

# $\varepsilon$ -STRONG SIMULATION FOR MULTIDIMENSIONAL STOCHASTIC DIFFERENTIAL EQUATIONS VIA ROUGH PATH ANALYSIS<sup>1</sup>

BY JOSE BLANCHET, XINYUN CHEN AND JING DONG

*Columbia University, Wuhan University and Northwestern University*

Consider a multidimensional diffusion process  $X = \{X(t) : t \in [0, 1]\}$ . Let  $\varepsilon > 0$  be a *deterministic*, user defined, tolerance error parameter. Under standard regularity conditions on the drift and diffusion coefficients of  $X$ , we construct a probability space, supporting both  $X$  and an explicit, piecewise constant, fully simulatable process  $X_\varepsilon$  such that

$$\sup_{0 \leq t \leq 1} \|X_\varepsilon(t) - X(t)\|_\infty < \varepsilon$$

with probability one. Moreover, the user can adaptively choose  $\varepsilon' \in (0, \varepsilon)$  so that  $X_{\varepsilon'}$  (also piecewise constant and fully simulatable) can be constructed conditional on  $X_\varepsilon$  to ensure an error smaller than  $\varepsilon'$  with probability one. Our construction requires a detailed study of continuity estimates of the Itô map using Lyons' theory of rough paths. We approximate the underlying Brownian motion, jointly with the Lévy areas with a deterministic  $\varepsilon$  error in the underlying rough path metric.

## CONTENTS

1. Introduction . . . . .	276
2. Main results . . . . .	279
2.1. On relaxing boundedness assumptions . . . . .	282
2.2. The evaluation of $G$ . . . . .	282
3. The main idea of the algorithmic development . . . . .	283
3.1. Infinite sum representation of Brownian motion and Lévy area . . . . .	284
3.2. The idea of record breakers . . . . .	286
4. Tolerance-enforced simulation of bounds on $\alpha$ -Hölder norms . . . . .	288
5. Tolerance-enforced simulation for bounds on $2\alpha$ -Hölder norms of Lévy areas . . . . .	289
5.1. Conditional moment generating functions and associated exponential tilting . . . . .	291
5.2. Conditional large deviations estimates for $L_{i,j}^n(k)$ . . . . .	295
5.3. Joint Tolerance-Enforced Simulation for $\alpha$ -Hölder norms and proof of Theorem 2.2 . . . . .	296
6. Rough differential equations, error analysis and the proof of Theorem 2.1 . . . . .	302
7. Numerical implementation . . . . .	314
Appendix A: Proofs of results in Section 3 . . . . .	315
Appendix B: Proofs of results in Section 5 . . . . .	324
B.1. Proof of results in Section 5.1 . . . . .	324

Received March 2014; revised January 2016.

<sup>1</sup>Supported by NSF Grants CMMI 1069064 and DMS-13-20550.

*MSC2010 subject classifications.* Primary 34K50, 65C05, 82B80; secondary 97K60.

*Key words and phrases.* Stochastic differential equation, Monte Carlo method, Brownian motion, Lévy area, rough path.

B.2. Proofs of results in Section 5.2 . . . . . 329  
 Acknowledgments . . . . . 336  
 References . . . . . 336

**1. Introduction.** Consider the Itô Stochastic Differential Equation (SDE)

$$(1.1) \quad dX(t) = \mu(X(t)) dt + \sigma(X(t)) dZ(t), \quad X(0) = x(0),$$

where  $Z(\cdot)$  is a  $d'$ -dimensional Brownian motion, and  $\mu(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$  satisfy suitable regularity conditions. We shall assume, in particular, that both  $\mu(\cdot)$  and  $\sigma(\cdot)$  are Lipschitz continuous so that a strong solution to the SDE is guaranteed to exist. Additional assumptions on the first- and second-order derivatives of  $\mu(\cdot)$  and  $\sigma(\cdot)$ , which are standard in the theory of rough paths, will be discussed in the sequel.

Our contribution in this paper is the joint construction of  $X = \{X(t) : t \in [0, 1]\}$  and a family of processes  $X_\varepsilon = \{X_\varepsilon(t) : t \in [0, 1]\}$ , for each  $\varepsilon \in (0, 1)$ , supported on a probability space  $(\Omega, \mathcal{F}, P)$ , and such that the following properties hold:

(T1) The process  $X_\varepsilon$  is piecewise constant, with finitely many discontinuities in  $[0, 1]$ .

(T2) The process  $X_\varepsilon$  can be simulated exactly and, since it takes only finitely many values, its path can be fully stored.

(T3) We have that with  $P$ -probability one

$$(1.2) \quad \sup_{t \in [0,1]} \|X_\varepsilon(t) - X(t)\|_\infty < \varepsilon.$$

(T4) For any  $m > 1$  and  $0 < \varepsilon_m < \dots < \varepsilon_1 < 1$  we can simulate  $X_{\varepsilon_m}$  conditional on  $X_{\varepsilon_1}, \dots, X_{\varepsilon_{m-1}}$ .

We refer to the class of procedures which achieve the construction of such family  $\{X_\varepsilon : \varepsilon \in (0, 1)\}$  as *Tolerance-Enforced Simulation* (TES) or  $\varepsilon$ -strong simulation methods. Throughout the paper, we use  $\|\cdot\|_\infty$  to denote the max-norm on  $\mathbb{R}^d$ .

This paper provides the first construction of a Tolerance-Enforced Simulation procedure for multidimensional SDEs in substantial generality. All other TES or  $\varepsilon$ -strong simulation procedures up to now are applicable to one dimensional processes or multidimensional processes with constant diffusion matrix [i.e.,  $\sigma(x) = \sigma$ ].

Let us discuss some considerations that motivate our study. We first discuss how this paper relates to the current literature on  $\varepsilon$ -strong simulation of stochastic processes, which is a recent area of research. Chen and Huang [6] provides the construction of  $X_\varepsilon$  satisfying only (T1) to (T3), in one dimension. In particular, bound (1.2) is satisfied for a given fixed  $\varepsilon_0 = \varepsilon > 0$  and it is not clear how to jointly simulate  $\{X_{\varepsilon_m}\}_{m \geq 1}$  as  $\varepsilon_m \searrow 0$  applying the technique in [6]. The motivation of constructing  $X_{\varepsilon_0}$  for [6] came from the desire to produce exact samples from a one-dimensional diffusion  $X(\cdot)$  satisfying (1.1), and also assuming  $\sigma(\cdot)$  constant.

Chen and Huang [6] were interested in extending the applicability of an algorithm introduced by Beskos and Roberts; see [4]. The procedure of Beskos and

Roberts, applicable to one-dimensional diffusions, imposed strong boundedness assumptions on the drift coefficient and its derivative. The technique in [6] enabled an extension which is free of such boundedness assumptions by using a localization technique that allowed to apply the ideas behind the algorithm in [4]; see also [2] for another approach which eliminates boundedness assumptions. All of these developments are in the one-dimensional case.

The assumption of a constant diffusion coefficient comes at basically no cost in generality when considering one-dimensional diffusions because one can always apply Lamperti's (one-to-one) transformation. Such transformation allows to recast the simulation problem to one involving a diffusion with constant  $\sigma(\cdot)$ . Lamperti's transformation cannot be generally applied in higher dimensions.

Beskos, Peluchetti and Roberts [3] extends the work of [6] in that their algorithms satisfy (T1) to (T4), but also in the context of one dimensional processes. Pollock, Johansen and Roberts [11] not only provides an additional extension which allows to deal with one-dimensional SDEs with jumps, but also contains a comprehensive discussion on exact and  $\varepsilon$ -strong simulation for SDEs. Property (T4) in the definition of TES is desirable because it provides another approach at constructing unbiased estimators for expectations of the form  $Ef(X)$ , where  $f(\cdot)$  is, say, a continuous function of the sample path  $X$ . In order to see this, let us assume for simplicity that  $f(\cdot)$  is positive and Lipschitz continuous in the uniform norm with Lipschitz constant  $K$ . Then let  $T$  be any positive random variable with a strictly positive density  $g(\cdot)$  on  $[0, \infty)$  and define

$$(1.3) \quad Z := I(f(X) > T)/g(T).$$

Observe that

$$E[Z] = E[E[Z|X]] = E\left[\int_0^\infty I(f(X) > t) \frac{g(t)}{g(T)} dt\right] = E[f(X)],$$

so  $Z$  is an unbiased estimator for  $Ef(X)$ . Therefore, if Properties T(1) to T(4) hold, it is possible to simulate  $Z$  by noting that  $f(X_\varepsilon) > T + K\varepsilon$  implies  $f(X) > T$  and if  $f(X_\varepsilon) < T - K\varepsilon$ , then  $f(X) \leq T$ . Since (T4) allows to keep simulating as  $\varepsilon$  becomes smaller and  $T$  is independent of  $X_\varepsilon$  with a positive density  $g(\cdot)$ , then one eventually is able to simulate  $Z$  exactly.

The major obstacle involved in developing exact sampling algorithms for multidimensional diffusions is the fact that  $\sigma(\cdot)$  cannot be assumed to be constant. Moreover, even in the case of multidimensional diffusions with constant  $\sigma(\cdot)$ , the one-dimensional algorithms developed so far can only be extended to the case in which the drift coefficient  $\mu(\cdot)$  is the gradient of some function, that is, if  $\mu(x) = \nabla v(x)$  for some  $v(\cdot)$ . The reason is that in this case one can represent the likelihood ratio  $L(t)$ , between the solution to (1.1) and Brownian motion (assuming  $\sigma = I$  for simplicity) involving a Riemann integral as follows:

$$(1.4) \quad \begin{aligned} L(t) &= \exp\left(\int_0^t \mu(X(s)) dX(s) - \frac{1}{2} \int_0^t \|\mu(X(s))\|_2^2 ds\right) \\ &= \frac{\exp(v(X(t)))}{\exp(v(X(0)))} \exp\left(-\frac{1}{2} \int_0^t \lambda(X(s)) ds\right), \end{aligned}$$

for  $\lambda(x) = \Delta v(x) + \|\nabla v(x)\|_2^2$ . The fact that the stochastic integral can be transformed into a Riemann integral facilitates the execution of acceptance-rejection because one can interpret (up to a constant and using localization as in [6]) the exponential of the integral of  $\lambda(\cdot)$  as the probability that no arrivals occur in a Poisson process with a stochastic intensity. Such event (i.e., no arrivals) can be simulated by thinning.

So, our motivation in this paper is to investigate a novel approach that allows to study  $\varepsilon$ -strong simulation for multidimensional diffusions in substantial generality, without imposing the assumption that  $\sigma(\cdot)$  is constant or that a Lamperti-type transformation can be applied. Given the previous discussion on the connections between exact sampling and  $\varepsilon$ -strong simulation, and the limitations of the current techniques, we believe that our results here provide an important step in the development of exact sampling algorithms for general multidimensional diffusions. For example, in contrast to existing techniques, which demand  $L(t)$  to be expressed in terms of a Riemann integral as indicated in (1.4), our results here allow to approximate directly  $L(t)$  in terms of the stochastic integral representation [and thus one does not need to assume that  $\mu(x) = \nabla v(x)$ ]. We plan to report on these implications in future papers.

Our results already allow to obtain unbiased estimator of expectations of sample path functionals via (1.3). However, it is noted in [3] that the expected number of random variables required to simulate  $Z$  is typically infinite. Pollock, Johansen and Roberts [11] discusses via numerical examples the practical limitations of these types of estimators. Rhee and Glynn [12], also proposes unbiased estimators for the expectation of Lipschitz continuous functions of  $X(1)$  using randomized multilevel Monte Carlo. Nevertheless, their algorithm also exhibits infinite expected termination time, except when one can simulate the Lévy areas exactly, which currently can be done only in the context of two-dimensional SDEs using the results in [9].

Bayer, Friz, Riedel and Schoenmakers [1] also use rough path analysis for Monte Carlo estimation, but their focus is on connections to multilevel techniques and not on  $\varepsilon$ -strong simulation.

In this paper, we concentrate only on what *is possible to do* in terms of  $\varepsilon$ -strong simulation procedures and how to enable the use of rough path theory for  $\varepsilon$ -strong simulation. We shall study efficient implementations of the algorithms proposed in a separate paper. Other research avenues that we plan to investigate, and which leverage off our development in this paper, involve quantification of model uncertainty using the fact that our  $\varepsilon$ -strong simulation algorithms in the end are uniform for cases with a large class of drift and diffusion coefficients.

Finally, we note that in order to build our Tolerance-Enforced Simulation procedure we had to obtain new tools for the analysis of Lévy areas and associated conditional large deviations results for Lévy areas given the increments of Brownian motion. We believe that these technical results might be of independent interest.

The rest of the paper is organized as follows. In Section 2 we describe the two main results of the paper. The first of them, Theorem 2.1, provides an error bound

between the solution to the SDE described in (1.1) and a suitable piecewise constant approximation. The second result, Theorem 2.2, refers to the procedures that are involved in simulating the bounds, jointly with the piecewise constant approximation, thereby yielding (1.2). Section 3 is divided into two subsections and it builds the elements behind the proof of Theorem 2.2. As it turns out, one needs to simulate bounds on the so-called Hölder norms of the underlying Brownian motion and the corresponding Lévy areas. Section 4 lays out the details of the simulation of the Brownian motion and an upper bound of its  $\alpha$ -Hölder norm and Section 5 lays out the details of the simulation of the Lévy areas and an upper bound of its  $2\alpha$ -Hölder norm. Section 6 is also divided in several parts, corresponding to the elements of rough path theory required to analyze the SDE described in (1.1) as a continuous map of Brownian motion under a suitable metric (described in Section 2). While the final form of the estimates in Section 6 might be somewhat different than those obtained in the literature on rough path analysis, the techniques that we use here are certainly standard in that literature. We have chosen to present the details because the techniques might not be well known to the Monte Carlo simulation community and also because our emphasis is in finding explicit constants (i.e., bounds) that are amenable to simulation.

**2. Main results.** Our approach consists in studying the process  $X$  as a transformation of the underlying Brownian motion  $Z$ . Such transformation is known as the Itô–Lyons map and its continuity properties are studied in the theory of rough paths, pioneered by T. Lyons, in [10]. A rough path is an effective way to summarize an irregular path information. The theory of rough paths allows to define the solution to an SDE such as (1.1) in a path-by-path basis (free of probability) by imposing constraints on the regularity of the iterated integrals of the underlying process  $Z$ . Namely, integrals of the form

$$(2.1) \quad A_{i,j}(s,t) = \int_s^t (Z_i(u) - Z_i(s)) dZ_j(u).$$

The theory results in different interpretations of the solution to (1.1) depending on how the iterated integrals of  $Z$  are interpreted. In this paper, we interpret the integral in (2.1) in the sense of Itô.

It turns out that the Itô–Lyons map is continuous under a suitable  $\alpha$ -Hölder metric defined in the space of rough paths. In particular, such metric can be expressed as the maximum of the following two quantities:

$$(2.2) \quad \|Z\|_\alpha := \sup_{0 \leq s < t \leq 1} \frac{\|Z(t) - Z(s)\|_\infty}{|t - s|^\alpha},$$

$$(2.3) \quad \|A\|_{2\alpha} := \sup_{0 \leq s < t \leq 1} \max_{1 \leq i, j \leq d'} \frac{|A_{i,j}(s,t)|}{|t - s|^{2\alpha}}.$$

As we shall discuss, continuity estimates of the Itô–Lyons map can be given explicitly in terms of these two quantities.

In the case of Brownian motion, as we consider here, we have that  $\alpha \in (1/3, 1/2)$ . It is shown in [7], that under suitable regularity conditions on  $\mu(\cdot)$  and  $\sigma(\cdot)$ , which we shall discuss momentarily, the Euler scheme provides an almost sure approximation in uniform norm to the solution to the SDE (1.1). Our first result provides an explicit characterization of all of the (path-dependent) quantities that are involved in the final error analysis (such as  $\|Z\|_\alpha$  and  $\|A\|_{2\alpha}$ ), the difference between our analysis and what has been done in previous developments is that ultimately we must be able to implement the Euler scheme jointly with the path-dependent quantities that are involved in the error analysis. So, it is not sufficient to argue that there exists a path-dependent constant that serves as a bound of some sort, we actually must provide a suitable representation that can be simulated in finite time.

In order to provide our first result, we introduce some notation. Let  $D_n$  denote the dyadic discretization of order  $n$  and  $\Delta_n$  denote the mesh of the discretization. Specifically,  $D_n := \{t_0^n, t_1^n, \dots, t_{2^n}^n\}$  where  $t_k^n = k/2^n$  for  $k = 0, 1, 2, \dots, 2^n$  and  $\Delta_n = 1/2^n$ .

Given  $\hat{X}^n(0) = x(0)$ , define  $\{\hat{X}^n(t) : t \in D_n\}$  by the following recursion:

$$\begin{aligned}
 \hat{X}_i^n(t_{k+1}^n) &= \hat{X}_i^n(t_k^n) + \mu_i(\hat{X}^n(t_k^n))\Delta_n + \sum_{j=1}^{d'} \sigma_{i,j}(\hat{X}^n(t_k^n))(Z_j(t_{k+1}^n) - Z_j(t_k^n)) \\
 (2.4) \quad &+ \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(\hat{X}^n(t_k^n)) \sigma_{l,m}(\hat{X}^n(t_k^n)) \tilde{A}_{m,j}^n(t_k^n, t_{k+1}^n),
 \end{aligned}$$

where  $\tilde{A}_{i,i}^n(t_k^n, t_{k+1}^n) = A_{i,i}(t_k^n, t_{k+1}^n) = (Z_i(t_{k+1}^n) - Z_i(t_k^n))^2/2 - \Delta_n/2$ , and  $\tilde{A}_{i,j}^n(t_k^n, t_{k+1}^n) = 0$  for  $i \neq j$ . We let  $\hat{X}^n(t) = \hat{X}^n(\lfloor t \rfloor)$  where  $\lfloor t \rfloor = \max\{t_k^n : t_k^n \leq t\}$  for  $t \in [0, 1]$ . We denote

$$R_{i,j}^n(t_l^n, t_m^n) := \sum_{k=l+1}^m \{A_{i,j}(t_{k-1}^n, t_k^n) - \tilde{A}_{i,i}^n(t_k^n, t_{k+1}^n)\},$$

and for fixed  $\beta \in (1 - \alpha, 2\alpha)$ , write

$$\Gamma_R := \sup_n \sup_{0 \leq s < t \leq 1, s, t \in D_n} \max_{1 \leq i, j \leq d'} \frac{|R_{i,j}^n(s, t)|}{|t - s|^\beta \Delta_n^{2\alpha - \beta}}.$$

We notice that when  $i = j$ ,  $R_{i,i}^n(t_l^n, t_m^n) = 0$ ; when  $i \neq j$ ,  $R_{i,j}^n(t_l^n, t_m^n) = \sum_{k=l+1}^m A_{i,j}(t_{k-1}^n, t_k^n)$ . We also redefine  $\|Z\|_\alpha$  and  $\|A\|_{2\alpha}$  as

$$\begin{aligned}
 \|Z\|_\alpha &:= \sup_n \sup_{0 \leq s < t \leq 1, s, t \in D_n} \frac{\|Z(t) - Z(s)\|_\infty}{|t - s|^\alpha}, \\
 \|A\|_{2\alpha} &:= \sup_n \sup_{0 \leq s < t \leq 1, s, t \in D_n} \max_{1 \leq i, j \leq d'} \frac{|A_{i,j}(s, t)|}{|t - s|^{2\alpha}}.
 \end{aligned}$$

The new definitions are equivalent to (2.2) and (2.3) since both  $Z$  and  $A$  are continuous processes. It is well known that a solution to  $X$  can be constructed path-by-path (see [7] and Section 6). The next result characterizes an explicit bound for the error obtained by approximating  $X$  using  $\hat{X}^n$ .

**THEOREM 2.1.** *Suppose that there exists a constant  $M$  such that  $\|\mu\|_\infty \leq M$ ,  $\|\nabla\mu\|_\infty \leq M$  and  $\|\sigma^{(i)}\|_\infty \leq M$  for  $i = 0, 1, 2, 3$ , where  $\sigma^{(i)}$  denotes the  $i$ th derivative of  $\sigma$ . If  $\|Z\|_\alpha \leq K_\alpha < \infty$ ,  $\|A\|_{2\alpha} \leq K_{2\alpha} < \infty$ , and  $\Gamma_R < K_R$ , we can compute  $G$  explicitly in terms of  $M$ ,  $K_\alpha$ ,  $K_{2\alpha}$  and  $K_R$ , such that*

$$\sup_{t \in [0,1]} \|\hat{X}^n(t) - X(t)\|_\infty \leq G \Delta_n^{2\alpha-\beta}.$$

**REMARK.** A recipe that explains step-by-step how to compute  $G$  in terms of algebraic expressions involving  $M$ ,  $K_\alpha$ ,  $K_{2\alpha}$  and  $K_R$  is given in Procedure A in Section 2.2.

Using Theorem 2.1, we can proceed to state the main contribution of this paper.

**THEOREM 2.2.** *In the context of Theorem 2.1, there is an explicit Monte Carlo procedure that allows us to simulate random variables  $K_\alpha$ ,  $K_{2\alpha}$  and  $K_R$  jointly with  $\{Z(t) : t \in D_n\}$  for any  $n \geq 1$ . Consequently, given any deterministic  $\varepsilon > 0$  we can select  $n(\varepsilon)$  such that  $G \Delta_{n(\varepsilon)}^{2\alpha-\beta} \leq \varepsilon$  and then set  $X_\varepsilon(t) = \hat{X}^n(t)$  so that*

$$(2.5) \quad \sup_{t \in [0,1]} \|X_\varepsilon(t) - X(t)\|_\infty \leq \varepsilon,$$

with probability one.

**REMARK.** An explicit description of the algorithm involved in the Monte Carlo procedure of Theorem 2.2 is given in Algorithm 2 at the end of Section 5.3, and the discussion that follows it.

Given  $\{Z(t) : t \in D_{n(\varepsilon)}\}$  so that (2.5) holds, the discussion in the remark that follows Algorithm 2 explains how to further simulate  $\{Z(t) : t \in D_{n'}\}$  for any  $n' > n(\varepsilon)$ . This refinement is useful in order to satisfy the important property (T4) given in the Introduction. In detail, once  $K_\alpha$ ,  $K_{2\alpha}$  and  $K_R$  have been simulated then  $G$  has also been simulated and evaluated. Consequently, given any sequence  $\varepsilon_m < \varepsilon_{m-1} < \dots < \varepsilon_1$  we just need to obtain  $n_i$  such that  $G \Delta_{n_i}^{2\alpha-\beta} \leq \varepsilon_i$ . Then simulate  $\{Z(t) : t \in D_{n_i}\}$  and construct  $\hat{X}^{n_i}(\cdot)$  according to (2.4). We let  $X_{\varepsilon_i}(t) = \hat{X}^{n_i}(t)$  and, owing to Theorem 2.1, we immediately obtain

$$\sup_{t \in [0,1]} \|X_{\varepsilon_i}(t) - X(t)\|_\infty \leq \varepsilon_i$$

with probability one, as desired.

2.1. *On relaxing boundedness assumptions.* The construction of  $\hat{X}^n(\cdot)$  in order to satisfy (2.5) assumes that  $\|\mu\|_\infty \leq M$ ,  $\|\mu^{(i)}\|_\infty \leq M$  and  $\|\sigma^{(i)}\|_\infty \leq M$  for  $i = 0, 1, 2, 3$ . Although these assumptions are strong, here we explain how to relax them. Theorem 2.2 extends directly to the case in which  $\mu$  and  $\sigma$  are Lipschitz continuous, with  $\mu$  differentiable and  $\sigma$  three times differentiable. Since  $\mu$  and  $\sigma$  are Lipschitz continuous, we know that  $X(\cdot)$  has a strong solution which is nonexplosive.

We can always construct  $\mu_M$  and  $\sigma_M$  so that  $\mu^{(i)}(x) = \mu_M^{(i)}(x)$  for  $\|x\|_\infty \leq c_M$  and  $i = 0, 1$  and  $\sigma^{(i)}(x) = \sigma_M^{(i)}(x)$  for  $\|x\|_\infty \leq c_M$  for  $i = 0, 1, 2, 3$ . Also we can construct  $c_M$ , where  $c_M \rightarrow \infty$  as  $M \rightarrow \infty$ , and  $\|\mu_M\|_\infty \leq M$ ,  $\|\mu_M^{(1)}\|_\infty \leq M$  and  $\|\sigma_M^{(i)}\|_\infty \leq M$  for  $i = 0, 1, 2, 3$ .

For  $M \geq 1$ , we consider the SDE (1.1) with  $\mu_M$  and  $\sigma_M$  as drift and diffusion coefficients, respectively, and let  $X_M(\cdot)$  be the corresponding solution to (1.1). We start by picking some  $M_0 \geq 1$  such that  $\varepsilon < c_{M_0}$  and let  $M = M_0$ . Then run Algorithm 2 to produce  $\{\hat{X}_M^n(t) : t \in [0, 1]\}$ , which according to Theorem 2.2 satisfies,

$$\sup_{t \in [0, 1]} \|\hat{X}_M^n(t) - X_M(t)\|_\infty \leq \varepsilon.$$

Note that only Steps 5 to 8 in Algorithm 2 depend on the SDE (1.1), through the evaluation of  $G$ , which depends on  $M$  and so we write  $G_M := G$ . If  $\sup_{t \in [0, 1]} \|\hat{X}_M^n(t)\|_\infty \leq c_M - \varepsilon$ , then we must have that  $X(t) = X_M(t)$  for  $t \in [0, 1]$  and we are done. Otherwise, we let  $M \leftarrow 2M$  and run again only Steps 5 to 8 of Algorithm 2. We repeat doubling  $M$  and re-running Steps 5 to 8 (updating  $G_M$ ) until we obtain a solution for which  $\sup_{t \in [0, 1]} \|\hat{X}_M^n(t)\|_\infty \leq c_M - \varepsilon$ . Eventually, this must occur because

$$\lim_{M \rightarrow \infty} \sup_{t \in [0, 1]} \|X_M(t) - X(t)\|_\infty = 0$$

almost surely and  $X(\cdot)$  is non explosive.

2.2. *The evaluation of  $G$ .* We next summarize the way to calculate  $G$  in terms of  $M$ ,  $K_\alpha$ ,  $K_{2\alpha}$  and  $K_R$ . We write  $\bar{d} = \max\{d, d'\}$ .

PROCEDURE A. 1. Find  $\delta$  and  $C_i(\delta) > 0$  for  $i = 1, 2, 3$  that satisfies the following relations:

$$\begin{aligned} C_1(\delta) &\geq C_3(\delta)\delta^{2\alpha} + M\delta^{1-\alpha} + \bar{d}MK_\alpha + \bar{d}^3M^2K_{2\alpha}\delta^\alpha, \\ C_2(\delta) &\geq C_3(\delta)\delta^\alpha + \bar{d}^3M^2K_{2\alpha}, \\ C_3(\delta) &\geq \frac{2}{1-2^{1-3\alpha}}\{MC_1(\delta) + \bar{d}MC_1(\delta)^2K_\alpha + \bar{d}^2MC_2(\delta)K_\alpha \\ &\quad + 2\bar{d}^3M^2C_1(\delta)K_{2\alpha}\}. \end{aligned}$$



[Refer to the proof of Lemma 6.1 for one particular method to find such  $C_i(\delta)$ 's.]

2. Set  $C_1 = \frac{2}{\delta} C_1(\delta)$ ,  $C_2 = \frac{2}{\delta}(C_2(\delta) + MC_1 + \bar{d}MC_1K_\alpha)$  and

$$C_3 = \frac{2}{1 - 2^{1-3\alpha}}(MC_1 + \bar{d}MC_1^2K_\alpha + \bar{d}^2MC_2K_\alpha + 2\bar{d}^3M^2C_1K_{2\alpha}).$$

3. Find  $\delta'$  and  $B_i(\delta')$  for  $i = 1, 2, 3$  that satisfies the following relations:

$$B_1(\delta') > B_3(\delta')\delta'^{2\alpha} + 2M\delta'^{1-\alpha} + 2MK_\alpha + 4M^2K_{2\alpha}\delta'^\alpha,$$

$$B_2(\delta') > B_3(\delta')\delta'^\alpha + 4M^2K_{2\alpha},$$

$$B_3(\delta') > \frac{4}{1 - 2^{1-3\alpha}}\{MB_1(\delta') + MB_1(\delta'^2)K_\alpha + MB_2(\delta')K_\alpha + 2M^2B_1(\delta')K_{2\alpha}\}.$$

4. Set  $B = \frac{2}{\delta'} B_1(\delta')$ .

5. Set  $G_1 = (1 + B)C_3$ .

6. Find  $\delta''$  and  $C_4(\delta'')$  such that

$$B\delta''^\alpha \leq 2^{\alpha+\beta} - 2,$$

$$C_4(\delta'') \geq 2\left(1 - \frac{2 + B\delta''^\alpha}{2^{\alpha+\beta}}\right)^{-1} (B\bar{d}^3M^2K_R + 2\bar{d}^3M^2C_1K_R).$$

7. Set  $C_4 = (1 + B\delta''^\alpha)C_4(\delta''^3M^2K_R + 2\bar{d}^3M^2C_1K_R)/\delta''$ .

8. Set  $G_2 = C_4 + \bar{d}^3M^2K_R$ .

9. Set  $G = G_1 + G_2$ .

LEMMA 2.1. *Given  $K_\alpha, K_{2\alpha}, K_R$  and  $M$ , Procedure A can be executed.*

PROOF. We prove the lemma by providing one particular method to find such  $\delta$  and  $C_i(\delta)$ 's,  $i = 1, 2, 3$ . The method to find  $\delta'$ ,  $B_i(\delta')$ 's, for  $i = 1, 2, 3$ , follows exactly the same rationale.

Set  $C_1(\delta) = \bar{d}M\|Z\|_\alpha + 1/2$ ,  $C_2(\delta) = \bar{d}^3M^2\|A\|_{2\alpha} + 1/2$  and  $C_3(\delta) = \frac{2}{1-2^{1-3\alpha}}(MC_1(\delta) + \bar{d}MC_1(\delta)^2\|Z\|_\alpha + \bar{d}^2MC_2(\delta)\|Z\|_\alpha + \bar{d}^2M^2\|Z\|_\alpha + 2\bar{d}^3M^2C_1(\delta)\|A\|_\alpha)$ . Then we can pick  $\delta$  small enough, such that  $C_3(\delta)\delta^{2\alpha} + M\delta^{1-\alpha} + \bar{d}^3M^2\|A\|_{2\alpha}\delta^\alpha < 1/2$  and  $C_3(\delta)\delta^\alpha < 1/2$ .  $\square$

**3. The main idea of the algorithmic development.** Based on Theorem 2.2, our main task is to calculate/simulate the upper bound for  $\|Z\|_\alpha$ ,  $\|A\|_{2\alpha}$  and  $\Gamma_R$ , respectively. In this section, we will introduce the main idea of our algorithmic development.

The development can be decomposed into two tasks. The first one is to find an infinite sum representation of the objects of interest. The second one is to truncate the infinite sum up to a finite but random level so that the error induced by the

remaining terms in the summation is suitably controlled. The second task calls for novel algorithmic constructions. Simulating infinitely many terms is impossible. We need to find an efficient way to extract enough information on the remaining terms after the truncation, so that we can obtain an almost sure bound on the contribution of the terms that are not simulated. We next carry out the two tasks one by one.

3.1. *Infinite sum representation of Brownian motion and Lévy area.* We start by introducing a wavelet synthesis of Brownian motion,  $\{Z(t) : 0 \leq t \leq 1\}$ , called the Lévy–Ciesielski construction of Brownian motion [13].

First, we need to define a step function  $H(\cdot)$  on  $[0, 1]$  by

$$H(t) = I(0 \leq t < 1/2) - I(1/2 \leq t \leq 1).$$

We then define a family of functions

$$H_k^n(t) = 2^{n/2} H(2^{n-1}t - k + 1)$$

for all  $n \geq 0$  and  $1 \leq k \leq 2^{n-1}$ . Set  $H_0^0(t) = 1$ . Then one obtains the following infinite sum representation of Brownian motion.

**THEOREM 3.1** (Lévy–Ciesielski construction). *If  $\{W_k^n : 1 \leq k \leq 2^{n-1}, n \geq 0\}$  is a sequence of independent standard normal random variables, then the series defined by*

$$(3.1) \quad Z(t) = W_0^0 \int_0^t H_0^0(s) ds + \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} \left( W_k^n \int_0^t H_k^n(s) ds \right)$$

*converges uniformly on  $[0, 1]$  with probability one. Moreover, the process  $\{Z(t) : t \in [0, 1]\}$  is a standard Brownian motion on  $[0, 1]$ .*

Figure 1 demonstrates the basic idea of the Lévy–Ciesielski construction using properties of the Brownian bridge. Specifically, as  $Z(1) \sim N(0, 1)$ , we set  $Z(1) = W_0^0$ . Conditional on the value of  $Z(0) = 0$  and  $Z(1)$ ,  $Z(1/2) \sim N(Z(1)/2, 1/4)$ .

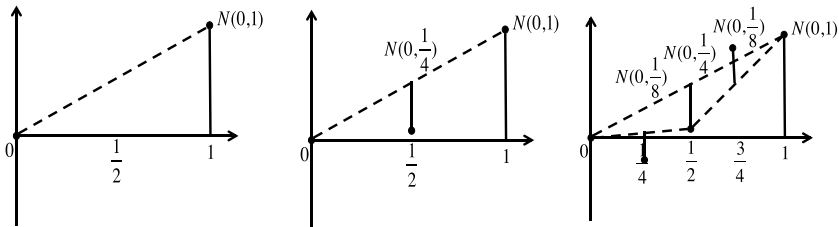


FIG. 1. Lévy–Ciesielski construction of Brownian motion on  $[0, 1]$ .

Thus, we set  $Z(1/2) = Z(1)/2 + 1/2W_1^1$ . In general, conditional on the value of  $Z(t_k^{n-1})$  and  $Z(t_{k+1}^{n-1})$ , for  $k = 0, 1, \dots, 2^{n-1}$ ,

$$Z(t_{2k+1}^n) \sim N((Z(t_k^{n-1}) + Z(t_{k+1}^{n-1}))/2, \Delta_{n+1}).$$

Thus, we set

$$Z(t_{2k+1}^n) = (Z(t_k^{n-1}) + Z(t_{k+1}^{n-1}))/2 + \Delta_{n+1}^{1/2} W_{k+1}^n.$$

Eventually, we will simulate the series up to a finite but random level  $N_1$  to be discussed later. By level, we mean the order of dyadic discretization. As we are simulating the discretization levels sequentially, we often refer to “time” when discussing levels.

We next analyze the Lévy area,  $A_{i,j}(t_k^n, t_{k+1}^n)$ , for  $1 \leq i, j \leq d', n \geq 1, 0 \leq k \leq 2^n - 1$ . Using the algebraic property,

$$\begin{aligned} A_{i,j}(t_k^n, t_{k+1}^n) &= A_{i,j}(t_{2k}^{n+1}, t_{2k+1}^{n+1}) + A_{i,j}(t_{2k+1}^{n+1}, t_{2k+2}^{n+1}) \\ &\quad + (Z_i(t_{2k+1}^{n+1}) - Z_i(t_{2k}^{n+1}))(Z_j(t_{2k+2}^{n+1}) - Z_j(t_{2k+1}^{n+1})), \end{aligned}$$

we have the following infinite sum representation of  $A_{i,j}(t_k^n, t_{k+1}^n)$ .

LEMMA 3.1. For  $n \geq 1, 0 \leq k \leq 2^n - 1$ ,

$$\begin{aligned} A_{i,j}(t_k^n, t_{k+1}^n) &= \sum_{h=n+1}^{\infty} \sum_{l=1}^{2^{h-n-1}} \{(Z_i(t_{2^{h-n}k+2l-1}^h) - Z_i(t_{2^{h-n}k+2l-2}^h)) \\ &\quad \times (Z_j(t_{2^{h-n}k+2l}^h) - Z_j(t_{2^{h-n}k+2l-1}^h))\}. \end{aligned}$$

The inner summation terms in the expression for  $A_{i,j}(t_k^n, t_{k+1}^n)$  motivate the definition of the following family of processes ( $L_{i,j}^n(k) : k = 0, 1, \dots, 2^{n-1}, n \geq 1$ ):

$$L_{i,j}^n(0) := 0,$$

$$L_{i,j}^n(k) := L_{i,j}^n(k-1) + (Z_i(t_{2k-1}^n) - Z_i(t_{2k-2}^n))(Z_j(t_{2k}^n) - Z_j(t_{2k-1}^n))$$

for  $k = 1, 2, \dots, 2^{n-1}$ .

Using this definition and Lemma 3.1, we can succinctly write  $A_{i,j}(t_k^n, t_{k+1}^n)$  as

$$(3.2) \quad A_{i,j}(t_k^n, t_{k+1}^n) = \sum_{h=n+1}^{\infty} (L_{i,j}^h(2^{h-n}(k+1)) - L_{i,j}^h(2^{h-n}k)).$$

3.2. *The idea of record breakers.* To truncate the infinite sum up to a finite but random level, we use a strategy called record breakers. Specifically, we first define a sequence of “record breakers.” We then formulate the “future” information we need to know as a sequence of “yes or no” questions. Specifically, the yes or no question is formulated as “will there be a new record breaker?” and answering the yes/no question is equivalent to simulating a properly defined Bernoulli random variable.

The definition of the record breakers need to satisfy the following two conditions:

- C1. The following event happens with probability one: beyond some random but finite time, there will be no more record breakers.
- C2. By knowing that there are no more record breakers, the contribution of the terms that we have not simulated yet are well under control (i.e., bounded by a user defined tolerance error).

We next explain how the above strategy is applied to the Brownian motion and the Lévy area, respectively.

We have  $d'$  independent Brownian motions and we will use  $W_{i,k}^n$  for  $i \in \{1, \dots, d'\}$  to denote the  $(n, k)$  coefficient in the expansion (3.1) for the  $i$ th Brownian motion.

For  $\|Z\|_\alpha$ , we say a record is broken at  $(i, n, k)$ , for  $1 \leq i \leq d'$ ,  $n \geq 0$  and  $1 \leq k \leq 2^{n-1}$ , if

$$|W_{i,k}^n| > 4\sqrt{n+1}.$$

Let  $\bar{N}_1 := \max\{n \geq 1 : |W_{i,k}^n| > 4\sqrt{n+1} \text{ for some } 1 \leq k \leq 2^{n-1}, 1 \leq i \leq d'\}$ . It is the last time the record breaker happens. The following lemma shows that condition C1 is satisfied ( $E[N_1] < \infty$  implies  $P(N_1 < \infty) = 1$ ).

LEMMA 3.2. *There exists an integer valued random variable  $N_1$ , with  $E[N_1] < \infty$ , such that for all  $n > N_1$ ,  $1 \leq k \leq 2^n - 1$  and  $1 \leq i \leq d'$   $W_{i,k}^n \leq 4\sqrt{n+1}$ .*

We next check condition C2. Define  $V^n = \max_{1 \leq k \leq 2^{n-1}} |W_k^n|$ . We have the following auxiliary lemma.

LEMMA 3.3.

$$\|Z\|_\alpha \leq 2^{2\alpha+1} \sum_{n=0}^{\infty} 2^{-n(\frac{1}{2}-\alpha)} V^n.$$

Once we found  $N_1$ , we have

$$\begin{aligned} \|Z\|_\alpha &\leq 2^{2\alpha+1} \sum_{n=0}^{N_1} 2^{-n(1/2-\alpha)} V^n + 2^{2\alpha+3} \sum_{n=N_1+1}^{\infty} 2^{-n(1/2-\alpha)} \sqrt{n+1} \\ &\leq 2^{2\alpha+1} \sum_{n=0}^{N_1} 2^{-n(1/2-\alpha)} V^n + 2^{2\alpha+3} C \frac{2^{-1/2(N_1+1)(1/2-\alpha)}}{1 - 2^{-1/2(1/2-\alpha)}}, \end{aligned}$$

where  $C = \max_{n \geq N_1+1} \{2^{-n/2(1/2-\alpha)} \sqrt{n+1}\}$ .

For the Lévy area, we first notice that when  $i = j$ ,

$$\begin{aligned} &\sup_n \sup_{0 \leq s < t \leq 1, s, t \in D_n} \frac{A_{i,i}(s, t)}{(t - s)^{2\alpha}} \\ &= \sup_n \sup_{0 \leq s < t \leq 1, s, t \in D_n} \frac{(B(t) - B(s))^2 - (t - s)^2}{2(t - s)^{2\alpha}} \\ &\leq \frac{\|Z\|_\alpha^2 + 1}{2}, \end{aligned}$$

and

$$R_{i,i}^n(t_l^n, t_m^n) = 0.$$

When  $i \neq j$ , the record breaker is defined for the random walk  $L_{i,j}^n$ 's. Specifically, for  $L$ , we say a record is broken at  $(i, j, n, k, k')$ , for  $1 \leq i, j \leq d', i \neq j, n \geq 1, 0 \leq k < k' < 2^{n-1}$ , if

$$|L_{i,j}^n(k') - L_{i,j}^n(k)| > (k' - k)^\beta \Delta_n^{2\alpha},$$

where  $\beta \in (1 - \alpha, 2\alpha)$ . Let  $\tilde{N}_2 := \max\{n \geq 1 : |L_{i,j}^n(k') - L_{i,j}^n(k)| > (k' - k)^\beta \Delta_n^{2\alpha} \text{ for some } 0 \leq k < k' \leq 2^{n-1}, 1 \leq i, j \leq d', i \neq j\}$ . It is the last time the record breaker happens. The following lemma shows that condition C1 is satisfied.

LEMMA 3.4. *There exists an integer valued random variable  $N_2$ , with  $E[N_2] = o((1 - 2\alpha)^{-2})$ , such that for all  $n > N_2$  and all  $0 \leq l < m \leq 2^{n-1}$  we have  $|L_{i,j}^n(m) - L_{i,j}^n(l)| \leq (m - l)^\beta \Delta_n^{2\alpha}$  for  $\alpha \in (1/3, 1/2)$  and  $\beta \in (1 - \alpha, 2\alpha)$ .*

We next check condition C2. The following corollary follows directly from (3.2) and the definition of  $R_{i,j}^n$ .

COROLLARY 3.1. *For  $i \neq j$ ,*

$$R_{i,j}^n(t_l^n, t_m^n) = \sum_{h=n+1}^{\infty} (L_{i,j}^h(2^{h-n}m) - L_{i,j}^h(2^{h-n}l)).$$

Then we have the following bounds for  $\|A\|_{2\alpha}$  and  $\Gamma_R$  based on the  $N_2$ .

LEMMA 3.5. *Suppose that  $N_2$  is chosen according to Lemma 3.4. We define*

$$\Gamma_L := \max \left\{ 1, \max_{1 \leq i, j \leq d', i \neq j} \max_{n < N_2} \max_{0 \leq l < m \leq 2^n - 1} \left\{ \frac{|L_{i,j}^n(m) - L_{i,j}^n(l)|}{(m - l)^\beta \Delta_n^{2\alpha}} \right\} \right\}.$$

Then

$$\Gamma_R \leq \frac{2^{-(2\alpha - \beta)}}{1 - 2^{-(2\alpha - \beta)}} \Gamma_L$$

and

$$\|A\|_{2\alpha} \leq \Gamma_R \frac{2}{1 - 2^{-2\alpha}} + \|Z\|_\alpha^2 \frac{2^{1-\alpha}}{1 - 2^{-\alpha}}.$$

In what follows, we shall explain how to simulate the random numbers ( $N_1$  and  $N_2$ ) jointly with the wavelet construction using the “record breaker” strategy introduced in the previous section. Specifically, we first find all the record breakers in sequence and then simulate the rest of the process conditional on the information obtained by knowing the location of all the (finitely many) record breakers. The challenge lies in the fact that the probability of success of the Bernoulli trials, which corresponds to the yes/no questions defined in terms of the record breakers, is not known to us. We start with the procedure to simulate  $N_1$  in Section 4, which is built on a sandwiching idea. Then conditional on the value of  $N_1$ , we introduce the procedure to simulation  $N_2$  in Section 5 based on an acceptance-rejection scheme, where the proposal distribution is built on some exponential tilting.

**4. Tolerance-enforced simulation of bounds on  $\alpha$ -Hölder norms.** We first note that  $N_1$  is not a stopping time with respect to the filtration generated by  $\{(W_{i,k}^n : 0 \leq k \leq 2^n - 1, 1 \leq i \leq d') : n \geq 1\}$ .

For the simplicity of demonstration, we shall focus on the 1-dimensional case. For  $d' > 1$ , we apply the same procedure for each Brownian motion. In what follows in this subsection, we shall drop the subscription  $i$ .

We call a pair  $(n, k)$  a record-broken-pair if  $|W_k^n| > 4\sqrt{n + 1}$ . All pairs (both record-broken-pairs and nonrecord-broken-pairs) can be totally ordered lexicographically, that is, using  $2^{n-1} + k$ . The distribution of subsequent pairs at which records are broken is not difficult to compute (because of the independence of  $W_k^n$ 's). So, using a sequential acceptance/rejection procedure we can simulate all of the record-broken-pairs. Conditional on these pairs, the distribution of the  $\{(W_k^n : 0 \leq k \leq 2^n - 1) : n \geq 1\}$  is straightforward to describe. Precisely, if  $(k, n)$  is a record-broken-pair, then  $W_k^n$  is conditioned on  $|W_k^n| > 4\sqrt{n + 1}$ , and thus is straightforward to simulate. Similarly, if  $(k, n)$  is not a record-broken-pair, then  $W_k^n$  is conditioned on  $|W_k^n| \leq 4\sqrt{n + 1}$ , and also can be easily simulated.

The simulation of the record-broken-pairs has been studied in [5]. The idea is to find all the record breakers sequentially until there are no more record breakers. The challenge lies in sampling the Bernoulli random variable corresponding to the question “whether there will be no more record breakers in the future.” We take sampling the first breaker as an example. The probability that there are no more record breakers beyond 1 is

$$p(1) := \prod_{n=1}^{\infty} \prod_{k=0}^{2^n-1} P(|W_k^n| \leq 4\sqrt{n+1}),$$

which involves evaluating the product of infinite many terms and we do not know its value in closed form. However, we can find a sequence of upper bound and lower bounds of  $p(1)$ , which are defined as

$$U_h(1) = \prod_{r=1}^h P(|W_{k(r)}^{n(r)}| \leq 4\sqrt{[\log_2 r] + 1}),$$

where  $r = 2^{n(r)-1} + k(r)$  and

$$D_h(1) = (1 - h^{1-4^2/2})U_h$$

respectively. The upper and lower bounds satisfy that  $D_h(1) < D_{h+1}(1) < p(1) < U_h(1) < U_{h+1}(1)$  and  $\lim_{h \rightarrow \infty} D_h(1) = p(1) = \lim_{h \rightarrow \infty} U_h(1)$ . We also have that  $U_h(1) - U_{h+1}(1)$  is equal to the probability that the first record breaker happens at position  $h$ . Thus, we can check whether the Bernoulli trial is a success or failure by updating the upper and lower bounds sequentially. Moreover, if the Bernoulli trial is a failure (there are more record breakers beyond the current index), we also know the index of the next record breaker. We synthesize algorithm 2W in [5] for our purposes next.

REMARK. Observe that for every  $l = 2^{n-1} + k \in S$ , we can generate  $W_k^n$  conditional on the event  $\{|W_k^n| > 4\sqrt{n+1}\}$ ; for other  $l$  (i.e.,  $l \notin S$ ), generate  $W_k^n$  given  $\{|W_k^n| \leq 4\sqrt{n+1}\}$ . Note that at the end of Algorithm 1 and after simulating  $W_k^n$  for  $n \leq N_1$  one can compute

$$K_\alpha = 2^{2\alpha+1} \sum_{n=0}^{N_1} 2^{-n(1/2-\alpha)} V^n + 2^{2\alpha+3} C \frac{2^{-1/2(N_1+1)(1/2-\alpha)}}{1 - 2^{-1/2(1/2-\alpha)}},$$

where  $C = \max_{n \geq N_1+1} \{2^{-n/2(1/2-\alpha)} \sqrt{n+1}\}$ .

**5. Tolerance-enforced simulation for bounds on  $2\alpha$ -Hölder norms of Lévy areas.** The simulation of  $N_2$ , is a lot more complicated, comparing to  $N_1$ , because there is fair amount of dependence on the structure of the  $L_{i,j}^n(k)$ 's as one varies  $n$ . Let us provide a general idea of our simulation procedure in order to set the stage for the definitions and estimates that must be studied first.

**Algorithm 1** Simulate  $N_1$  jointly with the record-broken-pairs

**Output:** A vector  $S$  which gives all the indices  $l = 2^n + k$  such that  $(n, k)$  is a broken-record-pair.

**Step 0:** Initialize  $R = 0$  and  $S$  to be an empty array.

**Step 1:** Set  $U = 1, D = 0$ . Simulate  $V \sim \text{Uniform}(0, 1)$ .

**Step 2:** While  $U > V > D$ , set  $R \leftarrow R + 1$  and  $U \leftarrow P(|W_k^n| \leq 4\sqrt{\lceil \log_2 R \rceil + 1}) \times U$  and  $D \leftarrow (1 - R^{1-4^2/2}) \times U$ .

**Step 3:** If  $V \geq U$ , add  $R$  to the end of  $S$ , that is,  $S = [S, R]$ , and return to Step 1.

**Step 4:** If  $V \leq D$ ,  $N_1 = \lceil \log_2 \max(S) \rceil$ .

**Step 5:** Output  $S$ .

**End of Algorithm 1**

Define

$$\mathcal{F}_n = \sigma\{(W_{i,k}^m : 1 \leq k \leq 2^{m-1}) : m \leq n\},$$

and for the conditional expectation given  $\mathcal{F}_n$  we write

$$E_n[\cdot] := E[\cdot | \mathcal{F}_n].$$

Suppose we have simulated  $\{(W_{i,k}^n : 0 \leq k \leq 2^n - 1, 1 \leq i \leq d') : n \leq N\}$  for some  $N$  and define

$$\tau_1(N) = \inf\{n \geq N + 1 : |L_{i,j}^n(m) - L_{i,j}^n(l)| > (m - l)^\beta \Delta_n^{2\alpha} \text{ for some } 0 \leq l < m \leq 2^{n-1}\}.$$

Because of Lemma 3.4 we have that the event  $\{\tau_1(N) = \infty\}$  has positive probability. In what follows, we will explain how to simulate a Bernoulli random variable with probability of success  $P(\tau_1(N) = \infty | \mathcal{F}_N)$ . If such Bernoulli is a success, then we have that  $N_2 = N$  and we would have basically concluded the difficult part of the simulation procedure (the rest of the process can be simulated under a series of conditioning events whose probability increases to one as  $n$  grows). If the Bernoulli is a failure (i.e., its value is zero), then we will find  $\tau_1(N)$  and simulate all the information up to  $\tau_1(N)$ . We repeat the above Bernoulli trial with updated probability of success until we obtain a successful Bernoulli trial.

Now, part of the problem is that Algorithm 1 has been already executed, so  $N \geq N_1$ , in other words, while the random variables  $\{W_{i,k}^n : 1 \leq k \leq 2^{n-1}\}$  are independent (for fixed  $n > N$ ), they are no longer identically distributed. Instead,  $W_{i,k}^n$  is standard Gaussian conditional on the event  $\{|W_{i,k}^n| \leq 4\sqrt{n+1}\}$ . Nevertheless, if  $n$  is large enough, all of the events  $\{|W_{i,k}^n| \leq 4\sqrt{n+1}\}$  will occur with high probability. So, we shall first proceed to explain how to simulate a Bernoulli random variable with probability of success  $P(\tau_1(n') = \infty | \mathcal{F}_{n'})$  assuming  $n'$  is a deterministic number. The procedure actually will produce both the outcome of the



Bernoulli trial and if such outcome is a failure [i.e.,  $\tau_1(n') < \infty$ ], also the sample path

$$\{W_{i,k}^m : 1 \leq k \leq 2^{m-1}, n' < m \leq \tau_1(n')\}.$$

Our procedure is based on acceptance / rejection using a carefully chosen proposal distribution for the  $W_{i,k}^n$ 's,  $n \geq n'$  based on exponential tilting of  $L_{i,j}^n(k)$ 's, conditional on  $\mathcal{F}_{n'}$ . To this end, we will need to compute the conditional moment generating function (conditional on  $\mathcal{F}_{n'}$ ) of  $L_{i,j}^n(k)$ 's and the family of distributions induced over  $W_{i,k}^n$ 's and  $W_{j,k}^n$ 's under the exponentially tilting. This will be done in Section 5.1. Then we need some large deviation estimates to bound the likelihood ratio of a certain randomization procedure. These bounds are developed in Section 5.2. These are the main elements needed to simulate  $N_2$  together with the wavelet construction. We introduce the actual randomization procedure and the details of the algorithm in Section 5.3.

5.1. *Conditional moment generating functions and associated exponential tilting.* In this section, we characterize the distribution of  $\{(W_{i,k}^{n+m} : 1 \leq k \leq 2^{n+m-1}) : m \geq 1\}$  under the exponential tilting conditional on  $\mathcal{F}_n$ .

In order to reduce the length of some of the equations that follow, we write, for each  $r \in \{1, 2, \dots, 2^n\}$ ,

$$(5.1) \quad \Lambda_i^n(t_r^n) := Z_i(t_r^n) - Z_i(t_{r-1}^n).$$

Then we have the following recursive relations for  $\Lambda_i^n(t_r^n)$ 's.

LEMMA 5.1. For  $k = 1, 2, \dots, 2^{n-1}$

$$\begin{aligned} \Lambda_i^n(t_{2k-1}^n) &= \frac{1}{2} \Lambda_i^{n-1}(t_k^{n-1}) + \Delta_{n+1}^{1/2} W_{i,k}^n, \\ \Lambda_i^n(t_{2k}^n) &= \frac{1}{2} \Lambda_i^{n-1}(t_k^{n-1}) - \Delta_{n+1}^{1/2} W_{i,k}^n. \end{aligned}$$

From Lemma 5.1, we can see that

$$\mathcal{F}_n = \sigma\{Z(t_{k'}^m) - Z(t_k^m) : 0 \leq k < k' \leq 2^{m-1}, m \leq n\}.$$

Assume that  $k < k'$ , we will iteratively compute the conditional moment generating function as

$$(5.2) \quad \begin{aligned} E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})] \\ = E_n[E_{n+1}[\dots E_{n+m-1}[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})] \dots]]. \end{aligned}$$

Recall that, for  $1 \leq k \leq 2^{n-1}$ ,

$$L_{i,j}^n(k) = \sum_{r=1}^k \Lambda_i^n(t_{2r-1}^n) \Lambda_j^n(t_{2r}^n).$$

We shall start from the expectation of  $\exp(\theta_0 \Lambda_i^{n+m}(t_{2r-1}^{n+m}) \Lambda_j^{n+m}(t_{2r}^{n+m}))$  conditional on  $\mathcal{F}_{n+m-1}$ .

**COROLLARY 5.1.** *For  $i \neq j$ ,*

$$\begin{aligned} E_{n+m-1}[\exp(\theta_0 \Lambda_i^{n+m}(t_{2r-1}^{n+m}) \Lambda_j^{n+m}(t_{2r}^{n+m}))] \\ = (1 - \theta_0^2 \Delta_{n+m}^2)^{-1/2} \exp(\theta_1 \Lambda_j^{n+m-1}(t_r^{n+m-1}) \Lambda_i^{n+m-1}(t_r^{n+m-1})) \\ \times \exp(\eta_1 \Lambda_j^{n+m-1}(t_r^{n+m-1})^2 + \eta_1 \Lambda_i^{n+m-1}(t_r^{n+m-1})^2), \end{aligned}$$

where

$$\theta_1 := \theta_0(1 - \theta_0^2 \Delta_{n+m+1}^2)^{-1}/4, \quad \eta_1 := \theta_0^2(1 - \theta_0^2 \Delta_{n+m+1}^2)^{-1} \Delta_{n+m}/8.$$

Moreover, define

$$\begin{aligned} P'_{n+m, t_r^{n+m}}(W_{i,r}^{n+m} \in A, W_{j,r}^{n+m} \in B) \\ = \frac{E_{n+m-1}[I(W_{i,r}^{n+m} \in A, W_{j,r}^{n+m} \in B) \exp(\theta_0 \Lambda_i^{n+m}(t_{2r-1}^{n+m}) \Lambda_j^{n+m}(t_{2r}^{n+m}))]}{E_{n+m-1}[\exp(\theta_0 \Lambda_i^{n+m}(t_{2r-1}^{n+m}) \Lambda_j^{n+m}(t_{2r}^{n+m}))]}, \end{aligned}$$

then under  $P'_{n+m, t_r^{n+m}}$ , and given  $\mathcal{F}_{n+m-1}$ , we have that  $(W_{i,r}^{n+m}, W_{j,r}^{n+m})$  follows a Gaussian distribution with covariance matrix

$$\Sigma_{n+m}^{i,j}(t_r^{n+m+1}) = \frac{1}{1 - \theta_0^2 \Delta_{n+m+1}^2} \begin{pmatrix} 1 & -\theta_0 \Delta_{n+m+1} \\ -\theta_0 \Delta_{n+m+1} & 1 \end{pmatrix},$$

and mean vector

$$\mu_{n+m}^{i,j}(t_r^{n+m}) = \Sigma_{n+m}^{i,j}(t_r^{n+m}) \begin{pmatrix} \theta_0 \Delta_{n+m+1}^{1/2} \Lambda_j^{n+m-1}(t_r^{n+m-1})/2 \\ -\theta_0 \Delta_{n+m+1}^{1/2} \Lambda_i^{n+m-1}(t_r^{n+m-1})/2 \end{pmatrix}.$$

So, from Corollary 5.1 we conclude that

$$\begin{aligned} E_{n+m-1} \left[ \exp \left( \theta_0 \sum_{r=k+1}^{k'} \Lambda_i^{n+m}(t_{2r-1}^{n+m}) \Lambda_j^{n+m}(t_{2r}^{n+m}) \right) \right] \\ (5.3) \quad = (1 - \theta_0^2 \Delta_{n+m+1}^2)^{-(k'-k)/2} \\ \times \exp \left( \theta_1 \sum_{r=k+1}^{k'} \Lambda_j^{n+m-1}(t_r^{n+m-1}) \Lambda_i^{n+m-1}(t_r^{n+m-1}) \right) \\ \times \exp \left( \eta_1 \sum_{r=k+1}^{k'} \Lambda_j^{n+m-1}(t_r^{n+m-1})^2 + \eta_1 \sum_{r=k+1}^{k'} \Lambda_i^{n+m-1}(t_r^{n+m-1})^2 \right). \end{aligned}$$

If  $m \geq 2$ , we can continue taking the corresponding conditional expectation given  $\mathcal{F}_{n+m-2}$ . Due to the recursive nature of (5.2) and the linear and quadratic terms that arise in (5.3), it is convenient to consider

$$(5.4) \quad \sum_{r=1}^{2^{n+m-1}} \theta_1(t_r^{n+m-1}) \Lambda_j^{n+m-1}(t_r^{n+m-1}) \Lambda_i^{n+m-1}(t_r^{n+m-1}) + \sum_{r=1}^{2^{n+m-1}} \eta_1(t_r^{n+m-1}) (\Lambda_j^{n+m-1}(t_r^{n+m-1})^2 + \Lambda_i^{n+m-1}(t_r^{n+m-1})^2),$$

where

$$\begin{aligned} \theta_1(t_r^{n+m-1}) &= \theta_1 \times I(r \in \{k+1, \dots, k'\}), \\ \eta_1(t_r^{n+m-1}) &= \eta_1 \times I(r \in \{k+1, \dots, k'\}). \end{aligned}$$

We also introduce the following notation to simply the presentation of our tilting parameters. Due to the difference in the recursive relation for  $\Lambda_i^n(t_r^n)$  between odd and even  $r$ 's, we recursively define for  $l = 2, \dots, m$

$$(5.5) \quad \begin{aligned} \theta_+^l(t_r^{n+m-l}) &= \theta_{l-1}(t_{2r-1}^{n+m-l+1}) + \theta_{l-1}(t_{2r}^{n+m-l+1}), \\ \theta_-^l(t_r^{n+m-l}) &= \theta_{l-1}(t_{2r-1}^{n+m-l+1}) - \theta_{l-1}(t_{2r}^{n+m-l+1}), \\ \eta_+^l(t_r^{n+m-l}) &= \eta_{l-1}(t_{2r-1}^{n+m-l+1}) + \eta_{l-1}(t_{2r}^{n+m-l+1}), \\ \eta_-^l(t_r^{n+m-l}) &= \eta_{l-1}(t_{2r-1}^{n+m-l+1}) - \eta_{l-1}(t_{2r}^{n+m-l+1}), \\ \rho_l(t_r^{n+m-l}) &= \frac{\Delta_{n+m-l+2} \theta_+^l(t_r^{n+m-l})}{1 - 2\Delta_{n+m-l+2} \eta_+^l(t_r^{m+n-l})}, \\ h_l(t_r^{n+m-l}) &= \frac{\Delta_{n+m-l+2}}{(1 - 2\Delta_{n+m-l+2} \eta_+^l(t_r^{m+n-l}))(1 - \rho_l(t_r^{n+m-l})^2)}, \end{aligned}$$

and set

$$\begin{aligned} \eta_l(t_r^{m+n-l}) &= \frac{\eta_+^l(t_r^{m+n-l})}{4} \\ &\quad + \frac{h_l(t_r^{m+n-l})}{8} \{ \theta_-^l(t_r^{m+n-l})^2 + 4\eta_-^l(t_r^{m+n-l})^2 \\ &\quad + 4\theta_-^l(t_r^{m+n-l})\eta_-^l(t_r^{m+n-l})\rho_l(t_r^{m+n-l}) \}, \\ \theta_l(t_r^{m+n-l}) &= \frac{\theta_+^l(t_r^{m+n-l})}{4} \\ &\quad + h_l(t_r^{m+n-l}) \left\{ \theta_-^l(t_r^{m+n-l})\eta_-^l(t_r^{m+n-l}) \right. \\ &\quad \left. + \frac{1}{4}\theta_-^l(t_r^{m+n-l})^2 g_l(t_r^{m+n-l}) + \eta_-^l(t_r^{m+n-l})^2 \rho_l(t_r^{m+n-l}) \right\}. \end{aligned}$$

Finally, we decompose (5.4) into two parts (the cross term and the quadratic term) by defining

$$\begin{aligned}
 A(t_r^{n+m-l}) &= \theta_{l-1}(t_{2r-1}^{n+m-l+1})\Lambda_j^{n+m-l+1}(t_{2r-1}^{n+m-l+1})\Lambda_i^{n+m-l+1}(t_{2r-1}^{n+m-l+1}) \\
 &\quad + \theta_{l-1}(t_{2r}^{n+m-l+1})\Lambda_j^{n+m-l+1}(t_{2r}^{n+m-l+1})\Lambda_i^{n+m-l+1}(t_{2r}^{n+m-l+1}), \\
 B(t_r^{n+m-l}) &= \eta_{l-1}(t_{2r-1}^{n+m-l+1})(\Lambda_j^{n+m-l+1}(t_{2r-1}^{n+m-l+1})^2 + \Lambda_j^{n+m-l+1}(t_{2r-1}^{n+m-l+1})^2) \\
 &\quad + \eta_{l-1}(t_{2r}^{n+m-l+1})(\Lambda_j^{n+m-l+1}(t_{2r}^{n+m-l+1})^2 \\
 &\quad + \Lambda_j^{n+m-l+1}(t_{2r}^{n+m-l+1})^2),
 \end{aligned}$$

and

$$C(t_r^{n+m-l}) = (1 - 2\Delta_{n+m-l+1}\eta_+^l(t_r^{m+n-l}))^{-1}(1 - \rho_l(t_r^{m+n-l})^2)^{-1/2}.$$

Then (5.4) can be written as

$$\sum_{r=1}^{2^{n+m-2}} (A(t_r^{n+m-2}) + B(t_r^{n+m-2})),$$

and the following result is key in evaluating (5.2).

**COROLLARY 5.2.** *For  $i \neq j, l = 2, 3, \dots, m$  and  $r = 1, 2, \dots, 2^{n+m-l}$*

$$\begin{aligned}
 E_{n+m-l}[\exp(A(t_r^{n+m-l}) + B(t_r^{n+m-l}))] \\
 &= C(t_r^{n+m-l}) \exp(\theta_l(t_r^{m+n-l})\Lambda_i(t_r^{m+n-l})\Lambda_j(t_r^{m+n-l})) \\
 &\quad \times \exp(\eta_l(t_r^{m+n-l})(\Lambda_i(t_r^{m+n-l})^2 + \Lambda_j(t_r^{m+n-l})^2)).
 \end{aligned}$$

Moreover, define

$$\begin{aligned}
 P'_{n+m-l+1, t_r^{n+m-l+1}}(W_{i,r}^{n+m-l+1} \in A, W_{j,r}^{n+m-l+1} \in B) \\
 = \frac{E_{n+m-l}[I(W_{i,r}^{n+m-l+1} \in A, W_{j,r}^{n+m-l+1} \in B) \exp(A(t_r^{n+m-l}) + B(t_r^{n+m-l}))]}{E_{n+m-l}[\exp(A(t_r^{n+m-l}) + B(t_r^{n+m-l}))]},
 \end{aligned}$$

then under  $P'_{n+m-l+1, t_r^{n+m-l+1}}$ , and given  $\mathcal{F}_{n+m-l}$ , we have that  $(W_{i,r}^{n+m-l+1}, W_{j,r}^{n+m-l+1})$  follows a Gaussian distribution with covariance matrix

$$\begin{aligned}
 &\Sigma_{n+m-l+1}^{i,j}(t_r^{n+m-l+1}) \\
 &= \frac{1}{1 - \rho_l(t_r^{m+n-l})^2} \\
 &\quad \times \begin{pmatrix} (1 - 2\Delta_{n+m-l+1}\eta_+^l(t_r^{m+n-l}))^{-1} & g_l(t_r^{m+n-l}) \\ g_l(t_r^{m+n-l}) & (1 - 2\Delta_{n+m-l+1}\eta_+^l(t_r^{m+n-l}))^{-1} \end{pmatrix},
 \end{aligned}$$

where  $g_l(t_r^{n+m-l}) = \Delta_{n+m-l+2}\theta_+^l(t_r^{n+m-l})(1 - 2\Delta_{n+m-l+2}\eta_+^l(t_r^{n+m-l}))^{-2}$ , and mean vector

$$\begin{aligned} \mu_{n+m}^{i,j}(t_r^{n+m-l+1}) &= \Delta_{n+m-l+1}^{1/2} \sum_{n+m-l+1}^{i,j} (t_r^{n+m-l+1}) \\ &\times \begin{pmatrix} \Lambda_i(t_r^{n+m-l})\eta_-^l(t_r^{n+m-l}) + \frac{1}{2}\Lambda_j(t_r^{n+m-l})\theta_-^l(t_r^{n+m-l}) \\ \Lambda_j(t_r^{n+m-l})\eta_-^l(t_r^{n+m-l}) + \frac{1}{2}\Lambda_i(t_r^{n+m-l})\theta_-^l(t_r^{n+m-l}) \end{pmatrix}. \end{aligned}$$

Using Corollary 5.2, we conclude that

$$\begin{aligned} E_{n+m-l} &\left[ \exp\left( \sum_{r=1}^{2^{n+m-l}} (A(t_r^{n+m-l}) + B(t_r^{n+m-l})) \right) \right] \\ &= \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}) \times \exp\left( \sum_{r=1}^{2^{n+m-l-1}} (A(t_r^{n+m-l-1}) + B(t_r^{n+m-l-1})) \right). \end{aligned}$$

Therefore, combining Corollary 5.1 and repeatedly iterating the previous expression we conclude that

$$\begin{aligned} E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})] \\ (5.6) \quad &= (1 - \theta_0^2 \Delta_{n+m}^2)^{-(k'-k)/2} \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}) \\ &\times \exp\left( \sum_{r=1}^{2^n} \theta_m(t_r^n) \Lambda_i(t_r^n) \Lambda_j(t_r^n) + \sum_{r=1}^{2^n} \eta_m(t_r^n) \{ \Lambda_i(t_r^n)^2 + \Lambda_j(t_r^n)^2 \} \right). \end{aligned}$$

5.2. *Conditional large deviations estimates for  $L_{i,j}^n(k)$ .* We wish to estimate, for  $1 \leq i, j \leq d', i \neq j, k' > k$  and  $k', k \in \{0, 1, \dots, 2^{n+m-1}\}$ ,

$$\begin{aligned} P_n(|L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k' - k)^\beta \Delta_{n+m}^{2\alpha}) \\ \leq \exp(-\theta_0(k' - k)^\beta \Delta_{n+m}^{2\alpha}) \times \{ E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})] \\ + E_n[\exp(-\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})] \}. \end{aligned}$$

We borrow some intuition from the proof of Lemma 3.4 and select

$$(5.7) \quad \theta_0(m, k', k) := \theta_0 = \frac{\gamma}{(k' - k)^{1/2} \Delta_n^{2\alpha'} \Delta_m}.$$

We will drop the dependence on  $(m, k', k)$  for brevity. In addition, we pick  $\gamma \leq 1/4$  and  $\alpha' \in (\alpha, 1/2)$  so that

$$\exp(-\theta_0(k' - k)^\beta \Delta_{n+m}^{2\alpha}) = \exp(-\gamma(k' - k)^{\beta-1/2} \Delta_n^{2(\alpha-\alpha')} \Delta_m^{2\alpha-1}).$$

Our next task is to control the  $E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})]$ , which is the purpose of the following result, proved in the [Appendix](#) to this section.

LEMMA 5.2. *For  $i \neq j$ , suppose that  $\theta_0$  is chosen according to (5.7), and  $n$  is chosen such that*

$$(5.8) \quad \max_{r \leq 2^n} \{|\Lambda_i(t_r^n)|, |\Lambda_j(t_r^n)|\} \leq \Delta_n^{\alpha'}$$

and for  $\varepsilon_0 \in (0, 1/2)$

$$(5.9) \quad \left| \sum_{r=l+1}^m \Lambda_i(t_r^n) \Lambda_j(t_r^n) \right| \leq \varepsilon_0(m-l)^\beta \Delta_n^{2\alpha'} \quad \text{for all } 0 \leq l < m \leq 2^n$$

with  $\alpha' \in (\alpha, 1/2)$ , then

$$E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})] \leq 4 \exp(\varepsilon_0 \gamma (k' - k)^{\beta-1/2}).$$

REMARK. It is very important to note that due to Lemma 3.2 we can always continue simulating the  $W_{i,k}^n$ 's (maybe conditional on  $\{|W_{i,k}^n| \leq 4\sqrt{n+1}\}$  in case  $n > N_1$ ) to make sure that (5.8) holds for some  $n$ . Similarly, condition (5.9) can be simultaneously enforced with (5.8) because of Lemma 3.4. Actually, Lemma 3.2 and Lemma 3.4 indicate that conditions (5.8) and (5.9) will occur eventually for all  $n$  larger than some random threshold. Our simulation algorithms will ultimately detect such threshold, but Lemma 5.2 does not require that we know that threshold.

As a consequence of Lemma 5.2, using Chernoff's bound, we obtain the following proposition.

PROPOSITION 5.1. *For  $i \neq j$ , if  $n$  is chosen such that (5.8) and (5.9) hold, then*

$$\begin{aligned} P_n(|L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k' - k)^\beta \Delta_{n+m}^{2\alpha}) \\ \leq 8 \exp\left(-\frac{1}{2} \gamma (k' - k)^{\beta-1/2} \Delta_n^{2(\alpha-\alpha')} \Delta_m^{2\alpha-1}\right). \end{aligned}$$

5.3. *Joint Tolerance-Enforced Simulation for  $\alpha$ -Hölder norms and proof of Theorem 2.2.* Define

$$C_n(m) = \{|L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k' - k)^\beta \Delta_{n+m}^{2\alpha}$$

$$\text{for some } 0 \leq k < k' < 2^{n+m-1}, 1 \leq i, j \leq d', i \neq j\},$$

and put  $\tau_1(n) = \inf\{m \geq 1 : C_n(m) \text{ occurs}\}$ . We write  $\bar{C}_n(m)$  for the complement of  $C_n(m)$ , so that

$$P_n(\tau_1(n) < \infty) = \sum_{m=1}^{\infty} P\left(C_n(m) \cap \bigcap_{l=1}^{m-1} \bar{C}_n(l)\right).$$

To facilitate the explanation, we next introduce a few more notation. Let

$$\omega_{n:n+m} := \{W_{i,k}^l : 0 \leq k \leq 2^n - 1, 1 \leq i \leq d', n < l \leq n + m\}.$$

In addition, define

$$v_n(k, k'|m) := 8 \exp\left(-\frac{1}{2}\gamma(k' - k)^{\beta-1/2} \Delta_n^{2(\alpha-\alpha')} \Delta_m^{2\alpha-1}\right) \times I(0 \leq k < k' \leq 2^{n+m-1})I(m \geq 1),$$

$$b_n(m) := \sum_{0 \leq k < k' \leq 2^{n+m-1}} v_n(k, k'|m),$$

$$q_n(k, k'|m) := \frac{v_n(k, k'|m)}{b_n(m)}$$

and

$$P_{n,m}^{i,j,k,k'}(\omega_{n:n+m} \in \cdot) = \frac{E_n[I(\omega_{n:n+m} \in \cdot) \exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})]}{E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})]}.$$

We also denote

$$\psi_n(m, i, j, k, k') := \log E_n[\exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\})].$$

Observe that

$$b_n(m) = \sum_{0 \leq k < k' \leq 2^{n+m-1}} 8 \exp\left(-\frac{1}{2}\gamma(k' - k)^{\beta-1/2} \Delta_n^{2(\alpha-\alpha')} \Delta_m^{2\alpha-1}\right) \leq 2^{2(m+n)+3} \exp\left(-\frac{1}{2}\gamma \Delta_n^{2(\alpha-\alpha')} \Delta_m^{2\alpha-1}\right).$$

Thus,  $b_n(m) \rightarrow 0$  as  $n \rightarrow \infty$ . Then we can select any probability mass function  $\{g(m) : m \geq 1\}$ , for example,  $g(m) = e^{-1}/(m - 1)!$  for  $m \geq 1$ , by assuming that  $n$  is sufficiently large,

$$g(m) \geq d'^2 b_n(m).$$

Now consider the following procedure which we called Procedure **Aux**, Aux for “auxiliar,” which is given for pedagogical purposes, because as we shall see shortly it is not directly applicable but useful to understand the nature of the method that we shall ultimately use.

PROCEDURE AUX.

**Input:** We assume that we have simulated  $\{(W_{i,k}^n : 0 \leq k < 2^l) : l \leq n\}$ .

**Output:** A Bernoulli  $F$  with parameter  $P_n(\tau_1(n) < \infty)$ , and if  $F = 1$ , also

$$\omega_{n:\tau_1(n)} = \{W_{i,k}^l : 1 \leq k \leq 2^l - 1, 1 \leq i \leq d', n < l \leq \tau_1(n)\}$$

conditional on the event  $\tau_1(n) < \infty$ .

**Step 1:** Sample  $M$  according to  $g(m)$ .

**Step 2:** Given  $M = m$  sample  $I$  and  $J$  ( $I \neq J$ ) uniformly over the set  $\{1, 2, \dots, d'\}$ . Then, sample  $K', K$  from  $q_n(k, k'|m)$ .

**Step 3:** Given  $M = m, I = i, J = j, K = k,$  and  $K' = k'$ , simulate  $\omega_{n:n+m}$  from  $P_{n,m}^{i,j,k,k'}(\cdot)$ . Note that simulation from  $P_{n,m}^{i,j,k,k'}(\cdot)$  can be done according to Corollary 5.2.

**Step 4:** Compute

$$\begin{aligned} \Xi_n(m, i, j, k, k', \omega_{n:n+m}) &= 1/(g(m)(d'(d' - 1))^{-1}q_n(k, k'|m) \\ &\quad \times \exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\} - \psi_n(m, i, j, k, k'))), \end{aligned}$$

and

$$\begin{aligned} \mathcal{N}_n(m) &= \sum_{1 \leq i, j \leq d', i \neq j} \sum_{1 \leq h < h' \leq 2^{n+m-1}} I(|L_{i,j}^{n+m}(h') - L_{i,j}^{n+m}(h)| \\ &\quad > (h - h')^\beta \Delta_{n+m}^{2\alpha}). \end{aligned}$$

**Step 5:** Simulate  $U$  uniformly distributed on  $[0, 1]$  independent of everything else and output

$$\begin{aligned} F &= I \left\{ U < I \left( \{|L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k - k')^\beta \Delta_{n+m}^{2\alpha}\} \cap \bigcap_{l=1}^{m-1} \bar{C}_n(l) \right) \right. \\ &\quad \left. \times \Xi_n(m, i, j, k, k', \omega_{n:n+m})/\mathcal{N}_n(m) \right\}. \end{aligned}$$

If  $F = 1$ , also output  $\omega_{n:n+m}$ .

**End of Procedure Aux**

We first notice that when  $|L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k - k')^\beta \Delta_{n+m}^{2\alpha}$ ,

$$\begin{aligned} g(m)(d'(d' - 1))^{-1}q_n(k, k'|m) \exp(\theta_0\{L_{i,j}^{n+m}(k') \\ - L_{i,j}^{n+m}(k)\} - \psi_n(m, i, j, k, k')) > 1. \end{aligned}$$

Thus,  $\Xi_n(m, i, j, k, k', \omega_{n:n+m}) < 1$ . That is to say the likelihood ration function is bounded and the Bernoulli random variable  $F$  is well defined.

We claim that the output  $F$  is distributed as a Bernoulli random variable with parameter  $P_n(\tau_1(n) < \infty)$ . Moreover, we claim that if  $F = 1$ , then,  $\omega_{n:n+M}$  is distributed according to  $P_n(\omega_{n:\tau_1(n)} \in \cdot | \tau_1(n) < \infty)$ . We first verify the claim that the outcome in Step 5 follows a Bernoulli with parameter  $P_n(\tau_1(n) < \infty)$ . In order



to see this, let  $Q_n$  denote the distribution induced by Procedure **Aux**. Note that

$$\begin{aligned} & Q_n \left( U < I \left( \{ |L_{i,j}^{n+M}(K') - L_{i,j}^{n+M}(K)| > (K - K')^\beta \Delta_{n+M}^{2\alpha} \} \cap \bigcap_{l=1}^{M-1} \bar{C}_n(l) \right) \right) \\ & \times \Xi_n(M, I, J, K, K', \omega_{n:n+M}) / \mathcal{N}_n(m) \\ & = E^{Q_n} \left[ I \left( \{ |L_{i,j}^{n+M}(K') - L_{i,j}^{n+M}(K)| > (K - K')^\beta \Delta_{n+M}^{2\alpha} \} \cap \bigcap_{l=1}^{M-1} \bar{C}_n(l) \right) \right. \\ & \quad \left. \times \Xi_n(M, I, J, K, K', \omega_{n:n+M}) / \mathcal{N}_n(m) \right] \\ & = \sum_{m=1}^{\infty} \sum_{1 \leq i, j \leq d'} \sum_{1 \leq k < k' \leq 2^{n+m-1}} E^{Q_n} \left[ I \left( \{ |L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| \right. \right. \\ & \quad \left. \left. > (k - k')^\beta \Delta_{n+m}^{2\alpha} \} \cap \bigcap_{l=1}^{m-1} \bar{C}_n(l) \right) \right. \\ & \quad \left. \times \frac{dP_n}{dP_{n,m}^{i,j,k,k'}}(\omega_{n:n+m}) \times \frac{1}{\mathcal{N}_n(m)} \right] \\ & = \sum_{m=1}^{\infty} \sum_{1 \leq i, j \leq d'} \sum_{1 \leq k < k' \leq 2^{n+m-1}} E_n \left( \left( I \left( \{ |L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| \right. \right. \right. \\ & \quad \left. \left. \left. > (k - k')^\beta \Delta_{n+m}^{2\alpha} \} \cap \bigcap_{l=1}^{m-1} \bar{C}_n(l) \right) \right) / \mathcal{N}_n(m) \right) \\ & = \sum_{m=1}^{\infty} P_n \left( \mathcal{C}_n(m) \cap \bigcap_{l=1}^{m-1} \bar{C}_n(l) \right) \\ & = P_n(\tau_1(n) < \infty). \end{aligned}$$

Similarly, for the second claim,

$$\begin{aligned} & Q_n \left( \omega_{n:n+M} \in A \mid U < I \left( \mathcal{C}_n(M) \cap \bigcap_{l=1}^{M-1} \bar{C}_n(l) \right) \right) \\ & \quad \times \Xi_n(M, I, J, K, K', \omega_{n:n+M}) \\ & = \sum_{m=1}^{\infty} E^{Q_n} \left( \omega_{n:n+m} \in A, \frac{dP_n^{I,J,K,K'}}{dP_n}(\omega_{n:n+m}) \right) \end{aligned}$$

$$\begin{aligned} & \times I\left(C_n(m) \cap \bigcap_{l=1}^{M-1} \bar{C}_n(l)\right) / P_n(\tau_1(n) < \infty) \\ &= \sum_{m=1}^{\infty} P_n(\omega_{n:n+m} \in A, \tau_1(n) = m) / P_n(\tau_1(n) < \infty) \\ &= P_n(\omega_{n:n+\tau_1(n)} \in A | \tau_1(n) < \infty). \end{aligned}$$

The deficiency of Procedure **Aux** is that it does not recognize that  $n > N_1$ . Let us now account for this fact and note that conditional on  $\mathcal{F}_{N_1}$  we have that  $W_{i,k}^n$ 's are i.i.d.  $N(0, 1)$  but conditional on  $\{|W_{i,k}^n| \leq 4\sqrt{n+1}\}$  for all  $n > N_1$ . Define

$$\mathcal{H}_m^n = \{|W_{i,k}^h| \leq 4\sqrt{h+1} : 0 \leq k \leq 2^h - 1, n < h \leq n+m\}.$$

In order to simulate  $P_{N_1}(\tau_1(N_1) < \infty)$  we modify step 3 of Procedure **Aux**. Specifically, we have

**PROCEDURE B.**

**Input:** We assume that we have simulated  $\{(W_{i,k}^l : 0 \leq k < 2^l) : l \leq n\}$ . So, the  $W_{i,k}^m$ 's are i.i.d.  $N(0, 1)$  but conditional on  $\{|W_{i,k}^m| < 4\sqrt{m+1}\}$  for all  $m > n$ . We also assume that conditions (5.8) and (5.9) hold in Lemma 5.2; note the discussion following Lemma 5.2 which notes that this can be assumed at the expense of simulating additional  $W_{i,k}^m$ 's (with  $\{|W_{i,k}^m| < 4\sqrt{m+1}\}$  if  $m > N_1$ ).

**Output:** A Bernoulli  $F$  with parameter  $P_n(\tau_1(n) < \infty, \mathcal{H}_\infty^n)$ , and if  $F = 1$ , also

$$\omega_{n:n+\tau_1(n)} = \{W_{i,k}^l : 1 \leq k \leq 2^n, 1 \leq i \leq d', n < l \leq n + \tau_1(n)\}$$

conditional on  $\tau_1(n) < \infty$  and on  $\mathcal{H}_\infty^n$ .

**Step 1:** Sample  $M$  according to  $g(m)$ .

**Step 2:** Given  $M = m$  sample  $I$  and  $J$  ( $I \neq J$ ) uniformly over the set  $\{1, 2, \dots, d'\}$ . Then sample  $K', K$  from  $q_n(k, k'|m)$ .

**Step 3:** Given  $M = m, I = i, J = j, K = k,$  and  $K' = k'$ , simulate  $\omega_{n:n+m}$  from  $P_{n,m}^{i,j,k,k'}(\cdot)$ . Note that simulation from  $P_{n,m}^{i,j,k,k'}(\cdot)$  can be done according to Corollary 5.2. Check if  $\mathcal{H}_m^n$  occurs. If yes, continue to Step 4; otherwise, go back to Step 1.

**Step 4:** Compute

$$\begin{aligned} & \Xi_n(m, i, j, k, k', \omega_{n:n+m}) \\ &= 1 / (g(m)(d'(d' - 1))^{-1} q_n(k, k'|m) \\ & \quad \times \exp(\theta_0\{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\} - \psi_n(m, i, j, k, k'))), \end{aligned}$$

and

$$\mathcal{N}_n(m) = \sum_{1 \leq i, j \leq d', i \neq j} \sum_{1 \leq k < k' \leq 2^{n+m-1}} I(|L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k - k')^\beta \Delta_{n+m}^{2\alpha}).$$

**Step 5:** Simulate  $U$  uniformly distributed on  $[0, 1]$  independent of everything else and output

$$F = I \left\{ U < \left( I \left( \mathcal{H}_m^n \cap \{ |L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k - k')^\beta \Delta_{n+m}^{2\alpha} \} \cap \bigcap_{l=1}^{M-1} \bar{C}_n(l) \right) P(\mathcal{H}_\infty^{n+m}) \right) / P(\mathcal{H}_\infty^n) \times \Xi_n(m, i, j, k, k', \omega_{n:n+m}) / \mathcal{N}_n(m) \right\}.$$

[Notice that  $P(\mathcal{H}_\infty^{n+m})/P(\mathcal{H}_\infty^n) = P(\mathcal{H}_{n+m}^n)$  and can be computed in finite steps.]

If  $F = 1$ , also output  $\omega_{n:n+m}$ .

**End of Procedure B**

Let  $\tilde{Q}_n$  denote the distribution induced by Procedure B. Following the same analysis as that given for Procedure Aux, we can verify that

$$\tilde{Q}_n \left( U < \left( I \left( \mathcal{H}_m^n \cap \{ |L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)| > (k - k')^\beta \Delta_{n+m}^{2\alpha} \} \cap \bigcap_{l=1}^{M-1} \bar{C}_n(l) \right) P(\mathcal{H}_\infty^{n+m}) \right) / P(\mathcal{H}_\infty^n) \times \Xi_n(m, i, j, k, k', \omega_{n:n+m}) / \mathcal{N}_n(m) \right) = P_n(\tau_1(n) < \infty | \mathcal{H}_\infty^n).$$

And if the Bernoulli trial is a success, then,  $\omega_{n:n+M}$  is distributed according to

$$P_n(\omega_{n:n+\tau_1(n)} \in \cdot | \tau_1(n) < \infty, \mathcal{H}_\infty^n).$$

Finally, if  $\tau_1(n) = \infty$ , we may still need to simulate  $\omega_{n:n+m}$  for any  $m \geq 1$ , but now, conditional on  $\{\tau_1(n) = \infty, \mathcal{H}_\infty^n\}$ . Note that

$$\begin{aligned} & P_n(\omega_{n:n+m} \in A | \tau_1(n) = \infty, \mathcal{H}_\infty^n) \\ &= \frac{P_n(\omega_{n:n+m} \in A, \tau_1(n) = \infty, \mathcal{H}_\infty^n)}{P_n(\tau_1(n) = \infty, \mathcal{H}_\infty^n)} \\ &= \frac{E_n I(\omega_{n:n+m} \in A, \tau_1(n) > m, \mathcal{H}_m^n) P_{n+m}(\tau_1(n+m) = \infty, \mathcal{H}_\infty^{n+m})}{P_n(\tau_1(n) = \infty, \mathcal{H}_\infty^n)}. \end{aligned}$$

Thus, we can sample  $\omega_{n:n+m}$  from  $P_n(\cdot)$  and accept the path with probability

$$I(\tau_1(n) > m, \mathcal{H}_m^n) P_{n+m}(\tau_1(n+m) = \infty, \mathcal{H}_\infty^{n+m}).$$

---

**Algorithm 2** Simulate  $N_1$  and  $N_2$  jointly with  $W_{i,k}^n$ 's for  $1 \leq n \leq N_0$ , where  $N_0$  is chosen such that  $\sup_{t \in [0,1]} \|\hat{X}^{N_0}(t) - X(t)\|_\infty \leq \varepsilon$

---

**Input:** The parameters required to run Algorithm 1, and Procedures A and B. These are the tilting parameters  $\theta_0$ 's.

**Step 1:** Simulate  $N_1$  jointly with  $W_{i,k}^m$ 's for  $0 \leq m \leq N_1$  using Algorithm 1 (see the remark that follows after Algorithm 1). Let  $n = N_1$ .

**Step 2:** If any of the conditions (5.8) and (5.9) from Lemma 5.2 are not satisfied keep simulating  $W_{i,k}^m$ 's for  $m > n$  until the first level  $m > n$  for which conditions (5.8) and (5.9) are satisfied. Redefine  $n$  to be such first level  $m$ .

**Step 3:** Run Procedure B and obtain as output  $F$  and if  $F = 1$  also obtain  $\omega_{n:n+\tau(n)}$ .

**Step 4:** If  $\tau(n) < \infty$  (i.e.,  $F = 1$ ) set  $n \leftarrow \tau(n)$  and go back to Step 2. Otherwise, go to Step 4.

**Step 5:** Calculate  $G$  according to Procedure A and solve for  $N_0$  such that  $G\Delta_{N_0}^{2\alpha-\beta} < \varepsilon$ .

**Step 6:** If  $N_0 > n$  sample  $\omega_{n:N_0}$  from  $P_n(\cdot)$  and sample a Bernoulli random variable,  $I$  with probability of success  $P_{N_0}(\tau(N_0) = \infty, \mathcal{H}_\infty^{N_0})$ .

**Step 7:** If  $I = 0$ , go back to Step 6.

**Step 8:** Output  $\omega_{0:N_0}$ .

**End of Algorithm 2**

---

This clearly can be done since we can easily simulate Bernoulli's with probability

$$P_{n+m}(\tau(n+m) = \infty, \mathcal{H}_\infty^{n+m}) = P_{n+m}(\tau_1(n+m) = \infty | \mathcal{H}_\infty^{n+m}) P_{n+m}(\mathcal{H}_\infty^{n+m}).$$

We summarize the above discussions in Algorithm 2.

We obtain  $\{W_{i,k}^l : 0 \leq k < 2^l, l \leq N_0, 1 \leq i \leq d\}$  from Algorithm 2. We have from recursions in Lemma 5.1 how to obtain

$$(5.10) \quad \{(Z_i(t_r^l) - Z_i(t_{r-1}^l)) : 1 \leq r \leq 2^l, 1 \leq l \leq N_0, 1 \leq i \leq d\}$$

and then we can compute  $\{\hat{X}^{N_0}(t) : t \in D_{N_0}\}$  using equation (2.4).

REMARK. Observe that after completion of Algorithm 2, one can actually continue the simulation of increments in order to obtain an approximation with an error  $\varepsilon' < \varepsilon$ . In particular, this is done by repeating Steps 4 to 8. Start from Step 4 with  $n = N_0$ . The value of  $G$  has been computed, it does not depend on  $\varepsilon$ . However, one needs to recompute  $N_0 := N_0(\varepsilon')$  such that  $G\Delta_{N_0}^{2\alpha-\beta} < \varepsilon'$ . Then we can implement Steps 5 to 8 without change. One obtains an output that, as before, can be transformed into (5.10) via the recursions (5.1), yielding  $\{\hat{X}^{N_0(\varepsilon')}(t) : t \in D_{N_0(\varepsilon')}\}$  with a guaranteed error smaller than  $\varepsilon'$  in uniform norm with probability 1.

**6. Rough differential equations, error analysis and the proof of Theorem 2.1.** The analysis in this section follows closely the discussion from [7]

Section 3 and Section 7; see also [8] Chapter 10. We made some modifications to account for the drift of the process and also to be able to explicitly calculate the constant  $G$ . Let us start with the definition of a solution to (1.1) using the theory of rough differential equations. We first provide a definition of the solution of (1.1) in a pathwise sense, following [7].

DEFINITION 6.1.  $X(\cdot)$  is a solution of (1.1) on  $[0, 1]$  if  $X(0) = x(0)$  and for almost every sample path  $\{Z_j(\cdot) : j = 1, 2, \dots, d\}$  it holds

$$\left| X_i(t) - X_i(s) - \mu_i(X(s))(t - s) - \sum_{j=1}^{d'} \sigma_{i,j}(X(s))(Z_j(t) - Z_j(s)) - \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X(s)) \sigma_{l,m}(X(s)) A_{m,j}(s, t) \right| = o(t - s)$$

for  $i = 1, 2, \dots, d$  and  $0 \leq s < t \leq 1$ , where  $A_{i,j}(\cdot)$  satisfies

(6.1)  $A_{i,j}(r, t) = A_{i,j}(r, s) + A_{i,j}(s, t) + (Z_i(s) - Z_i(r))(Z_j(t) - Z_j(s))$

for  $0 \leq r < s < t \leq 1$ .

The previous definition is motivated by the following Taylor-type development:

$$\begin{aligned} X_i(t+h) &= X_i(t) + \int_t^{t+h} \mu_i(X(u)) du + \sum_{j=1}^{d'} \int_t^{t+h} \sigma_{i,j}(X(u)) dZ_j(u) \\ &\approx X_i(t) + \int_t^{t+h} \mu_i(X(u)) du \\ &\quad + \sum_{j=1}^{d'} \int_t^{t+h} \sigma_{i,j}(X(t) + \mu(X(t))(u-t) + \sigma(X(t))(Z(u) - Z(t))) dZ_j(u) \\ &\approx X_i(t) + \mu_i(X(t))h + \sum_{j=1}^{d'} \sigma_{i,j}(X(t))(Z_j(t+h) - Z_j(t)) \\ &\quad + \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X(t)) \sigma_{l,m}(X(t)) \int_t^{t+h} (Z_m(u) - Z_m(t)) dZ_j(u). \end{aligned}$$

The previous Taylor development suggests defining  $A_{i,j}(s, t) := \int_s^t (Z_i(u) - Z_i(s)) dZ_j(u)$ . Depending on how one interprets  $A(s, t)$ , for example, via Itô or Stratonovich integrals, one obtains a solution  $X(\cdot)$  which is interpreted in the corresponding context.

In order to obtain the Itô interpretation of the solution to equation (1.1) via definition (6.1), we shall interpret the integrals in the sense of Itô. In addition, as we shall explain, some technical conditions (in addition to the standard Lipschitz continuity typically required to obtain a strong solution) must be imposed in order to enforce the existence of a unique solution to (6.1).

There are two sources of errors when using  $\hat{X}^n$  in equation (2.4) to approximate  $X$ . One is the discretization on the dyadic grid, but assuming that  $A_{i,j}(t_k^n, t_{k+1}^n)$  is known; this type of analysis is the one that is most common in the literature on rough paths (see [7]). The second source of error arises due to the fact that  $A_{i,j}(t_k^n, t_{k+1}^n)$  is not known for  $i \neq j$ . Thus, we divide the proof of Theorem 2.1 into two steps (two propositions), each dealing with one source of error.

Similar to  $\hat{X}^n(t)$ , we define  $\{X^n(t) : t \in D_n\}$  by the following recursion: given  $X^n(0) = X(0)$ ,

$$(6.2) \quad \begin{aligned} X_i^n(t_{k+1}^n) &= X_i^n(t_k^n) + \mu_i(X^n(t_k^n))\Delta_n + \sum_{j=1}^{d'} \sigma_{i,j}(X^n(t_k^n))(Z_j(t_{k+1}^n) - Z_j(t_k^n)) \\ &+ \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X^n(t_k^n)) \sigma_{l,m}(X^n(t_k^n)) A_{m,j}(t_k^n, t_{k+1}^n), \end{aligned}$$

and for  $t \in [0, 1]$ , we let  $X^n(t) = X^n(\lfloor t \rfloor)$ , where in this context  $\lfloor t \rfloor = \max\{s \in D_n : s \leq t\}$ .

**PROPOSITION 6.1.** *Under the conditions of Theorem 2.1, we can compute a constant  $G_1$  explicitly in terms of  $M$ ,  $\|Z\|_\alpha$  and  $\|A\|_{2\alpha}$ , such that for  $n$  large enough*

$$\|X^n(t) - X(t)\|_\infty \leq G_1 \Delta_n^{3\alpha-1}.$$

The proof of Proposition 6.1 will be given after introducing some definitions and key auxiliary results. We denote

$$I_i^n(r, t) := X_i^n(t) - X_i^n(r) - \mu_i(X^n(r))(t - r) - \sum_{j=1}^{d'} \sigma_{i,j}(X^n(r))(Z_j(t) - Z_j(r))$$

and

$$J_i^n(r, t) := I_i^n(r, t) - \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X^n(r)) \sigma_{l,m}(X^n(r)) A_{m,j}(r, t).$$

The following lemmas introduce the main technical results for the proof of Proposition 6.1.

LEMMA 6.1. *Under the conditions of Theorem 2.1, there exist constants  $C_1, C_2$  and  $C_3$  that depend only on  $M, \|Z\|_\alpha$  and  $\|A\|_{2\alpha}$ , such that for any large enough  $n$  and  $r, t \in D_n$ ,*

$$\begin{aligned} \|X^n(t) - X^n(r)\|_\infty &\leq C_1|t - r|^\alpha, \\ \|I^n(r, t)\|_\infty &\leq C_2|t - r|^{2\alpha} \end{aligned}$$

and

$$\|J^n(r, t)\|_\infty \leq C_3|t - r|^{3\alpha}.$$

PROOF. For  $r \leq s \leq t, r, s, t \in D_n$ , we have the following important recursions:

$$\begin{aligned} I_i^n(r, t) &= I_i^n(r, s) + I_i^n(s, t) + (\mu_i(X^n(s)) - \mu_i(X^n(r)))(t - s) \\ &\quad + \sum_{j=1}^{d'} (\sigma_{i,j}(X^n(s)) - \sigma_{i,j}(X^n(r)))(Z_j(t) - Z_j(s)) \end{aligned}$$

and

$$\begin{aligned} &J_i^n(r, t) \\ &= J_i^n(r, s) + J_i^n(s, t) + (\mu_i(X^n(s)) - \mu_i(X^n(r)))(t - s) \\ &\quad + \sum_{j=1}^{d'} \left[ \sigma_{i,j}(X^n(s)) - \sigma_{i,j}(X^n(r)) - \sum_{l=1}^d \partial_l \sigma_{i,j}(X^n(r))(X_l^n(s) - X_l^n(r)) \right. \\ (6.3) \quad &\quad + \sum_{l=1}^d \partial_l \sigma_{i,j}(X^n(r))I_l^n(r, s) \\ &\quad \left. + \sum_{l=1}^d \partial_l \sigma_{i,j}(X^n(r))\mu_i(X^n(r))(s - r) \right] (Z_j(t) - Z_j(s)) \\ &\quad + \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} [\partial_l \sigma_{i,j}(X^n(s))\sigma_{l,m}(X^n(s)) \\ &\quad - \partial_l \sigma_{i,j}(X^n(r))\sigma_{l,m}(X^n(r))] A_{m,j}(s, t). \end{aligned}$$

We next divide the proof into two parts. We first prove that there exists a small enough constant  $\delta > 0$  and three large enough constants  $C_1(\delta), C_2(\delta)$  and  $C_3(\delta)$ , all independent of  $n$ , such that for  $|t - r| < \delta, \|X^n(t) - X^n(r)\|_\infty \leq C_1(\delta)|t - r|^\alpha, \|I^n(r, t)\|_\infty \leq C_2(\delta)|t - r|^{2\alpha}$  and  $\|J^n(r, t)\|_\infty \leq C_3(\delta)|t - r|^{3\alpha}$ . We prove it by induction. First, we have  $J^n(r, r) = 0$  and  $J^n(r, r + \Delta_n) = 0$ . Suppose the result hold for all pairs of  $r_0, t_0 \in D_n$  with  $|t_0 - r_0| < |t - r|$ . We then pick  $s \in D_n$  as

the largest point between  $r$  and  $t$  such that  $|s - r| \leq |t - r|/2$ . Then we also have  $|s + \Delta_n - r| > |t - r|/2$  and  $|t - (s + \Delta_n)| < |t - r|/2$ . For simplicity of notation, we denote  $\bar{d} = \max\{d, d'\}$ .

As

$$X_i^n(t) - X_i^n(s) = J_i^n(s, t) + \mu_i(X^n(s))(t - s) + \sum_{j=1}^{d'} \sigma_{i,j}(X^n(s))(Z_j(t) - Z_j(s)) \\ + \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X^n(s)) \sigma_{l,m}(X^n(s)) A_{m,j}(s, t),$$

we have

$$|X_i^n(t) - X_i^n(s)| \\ \leq C_3(\delta) |t - s|^{3\alpha} + M |t - s| + \bar{d} M \|Z\|_\alpha |t - s|^\alpha + \bar{d}^3 M^2 \|A\|_{2\alpha} |t - s|^{2\alpha} \\ \leq (C_3(\delta) \delta^{2\alpha} + M \delta^{1-\alpha} + \bar{d} M \|Z\|_\alpha + \bar{d}^3 M^2 \|A\|_{2\alpha} \delta^\alpha) |t - s|^\alpha \\ \leq C_1(\delta) |t - s|^\alpha$$

for  $C_1(\delta) \geq C_3(\delta) \delta^{2\alpha} + M \delta^{1-\alpha} + \bar{d} M \|Z\|_\alpha + \bar{d}^3 M^2 \|A\|_{2\alpha} \delta^\alpha$ .

And as

$$I_i^n(s, t) = J_i^n(s, t) + \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X^n(s)) \sigma_{l,m}(X^n(s)) A_{m,j}(s, t),$$

we have

$$|I_i^n(s, t)| \leq C_3(\delta) |t - s|^{3\alpha} + \bar{d}^3 M^2 \|A\|_{2\alpha} |t - s|^{2\alpha} \\ \leq (C_3(\delta) \delta^\alpha + \bar{d}^3 M^2 \|A\|_{2\alpha}) |t - s|^{2\alpha} \leq C_2(\delta) |t - s|^{2\alpha}$$

for  $C_2(\delta) \geq C_3(\delta) \delta^\alpha + \bar{d}^3 M^2 \|A\|_{2\alpha}$ .

We now analyze the recursion (6.3) term by term. First,

$$|\mu_i(X^n(s)) - \mu_i(X^n(r))| \leq M C_1(\delta) |s - r|^\alpha, \\ \left| \sigma_{i,j}(X^n(s)) - \sigma_{i,j}(X^n(r)) - \sum_{l=1}^d \partial_l \sigma_{i,j}(X^n(r)) (X_l^n(s) - X_l^n(r)) \right| \\ \leq M C_1(\delta)^2 |s - r|^{2\alpha}, \\ \left| \sum_{l=1}^d \partial_l \sigma_{i,j}(X^n(r)) I_l^n(r, s) \right| \leq \bar{d} M C_2(\delta) |s - r|^{2\alpha}, \\ \sum_{l=1}^d \partial_l \sigma_{i,j}(X^n(r)) \mu_i(X^n(r)) (s - r) \leq \bar{d} M^2 |s - r|$$



and

$$|\partial_l \sigma_{i,j}(X^n(s)) \sigma_{l,m}(X^n(s)) - \partial_l \sigma_{i,j}(X^n(r)) \sigma_{l,m}(X^n(r))| \leq 2M^2 C_1(\delta) |s - r|^\alpha.$$

Then

$$\begin{aligned} & |J_i^n(r, t)| \\ & \leq |J_i^n(r, s)| + |J_i^n(s, t)| \\ & \quad + (MC_1(\delta) + \bar{d}MC_1(\delta)^2 \|Z\|_\alpha + \bar{d}^2 MC_2(\delta) \|Z\|_\alpha + \bar{d}^2 M^2 \|Z\|_\alpha \\ & \quad + 2\bar{d}^3 M^2 C_1(\delta) \|A\|_\alpha) |t - r|^{3\alpha}. \end{aligned}$$

Likewise, we have

$$\begin{aligned} & |J_i^n(s, t)| \\ & \leq |J_i^n(s, s + \Delta_n)| + |J_i^n(s + \Delta_n, t)| \\ & \quad + (MC_1(\delta) + \bar{d}MC_1(\delta)^2 \|Z\|_\alpha + \bar{d}^2 MC_2(\delta) \|Z\|_\alpha + \bar{d}^2 M^2 \|Z\|_\alpha \\ & \quad + 2\bar{d}^3 M^2 C_1(\delta) \|A\|_\alpha) |t - s|^{3\alpha} \\ & = |J_i^n(s + \Delta_n, t)| \\ & \quad + (MC_1(\delta) + \bar{d}MC_1(\delta)^2 \|Z\|_\alpha + \bar{d}^2 MC_2(\delta) \|Z\|_\alpha + \bar{d}^2 M^2 \|Z\|_\alpha \\ & \quad + 2\bar{d}^3 M^2 C_1(\delta) \|A\|_\alpha) |t - s|^{3\alpha}. \end{aligned}$$

Then

$$\begin{aligned} & |J_i^n(r, t)| \\ & \leq |J_i^n(r, s)| + |J_i^n(s + \Delta_n, t)| \\ & \quad + 2(MC_1(\delta) + \bar{d}MC_1(\delta)^2 \|Z\|_\alpha + \bar{d}^2 MC_2(\delta) \|Z\|_\alpha + \bar{d}^2 M^2 \|Z\|_\alpha \\ & \quad + 2\bar{d}^3 M^2 C_1(\delta) \|A\|_\alpha) |t - s|^{3\alpha} \\ & \leq \{2^{1-3\alpha} C_3(\delta) + 2(MC_1(\delta) + \bar{d}MC_1(\delta)^2 \|Z\|_\alpha + \bar{d}^2 MC_2(\delta) \|Z\|_\alpha \\ & \quad + \bar{d}^2 M^2 \|Z\|_\alpha + 2\bar{d}^3 M^2 C_1(\delta) \|A\|_\alpha)\} |t - s|^{3\alpha} \\ & \leq C_3(\delta) |t - s|^{3\alpha}, \end{aligned}$$

for

$$\begin{aligned} (1 - 2^{1-3\alpha}) C_3(\delta) & \geq 2(MC_1(\delta) + \bar{d}MC_1(\delta)^2 \|Z\|_\alpha + \bar{d}^2 MC_2(\delta) \|Z\|_\alpha \\ & \quad + \bar{d}^2 M^2 \|Z\|_\alpha + 2\bar{d}^3 M^2 C_1(\delta) \|A\|_\alpha). \end{aligned}$$

Therefore, if we deliberately choose  $\delta$ ,  $C_1(\delta)$ ,  $C_2(\delta)$  and  $C_3(\delta)$  such that

$$\begin{aligned}
 C_1(\delta) &\geq C_3(\delta)\delta^{2\alpha} + M\delta^{1-\alpha} + \bar{d}M\|Z\|_\alpha + \bar{d}^3M^2\|A\|_{2\alpha}\delta^\alpha, \\
 C_2(\delta) &\geq C_3(\delta)\delta^\alpha + \bar{d}^3M^2\|A\|_{2\alpha}, \\
 (6.4) \quad C_3(\delta) &\geq \frac{2}{1-2^{1-3\alpha}}(MC_1(\delta) + \bar{d}MC_1(\delta)^2\|Z\|_\alpha + \bar{d}^2MC_2(\delta)\|Z\|_\alpha \\
 &\quad + \bar{d}^2M^2\|Z\|_\alpha + 2\bar{d}^3M^2C_1(\delta)\|A\|_\alpha).
 \end{aligned}$$

Then we have, for  $|t - r| < \delta$ ,

$$\begin{aligned}
 \|X^n(t) - X^n(r)\|_\infty &\leq C_1(\delta)|t - r|^\alpha, \\
 \|I^n(r, t)\|_\infty &\leq C_2(\delta)|t - r|^{2\alpha}, \\
 \|J^n(r, t)\|_\infty &\leq C_3(\delta)|t - r|^{3\alpha}.
 \end{aligned}$$

The existence of  $\delta$ ,  $C_1(\delta)$ ,  $C_2(\delta)$  and  $C_3(\delta)$ , satisfying the system of inequalities (6.4), follows from Lemma 2.1.

We now extend the analysis to the case when  $|t - r| > \delta$ . For  $n$  large enough ( $\Delta_n < \delta/2$ ), if  $|t - r| > \delta$ , we can always find points  $s_i \in D_n$  and  $r = s_0 < s_1 < \dots < s_k = t$  such that  $\max_{1 \leq i \leq k} |s_i - s_{i-1}| < \delta$  and  $\min_{1 \leq i \leq k} |s_i - s_{i-1}| \geq \delta/2$ . Then

$$|X_i^n(t) - X_i^n(r)| \leq \sum_{l=1}^k |X_i^n(s_l) - X_i^n(s_{l-1})| \leq kC_1(\delta)|t - r|^\alpha \leq \frac{2}{\delta}C_1(\delta)|t - r|^\alpha.$$

Let  $C_1 = \frac{2}{\delta}C_1(\delta)$  and we can write  $\|X^n(t) - X^n(r)\|_\infty \leq C_1|t - r|^\alpha$ . Next,

$$\begin{aligned}
 |I_i^n(r, t)| &\leq \sum_{l=1}^k \left\{ |I_i^n(s_{l-1}, s_l)| + |(\mu_i(X^n(s_l)) - \mu_i(X^n(s_0)))(s_l - s_{l-1})| \right. \\
 &\quad \left. + \left| \sum_{j=1}^{d'} (\sigma_{ii}(X^n(s_l)) - \sigma_{ij}(X^n(s_0)))(Z_j(s_{l+1}) - Z_j(s_l)) \right| \right\} \\
 &\leq k[C_2(\delta)|t - r|^{2\alpha} + MC_1|t - r|^{1+\alpha} + dMC_1\|Z\|_\alpha|t - r|^{2\alpha}] \\
 &\leq \frac{2}{\delta}(C_2(\delta) + MC_1 + \bar{d}MC_1\|Z\|_\alpha)|t - r|^{2\alpha}.
 \end{aligned}$$

By setting  $C_2 = \frac{2}{\delta}(C_2(\delta) + MC_1 + \bar{d}MC_1\|Z\|_\alpha)$ , we have  $\|I^n(r, t)\|_\infty \leq C_2|t - r|^{2\alpha}$ .

Now following the same induction analysis on  $J_i^n(s, t)$  as we did in the case  $|t - s| < \delta$ , we have

$$\begin{aligned}
 |J_i^n(r, t)| &\leq \frac{2}{2^{3\alpha}}C_3|t - r|^{3\alpha} \\
 &\quad + 2(MC_1 + \bar{d}MC_1^2\|Z\|_\alpha \\
 &\quad + \bar{d}^2MC_2\|Z\|_\alpha + 2\bar{d}^3M^2C_1\|A\|_\alpha)|t - r|^{3\alpha}.
 \end{aligned}$$

If we choose

$$C_3 = \frac{2}{1 - 2^{1-3\alpha}} (MC_1 + \bar{d}MC_1^2 \|Z\|_\alpha + \bar{d}^2 MC_2 \|Z\|_\alpha + 2\bar{d}^3 M^2 C_1 \|A\|_\alpha),$$

then  $\|J^n(r, t)\|_\infty \leq C_3 |t - s|^{3\alpha}$ .  $\square$

LEMMA 6.2. *Let  $x(0)$  and  $\tilde{x}(0) \in R^d$  be two different vectors. We denote  $X^n(t)$  and  $\tilde{X}^n(t)$  for  $t \in D_n$  as the  $n$ th dyadic approximation defined by (6.2) with initial value  $x(0)$  and  $\tilde{x}(0)$ , respectively. Under the conditions of Theorem 2.1, there exists a constant  $B$ , independent of  $n$ , such that for  $t \in D_n$ ,*

$$\|X^n(t) - \tilde{X}^n(t) - (X^n(0) - \tilde{X}^n(0))\|_\infty \leq Bt^\alpha \|X^n(0) - \tilde{X}^n(0)\|_\infty.$$

Moreover,

$$\|X^n(t) - \tilde{X}^n(t)\|_\infty \leq (1 + B) \|X^n(0) - \tilde{X}^n(0)\|_\infty.$$

PROOF. Let

$$Y_{i,h}^n(t) = \frac{X_i^n(t) - \tilde{X}_i^n(t)}{\|X_h^n(0) - \tilde{X}_h^n(0)\|_\infty}.$$

We define  $0/0 = 0$ .

Then following the recursion (6.2), we have

$$\begin{aligned} & Y_i^n(t_{k+1}^n) \\ &= Y_i^n(t_k^n) + \frac{\mu_i(X^n(t_k^n)) - \mu_i(\tilde{X}^n(t_k^n))}{\|X^n(0) - \tilde{X}^n(0)\|_\infty} \Delta_n \\ (6.5) \quad & + \sum_{j=1}^{d'} \frac{\sigma_{i,j}(X^n(t_k^n)) - \sigma_{i,j}(\tilde{X}^n(t_k^n))}{\|X^n(0) - \tilde{X}^n(0)\|_\infty} (Z_j(t_{k+1}^n) - Z_j(t_k^n)) \\ & + \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \frac{\partial_l \sigma_{i,j}(X^n(t_k^n)) \sigma_{l,m}(X^n(t_k^n)) - \partial_l \sigma_{i,j}(\tilde{X}^n(t_k^n)) \sigma_{l,m}(\tilde{X}^n(t_k^n))}{\|X^n(0) - \tilde{X}^n(0)\|_\infty} \\ & \times A_{m,j}(t_k^n, t_{k+1}^n). \end{aligned}$$

Then (6.2) and (6.5) together define an recursion to generate  $X^n$ ,  $\tilde{X}^n$  and  $Y^n$ . Following Lemma 6.1, there exists a constant  $B$  that depends only on  $M$ ,  $\|Z\|_\alpha$  and  $\|A\|_{2\alpha}$ , such that

$$\|Y^n(t) - Y^n(0)\|_\infty \leq Bt^\alpha.$$

Thus,

$$\|X^n(t) - \tilde{X}^n(t) - (X^n(0) - \tilde{X}^n(0))\|_\infty \leq Bt^\alpha \|X^n(0) - \tilde{X}^n(0)\|_\infty,$$

and

$$\|X^n(t) - \tilde{X}^n(t)\|_\infty \leq (1 + B)\|X^n(0) - \tilde{X}^n(0)\|_\infty. \quad \square$$

We are now ready to prove Proposition 6.1.

**PROOF OF PROPOSITION 6.1.** From Lemma 6.1 we have  $\|X^n(t) - X^n(r)\|_\infty \leq C_1|t - r|^\alpha$ . By the Arzela–Ascoli theorem, there exists a subsequence of  $\{X^n\}$  that converges uniformly to some continuous function  $X$  on  $[0, 1]$ . Moreover, we have  $\|X(t) - X(r)\|_\infty \leq C_1|t - r|^\alpha$  and

$$\begin{aligned} & \left| X_i(t) - X_i(r) - \mu_i(X(r)) - \sum_{j=1}^{d'} \sigma_{i,j}(X(r))(Z_j(t) - Z_j(r)) \right. \\ & \quad \left. - \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X(r)) \sigma_{l,m}(X(r)) A_{m,j}(r, t) \right| \\ & \quad < C_2|t - r|^{3\alpha}. \end{aligned}$$

Therefore, the limit  $X$  is a solution to the SDE.

Let  $X^{n,(s)}(t; X(s)) := X^n(t - s) | X^n(0) = X(s)$ . Specifically, we have  $X^{n,(0)}(t; X(0)) = X^n(t)$  with  $X^n(0) = X(0)$ , and  $X^{n,(t)}(t; X(t)) = X(t)$ . Then we can write

$$X^n(t_m^n) - X(t_m^n) = \sum_{k=1}^m (X^{n,(t_k^n)}(t_m^n; X(t_k^n)) - X^{n,(t_{k-1}^n)}(t_m^n; X(t_{k-1}^n))).$$

By Lemma 6.2,  $\|X^{n,(t_k^n)}(t_m; X(t_k^n)) - X^{n,(t_{k-1}^n)}(t_m; X(t_{k-1}^n))\|_\infty \leq (1 + B)\|X(t_k^n) - X^{n,(t_{k-1}^n)}(t_k^n; X(t_{k-1}^n))\|_\infty$ . We also have

$$\begin{aligned} & |X_i(t_k^n) - X_i^{n,(t_{k-1}^n)}(t_k^n; X(t_{k-1}^n))| \\ & = \left| X_i(t_k^n) - X_i(t_{k-1}^n) - \mu_i(X(t_{k-1}^n))(t_k^n - t_{k-1}^n) \right. \\ & \quad \left. - \sum_{j=1}^{d'} \sigma_{i,j}(X(t_{k-1}^n))(Z_j(t_k^n) - Z_j(t_{k-1}^n)) \right. \\ & \quad \left. - \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(X(t_{k-1}^n)) \sigma_{l,m}(X(t_{k-1}^n)) A_{m,j}(t_{k-1}^n, t_k^n) \right| \\ & \leq C_3|t_k^n - t_{k-1}^n|^{3\alpha}. \end{aligned}$$

Thus,

$$\begin{aligned} \|X^n(t_m^n) - X(t_m^n)\|_\infty &\leq \sum_{k=1}^m \|X^{n,(t_k^n)}(t_m^n; X(t_k^n)) - X^{n,(t_{k-1}^n)}(t_m^n; X(t_{k-1}^n))\|_\infty \\ &\leq m(1+B)C_3\Delta_n^{3\alpha} \\ &\leq (1+B)C_3\Delta_n^{3\alpha-1}. \end{aligned} \quad \square$$

Next, we turn to the analysis of the error induced by approximating the Lévy area.

PROPOSITION 6.2. *Under the conditions of Theorem 2.1, we can compute a constant  $G_2$  explicitly in terms of  $M$ ,  $\|Z\|_\alpha$ ,  $\|A\|_{2\alpha}$  and  $\Gamma_R$ , such that for  $n$  large enough*

$$\|\hat{X}^n(t) - X^n(t)\|_\infty \leq G_2\Delta_n^{2\alpha-\beta},$$

where  $\beta \in (1 - \alpha, 2\alpha)$ .

The proof of Proposition 6.2 uses a similar technique as the proof of Proposition 6.1 and also relies on some auxiliary results. Let

$$\begin{aligned} U_i^n(s, t) &:= \hat{X}_i^n(t) - X_i^{n,(s)}(t; \hat{X}^n(s)) \\ &\quad + \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(\hat{X}^n(s)) \sigma_{l,m}(\hat{X}^n(s)) R_{m,j}^n(s, t). \end{aligned}$$

We first prove the following technical result.

LEMMA 6.3. *Under the conditions of Theorem 2.1, there exists a constant  $C_4$ , that depends only on  $M$ ,  $\|Z\|_\alpha$ ,  $\|A\|_{2\alpha}$  and  $\Gamma_R$ , such that*

$$\|U^n(r, t)\|_\infty \leq C_4|t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta}.$$

PROOF. For  $0 \leq r < s < t \leq 1$ ,  $r, s, t \in D_n$ , we have

$$\begin{aligned} U_i^n(r, t) &= U_i^n(r, s) + U_i^n(s, t) \\ &\quad + [X_i^{n,(s)}(t; \hat{X}^n(s)) - X_i^{n,(r)}(t; \hat{X}^n(r)) - (\hat{X}_i^n(s) - X_i^{n,(r)}(s; \hat{X}^n(r)))] \\ &\quad - \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} (\partial_l \sigma_{i,j}(\hat{X}^n(s)) \sigma_{l,m}(\hat{X}^n(s)) \\ &\quad - \partial_l \sigma_{i,j}(\hat{X}^n(r)) \sigma_{l,m}(\hat{X}^n(r))) R_{m,j}^n(s, t). \end{aligned}$$

From Lemma 6.2,

$$\begin{aligned} & |X_i^{n,(s)}(t; \hat{X}^n(s)) - X_i^{n,(r)}(t; \hat{X}^n(r)) - (\hat{X}_i^n(s) - X_i^{n,(r)}(s; \hat{X}^n(r)))| \\ & \leq B|t - s|^\alpha \|\hat{X}^n(s) - X^{n,(r)}(s; \hat{X}^n(r))\|_\infty. \end{aligned}$$

From Lemma 6.1,

$$\begin{aligned} & |(\partial_l \sigma_{i,j}(\hat{X}^n(s)) \sigma_{l,m}(\hat{X}^n(s)) - \partial_l \sigma_{i,j}(\hat{X}^n(r)) \sigma_{l,m}(\hat{X}^n(r))) R_{m,j}^n(s, t)| \\ & \leq 2M^2 C_1 |s - r|^\alpha \Gamma_R |t - s|^\beta \Delta_n^{2\alpha - \beta} \\ & \leq 2M^2 C_1 \Gamma_R |t - r|^{\alpha + \beta} \Delta_n^{2\alpha - \beta}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \|U^n(r, t)\|_\infty \\ & \leq \|U^n(r, s)\|_\infty + \|U^n(s, t)\|_\infty + B|t - s|^\alpha \|\hat{X}^n(s) - X^{n,(r)}(s; \hat{X}^n(r))\|_\infty \\ & \quad + 2\bar{d}^3 M^2 C_1 \Gamma_R |t - r|^{\alpha + \beta} \Delta_n^{2\alpha - \beta} \\ (6.6) \quad & \leq \|U^n(r, s)\|_\infty + \|U^n(s, t)\|_\infty + B|t - s|^\alpha \|U^n(r, s)\|_\infty \\ & \quad + B|t - s|^\alpha \max_i \left\{ \left| \sum_{j=1}^{d'} \sum_{l=1}^d \sum_{m=1}^{d'} \partial_l \sigma_{i,j}(\hat{X}^n(r)) \sigma_{l,m}(\hat{X}^n(r)) R_{m,j}^n(r, s) \right| \right\} \\ & \quad + 2\bar{d}^3 M^2 C_1 \Gamma_R |t - r|^{\alpha + \beta} \Delta_n^{2\alpha - \beta} \\ & \leq (1 + B|t - s|^\alpha) \|U^n(r, s)\|_\infty + \|U^n(s, t)\|_\infty \\ & \quad + (B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha + \beta} \Delta_n^{2\alpha - \beta}, \end{aligned}$$

where  $\bar{d} = \max\{d, d'\}$ .

Like the proof of Lemma 6.1, we divide the proof into two parts. We first prove that there exist a small enough constant  $\delta > 0$  and a large enough constant  $C_4(\delta)$ , both independent of  $n$ , such that for  $|t - r| < \delta$ ,  $|U^n(r, t)| \leq C_4(\delta) |t - r|^{\alpha + \beta} \Delta_n^{2\alpha - \beta}$ . And we prove it by induction. First we have  $U_{t_k^n, t_k^n}^n = 0$  and  $U_{t_k^n, t_{k+1}^n}^n = 0$ . Suppose the bound holds for all pairs  $r_0, t_0 \in D_n$  with  $|t_0 - r_0| < |t - r|$ . We pick  $s \in D_n$  as the largest point between  $r$  and  $t$  such that  $|s - r| \leq 1/2|t - r|$ . Then we also have  $|(s + \Delta_n) - r| > 1/2|t - r|$  and  $|t - (s + \Delta_n)| < 1/2|t - r|$ .

$$\begin{aligned} \|U^n(r, t)\|_\infty & \leq (1 + B|t - s|^\alpha) \|U^n(r, s)\|_\infty + \|U^n(s, t)\|_\infty \\ & \quad + (B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha + \beta} \Delta_n^{2\alpha - \beta} \end{aligned}$$

and

$$\begin{aligned} & \|U^n(s, t)\|_\infty \\ & \leq (1 + B\Delta_n^\alpha) \|U^n(s, s + \Delta_n)\|_\infty + \|U^n(s + \Delta_n, t)\|_\infty \end{aligned}$$

$$\begin{aligned}
 &+ (B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - s|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\leq \|U^n(s + \Delta_n, t)\|_\infty + (B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta}.
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 &\|U^n(r, t)\|_\infty \\
 &\leq (1 + B\delta^\alpha) \|U^n(r, s)\|_\infty + \|U^n(s + \Delta_n, t)\|_\infty \\
 &\quad + 2(B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\leq \frac{2 + B\delta^\alpha}{2^{\alpha+\beta}} C_4(\delta) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\quad + 2(B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta}.
 \end{aligned}$$

If we pick  $\delta$  and  $C_4(\delta)$  such that

$$B\delta^\alpha \leq 2^{\alpha+\beta} - 2$$

and

$$\left(1 - \frac{2 + B\delta^\alpha}{2^{\alpha+\beta}}\right) C_4(\delta) \geq 2(B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R).$$

Then  $\|U^n(r, t)\|_\infty \leq C(\delta) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta}$ . We next extend the result to the case when  $|t - r| > \delta$ . We can always divide the interval  $[r, t]$  into smaller intervals of length less than  $\delta$ , specifically, for  $n$  large enough, we consider  $r = s_0 < s_1 < \dots < s_k = t$  where  $s_i \in D_n$  and  $1/2\delta < |s_i - s_{i-1}| < \delta$  for  $i = 1, 2, \dots, k$ . Then  $k < 2|t - r|/\delta \leq 2/\delta$  and

$$\begin{aligned}
 &\|U^n(r, t)\|_\infty \leq (1 + B|s_1 - s_0|^\alpha) \|U^n(s_0, s_0)\|_\infty + \|U^n(s_1, s_2)\|_\infty \\
 &\quad + (B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\leq \sum_{i=1}^k (1 + B\delta^\alpha) \|U^n(s_{i-1}, s_i)\|_\infty \\
 &\quad + k(B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\leq (1 + B\delta^\alpha) C_4(\delta) \Delta_n^{2\alpha-\beta} \sum_{i=1}^k |s_i - s_{i-1}|^{\alpha+\beta} \\
 &\quad + k(B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\leq (1 + B\delta^\alpha) C_4(\delta) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\quad + \frac{2}{\delta} (B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R) |t - r|^{\alpha+\beta} \Delta_n^{2\alpha-\beta} \\
 &\leq C_4 |m - k|^{\alpha+\beta} \Delta_n^{2\alpha-\beta}
 \end{aligned}$$

for  $C_4 \geq (1 + B\delta^\alpha) C_4(\delta) + 2(B\bar{d}^3 M^2 \Gamma_R + 2\bar{d}^3 M^2 C_1 \Gamma_R)/\delta$ .  $\square$

We are now ready to prove Proposition 6.2.

PROOF OF PROPOSITION 6.2. From Lemma 6.3, we have

$$\|U^n(0, t)\|_\infty \leq C_4 t^{\alpha+\beta} \Delta_n^{2\alpha-\beta}.$$

Then

$$\begin{aligned} |\hat{X}_i^n(t) - X_i^n(t)| &\leq |U_i^n(0, t)| + \sum_{j=1}^d \sum_{l=1}^d \sum_{m=1}^d |\partial_l \sigma_{i,j}(X(0)) \sigma_{l,m}(X(0))| |R_{m,j}^n(0, t)| \\ &\leq C_4 t^{\alpha+\beta} \Delta_n^{2\alpha-\beta} + \bar{d}^3 M^2 \Gamma_R t^\beta \Delta_n^{2\alpha-\beta} \\ &\leq (C_4 + \bar{d}^3 M^2 \Gamma_R) \Delta_n^{2\alpha-\beta}. \quad \square \end{aligned}$$

**7. Numerical implementation.** We conducted some numerical experiments. The goal is to demonstrate that the algorithms are implementable and correct. We would also like to explain some limitations in our implementation process and hope these would provide directions for future improvement of the framework developed here.

1. For values of  $X(t) : 0 \leq t \leq 1$  which fluctuate around numerical values around, say 1 (assuming that drift and diffusion coefficients also take these values), Procedure A obtains a value of the parameter  $G$  of order  $10^3$ . Thus for a reasonable level of accuracy, doing the computations implied by this size of  $G$ , one would generate about 20 wavelet levels, which corresponds to about  $2^{20}$  normal random variables. This amount is manageable in a standard single processor, but the amount could go out of hand in a standard computing environment if  $G$  is of size, say 100. A potential way to mitigate this issue would be to simulate a properly scaled down version of the path and scale everything back once we have simulated the path, or, alternatively, to make this portion of the procedure run in parallel computing cores.

2. We have some freedom in picking the parameter  $\alpha \in (1/3, 1/2)$  and  $\beta \in (1 - \alpha, 2\alpha)$ , but there is a tradeoff. From Theorem 2.1, we want  $2\alpha - \beta$  as close to  $1/2$  as possible ( $\alpha$  close to  $1/2$  and  $\beta$  close to  $1 - \alpha$ ). On the other hand, for the upper bound of  $\|Z\|_\alpha$  and due to our procedure for finding  $N_2$  (Section 5.2), we want  $\alpha$  to be reasonably small and  $\beta$  to be reasonably large. The point is, even if the theoretical complexity as  $\varepsilon$  decreases is driven by Theorem 2.1, we observed that in practice, given a fixed  $\varepsilon$ , it might be better to choose  $\alpha$  somewhat small, but within the range  $(1/3, 1/2)$ .

For our numerical experiments, we simulated a 2-dimensional geometric Brownian motion:

$$\begin{aligned} dX_1(t) &= \mu_1 X_1(t) dt + \sigma_1 X_1(t) dB_1(t), \\ dX_2(t) &= \mu_2 X_2(t) dt + \rho \sigma_2 X_2(t) dB_1(t) + \sqrt{1 - \rho^2} \sigma_2 X_2(t) dB_2(t) \end{aligned}$$



with initial value  $X_1(0)$  and  $X_2(0)$ . We recognize that this SDE has a closed form solution; this is useful because we want to compare the output of our method and the output of an algorithm that does not take advantage of the Euler discretization. The previous SDE has the following closed form solution:

$$\begin{aligned}
 X_1(t) &= X_1(0) \exp((\mu_1 - \sigma_1^2/2)t + \sigma_1 B_1(t)), \\
 X_2(t) &= X_2(0) \exp((\mu_1 - \sigma_2^2/2)t + \rho\sigma_2 B_1(t) + \sqrt{1 - \rho^2}\sigma_2 B_2(t)).
 \end{aligned}$$

Note that, the solution to this SDE is a continuous function of the Brownian motion under the uniform topology; so a Tolerance Enforced Simulation procedure using the closed form expression is much easier to design and, therefore, it can be used as a benchmark. Note that continuity of solution of the SDE under uniform norm does not imply that by only controlling the error of the wavelet approximation to Brownian motion in uniform metric, one can approximate to a given (deterministic) tolerance the error of the solution to the SDE *when applying the Euler scheme*. In order to guarantee that the Euler scheme yields an error which is bounded by a user defined (deterministic) tolerance with probability one, one needs to apply our procedure.

Figure 2 provides one numerical illustration of the performance of our algorithm. The light color is the path produced by our algorithm using the Euler scheme with a random truncation (which captures enough information to enforce a deterministic error in path space). The dark color is the simulation obtained by using a TES in uniform norm for the closed form expression. We observe that the two are indeed very close to each other. In particular, the recursively constructed path is within  $\varepsilon$  ( $\varepsilon = 0.1$ ) error bound of the true path. In fact, it appears that the constants are probably pessimistic in the sense that the actual error is much smaller than the prescribed guaranteed error. It might be worth to optimize the various tuning parameters in the algorithm, due to its complexity, however, we prefer to leave this task for future research.

### APPENDIX A: PROOFS OF RESULTS IN SECTION 3

We start by recalling the following algebraic property of the Lévy areas: for each  $0 \leq r < s < t$

$$(A.1) \quad A_{i,j}(r, t) = A_{i,j}(r, s) + A_{i,j}(s, t) + (Z_i(s) - Z_i(r))(Z_j(t) - Z_j(s)).$$

Using this property and a simple use of the Borel–Cantelli lemma we can obtain the proof of Lemma 3.1.

**PROOF OF LEMMA 3.1.** We use (A.1) repeatedly. First, note that

$$\begin{aligned}
 A_{i,j}(t_k^n, t_{k+1}^n) &= A_{i,j}(t_{2k}^{n+1}, t_{2k+1}^{n+1}) + A_{i,j}(t_{2k+1}^{n+1}, t_{2k+2}^{n+1}) \\
 &\quad + (Z_i(t_{2k+1}^{n+1}) - Z_i(t_{2k}^{n+1}))(Z_j(t_{2k+2}^{n+1}) - Z_j(t_{2k+1}^{n+1})).
 \end{aligned}$$

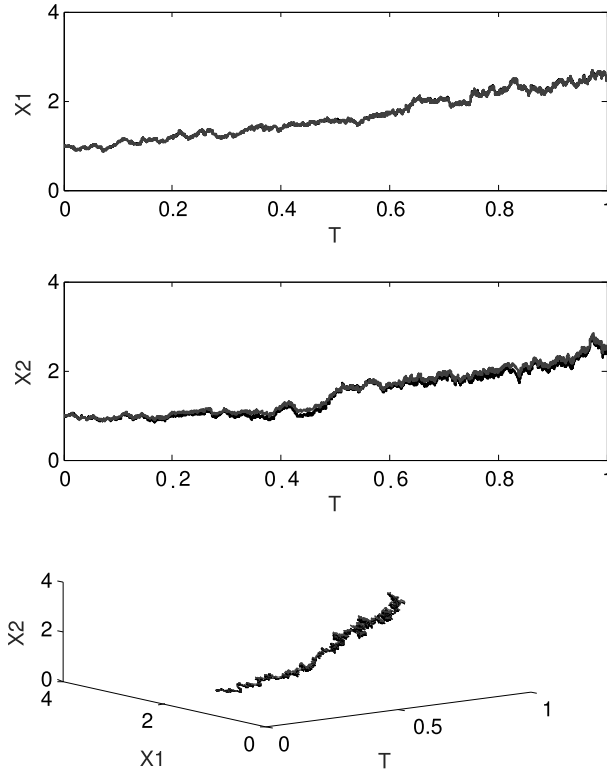


FIG. 2. Simulation of the Geometric Brownian Motion on  $[0, 1]$  [ $X_1(0) = X_2(0) = 1$ ,  $\mu_1 = \mu_2 = 1, \sigma_1 = \sigma_2 = 0.5, \rho = 0.25$ ].

We continue, this time splitting  $A_{i,j}(t_{2k}^{n+1}, t_{2k+1}^{n+1})$  and  $A_{i,j}(t_{2k+1}^{n+1}, t_{2k+2}^{n+1})$ , thereby obtaining

$$\begin{aligned}
 & A_{i,j}(t_k^n, t_{k+1}^n) \\
 &= (Z_i(t_{2k+1}^{n+1}) - Z_i(t_{2k}^{n+1}))(Z_j(t_{2k+2}^{n+1}) - Z_j(t_{2k+1}^{n+1})) \\
 &\quad + A_{i,j}(t_{2^{2k}}^{n+2}, t_{2^{2k+1}}^{n+2}) + A_{i,j}(t_{2^{2k+1}}^{n+2}, t_{2^{2k+2}}^{n+2}) \\
 &\quad + (Z_i(t_{2^{2k+1}}^{n+2}) - Z_i(t_{2^{2k}}^{n+2}))(Z_j(t_{2^{2k+2}}^{n+2}) - Z_j(t_{2^{2k+1}}^{n+2})) \\
 &\quad + A_{i,j}(t_{2^{2k+2}}^{n+2}, t_{2^{2k+3}}^{n+2}) + A_{i,j}(t_{2^{2k+3}}^{n+2}, t_{2^{2k+4}}^{n+2}) \\
 &\quad + (Z_i(t_{2^{2k+3}}^{n+2}) - Z_i(t_{2^{2k+2}}^{n+2}))(Z_j(t_{2^{2k+4}}^{n+2}) - Z_j(t_{2^{2k+3}}^{n+2})).
 \end{aligned}$$

Suppose by iterating the previous splitting procedure  $m$  times, we have

$$\begin{aligned}
 (A.2) \quad A_{i,j}(t_k^n, t_{k+1}^n) &= \sum_{h=n+1}^m \sum_{l=1}^{2^{h-n-1}} [Z_i(t_{2^{h-nk+2l-1}}^h) - Z_i(t_{2^{h-nk+2l-2}}^h)] \\
 &\quad \times [Z_j(t_{2^{h-nk+2l}}^h) - Z_j(t_{2^{h-nk+2l-1}}^h)]
 \end{aligned}$$

$$\begin{aligned}
 & + \sum_{l=1}^{2^{m-n-1}} A_{i,j}(t_{2^{m-n}k+2l-2}^m, t_{2^{m-n}k+2l-1}^m) \\
 & + \sum_{l=1}^{2^{m-n-1}} A_{i,j}(t_{2^{m-n}k+2l-1}^m, t_{2^{m-n}k+2l}^m).
 \end{aligned}$$

Then for the  $(m + 1)$ th iteration, we have

$$\begin{aligned}
 & A_{i,j}(t_k^n, t_{k+1}^n) \\
 & = \sum_{h=n+1}^m \sum_{l=1}^{2^{h-n-1}} [Z_i(t_{2^{h-n}k+2l-1}^h) - Z_i(t_{2^{h-n}k+2l-2}^h)] \\
 & \quad \times [Z_j(t_{2^{h-n}k+2l}^h) - Z_j(t_{2^{h-n}k+2l-1}^h)] \\
 & \quad + \sum_{l=1}^{2^{m-n-1}} \{A_{i,j}(t_{2(2^{m-n}k+2l-2)}^{m+1}, t_{2(2^{m-n}k+2l-2)+1}^{m+1}) \\
 & \quad + A_{i,j}(t_{2(2^{m-n}k+2l-2)+1}^{m+1}, t_{2(2^{m-n}k+2l-2)+2}^{m+1}) \\
 & \quad + [Z_i(t_{2(2^{m-n}k+2l-2)+1}^{m+1}) - Z_i(t_{2(2^{m-n}k+2l-2)}^{m+1})] \\
 & \quad \times [Z_i(t_{2(2^{m-n}k+2l-2)+2}^{m+1}) - Z_i(t_{2(2^{m-n}k+2l-2)+1}^{m+1})]\} \\
 & \quad + \sum_{l=1}^{2^{m-n-1}} \{A_{i,j}(t_{2(2^{m-n}k+2l-1)}^{m+1}, t_{2(2^{m-n}k+2l-1)+1}^{m+1}) \\
 & \quad + A_{i,j}(t_{2(2^{m-n}k+2l-1)+1}^{m+1}, t_{2(2^{m-n}k+2l-1)+2}^{m+1}) \\
 & \quad + [Z_i(t_{2(2^{m-n}k+2l-1)+1}^{m+1}) - Z_i(t_{2(2^{m-n}k+2l-1)}^{m+1})] \\
 & \quad \times [Z_i(t_{2(2^{m-n}k+2l-1)+2}^{m+1}) - Z_i(t_{2(2^{m-n}k+2l-1)+1}^{m+1})]\} \\
 & = \sum_{h=n+1}^{m+1} \sum_{l=1}^{2^{h-n-1}} [Z_i(t_{2^{h-n}k+2l-1}^h) - Z_i(t_{2^{h-n}k+2l-2}^h)] \\
 & \quad \times [Z_j(t_{2^{h-n}k+2l}^h) - Z_j(t_{2^{h-n}k+2l-1}^h)] \\
 & \quad + \sum_{l=1}^{2^{m-n}} A_{i,j}(t_{2^{m+1-n}k+2l-2}^{m+1}, t_{2^{m+1-n}k+2l-1}^{m+1}) \\
 & \quad + \sum_{l=1}^{2^{m+1-n-1}} A_{i,j}(t_{2^{m+1-n}k+2l-1}^{m+1}, t_{2^{m+1-n}k+2l}^{m+1}).
 \end{aligned}$$

Thus, (A.2) holds by induction.

We next claim that

$$(A.3) \quad \begin{aligned} & \sum_{l=1}^{2^{m-n-1}} A_{i,j}(t_{2^{m-n}k+2l-2}^m, t_{2^{m-n}k+2l-1}^m) \\ & + \sum_{l=1}^{2^{m-n-1}} A_{i,j}(t_{2^{m-n}k+2l-1}^m, t_{2^{m-n}k+2l}^m) \rightarrow 0 \end{aligned}$$

almost surely as  $m \rightarrow \infty$ . To see this note that

$$\begin{aligned} & P\left(\left|\sum_{l=1}^{2^{m-n-1}} A_{i,j}(t_{2^{h-n}k+2l-2}^h, t_{2^{h-n}k+2l-1}^h)\right| > 1/m\right) \\ & \leq m^2 \sum_{l=1}^{2^{m-n-1}} E[A_{i,j}^2(t_{2^{m-n}k+2l-2}^m, t_{2^{m-n}k+2l-1}^m)] = m^2 2^{m-n+1} E \int_0^{\Delta_m} Z_i^2(s) ds \\ & = m^2 2^{m-n} \Delta_m^2 = 2^{-n} m^2 \Delta_m. \end{aligned}$$

Since  $\sum_{m=1}^{\infty} m^2 \Delta_m < \infty$ , we conclude by Borel–Cantelli’s lemma that, almost surely, for  $m$  large enough

$$\left|\sum_{l=1}^{2^{m-n-1}} A_{i,j}(t_{2^{h-n}k+2l-2}^h, t_{2^{h-n}k+2l-1}^h)\right| < 1/m.$$

Thus, we have (A.3) holds almost surely and, therefore, from (A.2), by sending  $m \rightarrow \infty$  we obtain the conclusion of the lemma.  $\square$

**PROOF OF LEMMA 3.2.** Let  $N_{i,1} = \max\{n \geq 1 : |W_{i,k}^n| > 4\sqrt{n+1} \text{ for some } 1 \leq k \leq 2^{n-1}\}$ . Then  $N_1 = \max_i\{N_{i,1}\}$ .

$$\begin{aligned} E[N_{i,1}] &= \sum_{n=1}^{\infty} P(N_{i,1} \geq n) \\ &\leq \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} \sum_{k=1}^{2^{m-1}} P(|W_{i,k}^m| > 4\sqrt{m+1}) \\ &\leq \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} 2^{m-1} \exp(-8m) \\ &\leq \sum_{n=1}^{\infty} \frac{\exp(-(8 - \log 2)n)}{1 - \exp(-(8 - \log 2))} \\ &= \frac{\exp(-(8 - \log 2))}{(1 - \exp(-(8 - \log 2)))^2} < \infty. \end{aligned}$$

Thus,  $E[N] < \infty$ . We also notice that  $E[N_1]$  is independent of our choice of  $\alpha$  and  $\beta$ .  $\square$

**PROOF OF LEMMA 3.3.** For any interval  $[t, t + \delta] \subset [0, 1]$ , there exists  $m \in \mathbb{Z}^+$ , such that  $2^{-(m+1)} \leq \delta \leq 2^{-m}$ . We next divide the analysis into two cases.

Case 1. There exist two level  $m$  dyadic points  $t_k^m$  and  $t_{k+1}^m$ , such that  $[t, t + \delta] \subset [t_k^m, t_{k+1}^m]$ .

Case 2. There exist three level  $m$  dyadic points  $t_k^m, t_{k+1}^m$  and  $t_{k+2}^m$ , such that  $t \in [t_k^m, t_{k+1}^m]$  and  $t + \delta \in [t_1^m, t_{k+2}^m]$ .

In Case 1, using the Lévy–Ciesielski construction, we have

$$|Z(t + \delta) - Z(t)| \leq 2^{-m} V^0 + \sum_{n=1}^m 2^{-m+\frac{n-1}{2}} V^n + \sum_{n=m+1}^{\infty} 2^{-\frac{n+1}{2}} V^n.$$

Since  $\delta \geq 2^{-(m+1)}$ , we have

$$\begin{aligned} \frac{|Z(t + \delta) - Z(t)|}{\delta^\alpha} &\leq 2^{2\alpha} \left( 2^{-(1-\alpha)m-\alpha} V^0 + \sum_{n=1}^m 2^{-(1-\alpha)m-\alpha+\frac{n-1}{2}} V^n \right. \\ &\quad \left. + \sum_{n=m+1}^{\infty} 2^{-\frac{n+1}{2}+\alpha(m+1)} V^n \right) \\ &\leq 2^{2\alpha} \left( \sum_{n=0}^m 2^{-(1-\alpha)n+\frac{n}{2}} V^n + \sum_{n=m+1}^{\infty} 2^{-\frac{n}{2}+\alpha n} V^n \right) \\ &\leq 2^{2\alpha} \sum_{n=0}^{\infty} 2^{-n(\frac{1}{2}-\alpha)} V^n. \end{aligned}$$

Similar to Case 1, in Case 2 we have

$$\begin{aligned} |Z(t + \delta) - Z(t)| &\leq |Z(t_{k+1}^m) - Z(t)| + |Z(t + \delta) - Z(t_{k+1}^m)| \\ &\leq 2 \left( V^0 + \sum_{n=1}^m 2^{-m+\frac{n-1}{2}} V^n + \sum_{n=m+1}^{\infty} 2^{-\frac{n+1}{2}} V^n \right). \end{aligned}$$

Then

$$\frac{|Z(t + \delta) - Z(t)|}{\delta^\alpha} \leq 2^{2\alpha+1} \sum_{n=0}^{\infty} 2^{-n(\frac{1}{2}-\alpha)} V^n.$$

As the interval  $[t, t + \delta]$  is arbitrarily chosen, we obtain the result.  $\square$

PROOF OF LEMMA 3.4. For  $i \neq j$ , let  $N_{i,j,2} = \max\{n : |L_{i,j}^n(m) - L_{i,j}^n(l)| > (m-l)^\beta \Delta_n^{2\alpha}$  for some  $0 \leq l < m \leq 2^{n-1}\}$ . Then  $N_2 = \max_{1 \leq i, j \leq d', i \neq j} \{N_{i,j,2}\}$ .

Fix any  $(i, j)$  pair, define

$$C_n = \{|L_{i,j}^n(m) - L_{i,j}^n(l)| > (m-l)^\beta \Delta_n^{2\alpha} \text{ for some } 0 \leq l < m \leq 2^{n-1}\}.$$

We will show that the events  $\{C_n : n \geq 0\}$  occur finitely many times. Note that

$$(A.4) \quad P(C_n) \leq \sum_{0 \leq l < m \leq 2^{n-1}} 2P((L_{i,j}^n(m) - L_{i,j}^n(l)) > (m-l)^\beta \Delta_n^{2\alpha}).$$

Also observe that for fixed  $m$  and  $n$ ,  $L_{i,j}^n(m)$  is the sum of  $m$  i.i.d. random variables, each of which is distributed as  $(Z_i(t_1^n) - Z_i(t_0^n))(Z_j(t_2^n) - Z_j(t_1^n))$  and

$$E[\exp(\theta(Z_i(t_1^n) - Z_i(t_0^n))(Z_j(t_2^n) - Z_j(t_1^n)))] = (1 - \theta^2 \Delta_n^2)^{-1/2}.$$

We apply Chernoff's bound and have

$$\begin{aligned} P((L_{i,j}^n(m) - L_{i,j}^n(l)) > (m-l)^\beta \Delta_n^{2\alpha}) \\ \leq \exp\left(-\theta(m-l)^\beta \Delta_n^{2\alpha} - \frac{1}{2}(m-l) \log(1 - \theta^2 \Delta_n^2)\right). \end{aligned}$$

Select  $\theta = \theta'(m-l)^{-1/2} \Delta_n^{-1}$  for  $\theta' \in (0, 1/4)$

$$P((L_{i,j}^n(m) - L_{i,j}^n(l)) > (m-l)^\beta \Delta_n^{2\alpha}) \leq \exp(-\theta'(m-l)^{\beta-1/2} \Delta_n^{2\alpha-1} + 1).$$

Hence,

$$(A.5) \quad \begin{aligned} P(C_n) &\leq \sum_{0 \leq l < m \leq 2^{n-1}} 2 \exp(-\theta'(m-l)^{\beta-1/2} \Delta_n^{2\alpha-1} + 1) \\ &\leq 2^{2n} \exp(-\theta' 2^{n(1-2\alpha)}). \end{aligned}$$

We notice that  $2\alpha < 1$

$$E[N_{i,j,2}] = \sum_{n=1}^{\infty} P(N_{i,j,2} \geq n)$$

and

$$P(N_{i,j,2} \geq n) \leq \sum_{m=n}^{\infty} P(C_m).$$

From (A.5), we denote

$$M := \min\{n : 2^{2n} \exp(-\theta' 2^{n(1-2\alpha)}) < 1/4\}.$$

Then  $M = o((1 - 2\alpha)^{-2})$ . We also notice that for  $n \geq M$ ,

$$2^{2(n+1)} \exp(-\theta' 2^{(n+1)(1-2\alpha)}) < (2^{2n} \exp(-\theta' 2^{n(1-2\alpha)}))^2.$$

Thus,  $\sum_{m=M+k}^{\infty} P(\mathcal{C}_m) \leq (1/2)^k$  and

$$E[N_{i,j,2}] \leq M - 1 + \sum_{n=M}^{\infty} \sum_{m=n}^{\infty} P(\mathcal{C}_m) \leq M.$$

Thus,  $E[N_2] = o((1 - 2\alpha)^{-2})$ .  $\square$

The proof of Corollary 3.1 follows directly from Lemmas 3.1 and 3.4.

PROOF OF COROLLARY 3.1. Using Lemma 3.1, we obtain that

$$(A.6) \quad R_{i,j}^n(t_l^n, t_m^n) = \sum_{k=l+1}^m \sum_{h=n+1}^{\infty} (L_{i,j}^h(2^{h-n}(k+1)) - L_{i,j}^h(2^{h-n}k)).$$

On the other hand, due to Lemma 3.4 if  $n \geq N_2$

$$\begin{aligned} & \sum_{k=l+1}^m \sum_{h=n+1}^{\infty} |L_{i,j}^h(2^{h-n}(k+1)) - L_{i,j}^h(2^{h-n}k)| \\ & \leq \sum_{k=l+1}^m \sum_{h=n+1}^{\infty} (2^{-n}(k+1) - 2^{-n}k)^{\beta} \Delta_h^{2\alpha-\beta} < \infty \end{aligned}$$

because  $\beta < 2\alpha$ . Thus (by Fubini's theorem) the order of the summations in (A.6) can be exchanged and we obtain the result.  $\square$

PROOF OF LEMMA 3.5. We start by showing the bound on  $\Gamma_R$ . By the definition of  $\Gamma_L$ , for any  $n$

$$|L_{i,j}^n(m) - L_{i,j}^n(l)| \leq \Gamma_L(m-l)^{\beta} \Delta_n^{2\alpha}.$$

Consequently, for any  $0 \leq l < m \leq 2^{n-1}$ ,

$$\begin{aligned} |R_{i,j}^n(t_l^n, t_m^n)| & \leq \sum_{h=n+1}^{\infty} |L_{i,j}^h(2^{h-n}m) - L_{i,j}^h(2^{h-n}l)| \\ & \leq \sum_{h=n+1}^{\infty} \Gamma_L(m-l)^{\beta} 2^{(h-n)\beta} \Delta_h^{2\alpha} = \Gamma_L(m-l)^{\beta} \Delta_n^{\beta} \sum_{h=n+1}^{\infty} \Delta_h^{2\alpha-\beta} \\ & = \Gamma_L(t_m^n - t_l^n)^{\beta} \Delta_n^{2\alpha-\beta} \frac{2^{-(2\alpha-\beta)}}{1 - 2^{-(2\alpha-\beta)}}. \end{aligned}$$

Therefore, we conclude that

$$\begin{aligned} \Gamma_R & := \max_{1 \leq i, j \leq d'} \sup_{n \geq 0} \sup_{0 \leq s < t \leq 1, s, t \in D_n} \frac{|R_{i,j}^n(s, t)|}{|t - s|^{\beta} \Delta_n^{2\alpha-\beta}} \\ & \leq \Gamma_L \frac{2^{-(2\alpha-\beta)}}{1 - 2^{-(2\alpha-\beta)}}. \end{aligned}$$

Let  $r(n, l, m) = \min\{h : |t_m^n - t_l^n| \geq \Delta_h\}$ . For simplicity of notation, we define the following sequence of operators of time:

$$\begin{aligned} \underline{s}^h(t_l^n) &= \min\{t_k^h : t_k^h \geq t_l^n\}, \\ \bar{s}^h(t_m^n) &= \max\{t_k^h : t_k^h \leq t_m^n\} \end{aligned}$$

for  $r(n, l, m) \leq h \leq n$ .

Then

$$\begin{aligned} &|A_{i,j}(t_l^n, t_m^n)| \\ &\leq |A_{i,j}(t_l^n, \underline{s}^{n-1}(t_l^n))| + |A_{i,j}(\underline{s}^{n-1}(t_l^n), \bar{s}^{n-1}(t_m^n))| + |A_{i,j}(\bar{s}^{n-1}(t_m^n), t_m^n)| \\ &\quad + |Z_i(\underline{s}^{n-1}(t_l^n)) - Z_i(t_l^n)| |Z_j(\bar{s}^{n-1}(t_m^n)) - Z_j(\underline{s}^{n-1}(t_l^n))| \\ &\quad + |Z_i(\bar{s}^{n-1}(t_m^n)) - Z_i(t_m^n)| |Z_j(t_m^n) - Z_j(\bar{s}^{n-1}(t_m^n))|. \end{aligned}$$

Suppose by iterating the above procedure up to level  $\gamma$ , where  $r(n, l, m) < \gamma < n$ , we have

$$\begin{aligned} &|A_{i,j}(t_l^n, t_m^n)| \\ &\leq \sum_{h=\gamma+1}^n |A_{i,j}(\underline{s}^h(t_l^n), \underline{s}^{h-1}(t_l^n))| + |A_{i,j}(\underline{s}^\gamma(t_l^n), \bar{s}^\gamma(t_m^n))| \\ &\quad + \sum_{h=\gamma+1}^n |A_{i,j}(\bar{s}^h(t_m^n), \bar{s}^{h-1}(t_m^n))| \\ &\quad + \sum_{h=\gamma+1}^n |Z_i(\underline{s}^h(t_l^n)) - Z_i(\underline{s}^{h-1}(t_l^n))| |Z_j(\bar{s}^{h-1}(t_m^n)) - Z_j(\underline{s}^{h-1}(t_l^n))| \\ &\quad + \sum_{h=\gamma+1}^n |Z_i(\bar{s}^{h-1}(t_m^n)) - Z_i(\underline{s}^h(t_l^n))| |Z_j(\bar{s}^h(t_m^n)) - Z_j(\bar{s}^{h-1}(t_m^n))|. \end{aligned}$$

Then for level  $\gamma - 1$ , as  $\underline{s}^{h-1}(\underline{s}^h(t_l^n)) = \underline{s}^{h-1}(t_l^n)$  and  $\bar{s}^{h-1}(\bar{s}^h(t_m^n)) = \bar{s}^{h-1}(t_m^n)$  for  $h < n$ , we have

$$\begin{aligned} &|A_{i,j}(t_l^n, t_m^n)| \\ &\leq \sum_{h=\gamma+1}^n |A_{i,j}(\underline{s}^h(t_l^n), \underline{s}^{h-1}(t_l^n))| \\ &\quad + |A_{i,j}(\underline{s}^\gamma(t_l^n), \underline{s}^{\gamma-1}(\underline{s}^\gamma(t_l^n)))| \\ &\quad + |A_{i,j}(\underline{s}^{\gamma-1}(\underline{s}^\gamma(t_l^n)), \bar{s}^{\gamma-1}(\bar{s}^\gamma(t_m^n)))| + |A_{i,j}(\bar{s}^{\gamma-1}(\bar{s}^\gamma(t_m^n)), \bar{s}^\gamma(t_m^n))| \\ &\quad + |Z_i(\underline{s}^{\gamma-1}(\underline{s}^\gamma(t_l^n))) - Z_i(\underline{s}^\gamma(t_l^n))| \end{aligned}$$



$$\begin{aligned}
 & \times |Z_j(\bar{s}^{\gamma-1}(\bar{s}^\gamma(t_m^n))) - Z_j(\underline{s}^{\gamma-1}(\underline{s}^\gamma(t_l^n)))| \\
 & + |Z_i(\bar{s}^{\gamma-1}(\bar{s}^\gamma(t_m^n))) - Z_i(\underline{s}^\gamma(t_l^n))| |Z_j(\bar{s}^\gamma(t_m^n)) - Z_j(\bar{s}^{\gamma-1}(\bar{s}^\gamma(t_m^n)))| \\
 & + \sum_{h=r(n,l,m)+1}^n |A_{i,j}(\bar{s}^h(t_m^n), \bar{s}^{h-1}(t_m^n))| \\
 = & \sum_{h=\gamma}^n |A_{i,j}(\underline{s}^h(t_l^n), \underline{s}^{h-1}(t_l^n))| + |A_{i,j}(\underline{s}^{\gamma-1}(t_l^n), \bar{s}^{\gamma-1}(t_m^n))| \\
 & + \sum_{h=\gamma}^n |A_{i,j}(\bar{s}^h(t_m^n), \bar{s}^{h-1}(t_m^n))| \\
 & + \sum_{h=\gamma}^n |Z_i(\underline{s}^{h-1}(t_l^n)) - Z_i(\underline{s}^h(t_l^n))| |Z_j(\bar{s}^{h-1}(t_m^n)) - Z_j(\underline{s}^{h-1}(t_l^n))| \\
 & + \sum_{h=\gamma}^n |Z_i(\bar{s}^{h-1}(t_m^n)) - Z_i(\underline{s}^h(t_l^n))| |Z_j(\bar{s}^h(t_m^n)) - Z_j(\bar{s}^{h-1}(t_m^n))|.
 \end{aligned}$$

Thus, the following inequality holds by induction:

$$\begin{aligned}
 & |A_{i,j}(t_l^n, t_m^n)| \\
 \leq & \sum_{h=r(n,l,m)+1}^n |A_{i,j}(\underline{s}^h(t_l^n), \underline{s}^{h-1}(t_l^n))| + |A_{i,j}(\underline{s}^{r(n,l,m)}(t_l^n), \bar{s}^{r(n,l,m)}(t_m^n))| \\
 & + \sum_{h=r(n,l,m)+1}^n |A_{i,j}(\bar{s}^h(t_m^n), \bar{s}^{h-1}(t_m^n))| \\
 & + \sum_{h=r(n,l,m)+1}^n |Z_i(\underline{s}^{h-1}(t_l^n)) - Z_i(\underline{s}^h(t_l^n))| |Z_j(\bar{s}^{h-1}(t_m^n)) - Z_j(\underline{s}^{h-1}(t_l^n))| \\
 & + \sum_{h=r(n,l,m)+1}^n |Z_i(\bar{s}^{h-1}(t_m^n)) - Z_i(\underline{s}^h(t_l^n))| |Z_j(\bar{s}^h(t_m^n)) - Z_j(\bar{s}^{h-1}(t_m^n))|.
 \end{aligned}$$

We make the following important observations:

$$\begin{aligned}
 \underline{s}^{h-1}(t_l^n) - \underline{s}^h(t_l^n) &= \begin{cases} 0, & \text{if } \underline{s}^{h-1}(t_l^n) = \underline{s}^h(t_l^n), \\ \Delta_h, & \text{otherwise,} \end{cases} \\
 \bar{s}^h(t_m^n) - \bar{s}^{h-1}(t_m^n) &= \begin{cases} 0, & \text{if } \bar{s}^{h-1}(t_m^n) = \bar{s}^h(t_m^n), \\ \Delta_h, & \text{otherwise,} \end{cases} \\
 \bar{s}^{r(n,l,m)}(t_m^n) - \underline{s}^{r(n,l,m)}(t_l^n) &= \begin{cases} 0, & \text{if } \underline{s}^{r(n,l,m)}(t_l^n) = \bar{s}^{r(n,l,m)}(t_m^n), \\ \Delta_{r(n,l,m)}, & \text{otherwise.} \end{cases}
 \end{aligned}$$

Then

$$\begin{aligned} & \frac{|A_{i,j}(t_l^n, t_m^n)|}{(t_m^n - t_l^n)^{2\alpha}} \\ & \leq \sum_{h=r+1}^n \Gamma_R \frac{\Delta_h^{2\alpha}}{\Delta_{r(n,l,m)}^{2\alpha}} + \Gamma_R + \sum_{h=r+1}^n \Gamma_R \frac{\Delta_h^{2\alpha}}{\Delta_{r(n,l,m)}^{2\alpha}} \\ & \quad + \sum_{h=r+1}^n \|Z\|_\alpha^2 \frac{\Delta_h^\alpha}{\Delta_{r(n,l,m)}^\alpha} + \sum_{h=r+1}^n \|Z\|_\alpha^2 \frac{\Delta_h^\alpha}{\Delta_r^\alpha} \\ & \leq \Gamma_R \frac{2}{1 - 2^{-2\alpha}} + \|Z\|_\alpha^2 \frac{2^{1-\alpha}}{1 - 2^{-\alpha}}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|A\|_{2\alpha} & := \max_{1 \leq i \leq j \leq d'} \sup_{n \geq 1} \sup_{0 \leq s < t \leq 1; s, t \in D_n} \frac{|A_{i,j}(s)|}{|t - s|^{2\alpha}} \\ & \leq \Gamma_R \frac{2}{1 - 2^{-2\alpha}} + \|Z\|_\alpha^2 \frac{2^{1-\alpha}}{1 - 2^{-\alpha}}. \quad \square \end{aligned}$$

APPENDIX B: PROOFS OF RESULTS IN SECTION 5

**B.1. Proof of results in Section 5.1.**

PROOFS OF LEMMA 5.1. We first notice that  $t_{2k}^n = t_k^{n-1}$  for  $k = 0, 1, 2, \dots, 2^{n-1}$ . From the Lévy–Ciesielski construction, we have

$$Z_i(t_{2k-1}^n) = \frac{1}{2}(Z_i(t_{k-1}^{n-1}) + Z_i(t_k^{n-1})) + \Delta_{n+1}^{1/2} W_{i,k}^n.$$

Then

$$\Lambda_i^n(t_{2k-1}^n) = Z_i(t_{2k-1}^n) - Z_i(t_{k-1}^{n-1}) = \frac{1}{2}(Z_i(t_{k-1}^{n-1}) - Z_i(t_k^{n-1})) + \Delta_{n+1}^{1/2} W_{i,k}^n$$

and

$$\Lambda_i^n(t_{2k}^n) = Z_i(t_k^{n-1}) - Z_i(t_{2k-1}^n) = \frac{1}{2}(Z_i(t_{k-1}^{n-1}) - Z_i(t_k^{n-1})) - \Delta_{n+1}^{1/2} W_{i,k}^n. \quad \square$$

Before we prove Corollary 5.1, we first provide the following auxiliary result which summarizes basic computations of moment generating functions of quadratic forms of bivariate Gaussian random variables.

LEMMA B.1. *Suppose that  $Y$  and  $Z$  are i.i.d.  $N(0, 1)$  random variables, then for any numbers  $a_1, a_2, b, c_1, c_2 \in \mathbb{R}$  define*

$$\phi(a, b, c) := E \exp(a_1 Y + a_2 Z + b Y Z + c_1 Y^2 + c_2 Z^2),$$

then we have that if  $|2c_i| < 1$  for  $i = 1, 2$ , and  $|b| < (1 - 2c_1)(1 - 2c_2)$

$$\begin{aligned} \phi(a, b, c) &= (1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2}(1 - (b(1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2})^2)^{-1/2} \\ &\quad \times \exp\left(\frac{a_1^2(1 - 2c_1)^{-1} + a_2^2(1 - 2c_2)^{-1} + 2a_1a_2b(1 - 2c_1)^{-1}(1 - 2c_2)^{-1}}{2(1 - b^2(1 - 2c_1)^{-1}(1 - 2c_2)^{-1})}\right). \end{aligned}$$

Moreover, if we let

$$P'(Y \in dy, Z \in dz) = P(Y \in dy, Z \in dz) \frac{\exp(a_1y + a_2z + byz + c_1y^2 + c_2z^2)}{\phi(\theta; a, b, c)},$$

then under  $P'(\cdot)$  we have that  $(Y, Z)$  are distributed bivariate Gaussian with covariance matrix

$$\begin{aligned} \Sigma(a, b, c) &= \frac{1}{1 - b^2(1 - 2c_1)^{-1}(1 - 2c_2)^{-1}} \\ &\quad \times \begin{pmatrix} (1 - 2c_1)^{-1} & b(1 - 2c_1)^{-1}(1 - 2c_2)^{-1} \\ b(1 - 2c_1)^{-1}(1 - 2c_2)^{-1} & (1 - 2c_2)^{-1} \end{pmatrix}, \end{aligned}$$

and mean vector

$$\mu(a, b, c) = \Sigma(a, b, c) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

PROOF. First, it follows easily that  $E \exp(c_1Y^2 + c_2Z^2) = (1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2}$ , and under the probability measure

$$P_1(Y \in dy, Z \in dz) = \frac{\exp(c_1y^2 + c_2z^2)}{E[\exp(c_1Y^2 + c_2Z^2)]} P(Y \in dy)P(Z \in dz),$$

$Y$  and  $Z$  are independent with distributions  $N(1, (1 - 2c_1)^{-1})$  and  $N(1, (1 - 2c_2)^{-1})$ , respectively. Therefore,

$$\begin{aligned} \phi(a, b, c) &= (1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2} E_1 \exp(a_1Y + a_2Z + bYZ) \\ &= (1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2} \\ &\quad \times E[\exp\{a_1Y(1 - 2c_1)^{-1/2} + a_2Z(1 - 2c_2)^{-1/2} \\ &\quad + b(1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2}YZ\}]. \end{aligned}$$

Now, given  $|\theta| < 1$  define  $P_2(\cdot)$  via

$$P_2(Y \in dy, Z \in dz) = \frac{P(Y \in dy, Z \in dz) \exp(\chi yz)}{E[\exp(\chi YZ)]}.$$

Observe that

$$P(Y \in dy, Z \in dz) \exp(\chi yz) = \frac{1}{2\pi} \exp(-y^2/2 - z^2/2 + \chi yz)$$

and

$$-y^2/2 - z^2/2 + \chi yz = -(y, z) \Sigma^{-1} \begin{pmatrix} y \\ z \end{pmatrix} / 2,$$

where

$$\Sigma^{-1} = \begin{pmatrix} 1 & -\chi \\ -\chi & 1 \end{pmatrix},$$

and thus

$$\Sigma = \frac{1}{1 - \chi^2} \begin{pmatrix} 1 & \chi \\ \chi & 1 \end{pmatrix}.$$

Therefore, under  $P_2(\cdot)$ ,  $(Y, Z)$  is distributed bivariate normal with mean zero and covariance matrix  $\Sigma$ , with

$$\chi = b(1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2}$$

and we also must have that if  $|\chi| < 1$ ,

$$E[\exp(\phi YZ)] = (1 - \chi^2)^{-1/2} = (1 - (b(1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2})^2)^{-1/2}.$$

Consequently, we conclude that

$$\begin{aligned} \phi(a, b, c) &= (1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2} (1 - (b(1 - 2c_1)^{-1/2}(1 - 2c_2)^{-1/2})^2)^{-1/2} \\ &\quad \times E_2[\exp(a_1 Y(1 - 2c_1)^{-1/2} + a_2 Z(1 - 2c_2)^{-1/2})]. \end{aligned}$$

The final expression for  $\phi(a, b, c)$  is obtained from the fact that

$$\begin{aligned} E_2[\exp(a_1 Y(1 - 2c_1)^{-1/2} + a_2 Z(1 - 2c_2)^{-1/2})] \\ = \exp(\text{Var}_2(a_1 Y(1 - 2c_1)^{-1/2} + a_2 Z(1 - 2c_2)^{-1/2})/2). \end{aligned}$$

And  $P'(\cdot)$  is equivalent to a standard exponentially tilting to the measure  $P_2(\cdot)$  using as the natural parameter the vector

$$(a_1(1 - 2c_1)^{-1/2}, a_2(1 - 2c_2)^{-1/2}),$$

and thus under  $P'(\cdot)$  the covariance matrix is the same as under  $P_2(\cdot)$  and the mean vector is equal to  $\mu(a, b, c)$ .  $\square$

We now are ready to provide the proof of Corollary 5.1.

PROOF OF COROLLARY 5.1. Let us examine the term of the form

$$\Lambda_i^{n+m}(t_{2r-1}^{n+m})\Lambda_j(t_{2r}^{n+m}),$$

for  $i \neq j$ ,

$$\begin{aligned} & \Lambda_i^{n+m}(t_{2r-1}^{n+m})\Lambda_j(t_{2r}^{n+m}) \\ &= (\Lambda_i^{n+m-1}(t_r^{n+m-1})/2 + \Delta_{n+m+1}^{1/2} W_{i,r}^{n+m}) \\ & \quad \times (\Lambda_j^{n+m-1}(t_r^{n+m-1})/2 - \Delta_{n+m+1}^{1/2} W_{j,r}^{n+m}) \\ &= \Lambda_i^{n+m-1}(t_r^{n+m-1})\Lambda_j^{n+m-1}(t_r^{n+m-1})/4 - \Delta_{n+m+1} W_{i,r}^{n+m} W_{j,r}^{n+m} \\ & \quad + \Delta_{n+m+1}^{1/2} W_{i,r}^{n+m} \Lambda_j^{n+m-1}(t_r^{n+m-1})/2 \\ & \quad - \Delta_{n+m+1}^{1/2} W_{j,r}^{n+m} \Lambda_i^{n+m-1}(t_r^{n+m-1})/2. \end{aligned}$$

Then we have that Corollary 5.1 follows immediately from Lemma B.1.  $\square$

Finally, we provide the proof of Corollary 5.2.

PROOF OF COROLLARY 5.2. Recall that for each  $r \in \{1, 2, \dots, 2^n\}$ ,

$$\Lambda_i^n(t_r^n) := (Z_i(t_r^n) - Z_i(t_{r-1}^n)).$$

So

$$\begin{aligned} \Lambda_i^n(t_{2r-1}^n) &= \Lambda_i^n(t_r^{n-1})/2 + \Delta_{n+1}^{1/2} W_{i,r}^n, \\ \Lambda_i^n(t_{2r}^n) &= \Lambda_i^n(t_r^{n-1})/2 - \Delta_{n+1}^{1/2} W_{i,r}^n. \end{aligned}$$

We perform the first iteration in full detail; the rest are immediate just adjusting the notation. From Corollary 5.1, we obtain that, for  $i \neq j$ ,

$$\begin{aligned} & E_{n+m-1} \exp(\theta_0[L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)]) \\ &= \exp\left(\frac{1}{2} \sum_{r=k+1}^{k'} \frac{\theta_0^2 \Delta_{n+m+1}}{4(1 - \theta_0^2 \Delta_{n+m+1}^2)} \Lambda_i(t_r^{n+m-1})^2\right. \\ & \quad \left. + \frac{1}{2} \sum_{r=k+1}^{k'} \frac{\theta_0^2 \Delta_{n+m+1}}{4(1 - \theta_0^2 \Delta_{n+m}^2)} \Lambda_j(t_r^{n+m-1})^2\right) \\ & \quad \times \exp\left(\sum_{r=k+1}^{k'} \frac{\theta_0 \Delta_{n+m+1}}{4(1 - \theta_0^2 \Delta_{n+m+1}^2)} \Lambda_i(t_r^{n+m-1})\Lambda_j(t_r^{n+m-1})\right) \\ & \quad \times (1 - \theta_0^2 \Delta_{n+m+1}^2)^{-(k'-k)/2}. \end{aligned}$$

Using the definitions in (5.5), we have that the exponential component

$$\begin{aligned} & \frac{1}{2} \sum_{r=k+1}^{k'} \frac{\theta_0^2 \Delta_{n+m+1}}{4(1 - \theta_0^2 \Delta_{n+m+1}^2)} \Lambda_i(t_r^{n+m-1})^2 \\ & + \frac{1}{2} \sum_{r=k+1}^{k'} \frac{\theta_0^2 \Delta_{n+m+1}}{4(1 - \theta_0^2 \Delta_{n+m+1}^2)} \Lambda_j(t_r^{n+m-1})^2 \\ & + \sum_{r=k+1}^{k'} \frac{\theta_0 \Delta_{n+m+1}}{4(1 - \theta_0^2 \Delta_{n+m+1}^2)} \Lambda_i(t_r^{n+m-1}) \Lambda_j(t_r^{n+m-1}) \end{aligned}$$

is equal to

$$\begin{aligned} & \sum_{r=1}^{2^{n+m-2}} [\eta_1(t_{2r-1}^{n+m-1}) \Lambda_i(t_{2r-1}^{n+m-1})^2 + \eta_1(t_{2r}^{n+m-1}) \Lambda_i(t_{2r}^{n+m-1})^2] \\ & + \sum_{r=1}^{2^{n+m-2}} [\eta_1(t_{2r-1}^{n+m-1}) \Lambda_j(t_{2r-1}^{n+m-1})^2 + \eta_1(t_{2r}^{n+m-1}) \Lambda_j(t_{2r}^{n+m-1})^2] \\ & + \sum_{r=1}^{2^{n+m-2}} [\theta_1(t_{2r-1}^{n+m-1}) \Lambda_i(t_{2r-1}^{n+m-1}) \Lambda_j(t_{2r-1}^{n+m-1}) \\ & + \theta_1(t_{2r}^{n+m-1}) \Lambda_i(t_{2r}^{n+m-1}) \Lambda_j(t_{2r}^{n+m-1})]. \end{aligned}$$

We next expand each of the terms. To simplify the notation, we write

$$x = W_{i,r}^{n+m-1} \quad \text{and} \quad y = W_{j,r}^{n+m-1}.$$

Define  $\sqrt{\Delta} = \Delta_{n+m}^{1/2}$ , put  $u = \Lambda_i(t_r^{n+m-2})$  and  $v = \Lambda_j(t_r^{n+m-2})$

$$\begin{aligned} \Lambda_i(t_{2r-1}^{n+m-1}) &= u/2 + \sqrt{\Delta}x, & \Lambda_i(t_{2r}^{n+m-1}) &= u/2 - \sqrt{\Delta}x, \\ \Lambda_j(t_{2r-1}^{n+m-1}) &= v/2 + \sqrt{\Delta}y, & \Lambda_j(t_{2r}^{n+m-1}) &= v/2 - \sqrt{\Delta}y. \end{aligned}$$

Now, for brevity let us write  $\eta_o = \eta_1(t_{2r-1}^{n+m-1})$  and  $\eta_e = \eta_1(t_{2r}^{n+m-1})$  (“o” is used for odd, and “e” for even)

$$\begin{aligned} & (\eta_1(t_{2r-1}^{n+m-1}) \Lambda_i(t_{2r-1}^{n+m-1})^2 + \eta_1(t_{2r}^{n+m-1}) \Lambda_i(t_{2r}^{n+m-1})^2 \\ & + \eta_1(t_{2r-1}^{n+m-1}) \Lambda_j(t_{2r-1}^{n+m-1})^2 + \eta_1(t_{2r}^{n+m-1}) \Lambda_j(t_{2r}^{n+m-1})^2) \\ & = (\eta_o(u/2 + \sqrt{\Delta}x)^2 + \eta_e(u/2 - \sqrt{\Delta}x)^2 \\ & \quad + \eta_o(v/2 + \sqrt{\Delta}y)^2 + \eta_e(v/2 - \sqrt{\Delta}y)^2) \\ & = \frac{1}{4}u^2(\eta_e + \eta_o) + \frac{1}{4}v^2(\eta_e + \eta_o) + u(\eta_o - \eta_e)\sqrt{\Delta}x + v(\eta_o - \eta_e)\sqrt{\Delta}y \\ & \quad + (\eta_e + \Delta\eta_o)\Delta x^2 + (\eta_e + \eta_o)\Delta y^2. \end{aligned}$$

Likewise, put  $\theta_o = \theta_1(t_{2r-1}^{n+m-1})$  and  $\theta_e = \theta_1(t_{2r}^{n+m-1})$

$$\begin{aligned} &\theta_1(t_{2r-1}^{n+m-1})\Lambda_i(t_{2r-1}^{n+m-1})\Lambda_j(t_{2r-1}^{n+m-1}) + \theta_1(t_{2r}^{n+m-1})\Lambda_i(t_{2r}^{n+m-1})\Lambda_j(t_{2r}^{n+m-1}) \\ &= \theta_o(u/2 + \sqrt{\Delta x})(v/2 + \sqrt{\Delta y}) + \theta_e(u/2 - \sqrt{\Delta x})(v/2 - \sqrt{\Delta y}) \\ &= \frac{1}{4}uv(\theta_e + \theta_o) + (\theta_e + \theta_o)\Delta xy + \frac{1}{2}v(\theta_o - \theta_e)\sqrt{\Delta x} + \frac{1}{2}u(\theta_o - \theta_e)\sqrt{\Delta y}. \end{aligned}$$

We then collect the terms free of  $x$  and  $y$  and obtain

$$\frac{u^2}{4}(\eta_e + \eta_o) + \frac{v^2}{4}(\eta_e + \eta_o) + \frac{uv}{4}(\theta_e + \theta_o).$$

Now the coefficients of  $x, y, x^2, y^2$  and  $xy$

$$\begin{aligned} &\left\{u(\eta_o - \eta_e) + \frac{1}{2}v(\theta_o - \theta_e)\right\}\sqrt{\Delta x} + \left\{v(\eta_o - \eta_e) + \frac{1}{2}u(\theta_o - \theta_e)\right\}\sqrt{\Delta y} \\ &+ (\eta_e + \eta_o)\Delta x^2 + (\eta_e + \eta_o)\Delta y^2 \\ &+ (\theta_e + \theta_o)\Delta xy. \end{aligned}$$

And finally we can apply Lemma B.1 to get the corresponding results.  $\square$

**B.2. Proofs of results in Section 5.2.**

PROOF OF LEMMA 5.2. Recalling expression (5.6), we establish the bound for  $E_n[\exp(\theta_0\{L_{i,j}^{n+1}(k') - L_{i,j}^{n+m}(k)\})]$ , for  $i \neq j$ , by controlling the contribution of the term

$$(B.1) \quad \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}),$$

and the exponential term

$$(B.2) \quad \exp\left(\sum_{r=1}^{2^n} \theta_m(t_r^n)\Lambda_i(t_r^n)\Lambda_j(t_r^n) + \sum_{r=1}^{2^n} \eta_m(t_r^n)(\Lambda_i(t_r^n)^2 + \Lambda_j(t_r^n)^2)\right)$$

separately.

We start by analyzing  $\theta_l$  and  $\eta_l$ . From Corollary 5.1, we have

$$\theta_1 = \frac{\theta_0}{4(1 - \theta_0^2\Delta_{n+m+1}^2)} \quad \text{and} \quad \eta_1 = \frac{\theta_0^2\Delta_{n+m+1}}{8(1 - \theta_0^2\Delta_{n+m+1}^2)}.$$

We notice that  $2\eta_1 \leq \theta_1^2\Delta_{n+m+1} \leq (5/2)\eta_1$ .

Let

$$u = \max\{h : k' - k > 2^h\}.$$

We also denote

$$\underline{b}(l) := \min\{r : \theta_l(t_r^{n+m-l}) > 0\}$$

and

$$\bar{b}(l) := \max\{r : \theta_l(t_r^{n+m-l}) > 0\}.$$

The strategy throughout the rest of the proof proceeds as follows. We have that the  $\theta_l(t_r^{n+m-l})$ 's and  $\eta_l(t_r^{n+m-l})$ 's,  $r = 1, 2, \dots, 2^{n+m-l}$ , are nonnegative. We also have that for  $l \leq u \wedge m$ , the number of positive  $\theta_l(t_r^{n+m-l})$ 's and  $\eta_l(t_r^{n+m-l})$ 's reduces by about a half at each step  $l$  and also the actual value of the positive  $\theta_l(t_r^{n+m-l})$ 's and  $\eta_l(t_r^{n+m-l})$ 's shrinks by at least  $1/2$ . We will establish that if  $m > u$ , for  $u < l \leq m$ , there are at most two positive  $\theta_l(t_r^{n+m-l})$ 's and two positive  $\eta_l(t_r^{n+m-l})$ 's and at each step  $l$ , their values shrink by more than  $2^{-3/2}$ . Using these observations we will establish some facts and then use them to estimate (B.1) and finally (B.2). We now proceed to carry out this strategy.

We first verify the following claims.

**CLAIM 1.** *For  $l \leq u$ , we claim that  $\theta_l(t_r^{n+m-l}), \eta_l(t_r^{n+m-l}) \geq 0$  for all  $r = 1, 2, \dots, 2^{n+m-l}$  and  $\theta_l(t_r^{n+m-l})$ 's are equal for  $r \in (\underline{b}(l), \bar{b}(l))$  and we denote their values as  $\theta_l$ . So, following the recursion in (5.5) we have that  $\theta_l = \Delta_{l-1}\theta_1$ . If  $\theta_l(t_{\underline{b}(l)}^{n+m-l}) \neq \theta_l(t_{\underline{b}(l)+1}^{n+m-l})$ , then  $\theta_l(t_{\underline{b}(l)}^{n+m-l}) < \theta_l(t_{\underline{b}(l)+1}^{n+m-l}) = \theta_l$ , and if  $\theta_l(t_{\bar{b}(l)}^{n+m-l}) \neq \theta_l(t_{\bar{b}(l)-1}^{n+m-l})$ , then  $\theta_l(t_{\bar{b}(l)}^{n+m-l}) < \theta_l(t_{\bar{b}(l)-1}^{n+m-l}) = \theta_l$ . Likewise,  $\eta_l(t_r^{n+m-l})$ 's are equal for  $r \in (\underline{b}(l), \bar{b}(l))$ ; we denote their common values as  $\eta_l$  and we have from (5.5) that  $\eta_l = \Delta_{l-1}\eta_1$ . If  $\eta_l(t_{\underline{b}(l)}^{n+m-l}) \neq \eta_l(t_{\underline{b}(l)+1}^{n+m-l})$ , then  $\eta_l(t_{\underline{b}(l)}^{n+m-l}) < \eta_l(t_{\underline{b}(l)+1}^{n+m-l})$ , and if  $\eta_l(t_{\bar{b}(l)}^{n+m-l}) \neq \eta_l(t_{\bar{b}(l)-1}^{n+m-l})$ , then  $\eta_l(t_{\bar{b}(l)}^{n+m-l}) < \eta_l(t_{\bar{b}(l)-1}^{n+m-l})$ . In other words, at each step,  $l$  for  $l < u$ ,  $\theta_l(t_r^{n+m-l})$  and  $\eta_l(t_r^{n+m-l})$  decay at rate  $1/2$  if it is not at the boundary (i.e.  $r \in (\underline{b}(l), \bar{b}(l))$ ), and the boundary ones (i.e.  $\theta_l(t_{\underline{b}(l)}^{n+m-l}), \theta_l(t_{\bar{b}(l)}^{n+m-l})$  and  $\eta_l(t_{\underline{b}(l)}^{n+m-l}), \eta_l(t_{\bar{b}(l)}^{n+m-l})$ ), may decay at a faster rate.*

We now prove the claim by induction using the recursive relation in (5.5). The claim is immediate for  $\theta_1$  and  $\eta_1$ . Now suppose it holds for  $\theta_l(t_r^{n+m-l})$  and  $\eta_l(t_r^{n+m-l})$ ,  $r = 1, 2, \dots, 2^{n+m-l}$ . We next show that the claim holds for  $\theta_{l+1}(t_r^{n+m-l-1})$ ,  $r = 1, 2, \dots, 2^{n+m-l-1}$ , as well. We omit the proof of  $\eta_{l+1}(t_r^{n+m-l-1})$  here, as it follows exactly the same line of analysis as  $\theta_{l+1}(t_r^{n+m-l-1})$ .

We divide the analysis into five cases.

*Case 1.*  $\theta_l(t_{2r-1}^{m+n-l}) = \theta_l(t_{2r}^{m+n-l})$  and  $\eta_l(t_{2r-1}^{m+n-l}) = \eta_l(t_{2r}^{m+n-l})$ . Then  $\theta_+^{l+1}(t_r^{m+n-l}) = 2\theta_l(t_{2r-1}^{m+n-l+1})$  and  $\theta_-^{l+1}(t_r^{m+n-l}) = 0$ . Likewise  $\eta_+^{l+1}(t_r^{m+n-l}) = 2\eta_l(t_{2r-1}^{m+n-l+1})$  and  $\eta_-^{l+1}(t_r^{m+n-l}) = 0$ . From (5.5), we have  $\theta_l(t_r^{m+n-l-1}) = \theta_{l-1}(t_{2r-1}^{m+n-l+1})/2$  and  $\eta_l(t_r^{m+n-l-1}) = \eta_{l-1}(t_{2r-1}^{m+n-l+1})/2$ .



Case 2.  $\theta_l(t_{2r-1}^{m+n-l}) = 0$ ,  $\theta_l(t_{2r}^{m+n-l}) > 0$  and  $\eta_l(t_{2r-1}^{m+n-l}) = 0$ ,  $\eta_l(t_{2r}^{m+n-l}) > 0$ . Then we know that  $2r = \underline{b}(l)$ . We also have  $\theta_+^{l+1}(t_r^{m+n-l-1}) = \theta_l(t_{2r}^{m+n-l})$  and  $\theta_-^{l+1}(t_r^{m+n-l-1}) = -\theta_l(t_{2r}^{m+n-l})$ . Likewise,  $\eta_+^{l+1}(t_r^{m+n-l-1}) = \eta_l(t_{2r}^{m+n-l})$  and  $\eta_-^{l+1}(t_r^{m+n-l-1}) = -\eta_l(t_{2r}^{m+n-l})$ . We rewrite the expression for  $\theta_{l+1}(t_r^{m+n-l-1})$  in (5.5) as

$$\begin{aligned}
 & \theta_{l+1}(t_r^{m+n-l-1}) \\
 &= \theta_+^{l+1}(t_r^{m+n-l-1})\frac{1}{4} + |\theta_-^{l+1}(t_r^{m+n-l-1})| \\
 & \quad \times \left\{ h_{l+1}(t_r^{m+n-l-1})|\eta_-^{l+1}(t_r^{m+n-l-1})| \right. \\
 & \quad + \frac{1}{4}h_{l+1}(t_r^{m+n-l-1})|\theta_-^{l+1}(t_r^{m+n-l-1})|\rho_{l+1}(t_r^{m+n-l-1}) \\
 & \quad \left. + h_{l+1}(t_r^{m+n-l-1})\eta_-^{l+1}(t_r^{m+n-l-1})^2 \frac{\rho_{l+1}(t_r^{m+n-l-1})}{|\theta_-^{l+1}(t_r^{m+n-l-1})|} \right\} \\
 &= \theta_l(t_{2r}^{m+n-l}) \times \left\{ \frac{1}{4} + h_{l+1}(t_r^{m+n-l-1})\eta_l(t_{2r}^{m+n-l}) \right. \\
 & \quad + \frac{1}{4}h_{l+1}(t_r^{m+n-l-1})\theta_l(t_{2r}^{m+n-l})\rho_{l+1}(t_r^{m+n-l-1}) \\
 & \quad \left. + h_{l+1}(t_r^{m+n-l-1})\eta_l(t_{2r}^{m+n-l})^2 \frac{\rho_{l+1}(t_r^{m+n-l-1})}{\theta_l(t_{2r}^{m+n-l})} \right\}.
 \end{aligned}
 \tag{B.3}$$

As

$$\theta_l \Delta_{n+m-l} \leq \theta_1 \Delta_{n+m-1} \leq \frac{1}{4}$$

and

$$\eta_l \Delta_{n+m-l} \leq \eta_1 \Delta_{n+m-1} \leq \frac{1}{48},$$

then

$$\rho_{l+1}(t_r^{m+n-l-1}) \leq \frac{\frac{1}{4}\frac{1}{4}}{1 - \frac{1}{2}\frac{1}{48}} < \frac{1}{15}$$

and

$$h_{l+1}(t_r^{m+n-l-1})\theta_l(t_{2r}^{m+n-l}) \leq \frac{\frac{1}{4}}{(1 - \frac{1}{2}\frac{1}{48})(1 - \frac{1}{15^2})} < \frac{1}{3}.$$

Likewise,

$$h_{l+1}(t_r^{m+n-l-1})\eta_l(t_{2r}^{m+n-l}) < 1/95$$

and

$$\eta_l(t_{2r}^{m+n-l}) \frac{\rho_{l+1}(t_r^{m+n-l-1})}{\theta_l(t_{2r}^{m+n-l})} < 1/95.$$

Plug these in (B.3), we have

$$\frac{1}{4}\theta_l(t_{2r}^{m+n-l}) < \theta_{l+1}(t_r^{m+n-l-1}) < \frac{3}{10}\theta_l \leq \frac{3}{5}\theta_{l+1}.$$

Case 3.  $\theta_l(t_{2r-1}^{m+n-l}) > 0$ ,  $\theta_l(t_{2r}^{m+n-l}) = 0$  and  $\eta_l(t_{2r-1}^{m+n-l}) > 0$ ,  $\eta_l(t_{2r}^{m+n-l}) = 0$ . Then we know that  $2r - 1 = \bar{b}(l)$ . Following the same line of analysis as in Case 2, we have

$$\frac{1}{4}\theta_l(t_{2r}^{m+n-l}) < \theta_{l+1}(t_r^{m+n-l-1}) < \frac{3}{10}\theta_l \leq \frac{3}{5}\theta_{l+1}.$$

Case 4.  $0 < \theta_l(t_{2r-1}^{m+n-l}) < \theta_l(t_{2r}^{m+n-l})$  and  $0 < \eta_l(t_{2r-1}^{m+n-l}) < \theta_l(t_{2r}^{m+n-l})$ . Then we know that  $2r - 1 = \underline{b}(l)$ . There exist  $\xi < 1$ , such that  $\theta_l(t_{2r-1}^{m+n-l}) \leq \xi\theta_l(t_{2r}^{m+n-l}) = \xi\Delta_{l-1}\theta_l$  and  $\eta_l(t_{2r-1}^{m+n-l}) \leq \xi\eta_l(t_{2r}^{m+n-l}) = \xi\Delta_{l-1}\eta_l$ . From (5.5), we have

$$\begin{aligned} &\theta_{l+1}(t_r^{m+n-l-1}) \\ &\leq \theta_+^{l+1}(t_r^{m+n-l-1}) \left\{ \frac{1}{4} + h_{l+1}(t_r^{m+n-l-1}) \eta_-^{l+1}(t_r^{m+n-l-1})^2 \frac{\rho_{l+1}(t_r^{m+n-l-1})}{\theta_+^{l+1}(t_r^{m+n-l-1})} \right\} \\ &\quad + |\theta_-^{l+1}(t_r^{m+n-l-1})| \times \left\{ h_{l+1}(t_r^{m+n-l-1}) |\eta_-^{l+1}(t_r^{m+n-l-1})| \right. \\ &\quad \left. + \frac{1}{4} h_{l+1}(t_r^{m+n-l-1}) |\theta_-^{l+1}(t_r^{m+n-l-1})| \rho_{l+1}(t_r^{m+n-l-1}) \right\}. \end{aligned}$$

As  $|\theta_-^{l+1}(t_r^{m+n-l-1})| \leq \theta_l$  and  $|\eta_-^{l+1}(t_r^{m+n-l-1})| \leq \eta_l$ , following the same calculation as in Case 2, it is easy to check that

$$\theta_{l+1}(t_r^{m+n-l-1}) < \theta_+^{l+1}(t_r^{m+n-l-1}) \left( \frac{1}{4} + 0.01 \right) + |\theta_-^{l+1}(t_r^{m+n-l-1})| \times 0.05.$$

Since  $\theta_{l+1}(t_r^{m+n-l-1}) + |\theta_-^{l+1}(t_r^{m+n-l-1})| = \theta_l$ , we have

$$\begin{aligned} \theta_{l+1}(t_r^{m+n-l-1}) &< \theta_l \left( \left( \frac{1}{4} + 0.01 - 0.05 \right) (1 + \xi) + 0.05 \right) \\ &= \frac{1}{2}\theta_l \left( \frac{1}{2} + 0.02 + 0.42\xi \right) < \frac{\theta_l}{2} = \theta_{l+1}. \end{aligned}$$

Case 5.  $\theta_l(t_{2r-1}^{m+n-l}) > \theta_l(t_{2r}^{m+n-l}) > 0$  and  $\eta_l(t_{2r-1}^{m+n-l}) > \theta_l(t_{2r}^{m+n-l}) > 0$ . Then we know that  $2r = \bar{b}(l)$ . Following the same line of analysis as in Case 4, we have

$$\theta_{l+1}(t_r^{m+n-l-1}) < \theta_{l+1}.$$

We thus prove that the claim holds for  $\theta_{l+1}(t_r^{m+n-l-1})$ ,  $r = 1, 2, \dots, 2^{n+m-l-1}$ , as well.

We have established Claim 1. We now continue with a second claim.

CLAIM 2. For  $u < l < m$ , we have at most two positive  $\theta_l(t_r^{m+n-l})$ 's, namely  $\theta_l(t_{\underline{b}(l)}^{m+n-l})$  and  $\theta_l(t_{\bar{b}(l)}^{m+n-l})$ . Notice that it is possible that  $\underline{b}(l) = \bar{b}(l)$ . We then claim that if  $\underline{b} \neq \bar{b}$ ,  $\theta_l(t_{\underline{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\theta_1 2^{-(l-u-1)/2}$  and  $\theta_l(t_{\bar{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\theta_1 2^{-(l-u-1)/2}$ . Similarly,  $\eta_l(t_{\underline{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\eta_1 2^{-(l-u-1)/2}$  and  $\eta_l(t_{\bar{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\eta_1 2^{-(l-u-1)/2}$ . If  $\underline{b}(l) = \bar{b}(l)$ ,  $\theta_l(t_{\underline{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\theta_1 2^{-(l-u-2)/2}$ ,  $\theta_l(t_{\bar{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\theta_1 2^{-(l-u-2)/2}$  and  $\eta_l(t_{\underline{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\eta_1 2^{-(l-u-2)/2}$ ,  $\eta_l(t_{\bar{b}(l)}^{m+n-l}) \leq \Delta_{l-1}\eta_1 2^{-(l-u-2)/2}$ .

We prove the claim by induction. We shall give the proof of  $\theta_l(t_r^{m+n-l})$  only, as the proof of  $\eta_l(t_r^{m+n-l})$  follows exactly the same line of analysis. For  $l = u$ , we have the following cases:

(i)  $\bar{b}(l) = \underline{b}(l) + 2$ ,  $\underline{b}(l)$  is odd. In this case,  $\theta_{l+1}(t_{(\underline{b}(l)+1)/2}^{m+n-l-1}) < \Delta_l \theta_1$ , which follows from the analysis in Case 4 for  $l \leq u$ . And  $\theta_{l+1}(t_{(\bar{b}(l)+1)/2}^{m+n-l-1}) < (3/5)\Delta_l \theta_1$ , following the analysis in Case 3 for  $l \leq u$ .

(ii)  $\bar{b}(l) = \underline{b}(l) + 2$ ,  $\underline{b}(l)$  is even. In this case,  $\theta_{l+1}(t_{\underline{b}(l)/2}^{m+n-l-1}) < (3/5)\Delta_l \theta_1$ , which follows from the analysis in Case 2 for  $l \leq u$ . And  $\theta_{l+1}(t_{\bar{b}(l)/2}^{m+n-l-1}) < \Delta_l \theta_1$ , following the analysis in Case 5, for  $l \leq u$ .

(iii)  $\bar{b}(l) = \underline{b}(l) + 1$ ,  $\underline{b}(l)$  is odd. In this case, let  $\bar{\theta}_l = \max\{\theta_l(t_{\underline{b}(l)}^{m+n-l}), \theta_l(t_{\bar{b}(l)}^{m+n-l})\}$ . Then following the same analysis as in Case 4 or Case 5 for  $l \leq u$  [depending on which one of  $\theta_l(t_{\underline{b}(l)}^{m+n-l})$  and  $\theta_l(t_{\bar{b}(l)}^{m+n-l})$  is smaller], we have  $\theta_{l+1}(t_{\bar{b}(l)/2}^{m+n-l-1}) < \bar{\theta}_l/2 \leq \Delta_l \theta_1$ .

(iv)  $\bar{b}(l) = \underline{b}(l) + 1$ ,  $\underline{b}(l)$  is even. In this case,  $\theta_{l+1}(t_{\underline{b}(l)/2}^{m+n-l-1}) < (3/5)\Delta_l \theta_1$ , which follows from the analysis in Case 2 for  $l \leq u$ . And  $\theta_{l+1}(t_{(\bar{b}(l)+1)/2}^{m+n-l-1}) < (3/5)\Delta_l \theta_1$ , following the analysis in Case 3 for  $l \leq u$ .

Therefore, the claim holds for  $u + 1$ . Suppose the claim holds for  $l \geq u + 1$ . Then when moving from level  $l$  to level  $l + 1$ , one of the following three cases can happen:

(a)  $\bar{b}(l) = \underline{b}(l) + 1$  and  $\underline{b}(l)$  is even. In this case, following the analysis in Case 2 and Case 3 for  $l \leq u$ , we have

$$\theta_{l+1}(t_{\underline{b}(l)/2}^{m+n-l-1}) \leq \frac{3}{10}\theta_l(t_{\underline{b}(l)}^{m+n-l}) \leq \Delta_l \theta_1 2^{-(l-u)/2}$$

and

$$\theta_{l+1}(t_{(\bar{b}(l)+1)/2}^{m+n-l-1}) \leq \frac{3}{10}\theta_l(t_{\bar{b}(l)}^{m+n-l}) \leq \Delta_l\theta_1 2^{-(l-u)/2}.$$

(b)  $\bar{b}(l) = \underline{b}(l)$ . In this case, following the analysis in Case 2 or Case 3 for  $l \leq u$  [depending on whether  $\underline{b}(l)$  is odd or even], we have

$$\theta_{l+1}(t_{[\underline{b}(l)/2]}^{m+n-l-1}) \leq \frac{3}{10}\theta_l(t_{\underline{b}(l)}^{m+n-l}) \leq \Delta_l\theta_1 2^{-(l-u-1)/2}.$$

(c)  $\bar{b}(l) = \underline{b}(l) + 1$  and  $\underline{b}(l)$  is odd. In this case, we let  $\bar{\theta}_l = \max\{\theta_l(t_{\bar{b}(l)}^{m+n-l}), \theta_l(t_{\underline{b}(l)}^{m+n-l})\}$ . Then we can use the same analysis as in Case 4 or Case 5 for  $l \leq u$  [depending on which one of  $\theta_l(t_{\underline{b}(l)}^{m+n-l})$  and  $\theta_l(t_{\bar{b}(l)}^{m+n-l})$  is smaller] to conclude that

$$\theta_{l+1}(t_{\bar{b}(l)/2}^{m+n-l-1}) < \frac{1}{2}\bar{\theta}_l \leq \Delta_l\theta_1 2^{-(l-u-1)/2}.$$

We notice that case (c) can happen only once.

We are now ready to control the contribution of the term (B.1). As

$$\Delta_{n+m-l+2}\eta_+^l(t_r^{n+m-l}) \leq 1/30 \quad \text{and} \quad \rho_l(t_r^{n+m-l}) < 1/7,$$

we have when  $m \leq u$

$$\begin{aligned} & \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}) \\ & \leq \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} \exp(4\Delta_{n+m-l+2}\eta_+^l(t_r^{n+m-l}) + \rho_l(t_r^{n+m-l})^2) \\ & \leq \prod_{l=2}^m \exp\left(\left(16\Delta_{n+m}\eta_1 + \frac{(4\Delta_{n+m}\theta_1)^2}{(1 - 8\Delta_{n+m}\eta_1)^2}\right)((k' - k)\Delta_l + 2)\right) \\ & \leq \prod_{l=2}^m \exp\left(\left(\frac{11}{5} \frac{\gamma^2}{k' - k} \Delta_n^{1-2\alpha'} + \frac{6}{5} \frac{\gamma^2}{k' - k} \Delta_n^{2-4\alpha'}\right)((k' - k)\Delta_l + 2)\right). \end{aligned}$$

The last inequality follows from Corollary 5.1 that  $\theta_1 = \theta_0/4(1 - \theta_0^2\Delta_{n+m}^2)$ ,  $\eta_1 = \theta_0^2\Delta_{n+m}/2(1 - \theta_0^2\Delta_{n+m}^2)$ , and our choice of  $\theta_0 = \gamma/(k'^{1/2}\Delta_n^{2\alpha'}\Delta_m)$ . Then, as  $(k' - k)^{-1} \leq 2^{-m}$ ,

$$\begin{aligned} & \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}) \\ & \leq \exp\left(\frac{11}{5}\gamma^2\left(\sum_{l=2}^m \Delta_l + 2(m-1)\Delta_m\right) + \frac{6}{5}\gamma^2\left(\sum_{l=2}^u \Delta_l + 2(m-1)\Delta_m\right)\right) \\ & \leq \exp\left(\frac{8}{25}\right). \end{aligned}$$

When  $m > u$ ,

$$\begin{aligned} & \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}) \\ & \leq \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} \exp(4\Delta_{n+m-l+2}\eta_+^l(t_r^{n+m-l}) + \rho_l(t_r^{n+m-l})^2) \\ & \leq \prod_{l=2}^u \exp\left(\left(\frac{11}{5} \frac{\gamma^2}{k' - k} \Delta_n^{1-2\alpha'} + \frac{6}{5} \frac{\gamma^2}{k' - k} \Delta_n^{2-4\alpha'}\right)((k' - k)\Delta_l + 2)\right) \\ & \quad \times \prod_{l=u+1}^m \exp\left(\frac{11}{5} \frac{\gamma^2}{k' - k} \Delta_n^{1-2\alpha'} \Delta_{l-u-2}^{1/2} + \frac{6}{5} \frac{\gamma^2}{k' - k} \Delta_n^{2-4\alpha'} \Delta_{l-u-2}\right). \end{aligned}$$

As  $(k' - k)^{-1} \leq 2^{-u}$ ,

$$\begin{aligned} \prod_{l=2}^m \prod_{r=1}^{2^{n+m-l}} C(t_r^{n+m-l}) & \leq \exp\left\{\frac{11}{5} \gamma^2 \left(\sum_{l=2}^u \Delta_l + 2(u - 1)\Delta_u + \sum_{l=u+1}^m \Delta_{l-2}^{1/2}\right)\right. \\ & \quad \left. + \frac{6}{5} \gamma^2 \left(\sum_{l=2}^u \Delta_l + 2(u - 1)\Delta_u + \sum_{l=u+1}^m \Delta_{l-2}\right)\right\} \\ & \leq \exp\left(\frac{1}{2}\right). \end{aligned}$$

For (B.2), we notice that under condition (5.8) and (5.9), we have

$$\begin{aligned} \left| \sum_{r=1}^{2^n} \theta_m(t_r^n) \Lambda_i(t_r^n) \Lambda_j(t_r^n) \right| & \leq \theta_1 \Delta_{m-1} \varepsilon_0 ((k' - k) \Delta_m)^\beta \Delta_n^{2\alpha'} + 2\theta_1 \Delta_{m-1} \Delta_n^{2\alpha'} \\ & \leq \varepsilon_0 \gamma (k' - k)^{\beta-1/2} + 2\gamma \end{aligned}$$

and

$$\begin{aligned} \left| \sum_{r=1}^{2^n} \eta_m(t_r^n) (\Lambda_i(t_r^n)^2 + \Lambda_j(t_r^n)^2) \right| & \leq ((k' - k) \Delta_m + 2) \eta_1 \Delta_{m-1} 2 \Delta_n^{2\alpha'} \\ & \leq 2\gamma^2. \end{aligned}$$

Combining the analysis for (B.1) and (B.2), we have

$$\begin{aligned} & E_n \exp(\theta_0 \{L_{i,j}^{n+m}(k') - L_{i,j}^{n+m}(k)\}) \\ & \leq \exp\left(\theta_0^2 \Delta_{n+m}^2 (k' - k) + \frac{1}{2} + \varepsilon_0 \gamma (k' - k)^{\beta-1/2} + 2\gamma + 2\gamma^2\right) \\ & \leq 4 \exp(\varepsilon_0 \gamma (k' - k)^{\beta-1/2}). \end{aligned}$$

□

**Acknowledgments.** The authors would like to thank the Associate Editor and two referees for their detailed and insightful comments and suggestions. We also thank Yi Zhu for his help with the implementation of the algorithms.

## REFERENCES

- [1] BAYER, C., FRIZ, P. K., RIEDEL, S. and SCHOENMAKERS, J. (2016). From rough path estimates to multilevel Monte Carlo. *SIAM J. Numer. Anal.* **54** 1449–1483. [MR3501366](#)
- [2] BESKOS, A., PAPASPILIOPOULOS, O. and ROBERTS, G. O. (2006). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli* **12** 1077–1098. [MR2274855](#)
- [3] BESKOS, A., PELUCHETTI, S. and ROBERTS, G. (2012).  $\varepsilon$ -strong simulation of the Brownian path. *Bernoulli* **18** 1223–1248. [MR2995793](#)
- [4] BESKOS, A. and ROBERTS, G. O. (2005). Exact simulation of diffusions. *Ann. Appl. Probab.* **15** 2422–2444. [MR2187299](#)
- [5] BLANCHET, J. and CHEN, X. (2015). Steady-state simulation of reflected Brownian motion and related stochastic networks. *Ann. Appl. Probab.* **25** 3209–3250. [MR3404635](#)
- [6] CHEN, N. and HUANG, Z. (2013). Localization and exact simulation of Brownian motion-driven stochastic differential equations. *Math. Oper. Res.* **38** 591–616. [MR3092549](#)
- [7] DAVIE, A. M. (2007). Differential equations driven by rough paths: An approach via discrete approximation. *Appl. Math. Res. Express. AMRX 2* Art. ID abm009, 40. [MR2387018](#)
- [8] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. [MR2604669](#)
- [9] GAINES, J. G. and LYONS, T. J. (1994). Random generation of stochastic area integrals. *SIAM J. Appl. Math.* **54** 1132–1146. [MR1284705](#)
- [10] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** 215–310. [MR1654527](#)
- [11] POLLOCK, M., JOHANSEN, A. M. and ROBERTS, G. O. (2016). On the exact and  $\varepsilon$ -strong simulation of (jump) diffusions. *Bernoulli* **22** 794–856. [MR3449801](#)
- [12] RHEE, C. and GLYNN, P. (2012). A new approach to unbiased estimation for SDE's. In *Proceedings of the Winter Simulation Conference*. Winter Simulation Conference.
- [13] STEELE, J. M. (2001). *Stochastic Calculus and Financial Applications*. *Applications of Mathematics (New York)* **45**. Springer, New York. [MR1783083](#)

J. BLANCHET  
DEPARTMENT OF INDUSTRIAL ENGINEERING  
AND OPERATIONS  
COLUMBIA UNIVERSITY  
500 WEST 120TH STREET, RM 313  
NEW YORK, NEW YORK 10027  
USA  
E-MAIL: jose.blanchet@columbia.edu

X. CHEN  
ECONOMICS AND MANAGEMENT SCHOOL  
OF WUHAN UNIVERSITY  
LUOJIA HILL, WUHAN  
CHINA  
E-MAIL: xinyun.chen@whu.edu.cn

J. DONG  
DEPARTMENT OF INDUSTRIAL ENGINEERING  
AND MANAGEMENT SCIENCES  
McCORMICK SCHOOL OF ENGINEERING  
AND APPLIED SCIENCES  
NORTHWESTERN UNIVERSITY  
2145 SHERIDAN ROAD, M239  
EVANSTON, ILLINOIS 60208  
USA  
E-MAIL: jing.dong@northwestern.edu