha}_K] - \alpha = O(\Delta_K^{2\alpha-1}) = O(\epsilon)$ and $r_k = C_4 \Delta_k^{2(2\alpha-1)} \epsilon^{-2} \log(1/\epsilon)$. Then, the computational cost is

$$O \left(\epsilon^{-2} \log(1/\epsilon) \sum_{k=1}^K \Delta_k^{2(2\alpha-1)-1} \log(\Delta_k^{-1}) \right).$$

When $\alpha > 3/4$, the cost is $O(\epsilon^{-2} \log(1/\epsilon))$; otherwise the cost is $O(\epsilon^{-1/(2\alpha-1)} \log(1/\epsilon)^2)$.

When $\alpha > 3/4$, the MLMC achieves the near optimal complexity. For the $\alpha \in (1/2, 3/4)$, the computational complexity can get arbitrarily bad as α approaches $1/2$. The fundamental bottleneck here is the Euler scheme. If we use higher order discretization scheme like the Milstein scheme,

⁷ Here, we treat the cost of generating the last record-breaker as a constant. This is because N does not depend on ϵ and $N \leq N(\epsilon)$ in most cases

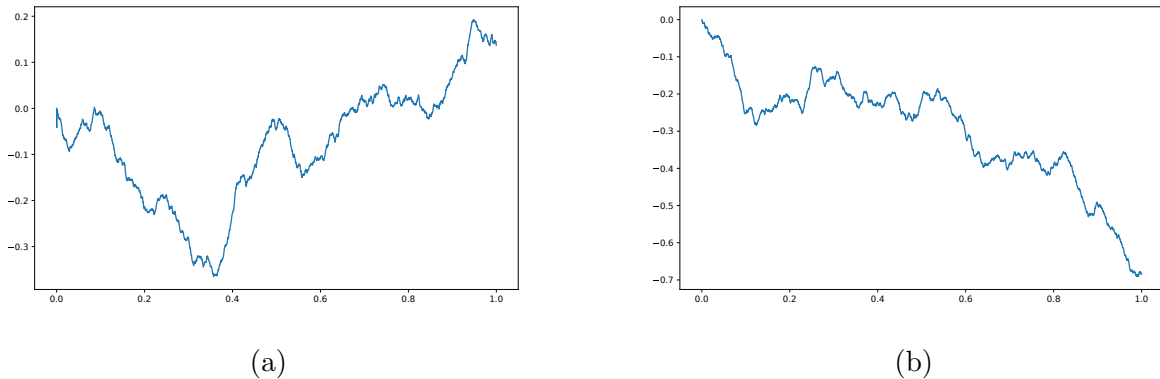


FIGURE 1. Two realizations of the ϵ -strong simulation of fBM where the parameters are $H = 0.8, \epsilon = 0.1, \rho = 5, \delta = 0.1$.

the convergence rate of the discretization error can be improved from $\Delta_n^{2\alpha-1}$ to $\Delta_n^{3\alpha-1}$. However, evaluating the iterated integrals of fBM (Lévy area) is itself a very challenging task. In this paper, we do not pursue more sophisticated discretization schemes, but we view this as an interesting future research direction.

7. Numerical experiments In this section, we conduct some numerical experiments as a sanity check of the algorithms we developed. We also provide some discussions about implementation issues of our algorithm.

Figure 1 displays two realizations of a fBM with $H = 0.8$ using Algorithm 1 with $\epsilon = 0.1$.

We next provide a brief discussion about the choices of record-breaker parameters ρ and δ in practice. Recall that the record-breaking threshold takes the form $\rho\Delta_n^{H-\delta}$ for level n . δ determines the asymptotic decay rate of the bound and ρ determines the scale of the bound. Thus, we would want both to be as small as possible when only considering truncation level, $N(\epsilon)$. Now, when also taking into account the time of the last record-breaker, N , we notice that larger values of ρ and δ result in larger thresholds, under which, the records are less likely to be broken. These in turn lead to a smaller value of N . We also note that larger values of ρ and δ lead to smaller starting level $N^*(\rho, \delta)$ in Algorithm 2. In the asymptotic sense, as N and $N^*(\rho, \delta)$ do not depend on ϵ , the values of N and $N^*(\rho, \delta)$ do not matter. This indicates that in theory, we should set ρ and δ as small as possible. However, in practice, we do care about the “cost” of sampling N . Thus, in actual implementations, we will tune ρ and δ to balance $N^*(\rho, \delta)$ and $N(\epsilon)$. Table 1 and Table 2 show the truncation level, the starting level, and the average level of the last record-breaker, $(N(\epsilon), N^*(\rho, \delta), \mathbb{E}[N])$, for different choices of ρ and δ . Table 1 is for fBM with $H = 0.8$ and Table 2 is for fBM with $H = 0.45$. We make two observations from the tables. First, the level of the last record-breaker, N , is quite sensitive to our choice of ρ . For reasonably large values of ρ , e.g. $\rho \geq 2.5$, the record breaker rarely happens beyond level $n = 1$. On the other hand, as we have discussed above, smaller values of ρ lead to smaller values of truncation level $N(\epsilon)$. Second, the starting level $N^*(\rho, \delta)$ can be unreasonably large if ρ and δ are not properly chosen. In what follows, we shall provide more discussions about $N^*(\rho, \delta)$, including an alternative algorithm to get rid of the starting level requirement.

Recall from the development of Algorithm 2 that the starting level is required such that the normalizing constant Z_n in (16) is smaller than one, and hence, the weighted likelihood ratios $\Theta_n^+(m, k)$ and $\Theta_n^-(m, k)$ conditional on the proper record-breaking event are also bounded by one. This property (bounded by one) is desirable as we can generate Bernoulli random variable with probability of success $\mathbb{E}[\Theta_n^+(m, k)]$ by generating $1\{U \leq \Theta_n^+(m, k)\}$, where U is a Uniform random variable independent of everything else. Note that when $Z_n > 1$, with the change-of-measure

$\delta \backslash \rho$	1	2.5	5
0.1	(7, 38, 14)	(9, 21, 1)	(11, 1, 1)
0.2	(9, 16, 6)	(11, 6, 1)	(12, 1, 1)

TABLE 1. $(N(\epsilon), N^*(\rho, \delta), \mathbb{E}[N])$ under different choices of ρ and δ when $H = 0.8, \epsilon = 0.1$.

$\delta \backslash \rho$	1	2.5	5
0.1	(16, 38, 15+)	(20, 21, 1)	(23, 1, 1)
0.2	(24, 16, 8)	(30, 6, 1)	(31, 1, 1)

TABLE 2. $(N(\epsilon), N^*(\rho, \delta), \mathbb{E}[N])$ under different choices of ρ and δ when $H = 0.45, \epsilon = 0.1$. (15+ means even at level 15 we still see record-breaker happening.)

technique we used in Algorithm 4, we are only able to generate Bernoulli random variables with success probability $\mathbb{P}_n(\tau = n + m)/Z_n$. But our initial objective is to generate Bernoulli random variables with success probability $\mathbb{P}_n(\tau = n + m)$. We can use a technique called the *Bernoulli factory* to overcome this gap. We next introduce the basic idea of Bernoulli factory and explain how it applies to our setting. The main objective of introducing this alternative to Algorithm 4 is to get rid of the starting level requirement in Algorithm 2.

Suppose that X_1, X_2, \dots are i.i.d. Bernoulli random variables with unknown success probability p . Given a known function f , a Bernoulli factory takes X_1, X_2, \dots as input and outputs a Bernoulli random variable with success probability $f(p)$. In our case, by sampling the fBM under the change of measure up to level $(n + m)$ and check whether the record is broken at level $n + m$, we are able to “generate” a Bernoulli random variable with success probability $\mathbb{P}_n(\tau = n + m)/Z_n$. Then our $f(p)$ function is a linear function in p , i.e. $f(p) = Z_n p$. This is also known as a linear Bernoulli factory. We refer to Huber (2016) for a nearly optimal linear Bernoulli factory that can be directly applied to our setting.

To sum up the discussion here, as record-breaking is a rare event ⁸, the benefits of shrinking the truncation level by choosing small ρ and δ is very appealing. However, in practice, small values of ρ and δ may lead to a large value of $N^*(\rho, \delta)$. When $N^*(\rho, \delta)$ is impractically large, we may consider using the Bernoulli factory to get rid of the starting level requirement. Even though, in theory, our algorithm achieves near optimal complexity for fBM, there are still a lot of rooms for improvements in practical implementations. Moreover, as we have discussed in Section 6, for fBM driven SDEs, even though our algorithm achieves near optimal complexity under the Euler scheme, there are a lot of rooms for improvements if higher order discretization schemes can be efficiently implemented. We view all these as interesting future research directions.

Acknowledgments. Support from NSF Grant DMS-1720433 is gratefully acknowledged.

References

- Abry P, Sellan F (1996) The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation.
- Ayache A, Taqqu M (2003) Approximating fractional Brownian motion by a random wavelet series: the rate optimality problem. *J. Fourier Anal. Appl.*
- Bayer C, Friz P, Riedel S, JSchoenmakers (2016) From rough path estimates to multilevel Monte Carlo. *SIAM Journal on Numerical Analysis* 54(3):1449–1483.

⁸ The corresponding probability decays double exponentially fast in n .

-
- Beskos A, Peluchetti S, Roberts G (2012) ε -Strong simulation of the Brownian path. *Bernoulli* 18(4):1223–1248.
- Beskos A, Roberts G (2005) Exact simulation of diffusions. *The Annals of Applied Probability* 15(4):2422–2444.
- Blanchet J, Chen X (2015) Steady-state simulation for reflected Brownian motion and related networks. *Ann. Appl. Prob.* 25(6):3209–3250.
- Blanchet J, Chen X, Dong J (2017) ε -Strong simulation for multidimensional stochastic differential equations via rough path analysis. *The Annals of Applied Probability* 27(1):275–336.
- Blanchet J, Dong J (2015) Perfect sampling for infinite server and loss systems. *Advances in Applied Probability* 47(3):761–786.
- Blanchet J, Sigman K (2011) On exact sampling of stochastic perpetuities. *Journal of Applied Probability* 48(A):165–182.
- Blanchet J, Zhang F (2017) Exact simulation for multivariate Itô diffusions. *arXiv preprint arXiv:1706.05124* .
- Chen N, Huang Z (2013) Localization and exact simulation of Brownian motion-driven stochastic differential equations. *Mathematics of Operations Research* 38(3):591–616.
- Coutin L, Qian Z (2002) Stochastic analysis, rough path analysis and fractional Brownian motions. *Probability theory and related fields* 122(1):108–140.
- Davie A (2008) Differential equations driven by rough paths: an approach via discrete approximation. *Applied Mathematics Research eXpress* 2008.
- Davies R, Harte D (1987) Tests for Hurst effect. *Biometrika* 74(1):95–101.
- Dereich S, Heidenreich F (2011) A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. *Stochastic Processes and their Applications* 121(7):1565–1587.
- Deya A, Neuenkirch A, Tindel S (2012) A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion. *Annales de l’IHP Probabilités et statistiques*, volume 48, 518–550.
- Dieker A (2004) *Simulation of fractional Brownian motion*. Ph.D. thesis, Masters Thesis, Department of Mathematical Sciences, University of Twente
- Dieker A, Mandjes M (2003) On spectral simulation of fractional Brownian motion. *Probability in the Engineering and Informational Sciences* 17(3):417–434.
- Dietrich C, Newsam G (1997) Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM Journal on Scientific Computing* 18(4):1088–1107.
- Durbin J (1960) The fitting of time-series models. *Revue de l’Institut International de Statistique* 233–244.
- Dzhaparidze K, Zanten HV (2004) A series expansion of fractional Brownian motion. *Probability theory and related fields* 130(1):39–55.
- Giles M (2008a) Improved multilevel Monte Carlo convergence using the Milstein scheme. *Monte Carlo and Quasi-Monte Carlo Methods 2006*, 343–358 (Springer).
- Giles M (2008b) Multilevel Monte Carlo path simulation. *Operations Research* 56(3):607–617.
- Glynn P (2016) Exact simulation vs exact estimation. *2016 Winter Simulation Conference*, 193–205.
- Höfling F, Franosch T (2013) Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics* 76(4):046602.
- Hu Y, Liu Y, Nualart D (2016) Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions. *The Annals of Applied Probability* 26(2):1147–1207.
- Huber M (2016) Nearly optimal Bernoulli factories for linear functions. *Combinatorics, Probability and Computing* 25(4):577–591.
- Kühn T, Linde W (2002) Optimal series representation of fractional Brownian sheets. *Bernoulli* 8(5):669–696.

- Laskin N (2000) Fractional market dynamics. *Physica A: Statistical Mechanics and its Applications* 287(3-4):482–492.
- Lau W, A Erramilli JW, Willinger W (1995) Self-similar traffic generation: the random midpoint displacement algorithm and its properties. *Proceedings IEEE International Conference on Communications ICC'95*, volume 1, 466–472 (IEEE).
- Lejay A (2010) Controlled differential equations as Young integrals: a simple approach. *Journal of Differential Equations* 249(8):1777–1798.
- Levinson N (1946) The Wiener root mean square error criterion in filter design and prediction. *Journal of Mathematics and Physics* 25(1-4):261–278.
- Liu Z, Blanchet J, Dieker A, Mikosch T (2016) On optimal exact simulation of max-stable and related random fields. *arXiv preprint arXiv:1609.06001* .
- Lyons T (1998) Differential equations driven by rough signals. *Revista Matemática Iberoamericana* 14(2):215–310.
- Mandelbrot B, Ness JV (1968) Fractional Brownian motions, fractional noises and applications. *SIAM review* 10(4):422–437.
- Meyer Y, Sellan F, Taqqu M (1999) Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. *Journal of Fourier Analysis and Applications* 5(5):465–494.
- Mishura I (2008) *Stochastic calculus for fractional Brownian motion and related processes*, volume 1929 (Springer Science & Business Media).
- Neuenkirch A, Nourdin I (2007) Exact rate of convergence of some approximation schemes associated to SDEs driven by a fractional Brownian motion. *Journal of Theoretical Probability* 20(4):871–899.
- Norros I (1995) On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE Journal on selected areas in communications* 13(6):953–962.
- Norros I, Mannersalo P, Wang J (1999) Simulation of fractional Brownian motion with conditionalized random midpoint displacement. *Advances in Performance Analysis* 2(1):77–101.
- Nualart D (2006) Fractional Brownian motion: stochastic calculus and applications. *International Congress of Mathematicians*, volume 3, 1541–1562 (Eur. Math. Soc.).
- Nualart D, Tindel S (2011) A construction of the rough path above fractional Brownian motion using Volterra’s representation. *The Annals of Probability* 39(3):1061–1096.
- Paxson V (1997) Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic. *ACM SIGCOMM Computer Communication Review* 27(5):5–18.
- Pollock M, Johansen A, Roberts G (2013) On the exact and ε -strong simulation of (jump) diffusions. *ArXiv Mathematics e-print* .
- Rhee C, Glynn P (2012) A new approach to unbiased estimation for SDE’s. *Proceedings of the Winter Simulation Conference*, 17.
- Sottinen T, Yazigi A (2014) Generalized Gaussian bridges. *Stochastic Processes and Their Applications* 124(9):3084–3105.
- Wood A, Chan G (1994) Simulation of stationary Gaussian processes in $[0, 1]^d$. *Journal of computational and graphical statistics* 3(4):409–432.

Appendix A: Detailed Proofs for Technical Lemmas in Section 3

A.1. Proof of Lemma 1 According to the definitions of a_j^k, c_j^k and \mathbf{B}_k^H , we have

$$\begin{aligned} \mathbb{V}_{k-1}(a_j^k - c_j^k) &= \mathbb{V}(B^H(t_{2j+1}^k) | \mathbf{B}_{k-1}^H) = \mathbb{V}(B^H(t_{2j+1}^k) | B^H(t_0^{k-1}), B^H(t_1^{k-1}), \dots, B^H(t_{2k-1}^{k-1})) \\ &\leq \mathbb{V}(B^H(t_{2j+1}^k) | B^H(t_j^{k-1})). \end{aligned}$$

where the inequality follows from the fact that for random variables X, Y and Z , we have $\mathbb{E}[\mathbb{V}(X|Y, Z)] \leq \mathbb{E}[\mathbb{V}(X|Y)]$ and when the joint distribution of (X, Y, Z) is multivariate Gaussian, the conditional variance $\mathbb{V}(X|Y, Z)$ and $\mathbb{V}(X|Y)$ are both constants.

The upper bound of conditional variance $\mathbb{V}(B^H(t_{2j+1}^k)|B^H(t_j^{k-1}))$ is based on the orthogonal bridge decomposition of Gaussian bridge (Sottinen and Yazigi (2014)). Specifically, let $r(\cdot, \cdot)$ denote the covariance function of B^H , then the distribution of $B^H(t)$ conditional on $B^H(t_j^{k-1}) = y^*$ is same with that of

$$B_{t_j^{k-1}, y^*}^H(t) = B^H(t) - \frac{r(t, t_j^{k-1})}{r(t_j^{k-1}, t_j^{k-1})} \cdot (B^H(t_j^{k-1}) - y^*).$$

By simple calculation, the covariance function of $B_{t_j^{k-1}, y^*}^H(t)$ is given by

$$r_{t_j^{k-1}, y^*}^{k-1}(s, t) = r(s, t) - \frac{r(t, t_j^{k-1}) \cdot r(t_j^{k-1}, s)}{r(t_j^{k-1}, t_j^{k-1})}.$$

Recall that $\Delta_k = t_{2j+1}^k - t_j^{k-1} = 2^{-k}$, then we have

$$\begin{aligned} \mathbb{V}(B^H(t_{2j+1}^k)|B^H(t_j^{k-1})) &= r_{t_j^{k-1}, y^*}^{k-1}(t_{2j+1}^k, t_{2j+1}^k) \\ &= |t_j^{k-1} + \Delta_k|^{2H} - 1/4 \cdot (|t_j^{k-1} + \Delta_k|^{2H} + |t_j^{k-1}|^{2H} - |\Delta_k|^{2H})^2 \cdot |t_j^{k-1}|^{-2H}. \end{aligned}$$

In the next, we show that $\mathbb{V}(B^H(t_{2j+1}^k)|B^H(t_j^{k-1}))$ is no greater than $2 \cdot \Delta_k^{2H}$, which is equivalent to

$$|2j+1|^{2H} - 1/4 \cdot (|2j+1|^{2H} + |2j|^{2H} - 1)^2 \cdot |2j|^{-2H} \leq 2,$$

for arbitrary positive integers j . By calculation, the left-hand side of above inequality becomes

$$\frac{1}{4} \cdot \left[2 \cdot |2j+1|^{2H} - \left(\frac{|2j+1|^{4H}}{|2j|^{2H}} + |2j|^{2H} \right) \right] - \frac{1}{4} \cdot |2j|^{-2H} + \frac{1}{2} \cdot \left(\frac{|2j+1|^{2H}}{|2j|^{2H}} + 1 \right).$$

Obviously, this quantity is no greater than 2. Thus, above all, we obtain

$$\mathbb{V}_{k-1}(a_j^k - c_j^k) \leq \mathbb{V}(B^H(t_{2j+1}^k)|B^H(t_j^{k-1})) \leq 2 \cdot \Delta_k^{2H} = 2 \cdot 2^{-2kH},$$

which concludes the proof.

A.2. Proof of Lemma 2 Let $Y = \max_{1 \leq i \leq n} |X_i|$. For any $t > 0$, by Jensen's inequality, we have

$$\begin{aligned} \exp\{t\mathbb{E}[Y]\} &\leq \mathbb{E}[\exp\{tY\}] = \mathbb{E}\left[\max_{1 \leq i \leq n} \exp\{t|X_i|\}\right] \leq \sum_{i=1}^n \mathbb{E}[\exp\{t|X_i|\}] \\ &\leq \sum_{i=1}^n \mathbb{E}[\exp\{tX_i\}] + \mathbb{E}[\exp\{-tX_i\}] \leq 2n \cdot \exp\{t^2\sigma^2/2\}. \end{aligned}$$

Hence, we obtain $\mathbb{E}[Y] \leq \log(2n)/t + t\sigma^2/2$ and the result follows by setting $t = \sqrt{2\log(2n)}/\sigma$.

A.3. Proof of Lemma 3 In the following, we denote by

$$\Sigma_{11}^{(k-1)} = \begin{bmatrix} 1^{2H} & (1^{2H} + 3^{2H} - 2^{2H})/2 & (1^{2H} + 5^{2H} - 4^{2H})/2 & \dots & \dots \\ (3^{2H} + 1^{2H} - 2^{2H})/2 & 3^{2H} & (3^{2H} + 5^{2H} - 2^{2H})/2 & \dots & \dots \\ (5^{2H} + 1^{2H} - 4^{2H})/2 & (5^{2H} + 3^{2H} - 2^{2H})/2 & 5^{2H} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}_{2^{k-1} \times 2^{k-1}},$$

which is a 2^{k-1} by 2^{k-1} matrix with (i, j) -th entry $(|2i-1|^{2H} + |2j-1|^{2H} - |2i-2j|^{2H})/2$. Recall the definition of \mathbf{N}_{k-1} and \mathbf{M}_{k-1} , we have

$$\begin{aligned} & (\mathbf{N}_{k-1} - \mathbf{M}_{k-1}) \boldsymbol{\Sigma}_{22}^{(k-1)} (\mathbf{N}_{k-1} - \mathbf{M}_{k-1})^\top \\ &= \mathbf{M}_{k-1} \boldsymbol{\Sigma}_{22}^{(k-1)} \mathbf{M}_{k-1}^\top - \mathbf{M}_{k-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top - \boldsymbol{\Sigma}_{12}^{(k-1)} \mathbf{M}_{k-1}^\top + \boldsymbol{\Sigma}_{12}^{(k-1)} \cdot [\boldsymbol{\Sigma}_{22}^{(k-1)}]^{-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top \\ &= (\mathbf{M}_{k-1} \boldsymbol{\Sigma}_{22}^{(k-1)} \mathbf{M}_{k-1}^\top - \mathbf{M}_{k-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top - \boldsymbol{\Sigma}_{12}^{(k-1)} \mathbf{M}_{k-1}^\top + \boldsymbol{\Sigma}_{11}^{(k-1)}) \\ &\quad - (\boldsymbol{\Sigma}_{11}^{(k-1)} - \boldsymbol{\Sigma}_{12}^{(k-1)} \cdot [\boldsymbol{\Sigma}_{22}^{(k-1)}]^{-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top). \end{aligned} \quad (30)$$

Note that $(\boldsymbol{\Sigma}_{11}^{(k-1)} - \boldsymbol{\Sigma}_{12}^{(k-1)} \cdot [\boldsymbol{\Sigma}_{22}^{(k-1)}]^{-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top) \cdot \Delta_k^{2H}$ is the conditional covariance matrix

$$\text{Cov} \left[(B^H(t_1^k), B^H(t_3^k), \dots, B^H(t_{2^{k-1}}^k)) \mid \mathbf{B}_{k-1}^H \right].$$

Hence, the diagonal entries in second part in (30) are nonnegative. In the next, we compute the diagonal entries of the first part in (30). We use $\xi_{i,j}$ and $\eta_{i,j}$ to denote the (i, j) -th entry of $\boldsymbol{\Sigma}_{22}^{(k-1)}$ and $\boldsymbol{\Sigma}_{12}^{(k-1)}$, respectively. After some calculation, we obtain that the j -th diagonal entries of $\mathbf{M}_{k-1} \boldsymbol{\Sigma}_{22}^{(k-1)} \mathbf{M}_{k-1}^\top$ and $\mathbf{M}_{k-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top$ are

$$\begin{aligned} [\mathbf{M}_{k-1} \boldsymbol{\Sigma}_{22}^{(k-1)} \mathbf{M}_{k-1}^\top]_{j,j} &= 1/4 \cdot (\xi_{j,j} + \xi_{j,j+1} + \xi_{j+1,j} + \xi_{j+1,j+1}), \\ [\mathbf{M}_{k-1} [\boldsymbol{\Sigma}_{12}^{(k-1)}]^\top]_{j,j} &= 1/2 \cdot (\eta_{j,j} + \eta_{j+1,j}). \end{aligned}$$

By plugging in expression of $\xi_{i,j}, \eta_{i,j}$ and $\boldsymbol{\Sigma}_{11}^{(k-1)}$, we obtain that the j -th diagonal entry of the first part in (30) is $1 - 2^{2H-2}$, which is a constant. As a result, the j -th diagonal entry of $\boldsymbol{\Sigma}^{(k)}$ is upper bounded by

$$[\boldsymbol{\Sigma}^{(k)}]_{j,j} \leq (1 - 2^{2H-2}) \cdot \Delta_k^{2H} < 2 \cdot 2^{-2kH}.$$

A.4. Proof of Lemma 4 We use $f_k(t)$ to denote $B_k^H(t) - B_{k-1}^H(t)$ and then by definition,

$$\|B_k^H - B_{k-1}^H\|_\alpha = \|f_k\|_\alpha = \sup_{0 \leq s < t \leq 1} \frac{|f_k(s) - f_k(t)|}{|s - t|^\alpha}.$$

Recall that $f_k(t_j^{k-1}) = 0, f_k(t_{2j+1}^k) = a_j^k - b_j^k$ and $f_k(t)$ is linear over intervals $[t_{2j}^k, t_{2j+1}^k]$ and $[t_{2j+1}^k, t_{2j+2}^k]$, where $j = 0, \dots, 2^{k-1} - 1$. Note that $t_{2j}^k = t_j^{k-1}$. Let

$$\kappa = 2^k \cdot \max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - b_j^k|.$$

Then κ is the maximal slope of all linear pieces of $f_k(t)$. We make a discussion based on the locations of s and t .

case 1: $|s - t| \leq 2^{-(k-1)}$. If there exists some j^* such that $s, t \in [t_{2j^*}^k, t_{2j^*+2}^k]$, since κ is the maximal slope, it is easy to show that $|f_k(s) - f_k(t)| \leq \kappa \cdot |s - t|$. Otherwise, there exists some j^* such that $t_{2j^*-1}^k \leq s < t_{2j^*}^k < t \leq t_{2j^*+1}^k$. Then we have

$$\begin{aligned} |f_k(s) - f_k(t)| &= |f_k(s) - f_k(t_{2j^*}^k) + f_k(t_{2j^*}^k) - f_k(t)| \leq |f_k(s) - f_k(t_{2j^*}^k)| + |f_k(t_{2j^*}^k) - f_k(t)| \\ &\leq 2\kappa \cdot (|s - t_{2j^*}^k| + |t - t_{2j^*}^k|) = 2\kappa \cdot |s - t|. \end{aligned}$$

Hence, by definition, we obtain

$$\frac{|f_k(s) - f_k(t)|}{|s - t|^\alpha} \leq 2\kappa \cdot |s - t|^{1-\alpha} \leq 2\kappa \cdot 2^{-(1-\alpha)k} = 2^{\alpha(k-1)+2} \cdot \max_{0 \leq j \leq 2^{k-1} - 1} |a_j^k - b_j^k|.$$

case 2: $|s - t| > 2^{-(k-1)}$. In this case, there exist some $i < j$ such that $s \in [t_i^k, t_{i+1}^k]$ and $t \in [t_j^k, t_{j+1}^k]$. Then we have

$$\begin{aligned} |f_k(s) - f_k(t)| &= |f_k(s) - f_k(t_{i+1}^k) + f_k(t_j^k) - f_k(t)| \leq |f_k(s) - f_k(t_{i+1}^k)| + |f_k(t_j^k) - f_k(t)| \\ &\leq 2\kappa \cdot (|s - t_{i+1}^k| + |t - t_j^k|) \leq 2^{-k+2} \cdot \kappa. \end{aligned}$$

Moreover, we have

$$\frac{|f_k(s) - f_k(t)|}{|s - t|^\alpha} \leq \frac{2^{-k+2} \cdot \kappa}{2^{-\alpha(k-1)}} = 2^{\alpha(k-1)+2} \cdot \max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k|.$$

Above all, we obtain $\|B_k^H - B_{k-1}^H\|_\alpha \leq 2^{\alpha(k-1)+2} \cdot \max_{0 \leq j \leq 2^{k-1}-1} |a_j^k - b_j^k|$, which concludes the proof of Lemma 4.

Appendix B: Detailed Proofs for Technical Lemmas in Section 4

B.1. Proof of Lemma 5 Before we prove Lemma 5, we first establish an upper bound on $\Xi_n^{(m,k)}$, i.e. the conditional moment generating function of $\beta^\top \alpha_n(m, k)$.

LEMMA 12. *Under the BCE condition, for $m \geq 1$, $1 \leq k \leq 2^{n+m-1}$ and $\theta > 0$,*

$$\Xi_n^{(m,k)}(\theta) \leq \exp\left\{1/2 \cdot (\rho \ell_{\tau_n+m} \cdot \theta + \theta^2 \cdot 2^{-2(\tau_n+m)H})\right\}.$$

Proof of Lemma 12. We denote by

$$\mu_n(m, k) = \mathbb{E}[\beta^\top \alpha_n(m, k) | \mathbf{B}_n^H], \quad \sigma_n^2(m, k) = \mathbb{V}(\beta^\top \alpha_n(m, k) | \mathbf{B}_n^H).$$

The key of the proof is to establish proper upper bounds for $\mu_n(m, k)$ and $\sigma_n^2(m, k)$. We will first show that

$$\sigma_n^2(m, k) \leq 2^{-2(n+m)H}.$$

Similar to the proof of Lemma 1, for Gaussian random vectors, taking condition reduces the variance. Hence, we have

$$\begin{aligned} \mathbb{V}(\beta^\top \alpha_n(m, k) | \mathbf{B}_n^H) &\leq \mathbb{V}\left(1/2 \cdot (B^H(t_{2k-2}^{n+m}) + B^H(t_{2k}^{n+m})) - B^H(t_{2k-1}^{n+m})\right) \\ &\leq 2 \cdot \mathbb{V}\left(1/2 \cdot (B^H(t_{2k-2}^{n+m}) - B^H(t_{2k-1}^{n+m}))\right) + 2 \cdot \mathbb{V}\left(1/2 \cdot (B^H(t_{2k}^{n+m}) - B^H(t_{2k-1}^{n+m}))\right). \end{aligned}$$

Since the fBM is a stationary process, we have

$$\mathbb{V}(B^H(t_{2k-1}^{n+m}) - B^H(t_{2k-2}^{n+m})) = \mathbb{V}(B^H(t_{2k}^{n+m}) - B^H(t_{2k-1}^{n+m})) = \mathbb{V}(B^H(\Delta_{n+m})) = \Delta_{n+m}^{2H},$$

which further implies that

$$\sigma_n^2(m, k) \leq 2^{-2(n+m)H}.$$

Under the BCE condition, we have $\mu_n(m, k) \leq \rho/2 \cdot \ell_{n+m}$. Then the upper bound of the moment generating function $\Xi_n^{(m,k)}(\theta)$ holds. \square

Proof of Lemma 5. We prove the case when $\pi = +$ only. The case when $\pi = -$ follows analogously. Note that if there is a up-crossing record-breaker at level $n + m$ position k , then $\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) > \rho \ell_{n+m}$. In this case, we have

$$\begin{aligned} \Theta_n^+(m, k) &= g_n(m)^{-1} \cdot 2^{n+m} \cdot \exp\left\{-\theta_n^+(m) \cdot \boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) + \log(\Xi_n^{(m,k)}(\theta_n^+(m)))\right\} \\ &\leq g_n(m)^{-1} \cdot 2^{n+m} \cdot \exp\{-\theta_n^+(m) \cdot \rho \ell_{n+m}\} \cdot \exp\left\{1/2 \cdot (\rho \theta_n^+(m) \ell_{n+m} + \theta_n^+(m)^2 \cdot 2^{-2(n+m)H})\right\} \\ &\leq g_n(m)^{-1} \cdot 2^{n+m} \cdot \exp\left\{-\rho/2 \cdot \theta_n^+(m) \cdot 2^{-(n+m)(H-\delta)} + 1/2 \cdot \theta_n^+(m)^2 \cdot 2^{-2(n+m)H}\right\}, \end{aligned}$$

where the second inequality follows from Lemma 12. Hence, by our choice of

$$\theta_n^+(m) = \rho/2 \cdot 2^{(n+m)(H+\delta)},$$

we obtain

$$\Theta_n^+(m, k) \leq g_n(m)^{-1} \cdot 2^{n+m} \cdot \exp\{-\rho^2/8 \cdot 2^{2(n+m)\delta}\} = Z_n \leq 1,$$

for all $n > N^*(\rho, \delta)$. \square

B.2. Proof of Lemma 6 We first consider the conditional expectation. Note that

$$\mu_n(m, k) = \mathbb{E}[\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) | \mathbf{B}_n^H] = (1/2, -1, 1/2)^\top \boldsymbol{\Sigma}_n^{(m,k)} \boldsymbol{\Sigma}_n^{-1} \mathbf{B}_n^H.$$

Since $\mu_n(m, k)$ is the linear combination of Gaussian random variables, itself is also Gaussian. It is easy to see that $\mathbb{E}[\mu_n(m, k)] = 0$. For the variance, according to the decomposition of conditional variance, we have

$$\mathbb{V}(\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k)) = \mathbb{E}[\mathbb{V}(\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) | \mathbf{B}_n^H)] + \mathbb{V}(\mathbb{E}[\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) | \mathbf{B}_n^H]).$$

Thus we obtain

$$\mathbb{V}(\mu_n(m, k)) \leq \mathbb{V}(\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k)) + \mathbb{E}[\mathbb{V}(\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) | \mathbf{B}_n^H)] \leq 2 \cdot 2^{-2(n+m)H},$$

where the inequality follows from the proof of Lemma 12. For fixed n, m, k , we define the event

$$\mathcal{E}_n(m, k) = \{|\mu_n(m, k)| > \rho/2 \cdot 2^{-(n+m)(H-\delta)}\}.$$

Then we have

$$\mathbb{P}(\mathcal{E}_n(m, k)) \leq C \exp\{-\rho^2/8 \cdot 2^{2(n+m)\delta}\},$$

where C is some constant. Note that $\mathcal{E}_n \subseteq \bigcup_{m=1}^{\infty} \bigcup_{k=1}^{2^{n+m-1}} \mathcal{E}_n(m, k)$. Then we have

$$\begin{aligned} \mathbb{P}(\mathcal{E}_n) &\leq \sum_{m=1}^{\infty} 2^{n+m-1} \cdot \mathbb{P}(\mathcal{E}_n(m, k)) \\ &\leq C \cdot \sum_{m=1}^{\infty} 2^{n+m-1} \cdot \exp\{-\rho^2/8 \cdot 2^{2(n+m)\delta}\} \leq C' \cdot \exp\{-C'' \cdot 2^{2(n+m)\delta}\}, \end{aligned}$$

where C' and C'' are some constants. Since

$$\sum_{n=1}^{\infty} C' \cdot \exp\{-C'' \cdot 2^{2(n+m)\delta}\} < \infty,$$

by Borel-Cantelli Lemma, events $\{\mathcal{E}_n\}_{n \geq 1}$ happen finite times almost surely. Moreover, similar to the proof of Theorem 3, we further have that the moment generating function of \mathcal{E} , the last level where the BCE condition is broken, exists everywhere.

B.3. Proof of Lemma 7 Recall that we have

$$\mu_n(m, k) = \mathbb{E}[\boldsymbol{\beta}^\top \boldsymbol{\alpha}_n(m, k) | \mathbf{B}_n^H] = (1/2, -1, 1/2)^\top \boldsymbol{\Sigma}_n^{(m, k)} \boldsymbol{\Sigma}_n^{-1} \mathbf{B}_n^H.$$

Here $\boldsymbol{\Sigma}_n$ is the covariance matrix of \mathbf{B}_n^H and $\boldsymbol{\Sigma}_n^{(m, k)}$ is the covariance matrix of $\boldsymbol{\alpha}_n(m, k)$ and \mathbf{B}_n^H , which has form

$$\boldsymbol{\Sigma}_n^{(m, k)} = \begin{bmatrix} r((2k-2)/2^{n+m}, 0) & r((2k-2)/2^{n+m}, 1/2^n) & \dots & \dots & r((2k-2)/2^{n+m}, 1) \\ r((2k-1)/2^{n+m}, 0) & r((2k-1)/2^{n+m}, 1/2^n) & \dots & \dots & r((2k-1)/2^{n+m}, 1) \\ r(2k/2^{n+m}, 0) & r(2k/2^{n+m}, 1/2^n) & \dots & \dots & r(2k/2^{n+m}, 1) \end{bmatrix}_{3 \times (2^n+1)},$$

where $r(\cdot, \cdot)$ denotes the covariance function of fBM. By calculation, for $1 \leq j \leq 2^n + 1$, the j -th entry of $(1/2, -1, 1/2)^\top \boldsymbol{\Sigma}_n^{(m, k)}$ is

$$\frac{\Delta_{m+n}^{2H}}{4} \cdot \left(|2k-2|^{2H} + |2k|^{2H} - 2|2k-1|^{2H} + |2k-2-j2^m|^{2H} + |2k-j2^m|^{2H} - 2|2k-1-j2^m|^{2H} \right).$$

We consider the function $\phi(x) = (x+2)^{2H} + x^{2H} - 2 \cdot (x+1)^{2H}$. Then we have $\phi'(x) = 2H \cdot ((x+2)^{2H-1} + x^{2H-1} - 2 \cdot (x+1)^{2H-1})$. When $1/2 < H < 1$, by Jensen's inequality, $\phi'(x) \leq 0$ for $x \geq 0$. Hence $\phi(x)$ is decreasing on $[0, \infty]$. Moreover, $\phi(x)$ is convex on $[0, \infty]$. As a result, for all positive integer i , we have

$$0 \leq (i+2)^{2H} + i^{2H} - 2 \cdot (i+1)^{2H} \leq 2^{2H} - 2 < 2.$$

When $0 < H < 1/2$, by Jensen's inequality, $\phi'(x) \geq 0$ for $x \geq 0$. Hence $\phi(x)$ is increasing on $[0, \infty]$. Moreover, $\phi(x)$ is concave on $[0, \infty]$. As a result, for all positive integer i , we have

$$-2 < 2^{2H} - 2 \leq (i+2)^{2H} + i^{2H} - 2 \cdot (i+1)^{2H} \leq 0.$$

Hence, for all H and i , we have $||2i-2|^{2H} + |2i|^{2H} - 2|2i-1|^{2H}| \leq 2$, which implies that the absolute value of any entry of $(1/2, -1, 1/2)^\top \boldsymbol{\Sigma}_n^{(m, k)}$ is bounded by $\Delta_{n+m}^{2H} = 2^{-2(n+m)H}$. Recall that γ_n denotes the maximal absolute value of the entries in vector $\boldsymbol{\Sigma}_n^{-1} \mathbf{B}_n^H$. Hence, we have

$$|\mu_n(m, k)| = |(1/2, -1, 1/2)^\top \boldsymbol{\Sigma}_n^{(m, k)} \boldsymbol{\Sigma}_n^{-1} \mathbf{B}_n^H| \leq \gamma_n \cdot (2^n + 1) \cdot 2^{-2(n+m)H}.$$

Appendix C: Detailed Proofs for Technical Lemmas in Section 5

C.1. Proof of Lemma 9 Let $\omega = (2dhC_\alpha K(2\alpha)|\nabla \mathbf{f}|)^{-1/\alpha}$ and $T_k = k\omega$, $k = 0, 1, 2, \dots$. Then the union of $[T_k, T_{k+1}]$ where $k = 0, \dots, \lceil \omega^{-1} \rceil$, covers $[0, 1]$. According to the property of Young integral, we have that for all $s, t \in [T_k, T_{k+1}]$,

$$\begin{aligned} \left| \int_s^t \mathbf{f}(\mathbf{y}(u)) d\mathbf{x}(u) \right| &\leq |\mathbf{f}(\mathbf{y}(s)) \cdot (\mathbf{x}(t) - \mathbf{x}(s))| + d \cdot K(2\alpha) \cdot H_{\alpha, [T_k, T_{k+1}]}(\mathbf{f}(\mathbf{y})) \cdot C_\alpha \cdot |s - t|^\alpha \\ &\leq \left(h \cdot |\mathbf{f}| \cdot C_\alpha + dh \cdot K(2\alpha) \cdot C_\alpha \cdot |\nabla \mathbf{f}| \cdot H_{\alpha, [T_k, T_{k+1}]}(\mathbf{y}) \cdot |s - t|^\alpha \right) \cdot |s - t|^\alpha. \end{aligned}$$

Since \mathbf{y} is the solution of equation (25), it is easy to see that

$$H_{\alpha, [T_k, T_{k+1}]}(\mathbf{y}) \leq h \cdot |\mathbf{f}| \cdot C_\alpha + dh \cdot K(2\alpha) \cdot C_\alpha \cdot |\nabla \mathbf{f}| \cdot H_{\alpha, [T_k, T_{k+1}]}(\mathbf{y}) \cdot \omega^\alpha,$$

which further implies that

$$\sup_{0 \leq k \leq \lceil \omega^{-1} \rceil} H_{\alpha, [T_k, T_{k+1}]}(\mathbf{y}) \leq 2h \cdot |\mathbf{f}| \cdot C_\alpha.$$

Now we turn to bound $H_{\alpha,[0,1]}(\mathbf{y})$. For any $s, t \in [0, 1]$, we assume that $T_i \leq s < T_{i+1} \leq \dots \leq T_j \leq t < T_{j+1}$. Then we have

$$\begin{aligned} |\mathbf{y}(s) - \mathbf{y}(t)| &\leq |\mathbf{y}(s) - \mathbf{y}(T_{i+1})| + \sum_{\ell=i+1}^{j-1} |\mathbf{y}(T_\ell) - \mathbf{y}(T_{\ell+1})| + |\mathbf{y}(T_j) - \mathbf{y}(t)| \\ &\leq \sup_{0 \leq k \leq \lfloor \omega^{-1} \rfloor} H_{\alpha,[T_k, T_{k+1}]}(\mathbf{y}) \cdot \left(|s - T_{i+1}|^\alpha + \sum_{\ell=i+1}^{j-1} |T_\ell - T_{\ell+1}|^\alpha + |T_j - t|^\alpha \right). \end{aligned}$$

Since $\alpha < 1$, using Jensen's inequality and the number of covering subintervals, we have

$$|s - T_{i+1}|^\alpha + \sum_{\ell=i+1}^{j-1} |T_\ell - T_{\ell+1}|^\alpha + |T_j - t|^\alpha \leq \lceil \omega^{-1} \rceil^{1-\alpha} \cdot |s - t|^\alpha.$$

So we have $H_{\alpha,[0,1]}(\mathbf{y}) \leq 2h \cdot \lceil \omega^{-1} \rceil^{1-\alpha} \cdot |\mathbf{f}| \cdot C_\alpha$, which is the first bound.

For another bound, note that according to Young-Lóeve estimate, for any partition Π of $[s, t]$, we have

$$\left| \sum_{t_i \in \Pi} \mathbf{f}(\mathbf{y}(t_i)) \cdot (\mathbf{x}(t_{i+1}) - \mathbf{x}(t_i)) - \mathbf{f}(\mathbf{y}(s)) \cdot (\mathbf{x}(t) - \mathbf{x}(s)) \right| \leq dh \cdot K(2\alpha) \cdot |\nabla \mathbf{f}| \cdot H_{\alpha,[0,1]}(\mathbf{y}) \cdot C_\alpha |s - t|^{2\alpha}.$$

Let the partition mesh goes to zero, we obtain the result.

C.2. Proof of Lemma 10 For convenience, in this proof, we use \mathbf{x}_k to denote $\mathbf{x}(t_k^n)$ and \mathbf{y}_k to denote $\mathbf{y}_n(t_k^n)$. For each $0 \leq j \leq \ell \leq 2^n$, let $\mathbf{I}_{j\ell} = \mathbf{y}_\ell - \mathbf{y}_j - \mathbf{f}(\mathbf{y}_j) \cdot (\mathbf{x}_{j+1} - \mathbf{x}_j)$. We first show that for all $t_r^n, t_j^n \in [0, 1]$, whenever $|t_r^n - t_j^n| \leq \omega$, then $|\mathbf{I}_{jr}| \leq L|t_r^n - t_j^n|^{2\alpha}$. By the definition of Euler scheme, $\mathbf{I}_{jr} = 0$, if $r - j = 0, 1$. For $r - j \geq 2$, we prove this lemma via induction. Suppose that the claim holds true for all pairs p, q with $q - p < r - j$. Let $\ell \in [j, r)$ be the largest integer such that $|t_j^n - t_\ell^n| \leq 1/2 \cdot |t_j^n - t_r^n|$. Then we have $|t_{\ell+1}^n - t_r^n| \leq 1/2 \cdot |t_j^n - t_r^n|$. By the inductive hypothesis, $|\mathbf{I}_{j\ell}| \leq L|t_\ell^n - t_j^n|^{2\alpha}$ and hence, we have

$$|\mathbf{y}_\ell^i - \mathbf{y}_j^i| \leq |\mathbf{I}_{j\ell}^i| + \left| \sum_{k=1}^h f_{ik}(\mathbf{y}_j)(x_\ell^k - x_j^k) \right| \leq L|t_\ell^n - t_j^n|^{2\alpha} + h \cdot |\mathbf{f}| \cdot C_\alpha |t_\ell^n - t_j^n|^\alpha.$$

Furthermore, since $|t_\ell^n - t_j^n| \leq \omega = (h|\mathbf{f}|C_\alpha/L)^{1/\alpha}$, we have $|\mathbf{y}_\ell - \mathbf{y}_j| \leq 2h \cdot |\mathbf{f}| \cdot C_\alpha |t_\ell - t_j|^\alpha$. Note that for $j \leq \ell \leq r$,

$$\mathbf{I}_{jr}^i = \mathbf{I}_{j\ell}^i + \mathbf{I}_{\ell r}^i + \sum_{k=1}^h (f_{ik}(\mathbf{y}_\ell) - f_{ik}(\mathbf{y}_j))(x_r^k - x_\ell^k).$$

Then we have

$$\begin{aligned} |\mathbf{I}_{jr}^i| &\leq |\mathbf{I}_{j\ell}^i| + |\mathbf{I}_{\ell r}^i| + h|\nabla \mathbf{f}| \cdot |\mathbf{y}_\ell - \mathbf{y}_j| \cdot C_\alpha |t_r^n - t_\ell^n|^\alpha \\ &\leq |\mathbf{I}_{j\ell}^i| + |\mathbf{I}_{\ell r}^i| + 2h^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}| \cdot C_\alpha^2 \cdot |t_r^n - t_j^n|^{2\alpha}. \end{aligned}$$

Similarly, we have

$$|\mathbf{I}_{\ell r}^i| \leq |\mathbf{I}_{\ell, \ell+1}^i| + |\mathbf{I}_{\ell+1, r}^i| + 2h^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}| \cdot C_\alpha^2 \cdot |t_r^n - t_j^n|^{2\alpha}.$$

Since $\mathbf{I}_{\ell, \ell+1}^i = 0$, we get

$$|\mathbf{I}_{jr}^i| \leq |\mathbf{I}_{j\ell}^i| + |\mathbf{I}_{\ell+1, r}^i| + (2hC_\alpha)^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}| \cdot |t_r^n - t_j^n|^{2\alpha}.$$

By applying the inductive hypothesis again, we obtain

$$\begin{aligned} |I_{jr}^i| &\leq L(|t_j - t_\ell|^{2\alpha} + |t_{\ell+1} - t_r|^{2\alpha}) + (2hC_\alpha)^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}| \cdot |t_r^n - t_j^n|^{2\alpha} \\ &\leq (2^{1-2\alpha}L + (2hC_\alpha)^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}|) \cdot |t_r^n - t_j^n|^{2\alpha} \\ &= L \cdot |t_r^n - t_j^n|^{2\alpha}, \end{aligned}$$

where $L = (1 - 2^{1-2\alpha})^{-1} \cdot (2hC_\alpha)^2 |\nabla \mathbf{f}| \cdot |\mathbf{f}|$. Hence, by induction, we finish the proof. Recall the definition of \mathbf{I}_{jr} , we also have that if $|t_r^n - t_j^n| \leq \omega$,

$$|\mathbf{y}_r - \mathbf{y}_j| \leq |\mathbf{I}_{jr}| + h|\mathbf{f}|C_\alpha |t_r^n - t_j^n|^\alpha \leq (L + h|\mathbf{f}|C_\alpha) \cdot |t_r^n - t_j^n|^\alpha.$$

Now we come back to the proof of Lemma 10. For $t_r^n, t_j^n \in [0, 1]$ with $|t_r^n - t_j^n| \leq \omega$, we already obtain the conclusion. Otherwise, we can decompose the interval $[t_r^n, t_j^n]$ as

$$t_r^n = t_{k_0}^n < t_{k_1}^n < t_{k_2}^n < \cdots < t_{k_m}^n = t_j^n,$$

such that $|t_{k_{i+1}}^n - t_{k_i}^n| \leq \omega$ or $k_{i+1} - k_i = 1$, $i = 0, 1, \dots, m-1$. In either case, it is easy to see that

$$|\mathbf{y}_{k_{i+1}} - \mathbf{y}_{k_i}| \leq (L + h|\mathbf{f}|C_\alpha) \cdot |t_{k_{i+1}}^n - t_{k_i}^n|^\alpha,$$

and hence

$$|\mathbf{y}_j - \mathbf{y}_r| \leq (L + h|\mathbf{f}|C_\alpha) \cdot |t_j^n - t_r^n|^\alpha \cdot m \leq (L + h|\mathbf{f}|C_\alpha) \cdot (1 + \omega^{-1}) \cdot |t_j^n - t_r^n|^\alpha.$$

Furthermore, we have

$$\begin{aligned} |\mathbf{I}_{jr}| &\leq |\mathbf{y}_j - \mathbf{y}_r| + h|\mathbf{f}|C_\alpha |t_j^n - t_r^n|^\alpha \\ &\leq (2 + \omega^{-1}) \cdot (L + h|\mathbf{f}|C_\alpha) \cdot |t_j^n - t_r^n|^\alpha \\ &\leq (2\omega^{-\alpha} + \omega^{-1-\alpha}) \cdot (L + h|\mathbf{f}|C_\alpha) \cdot |t_j^n - t_r^n|^{2\alpha}, \end{aligned}$$

where we use $|t_j^n - t_r^n| \geq \omega$ in the last inequality. Hence we conclude the proof of Lemma 10.

C.3. Proof of Lemma 11 Based on Taylor's expansion, for $t_i^n, t_j^n \in D_n$, we have

$$\begin{aligned} &\left| [\mathbf{f}(\mathbf{y}(t_i^n)) - \mathbf{f}(\mathbf{y}(t_j^n))] - [\mathbf{f}(\mathbf{y}_n(t_i^n)) - \mathbf{f}(\mathbf{y}_n(t_j^n))] \right| \\ &= \left| \int_0^1 \nabla \mathbf{f}(\mathbf{y}(t_i^n) + \tau \cdot [\mathbf{y}_n(t_i^n) - \mathbf{y}(t_i^n)])^\top [\mathbf{y}(t_i^n) - \mathbf{y}_n(t_i^n)] d\tau - \right. \\ &\quad \left. \int_0^1 \nabla \mathbf{f}(\mathbf{y}(t_j^n) + \tau \cdot [\mathbf{y}_n(t_j^n) - \mathbf{y}(t_j^n)])^\top [\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n)] d\tau \right| \\ &\leq \left| \int_0^1 \nabla \mathbf{f}(\mathbf{y}(t_i^n) + \tau \cdot [\mathbf{y}_n(t_i^n) - \mathbf{y}(t_i^n)])^\top [(\mathbf{y}(t_i^n) - \mathbf{y}_n(t_i^n)) - (\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n))] d\tau \right| + \\ &\quad \left| \int_0^1 \nabla \mathbf{f}(\mathbf{y}(t_j^n) + \tau \cdot [\mathbf{y}_n(t_j^n) - \mathbf{y}(t_j^n)]) - \nabla \mathbf{f}(\mathbf{y}(t_i^n) + \tau \cdot [\mathbf{y}_n(t_i^n) - \mathbf{y}(t_i^n)])^\top [\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n)] d\tau \right|. \end{aligned}$$

We use A and B to denote the two parts in above inequality. Recall the definition of restricted α -Hölder norm, we have

$$A \leq d \cdot |\nabla \mathbf{f}| \cdot H_\alpha(\mathbf{y} - \mathbf{y}_n | D_n) \cdot |t_i^n - t_j^n|^\alpha.$$

For the second term, by mean value theorem, it is easy to have

$$\begin{aligned} B &\leq d^2 \cdot |\nabla^2 \mathbf{f}| \cdot \left(H_\alpha(\mathbf{y} | D_n) + H_\alpha(\mathbf{y}_n | D_n) \right) \cdot |t_i^n - t_j^n|^\alpha \cdot |\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n)| \\ &\leq d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1) \cdot |t_i^n - t_j^n|^\alpha \cdot |\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n)|, \end{aligned}$$

where the second inequality follows from Lemma 9 and 10. Note that

$$\begin{aligned} |\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n)| &\leq |\mathbf{y}(0) - \mathbf{y}_n(0)| + |(\mathbf{y}(t_j^n) - \mathbf{y}_n(t_j^n)) - (\mathbf{y}(0) - \mathbf{y}_n(0))| \\ &\leq |\mathbf{y}(0) - \mathbf{y}_n(0)| + H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) \cdot |t_j^n|^\alpha. \end{aligned}$$

Then we have

$$H_\alpha(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{y}_n)|D_n) \leq (d \cdot |\nabla \mathbf{f}| + d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1)) \cdot H_\alpha(\mathbf{y} - \mathbf{y}_n|D_n) + d^2 \cdot |\nabla^2 \mathbf{f}| \cdot (G_1^* + G_1) \cdot |\mathbf{y}(0) - \mathbf{y}_n(0)|,$$

which concludes the proof of Lemma 11.